

# **PORADNIK EWALUACJI DLA ADMINISTRACJI PUBLICZNEJ**

**WARSZAWA, 2024**

## Poradnik ewaluacji dla administracji publicznej

Redakcja: Stanisław Bienias, Tomasz Kupiec, Piotr Strzęboszewski

Opracowanie wykonane przez zespół autorów:

Monika Bartosiewicz-Niziołek (Polskie Towarzystwo Ewaluacyjne)  
Stanisław Bienias (Fundacja Idea Rozwoju)  
Tomasz Gapski (Fundacja Idea Rozwoju)  
Szymon Kajma (Ministerstwo Funduszy i Polityki Regionalnej)  
Tomasz Kupiec (Uniwersytet Warszawski, EUROREG)  
Jacek Pokorski (Polskie Towarzystwo Ewaluacyjne/Polska Agencja Rozwoju  
Przedsiębiorczości)  
Paulina Skórska (Uniwersytet Jagielloński/Fundacja Idea Rozwoju)  
Piotr Strzęboszewski (Ministerstwo Funduszy i Polityki Regionalnej)  
Rafał Trzeciński (Fundacja Idea Rozwoju)  
Jakub Wróblewski (IBC Advisory S.A./Polskie Towarzystwo Ewaluacyjne)

Redakcja techniczna: Jakub Jąkałski (Ministerstwo Funduszy i Polityki Regionalnej)

ISBN: 978-83-7610-737-0

Warszawa 2024

Wydawca: Ministerstwo Funduszy i Polityki Regionalnej

Ul. Wspólna 2/4

00-926 Warszawa

[www.mfipr.gov.pl](http://www.mfipr.gov.pl)

[www.ewaluacja.gov.pl](http://www.ewaluacja.gov.pl)

[www.funduszeuropejskie.gov.pl](http://www.funduszeuropejskie.gov.pl)

Departament Strategii

Krajowa Jednostka Ewaluacji

Tel. 22 273 76 00, 22 273 76 01

Fax: 22 273 89 08

Email: [ewaluacja@mfipr.gov.pl](mailto:ewaluacja@mfipr.gov.pl)

Email: [sekretariatDSR@mfipr.gov.pl](mailto:sekretariatDSR@mfipr.gov.pl)

Egzemplarz bezpłatny



Pomoc Techniczna  
dla Funduszy Europejskich



Rzeczpospolita  
Polska

Dofinansowane przez  
Unię Europejską



# Spis treści

|                                                                                                         |           |
|---------------------------------------------------------------------------------------------------------|-----------|
| <b>Wykaz skrótów</b>                                                                                    | <b>5</b>  |
| <b>Wstęp</b>                                                                                            | <b>7</b>  |
| <b>Rozdział 1. Czym jest ewaluacja?</b>                                                                 | <b>9</b>  |
| Jak definiujemy ewaluację?                                                                              | 9         |
| Jakie funkcje pełni ewaluacja i w jakim celu się ją przeprowadza?                                       | 12        |
| Jakie wyróżniamy rodzaje ewaluacji?                                                                     | 13        |
| Czym są kryteria ewaluacyjne?                                                                           | 15        |
| Ewaluacja w szerszym kontekście polityk opartych na dowodach                                            | 17        |
| Jakie są uwarunkowania prawne ewaluacji w Polsce?                                                       | 18        |
| <b>Rozdział 2. Jaka jest rola ewaluacji we wdrażaniu interwencji publicznej?</b>                        | <b>21</b> |
| Teoria interwencji oraz jej weryfikacja w ramach ewaluacji                                              | 21        |
| Ewaluacja w cyklu wdrażania interwencji publicznej                                                      | 30        |
| Ewaluacja ex ante                                                                                       | 31        |
| Ewaluacja on-going                                                                                      | 34        |
| Ewaluacja ex post                                                                                       | 35        |
| <b>Rozdział 3. Co jest potrzebne, aby przeprowadzić ewaluację i w jaki sposób możemy ją realizować?</b> | <b>37</b> |
| Wprowadzenie                                                                                            | 37        |
| Wybór ekonomiczny w procesie planowania ewaluacji                                                       | 37        |
| Przegląd zasobów a wybór modelu realizacji ewaluacji                                                    | 39        |
| Kompetencje (wiedza, umiejętności, <i>know-how</i> )                                                    | 42        |
| Źródła wiedzy (dane, informatorzy)                                                                      | 44        |
| Czas (harmonogram prac badawczych, termin wdrożenia wyników)                                            | 47        |
| Zaplecze organizacyjno-techniczne (infrastruktura IT, licencje, technologia)                            | 50        |
| <b>Rozdział 4. Planowanie ewaluacji</b>                                                                 | <b>52</b> |
| Krok 1: Zrozumienie potrzeb i podjęcie decyzji o realizacji ewaluacji                                   | 52        |
| Krok 2: Definiujemy cel/e badania, stawiamy pytania badawcze<br>oraz określamy zakres ewaluacji         | 53        |
| Krok 3: Określamy sposób realizacji badania                                                             | 55        |
| Krok 4: Określamy zasoby niezbędne do realizacji badania                                                | 56        |
| Krok 5: Ustalamy harmonogram badania                                                                    | 61        |

|                                                                                                                                 |            |
|---------------------------------------------------------------------------------------------------------------------------------|------------|
| Podsumowanie                                                                                                                    | 62         |
| <b>Rozdział 5. Realizacja badania ewaluacyjnego</b>                                                                             | <b>63</b>  |
| Krok 6: Przygotowujemy i przeprowadzamy zamówienie na ewaluację                                                                 | 63         |
| Krok 7a: Współpracujemy z wykonawcą w trakcie realizacji badania                                                                | 71         |
| Krok 7b: Realizacja badania ewaluacyjnego siłami wewnętrznymi                                                                   | 75         |
| <b>Rozdział 6. Produkty badania oraz wykorzystanie wyników ewaluacji</b>                                                        | <b>80</b>  |
| Krok 8: Opracowujemy raport końcowy z badania                                                                                   | 80         |
| Krok 9: Rozpowszechniamy wyniki badania                                                                                         | 87         |
| Krok 10: Wykorzystujemy wnioski i rekomendacje z ewaluacji                                                                      | 92         |
| <b>Rozdział 7. Metodyka badań</b>                                                                                               | <b>96</b>  |
| Słowo wstępu                                                                                                                    | 96         |
| Najpopularniejsze metody i techniki badań wykorzystywanych w ewaluacji                                                          | 97         |
| Podejścia analityczne, które wzbogacą badanie ewaluacyjne                                                                       | 108        |
| Wybór i realizacja adekwatnej metodyki: kryteria i czynniki kontekstowe                                                         | 113        |
| <b>Załącznik 1. Jakimi ogólnymi zasadami należy się kierować przy realizacji ewaluacji? Słów kilka o standardach ewaluacji.</b> | <b>117</b> |
| Wprowadzenie                                                                                                                    | 117        |
| Jak Standardy Ewaluacji powinny wspierać proces ewaluacji?                                                                      | 119        |
| <b>Załącznik 2. Przykładowy Opis Przedmiotu Zamówienia (OPZ) dla ewaluacji ex ante strategii lokalnej</b>                       | <b>126</b> |
| <b>Bibliografia</b>                                                                                                             | <b>131</b> |
| <b>Spis tabel</b>                                                                                                               | <b>137</b> |
| <b>Spis ilustracji</b>                                                                                                          | <b>137</b> |

## Wykaz skrótów

|      |                                                                                                                |
|------|----------------------------------------------------------------------------------------------------------------|
| CAPI | Computer Assisted Personal Interview – Wywiad Osobisty Wspomagany Komputerowo                                  |
| CATI | Computer Assisted Telephone Interview – Wywiad Telefoniczny Wspomagany Komputerowo                             |
| CAWI | Computer Assisted Web Interview – Wywiad Internetowy Wspomagany Komputerowo                                    |
| CBA  | Cost-Benefit Analysis – Analiza kosztów i korzyści                                                             |
| CIE  | Counterfactual Impact Evaluation – Kontrfaktyczna Ewaluacja Wpływu                                             |
| DID  | Difference – in – difference – metoda podwójnej różnicy                                                        |
| DT   | design thinking                                                                                                |
| FGI  | Focus Group Interview – Zogniskowany Wywiad Grupowy                                                            |
| GUS  | Główny Urząd Statystyczny                                                                                      |
| IDI  | Individual Indepth Interview – Indywidualny Wywiad Pogłębiony                                                  |
| KJE  | Krajowa Jednostka Ewaluacji                                                                                    |
| KJO  | Krajowa Jednostka Oceny. Nazwa KJE używana przed 2014 rokiem.                                                  |
| OECD | <i>Organisation for Economic Co-operation and Development</i> – Organizacja Współpracy Gospodarczej i Rozwoju, |
| OOS  | Ocena Oddziaływania na Środowisko                                                                              |
| OPZ  | Opis Przedmiotu Zamówienia                                                                                     |
| OSR  | Ocena Skutków Regulacji                                                                                        |
| PAPI | Paper and Pencil Interview – Tradycyjny wywiad osobisty w formie papierowej                                    |
| PS   | Propensity Score – indeks skłonności                                                                           |
| PSM  | Propensity Score Matching – dopasowanie indeksów skłonności                                                    |
| PTE  | Polskie Towarzystwo Ewaluacyjne                                                                                |
| RK   | raport końcowy                                                                                                 |
| RM   | raport metodyczny                                                                                              |
| SWZ  | Specyfikacja Warunków Zamówienia                                                                               |
| TBE  | Theory Based Evaluation – Ewaluacja oparta na teorii                                                           |
| TBIE | Theory Based Impact Evaluation – Ewaluacja wpływu oparta na teorii                                             |

|      |                                                                                                                                                            |
|------|------------------------------------------------------------------------------------------------------------------------------------------------------------|
| TDI  | Telephone Indepth Interview – Telefoniczny Wywiad Pogłębiony                                                                                               |
| UE   | Unia Europejska                                                                                                                                            |
| UNDP | United Nations Development Programme – Program Narodów Zjednoczonych ds. Rozwoju – agenda ONZ powstała w 1965, która tworzy globalną sieć na rzecz rozwoju |
| ZUS  | Zakład Ubezpieczeń Społecznych                                                                                                                             |

# Wstęp

Publikacja, którą oddajemy w Państwa ręce jest praktycznym przewodnikiem po procesie ewaluacji dla administracji publicznej. Opracowanie to jest w dużej mierze aktualizacją poradnika, który został opracowany w ramach Krajowej Jednostki Oceny (obecnie Krajowej Jednostki Ewaluacji) w 2012r. Choć fundamentalne zasady i istota ewaluacji nie zmieniły się w ciągu ostatnich 10 lat, światowa i szczególnie polska praktyka ewaluacji rozwijała się w tym okresie niezwykle dynamicznie. Nowe podejścia, metody, wyzwania, zmieniające się postrzeganie tego, po co ewaluacja jest realizowana i jak ją wykorzystywać w kontekście zarządzania w sektorze publicznym, to tylko niektóre elementy, które uzasadniały aktualizację. Swoistym novum prezentowanego poradnika jest również szeroka gama uzupełniających się perspektyw. Wśród autorów poszczególnych rozdziałów znajdują się przedstawiciele organizacji zamawiających i wykorzystujących ewaluację, wykonawców, jak i akademików zajmujących się ewaluacją naukowo.

Poradnik został opracowany z myślą o osobach pracujących w sektorze publicznym, którzy zajmują się lub planują zająć się tematem ewaluacji. Celem publikacji jest wesprzeć te osoby w zdobyciu podstawowych umiejętności i narzędzi potrzebnych do wykonania pierwszej ewaluacji. Naszą ambicją było zebranie w jednym miejscu podstawowych informacji oraz praktycznych porad dotyczących realizacji badań, opracowanych w przyjaznej i przejrzystej formie.

Poradnik zorganizowano w siedem rozdziałów. Pierwszy z nich przedstawia definicje, rodzaje i funkcje ewaluacji. W rozdziale drugim omówiona została rola ewaluacji w cyklu interwencji publicznej. Rozdział trzeci jest refleksją na temat zasobów niezbędnych do prawidłowej realizacji badań. Kolejne rozdziały, aż do szóstego włącznie, zawierają swoisty przepis jak zrealizować ewaluację. Został on zorganizowany w 10 chronologicznie ułożonych kroków – od identyfikacji tematu badania ewaluacyjnego do wykorzystania wniosków i rekomendacji z ewaluacji. W opisie tym staraliśmy się uwzględnić specyfikę ewaluacji zewnętrznych, tj. powierzanych do wykonania zespołom działającym poza administracją publiczną, jak również specyfikę badań realizowanych własnymi siłami. Ostatni rozdział zawiera syntetyczne opisy metod i technik najczęściej używanych w badaniach ewaluacyjnych. W materiałach dodatkowych przekazujemy opis zasad realizacji ewaluacji wynikających ze standardów Polskiego Towarzystwa Ewaluacji oraz podstawowy szablon opisu badania ewaluacyjnego, który może stanowić materiał wyjściowy do przygotowania opisu przedmiotu zamówienia ewaluacji ex ante strategii lokalnej.

Ustawa o zasadach prowadzenia polityki rozwoju z dnia 6.12.2006 r. nakłada na jednostki administracji publicznej przygotowujące dokumenty strategiczne obowiązek zrealizowania ewaluacji ex ante. Mamy nadzieję, że ta publikacja będzie szczególnie użyteczna dla grupy urzędników, którzy są odpowiedzialni za realizację tego obowiązku. Niemniej, biorąc pod uwagę szeroki zakres niniejszej publikacji wierzymy, że będzie ona przydatna również dla urzędników zajmujących się ewaluacjami innych interwencji publicznych, jak np. badaniami projektów współfinansowanych ze środków europejskich w ramach polityki spójności.



# Rozdział 1. Czym jest ewaluacja?

Tomasz Kupiec

## Jak definiujemy ewaluację?

Słowo *ewaluacja* jest zapożyczeniem z angielskiego *evaluation* lub francuskiego *évaluation*, te z kolei pochodzą od łacińskiego *valeo -ere* oznaczającego *mieć się dobrze, być coś wartym* (Mark, Greene i Shaw, 2006, s. 6). W słowniku języka polskiego PWN termin *ewaluacja* jest aktualnie wyjaśniany jako *określenie wartości czegoś* i definicja ta jest bardzo bliska tym, które możemy znaleźć w uznanych podręcznikach. Przykładowo Rossi i in. (2004, s. 2) proponują przez ewaluację rozumieć *ustalenie wartości lub określenie jakości danego obiektu*.

W powszechnym użyciu są również bardziej rozbudowane definicje precyzyjniej charakteryzujące ewaluację, jako odrębną dyscyplinę i profesję. W pierwszym polskim podręczniku akademickim ewaluacji interwencji publicznych (Olejniczak 2008, s. 22) wskazano, że:



**Ewaluacja to systematyczne badanie prowadzone z użyciem zróżnicowanych metod, złożone ze zbierania danych, analizy, oceny oraz informowania o wynikach. Jego celem jest oszacowanie (w odniesieniu do jasno sformułowanych kryteriów) jakości i wartości procesu oraz efektów wdrażania interwencji publicznych.**

Bardzo podobną, choć nieco okrojoną definicję można znaleźć w dokumencie opracowanym w ramach administracji publicznej zajmującej się środkami europejskimi, tj. w wytycznych dotyczących ewaluacji polityki spójności (MFiPR, 2022). W obu przypadkach dobrze ujęte zostały **wyróżniające cechy ewaluacji**. Pierwszą z nich jest **analityczna, systematyczna** natura badania ewaluacyjnego. Wymaga ona naukowego podejścia popartego danymi, opartego o rygorystyczne stosowanie metod zapożyczanych zwykle z nauk społecznych. Po drugie, ewaluacja jest **normatywna, oceniająca**, co odróżnia ją od innych form badań. Jej istotą nie jest (a przynajmniej nie tylko) zebranie, podsumowanie danych i opis sytuacji, ale ocena, określenie, oszacowanie, takich charakterystyk ewaluowanego przedmiotu jak wartość, jakość, znaczenie, ważność, użyteczność. Jak ujął to Scriven (1991), ewaluacja ma dwa ramiona, z których jedno służy gromadzeniu danych, a drugie weryfikuje osiągniętą wartość, jakość względem przyjętych celów, standardów, kryteriów. Po trzecie, ewaluację wyróżnia jej **praktyczna użyteczność**. Zadaniem

ewaluacji jest wspomagać konkretnych odbiorców w konkretnej sytuacji (Stufflebeam 2001), usprawnić, zapewnić rozwój bądź lepsze zrozumienie konkretnej interwencji (Korporowicz, 1997). Ewaluacja powinna wspierać podejmowanie decyzji, dostarczając informacje niezbędne do ich podjęcia (Patton 1997). Jako podstawowy i ostateczny cel ewaluacji często wskazywana jest ogólna poprawa sytuacji społecznej (Henry, Mark 2003). Jak ujęli to Stufflebeam i Shinkfield (2012) najważniejszym celem ewaluacji jest nie udowodnienie, a udoskonalenie.

Praktyczny, utylitarny charakter ewaluacji jest wyraźnie podkreślany w regulacjach i wytycznych kluczowych organizacji międzynarodowych korzystających z ewaluacji jako narzędzia efektywnego zarządzania. Z ich perspektywy ewaluacja ma na celu:

- poprawę jakości, skuteczności i spójności pomocy (UE),
- powinna zapewniać wykorzystanie dotychczasowych doświadczeń w procesie podejmowania (OECD),
- upowszechniać i umożliwiać wykorzystanie nauki płynącej z dotychczasowych działań (Bank Światowy),
- ułatwić podejmowanie decyzji opartych na wiedzy (UNDP).

Szereg autorów wyjaśniając czym jest ewaluacja zwraca uwagę na to, czym różni się ona od innych, podobnych procesów (najczęściej monitoringu, kontroli i audytu). I tak, utylitarny charakter ewaluacji sprawia, że decyzja o jej realizacji zapada często w instytucji realizującej interwencję i jest odpowiedzią na potrzeby własne bądź innych interesariuszy interwencji. Instytucja ta ma również decydujący wpływ na zakres ewaluacji. Tymczasem kontrola jest często niezależna od woli instytucji wdrażającej interwencję, zarówno jeżeli chodzi o fakt realizacji, jak i zakres czy metody. W odróżnieniu od monitoringu, który prowadzony jest w sposób ciągły i służy „rejestrowaniu” bezpośrednich efektów – produktów i rezultatów<sup>1</sup> – interwencji, ewaluacja to zwykle projekt badawczy realizowany w określonym momencie i skupiający się na szerszym oddziaływaniu interwencji i mechanizmach przez nią uruchamianych. Więcej różnic prezentuje poniższa tabela.

Równocześnie, mimo niewątpliwych różnic między tymi procesami, warto mieć na uwadze, że wskazanie jednoznacznych granic między ewaluacją, monitoringiem i kontrolą nie jest możliwe, bowiem wzajemnie się one przenikają i uzupełniają. Przede wszystkim dane monitoringowe są podstawą, bez której realizacja wielu ewaluacji nie byłaby możliwa. Z drugiej strony, ewaluacja – przede wszystkim on-going – uzupełnia monitoring, identyfikując przyczyny zjawisk sygnalizowanych wskaźnikami. Podobnie kontrola czy audyt, służące rozliczaniu działania z perspektywy legalności, uzupełniane są ewaluacją,

---

<sup>1</sup> Przez produkty rozumiemy tu wszystkie policzalne dobra i usługi, które powstaną w ramach naszej interwencji. Natomiast pod pojęciem rezultatu rozumiemy wszystkie natychmiastowe i bezpośrednie korzyści, jakie beneficjenci odnieśli z uczestniczenia w interwencji.

przede wszystkim ex post, której jedną z funkcji jest rozliczanie zarządzających interwencją w wymiarze merytorycznym czy politycznym.

**Tabela 1.** Zestawienie różnic między ewaluacją, monitoringiem i kontrolą/audytem

|                            | EWALUACJA                                                                                                                                                                       | MONITORING                                                                                                | KONTROLA/AUDYT                                                                                                                                        |
|----------------------------|---------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|-----------------------------------------------------------------------------------------------------------|-------------------------------------------------------------------------------------------------------------------------------------------------------|
| <b>CEL</b>                 | udoskonalenie bieżącej lub przyszłych interwencji, zweryfikowanie ich zasadności                                                                                                | zweryfikowanie, czy interwencja postępuje zgodnie z planem, celem ewentualnego dokonania bieżących korekt | zapewnienie zgodności prowadzonych działań z obowiązującymi procedurami i prawem                                                                      |
| <b>ZAKRES I PRZEDMIOT</b>  | szeroki: zamierzone i niezamierzone efekty – oddziaływania interwencji efektywność, użyteczność interwencji w określonym kontekście społeczno-ekonomicznym                      | wąski: oczekiwane, bezpośrednie produkty i rezultaty interwencji                                          | wąski: głównie procedury, dokumentacja, w tym finansowo-księgowa, dowody wytworzenia produktów/usług                                                  |
| <b>PODMIOT REALIZUJĄCY</b> | eksperti zewnętrzni, naukowcy, konsultanci (ewaluacja zewnętrzna), lub z wewnętrznej jednostki ewaluacyjnej – zwykle osoby niezaangażowane w interwencję (ewaluacja wewnętrzna) | pracownicy organizacji, często zaangażowani we wdrażanie interwencji                                      | instytucje posiadające kompetencje kontrolne (kontrola zewnętrzna), pracownicy z organizacji z niezależnych komórek kontrolnych (kontrola wewnętrzna) |
| <b>MOMENT REALIZACJI</b>   | punktowy: przed, w trakcie i po zakończeniu interwencji (ewaluacje ex ante, on-going i ex post)                                                                                 | ciągły: przez cały czas wdrażania interwencji (czasami również po zakończeniu interwencji)                | punktowy: przed (kontrola dokumentów, zaplanowanych procedur), w trakcie i głównie po zakończeniu interwencji                                         |
| <b>SPOSÓB REALIZACJI</b>   | multidyscyplinarne, twórcze, pragmatyczne podejścia łączące szereg metod, zwykle z zakresu nauk społecznych                                                                     | gromadzenie i prosta analiza danych w oparciu o przyjęte procedury administracyjne                        | analiza dokumentacji i stanów faktycznych zgodnie z obowiązującymi procedurami                                                                        |

Źródło: Opracowanie własne na podstawie Kierzkowski (2002) i Olejniczak (2007)

## Jakie funkcje pełni ewaluacja i w jakim celu się ją przeprowadza?

Jak wskazano wyżej, istotą ewaluacji jest identyfikacja i ocena osiągniętych lub spodziewanych efektów interwencji, zaś jej ostatecznym celem zwiększanie dobrostanu

społecznego. Ten ogólny cel realizuje się przez szereg innych, bardziej szczegółowych, często określanych funkcjami ewaluacji.

Powszechnie wymieniane funkcje ewaluacji w obejmującym Polskę kontekście Unii Europejskiej to (np. Batterbury, 2006; Olejniczak, 2008; EVALSED, 2013):

- racjonalizacja planowania, poprawa efektywności – zapewnianie, że interwencja jest uzasadniona, a środki są efektywnie alokowane;
- rozliczanie i wzmacnianie odpowiedzialności – wskazywanie, w jakim stopniu interwencja osiągnęła swoje cele, jak dobrze zostały wykorzystane środki i jakie było oddziaływanie interwencji;
- poprawa wdrażania i bieżącej jakości – poprawianie wydajności interwencji, skuteczności wdrażania i zarządzania;
- poszerzanie wiedzy, poznanie – budowanie ogólnych zasad funkcjonowania i identyfikacja mechanizmów osiągania efektów przez interwencje publiczne;
- wzmocnienie instytucjonalne – poprawienie i rozwój potencjału uczestników interwencji, ich sieci i instytucji;
- budowanie partnerstwa – zwiększanie poczucia współwłasności interwencji wśród jej interesariuszy, włączanie grup dotychczas marginalizowanych we współdecydowanie o interwencji.

Innym ciekawym spojrzeniem jest wyróżnienie dwóch do pewnego stopnia rozbieżnych orientacji ewaluacji, zgodnie z którymi może ona służyć:

- rozliczaniu się przed społeczeństwem, interesariuszami, instytucjami nadrzędnymi ze skuteczności, efektywności działania (orientacja na zewnątrz);
- uczeniu się przez podmioty zaangażowane w zarządzanie interwencją, modyfikacje interwencji w oparciu o wnioski z ewaluacji (orientacja do wewnątrz).

Świadomość roli, jaką pełni ewaluacja w naszej organizacji oraz funkcji, jakie realizować ma konkretne badanie ewaluacyjne jest ważna dla jakości i użyteczności ewaluacji. Po pierwsze, chociaż jedno badanie ewaluacyjne może z powodzeniem realizować więcej, niż jedną funkcję, to nie dotyczy to wzmacniania odpowiedzialności, stojącego niejako w opozycji do pozostałych funkcji. Wynika to z odmiennej w jego przypadku relacji ewaluatorów i ewaluowanych, tzn. oceniających i ocenianych zamiast pomagających i wspomaganych (Olejniczak, 2008). Po drugie, różne funkcje ewaluacji związane są z różnymi pytaniami badawczymi oraz są istotne dla różnych odbiorców, np. poprawa wdrażania interesuje głównie osoby bezpośrednio zarządzające interwencją, a rozliczanie efektów decydentów, polityków czy też media.

## Jakie wyróżniamy rodzaje ewaluacji?

Przedstawione na wstępie definicje i próby delimitacji ewaluacji względem monitoringu czy kontroli mogły zbudować wrażenie, że jest to dziedzina wewnętrznie spójna i jednolita. Tymczasem, choć pewne fundamentalne cechy są wspólne dla wszystkich ewaluacji, pod wieloma innymi względami badania ewaluacyjne mogą się znacznie różnić, co oddaje kilka przedstawionych poniżej typologii.

Jednym z podstawowych podziałów jest ten osadzający ewaluację w cyklu życia interwencji i wyróżniający ewaluacje realizowane przed (**ex ante**), w trakcie (**on-going**), w połowie okresu programowania (**mid-term**) i po zakończeniu (**ex post**) realizacji interwencji. Zagadnieniu temu szerzej poświęcony jest rozdział 2 poradnika.

Tradycyjnie, ewaluacje dzielono na **formatywne** – wspierające proces udoskonalania ewaluowanej interwencji i kierowane zwykle do wąskiego grona osób zarządzających interwencją oraz **podsumowujące** – oceniające, analizujące skutki i oddziaływania interwencji, kierowane do szerszego grona odbiorców (Scriven, 1991). Podział ten jest stosunkowo bliski klasyfikacji proponowanej swego czasu w regulacjach UE (Council Regulation (CE) No 1083/2006) wyróżniającej ewaluacje **operacyjne** – wspierające proces monitoringu i wdrażania interwencji oraz **strategiczne** – oceniające interwencję w kontekście założonych celów. Trzecią uzupełniającą dychotomią jest podział na ewaluacje **procesów** i **efektów**. Te pierwsze służą analizie organizacyjnego kontekstu wdrażania interwencji, i sprawdzaniu, czy jest ona wdrażana w sposób zapewniający osiągnięcie celów. Te drugie służą oszacowaniu efektów i niekiedy ich ocenie przez pryzmat określonych kryteriów.

Charakterystyczne dla organizacji zarządzających dużymi, wielopoziomowymi interwencjami (np. OECD, UNDP, UE) jest rozróżnianie ewaluacji ze względu na przedmiot. Podstawowym przedmiotem w tych klasyfikacjach jest najczęściej **projekt**. Następnie wskazywane są **programy** (lub ich elementy – priorytety, działania), **polityki**, **strategie**. Osobną kategorię stanowi ewaluacja **horyzontalna/przekrojowa**, obejmująca zwykle grupę programów, które łączy wspólny temat – wspierana dziedzina, zagadnienie z zakresu zarządzania, uniwersalna zasada (np. równości szans, poszanowania środowiska). W dalszej części poradnika, odnosząc się do przedmiotu ewaluacji, będziemy posługiwać się często ogólnym pojęciem **interwencji publicznej**, które obejmuje wszystkie wskazane wyżej przedmioty.

Jak zasugerowano, klasyfikowanie ewaluacji ze względu na przedmiot jest osadzone w kontekście funkcjonowania danej organizacji lub sieci organizacji. Zatem w organizacjach nie funkcjonujących w oparciu o projekty czy programy, przedmiotem ewaluacji

mogą być po prostu inaczej nazywane działania czy aktywności. Ponadto przedmiotem ewaluacji może być też sama organizacja, np. jej struktura czy procedury.

Jeszcze jednym specyficznym ze względu na przedmiot, rodzajem ewaluacji jest **meta-ewaluacja**. Jej przedmiotem są inne ewaluacje – pojedyncze badania, albo grupy badań (Cooksy, Caracelli, 2005). Celem metaewaluacji może być identyfikacja słabości i mocnych stron w praktyce ewaluacji, lub częściej ocena jakości poszczególnych badań na potrzeby decyzji o ich uwzględnieniu bądź nie w **syntezie ewaluacji**. Syntezy te łączą informacje z więcej niż jednego badania, by wypracować generalne wnioski dotyczące określonej interwencji czy typu interwencji (Cooper, Hedges, 1994)<sup>2</sup>. Syntezy są niezwykle użytecznym sposobem generowania nowych i wiarygodnych wniosków w obszarach, w których wcześniej zrealizowano już sporo innych ewaluacji.

Bardzo istotną z praktycznego punktu widzenia jest też klasyfikacja ewaluacji ze względu na administracyjną/organizacyjną stronę badania. Z tej perspektywy wyróżniamy przede wszystkim ewaluacje **wewnętrzne** – przeprowadzane przez pracowników w ramach organizacji zarządzającej analizowaną interwencją oraz **zewnętrzne** – zlecane przez organizację zarządzającą badaną interwencją zewnętrznemu wykonawcy. Specyfika obu podejść jest odmienna i wiąże się z innymi wyzwaniami (w dalszej części poradnika wskazujemy najważniejsze różnice, zob. np. rozdział 5, kroki 7a i 7b). Warto dodać, że ewaluacja zewnętrzna nie spełnia jeszcze kryterium ewaluacji w pełni **niezależnej**, bowiem o takiej możemy mówić, gdy również podmiot zlecający i finansujący badanie jest odrębny od podmiotu ocenianego, czyli zarządzającego interwencją. Z taką sytuacją mamy miejsce np. gdy ewaluację działań rządu zleca komisja parlamentarna lub Najwyższa Izba Kontroli (Picciotto, 2016). Jest ona szczególnie zasadna, kiedy ewaluacja ma służyć rozliczaniu i wzmacnianiu odpowiedzialności, a mniej konieczna (a wręcz może przeszkadzać), gdy celem jest np. wsparcie w procesie wdrażania interwencji<sup>3</sup>.

<sup>2</sup> Przykładowa synteza ewaluacji z elementem meta-ewaluacji w obszarze polityki spójności w Polsce [https://www.poir.gov.pl/media/99950/RC\\_Metaewaluacja\\_08092020\\_wysylka\\_WCAG.pdf](https://www.poir.gov.pl/media/99950/RC_Metaewaluacja_08092020_wysylka_WCAG.pdf)

<sup>3</sup> W przypadku braku faktycznej niezależności warto zadbać przynajmniej o tzw. niezależność funkcjonalną, tak jak rozumiana jest ona w wytycznych dotyczących ewaluacji polityki spójności (MFIPR, 2022), tzn. odrębność faktycznego wykonawcy ewaluacji od osób odpowiedzialnych za przedmiot ewaluacji, przynajmniej na poziomie komórek organizacyjnych.

## Czym są kryteria ewaluacyjne?

Kryterium w najprostszym ujęciu jest pryzmatem (aspektem), przez który oceniamy badaną interwencję. Kryteria najczęściej są prezentowane w całościowych zestawach i w takiej formie mogą być traktowane jako uzupełniające się perspektywy, dające całościowy obraz interwencji. Ustalony i wystandardyzowany zestaw kryteriów ewaluacyjnych pełni szereg istotnych zadań w procesie ewaluacji. Kryteria:

- w zrozumiały sposób wyrażają, co ma podlegać ocenie,
- operacjonalizują istotne dla zlecających/prowadzących ewaluację wartości,
- priorytetyzują to co ważne,
- ukierunkowują stawiane pytania i proces zbierania danych,
- orientują prezentację wniosków.

Posiadanie wspólnie zdefiniowanych kryteriów dla większej grupy organizacji ułatwia porównywanie różnych interwencji i wzajemne uczenie się.

Najbardziej rozpowszechnionym zestawem kryteriów ewaluacyjnych w Polsce jest ten przyjęty na potrzeby początkowo polityki regionalnej, a następnie polityki spójności UE. Jest on prawie tożsamy z przyjętym w 1999 zestawem OECD/DAC (tabela 2).

**Tabela 2.** Kryteria ewaluacyjne

| Kryterium                | Opis                                                                                                                                             |
|--------------------------|--------------------------------------------------------------------------------------------------------------------------------------------------|
| <b>Trafność*</b>         | adekwatność planowanych działań do problemów i wyzwań, potrzeb w otoczeniu                                                                       |
| <b>Skuteczność</b>       | stopień realizacji zakładanych celów, skuteczność przyjętych rozwiązań, instytucji                                                               |
| <b>Efektywność</b>       | relacja między nakładami, kosztami, zasobami (np. finansowymi, ludzkimi) a osiągniętymi efektami                                                 |
| <b>Użyteczność*</b>      | adekwatność osiągniętych efektów do problemów i wyzwań, potrzeb w otoczeniu                                                                      |
| <b>Wpływ<sup>4</sup></b> | skala efektów zamierzonych i niezamierzonych, pozytywnych i negatywnych, przede wszystkim wykraczających poza bezpośrednie rezultaty interwencji |
| <b>Trwałość</b>          | ciągłość spodziewanych efektów po ustaniu interwencji w perspektywie średnio- i długookresowej                                                   |

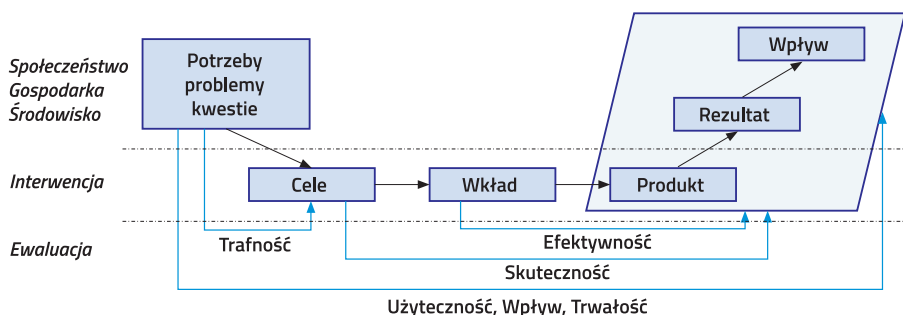
\* adekwatność założeń interwencji, jej celów, podejmowanych działań zmienia się w czasie wraz ze zmianami sytuacji w otoczeniu, stąd konieczność jej weryfikacji na etapie planowania (trafność) i po zakończeniu interwencji (użyteczność).

Źródło: Opracowanie własne na podstawie EVALSED (2013) i Olejniczak (2008)

<sup>4</sup> Należy zwrócić uwagę, że termin „wpływ” ma w praktyce ewaluacji jeszcze drugie znaczenie. Jest rozumiany jako ta część obserwowanej zmiany społeczno-gospodarczej, którą możemy przypisać działaniu badanej interwencji publicznej, a nie działaniu innych czynników.

By zrozumieć istotę i różnice między sześcioma podstawowymi kryteriami ewaluacyjnymi, warto zobrazować je na tle schematu logiki interwencji (Rysunek 1). W ten sposób łatwiej zapamiętać co i z perspektywy czego oceniają poszczególne kryteria. Przykładowo, zarówno efektywność, jak i skuteczność oceniają efekty, ale efektywność czyni to z perspektywy poniesionych nakładów, a skuteczność przyjętych celów. Trafność i użyteczność oceniają interwencję z tej samej perspektywy potrzeb i wyzwań w otoczeniu, ale trafność weryfikuje założenia, cele, plany, zaś użyteczność osiągnięte efekty.

**Rysunek 1. Kryteria ewaluacyjne na tle logiki interwencji**



Źródło: Opracowanie własne na podstawie DG Regio (2006)

Oczywiście przytoczona lista kryteriów nie jest wyczerpującą i jedyną możliwą, a skoro kryteria odzwierciedlają wartości, to razem z nimi się zmieniają, sygnalizując aktualne priorytety. W polu polityki spójności w perspektywie 2021-27 obok już wymienionych trafności, skuteczności i efektywności pojawiają się ważne kryteria spójności i unijnej wartości dodanej (rozumianej jako ta część efektów interwencji UE, która nie została osiągnięta w przypadku interwencji jedynie krajowej), a dodatkowo proponuje się kryteria włączenia społecznego, niedyskryminacji i widoczności (Council Regulation (CE) No 2021/1060, art. 44). W Europejskim Towarzystwie Ewaluacyjnym toczono natomiast debatę o zasadności wykorzystania jakości życia, jako kryterium ewaluacji (Potluka, 2023). Te przykłady zostały przytoczone by pokazać, że choć przyjęte przez uznane organizacje zestawy kryteriów ewaluacyjnych są dobrym wstępem do rozważań, to decyzja o faktycznie stosowanych kryteriach w danej ewaluacji zależy od osób projektujących ją i powinna odpowiadać specyfice danej interwencji i potrzebom jej interesariuszy. Oznacza to oczywiście również potrzebę świadomej selekcji i **ograniczenia** się do **kryteriów faktycznie niezbędnych w danym badaniu**.



## Ewaluacja w szerszym kontekście polityk opartych na dowodach

Praktyka ewaluacji rozwijana jest w administracji w ramach szerzej zakrojonych reform, mających sprawić, by decyzje dotyczące polityk publicznych opierały się na wiarygodnych przesłankach o skuteczności lub nieskuteczności określonych działań. Ten pożądany stan określany jest często terminem **polityk opartych na dowodach**.

Ewaluacja jest bardzo ważnym, ale nie jedynym źródłem dowodów wspierających decyzje dotyczące polityk publicznych. Inne istotne uwzględniają: a) badania naukowe, przedstawiane w formie artykułów, książek, czy raportów; b) dane statystyczne gromadzone np. w ramach statystyki publicznej, czy rejestrów urzędowych; c) diagnozy – badania zlecane przez administrację, opisujące sytuację w obszarze zainteresowania, ale nie oceniające konkretnej interwencji; d) konsultacje społeczne – różne formy zapytań skierowanych do społeczności, podmiotów w obszarze zainteresowania; e) opinie ekspertów – zarówno naukowców, teoretyków, jak i praktyków w branży; f) doświadczenia, niekoniecznie poparte rygorystycznym badaniem, zarówno własne z dotychczasowych działań, jak i cudze, z organizacji, podmiotów, które mogą służyć nam za przykład; g) pilotaże, eksperymenty, prototypowanie – wdrażanie interwencji w małej skali i możliwie kontrolowanych warunkach, tak by zrozumieć jej potencjalne skutki.

Różnorodność dowodów można rozpatrywać też od strony podmiotowej. Dostarczają ich m.in.: a) analitycy wewnętrzni, zatrudnieni w administracji; b) ciała doradcze powołane przez władze, c) think tanki, finansowane niezależnie, ze źródeł prywatnych lub współpracujące z partiami politycznymi, d) akademicy.

Wielkość źródeł dowodów sprawia, że praca ewaluatora w administracji często nie ogranicza się do zrealizowania czy zamówienia badania ewaluacyjnego, ale wymaga też osadzenia jego wyników w kontekście innych dowodów, tzn. zbudowania przekonującej argumentacji w oparciu o dowody z różnych źródeł. Dlatego też spotykanym rozwiązaniem, szczególnie w organizacjach o rozwiniętej kulturze ewaluacji, jest opracowanie tzw. planów ewaluacji, czy szerzej: planów uczenia się, które obejmują okres kilku lat i precyzują jaka wiedza powinna być pozyskana z jakichś źródeł, kiedy i dla kogo, by decyzje w organizacji mogły być podejmowane w sposób racjonalny.



### Dobra praktyka

Wieloletnią praktykę przygotowania planów ewaluacji zaobserwować można w polskiej administracji zarządzającej programami polityki spójności: <https://www.ewaluacja.gov.pl/strony/system-ewaluacji/plany-ewaluacji>

## Jakie są uwarunkowania prawne ewaluacji w Polsce?

Z perspektywy źródeł praktyk ewaluacji, kraje można podzielić na dwie zasadnicze grupy. Pierwsze to te, w których ewaluacja pojawiła się jako element krajowych reform administracji wywołanych czynnikami wewnętrznymi. W Europie są to np. Niemcy, Szwecja, Dania, Holandia (Derlien, 1990). Druga grupa, to kraje, w których ewaluacja pojawiła się w wyniku zewnętrznej presji, najczęściej wywieranej przez organizacje międzynarodowe. W przypadku Polski, niemającej wcześniejszych tradycji ewaluacyjnych, tym impulsem było przystąpienie w 2004 r. do UE i wiążący się z tym obowiązek dostosowania do wymogów ewaluacji polityki spójności. Od tego momentu systematyczna praktyka ewaluacji dynamicznie rozwija się w administracji centralnej i regionalnej zarządzającej programami współfinansowanymi ze środków UE. Dużo wolniej przebiega przenoszenie praktyki ewaluacji na grunt krajowych polityk publicznych.

Podstawowym aktem prawnym regulującym kwestię ewaluacji w Polsce, również poza polityką spójności jest **ustawa z dnia 6 grudnia 2006 r. o zasadach prowadzenia polityki rozwoju** z późn.zm<sup>5</sup>. Obecnie (2024) nakłada ona obowiązek ewaluacji uprzedniej (ex ante) strategii rozwoju, w tym rozwoju kraju, województw, gmin<sup>6</sup>, rozwoju ponadlokalnego. Obowiązkowa jest również ewaluacja podsumowująca efekty (ex post) dużych programów rozwoju (realizujących cele strategii rozwoju, o wartości powyżej 300 mln zł).

W Polsce obowiązuje również przepis o konieczności przygotowania planu ewaluacji i następnie ewaluacji programów pomocy publicznej, których średnioroczny budżet

<sup>5</sup> Dz. U. 2006 Nr 227 poz. 1658, z późn. zm. Pierwsze zapisy o ewaluacji poza programami polityki spójności pojawiły się w ustawie w 2008 r.

<sup>6</sup> Strategie rozwoju gminy zostały dodane do katalogu strategii rozwoju w ramach nowelizacji z 2020 i stanowi to bardzo istotne rozszerzenie obowiązku ewaluacji.

przekracza 150 mln EUR. Zapis ten wynika z rozporządzenia KE<sup>7</sup>, ale dotyczy również krajowej pomocy publicznej.

W szeregu innych ustaw znaleźć można zapisy o konieczności prowadzenia oceny, np. programów. Rosnąca świadomość ewaluacji sprawia, że zapisy dotyczące oceny są coraz częściej utożsamiane z obowiązkiem ewaluacji i będą skutkować prowadzeniem badań ewaluacyjnych. Sztandarowym przykładem tego zjawiska jest ustawa z dnia 9 października 2015 r. o rewitalizacji<sup>8</sup>.

Za elementy uwarunkowań prawnych dotyczących ewaluacji uznać też można regulacje dotyczące innych, ale pokrewnych form analizy interwencji publicznych. Podstawowe z nich to prognozy w ramach ocen oddziaływania na środowisko (OOS) i oceny skutków regulacji (OSR). Przeprowadzenie OOS wymagane jest ustawowo na etapie przygotowania dokumentów z zakresy zagospodarowania przestrzennego oraz polityk, programów, strategii, planów w szeregu dziedzin<sup>9</sup>. Opracowania te często traktowane są jako rodzaj ewaluacji ex ante. OSR to w uproszczeniu ewaluacja interwencji regulacyjnej – aktu normatywnego. Kwestia przygotowania OSR określona jest w Regulaminie pracy Rady Ministrów<sup>10</sup>. Przygotowanie oceny przewidywanych skutków (odpowiednik ewaluacji ex ante) jest obowiązkowe dla projektów dokumentów rządowych. OSR ex post jest przygotowywana tylko jeżeli wystąpi o to ktoś z wąskiego grona uprawnionych organów<sup>11</sup> i w praktyce dzieje się to niezwykle rzadko. Choć praktyki ewaluacji i OSR rozwinęły się niezależnie od siebie, to stanowią uzupełniające się elementy analizy i doskonalenia finansowych i regulacyjnych interwencji państwa. Od strony praktycznej realizacji wykazują wiele podobieństw, co potwierdzają inicjatywy usiłujące przenieść doświadczenia z pola ewaluacji do OSR (np. Bienias, Herman-Pawłowska, 2015).

Na koniec warto wspomnieć, że pojęcie ewaluacji związane jest też ściśle z polem **edukacji i nauki**. Istotne różnice, m.in. w zakresie organizacji procesu, przedmiotu oceny, sprawiają, że zwykło się ewaluację w tym polu traktować jako zupełnie odrębną praktykę

<sup>7</sup> Rozporządzenie Komisji (UE) nr 651/2014 z dnia 17 czerwca 2014 r. uznające niektóre rodzaje pomocy za zgodne z rynkiem wewnętrznym w zastosowaniu art. 107 i 108 Traktatu.

<sup>8</sup> Dz.U. 2015 poz. 1777, z późn. zm. Choć art. 22 ustawy mówi, że gminny program rewitalizacji podlega ocenie, to samorządy przygotowujące się do wypełnienia tego obowiązku organizują szkolenia dotyczące ewaluacji programów rewitalizacji, a następnie w materiałach z takich szkoleń można zauważyć że ustawowe pojęcie oceny jest interpretowane jako ewaluacja.

<sup>9</sup> Ustawa z dnia 3 października 2008 r. o udostępnianiu informacji o środowisku i jego ochronie, udziale społeczeństwa w ochronie środowiska oraz o ocenach oddziaływania na środowisko (Dz. U. 2008 Nr 199 poz. 1227, z późn. zm.).

<sup>10</sup> Uchwała Nr 190 Rady Ministrów z dnia 29 października 2013 r. – Regulamin pracy Rady Ministrów, z późn. zm.

<sup>11</sup> Należą do niego m.in. Rada Ministrów, Prezes Rządowego Centrum Legislacji, Szef Rządowego Centrum Analiz.

od ewaluacji pozostałych interwencji publicznych. Ewaluacja w polskim systemie oświaty stanowiła element nadzoru pedagogicznego. Wprowadzona została odpowiednim rozporządzeniem w 2009 r.<sup>12</sup> i zniesiona w 2021 r. Obowiązek ewaluacji działalności naukowej w obecnym kształcie wprowadzono w 2018 r.<sup>13</sup> Jest to swoisty algorytm tworzący ranking dyscyplin uprawianych w poszczególnych jednostkach naukowych w oparciu o osiągnięcia pracowników naukowych (ich publikacje). Procedura ta przeprowadzana jest raz na 4 lata i od jej wyników zależą m.in. uprawnienia do nadawania stopni i tytułów oraz kwota subwencji przyznawanej z budżetu państwa.

---

<sup>12</sup> Rozporządzenie Ministra Edukacji Narodowej z dnia 7 października 2009 r. w sprawie nadzoru pedagogicznego (Dz.U. 2009 nr 168 poz. 1324)

<sup>13</sup> Ustawa z dnia 20 lipca 2018 r. – Prawo o szkolnictwie wyższym i nauce (Dz.U. 2018 poz. 1668)

## Rozdział 2. Jaka jest rola ewaluacji we wdrażaniu interwencji publicznej?

Tomasz Gapski, Rafał Trzciński

### Teoria interwencji oraz jej weryfikacja w ramach ewaluacji

#### Czym jest teoria interwencji?

Teoria interwencji (polityki, strategii, programu, działania, projektu itp.)<sup>14</sup> to pojęcie stosowane w ramach szerokiej rodziny podejść badawczych, określanych mianem ewaluacji opartej na teorii (*theory-based evaluation* – TBE)<sup>15</sup>.

Dlaczego mówimy o interwencji? Ponieważ każde działanie zmierzające do zmiany zastanej sytuacji społeczno-ekonomicznej jest interwencją. Państwo ingeruje (interweniuje) w sytuację, aby osiągnąć pożądane efekty, np. spadek bezrobocia, zmniejszenie liczby wypadków drogowych, wzrost przedsiębiorczości itp. Założenia takich interwencji opisywać będą polityki, programy, działania i projekty.

Z perspektywy TBE, każda interwencja to swego rodzaju teoria (Weiss 1995, Pawson i Tilley 1997, Górniak, 2007). Składa się na nią zestaw hipotez na temat tego jak oraz przy spełnieniu jakich warunków uda się osiągnąć pożądaną zmianę (np. obniżyć poziom bezrobocia w wyróżnionej grupie osób lub zwiększyć poziom innowacyjności przedsiębiorstw z sektora MŚP itp.). Każdą taką teorię można przedstawić jako ciąg przyczynowy zdarzeń poczynawszy od uruchomienia interwencji, poprzez rozmaite powiązania łączące działania programowe z poszczególnymi rezultatami, aż do osiągnięcia oczekiwanych efektów końcowych (Stern i in. 2012).

<sup>14</sup> W literaturze przedmiotu teoria interwencji występuje pod wieloma nazwami, takimi jak np. teoria programu, teoria zmiany, logika interwencji. Szersze zestawienie tych określeń znaleźć można np. u Funnell, Rogers (2011). W niniejszym poradniku posługiwać będziemy się zamiennie pojęciem teorii interwencji oraz teorii programu.

<sup>15</sup> Analogicznie w literaturze spotkamy się również z innymi określeniami dla ewaluacji opartej na teorii – np. ewaluacja wiedzona teorią (*theory-driven evaluation*), ewaluacja kierowana teorią (*theory-guided evaluation*) oraz innymi (Hermann-Pawłowska, Skórska, 2017). W poradniku posługiwać będziemy się pojęciem ewaluacji opartej na teorii (TBE), podobnie jak ma to miejsce w poradniku Komisji Europejskiej – *Evalsed Sourcebook: Methods and Technics* (2013).

Jak zauważa Weiss (1997), teoria będąca podstawą realizacji danej interwencji nie musi być powszechnie zaakceptowana. Nie musi być również poprawna. Często nie będzie nawet wprost wyrażona (EVALSED, 2013). Rolą ewaluacji jest jednak możliwe dokładne jej przedstawienie (odtworzenie), w tym identyfikacja wszystkich założeń, które potencjalnie warunkują jej materializację. Następnie konieczne jest zgromadzenie i analiza dowodów (danych), które będą empirycznym testem przyjętej teorii interwencji. Wynik tego testu, a więc ustalenia ewaluacji, powinny wykazać, czy dana teoria faktycznie daje się utrzymać, czy należy ją odrzucić (Weiss, 1995; Górniak, 2007). Najkrócej mówiąc test wykazuje czy teoria działa, czy nie. W praktyce, z uwagi na cykliczny charakter wielu typów interwencji, szczególnie interesujące są te jej elementy, które mogą zostać poprawione (EVALSED, 2013). Kluczowa w tym względzie jest identyfikacja i weryfikacja założeń leżących u podstaw danej interwencji – które z nich okazały się trafne i pozwoliły osiągnąć zakładane rezultaty, a które błędne i ograniczyły skuteczność interwencji.

Teorię interwencji należy odróżnić od **teorii wdrażania** (Rogers i Weiss, 2007). Teoria interwencji skupia się na reakcji beneficjentów na podjęte działania oraz na mechanizmach, które odpowiadają za to, że dana interwencja działa (Pawson i Tilley, 1997; Leeuw, 2012, Weiss, 1997). Natomiast **teoria wdrażania** dotyczy przede wszystkim działań podejmowanych przez autorów interwencji. Może być elementem teorii interwencji, ponieważ działania wdrożeniowe mogą naturalnie warunkować skuteczność interwencji. Jednak to mechanizmy uruchomione na poziomie odbiorców interwencji oraz ich zachowania prowadzące do zmiany, odpowiadają za finalne efekty działań.

Pawson i Tilley (1997) zwracają dodatkowo uwagę na znaczenie kontekstu dla działania poszczególnych interwencji i tym samym konieczność uwzględnienia go w procesie ewaluacji. Te same interwencje mogą inaczej działać w różnych kontekstach, w tym np. dla różnych osób, w innych miejscach itp. Można zatem powiedzieć, że skuteczność danej interwencji będzie dodatkowo warunkowana kontekstem w jakim jest ona wdrażana (EVALSED, 2013).

## Dlaczego wykorzystanie koncepcji teorii interwencji w ewaluacji jest ważne?

Ewaluacja oparta na teorii – TBE ma swoje korzenie w próbie przezwyciężenia ograniczeń występujących typowo w badaniach koncentrujących się na ustaleniu efektów interwencji, w tym ewaluacjach wpływu (*impact evaluation*)<sup>16</sup>. Badania te, o ile przeprowadzone

---

<sup>16</sup> Chodzi tu w szczególności o badania eksperymentalne i quasi-eksperymentalne. Więcej na temat tych podejść znaleźć można w rozdziale 7.

zostaną poprawnie, pozwalają odpowiedzieć na pytanie „czy dana interwencja działa?”. Jednocześnie w bardzo ograniczony sposób lub w ogóle nie pozwalają odpowiedzieć na pytanie „dlaczego dana interwencja działa (lub nie działa)?”. W tym względzie ujawnia się problem tzw. „czarnej skrzynki” (Pawson i Tilley, 1997) – wiemy, że zrealizowany został pewien zestaw działań i że wywołał on pewne efekty, lecz nie wiemy jak do tego doszło<sup>17</sup>. Ten stan rzeczy przekłada się naturalnie na użyteczność wyników ewaluacji. Jak bowiem możemy usprawnić daną interwencję, jeśli nie znamy źródła jej sukcesu lub porażki? Podejścia z rodziny TBE są bezpośrednią odpowiedzią na tę lukę. Robią to poprzez staranne „rozpakowanie” planowanej teorii interwencji oraz systematyczną weryfikację jej poszczególnych elementów.



### Metafora „czarnej skrzynki”.

**Wyobraźmy sobie urządzenie, które z jednej strony wciąga dostarczone klocki drewna, z drugiej wydaje zabawki dla dzieci. Zakładamy, że skrzynka przerabia drewno na zabawki, ale bez zajrzenia do środka nie wiemy, jak to się odbywa. Podobnie może być z mechanizmem interwencji, który wymaga analiz aby sprawdzić jak dokładnie działa.**

Rekomendowaną i coraz częstszą praktyką jest łączenie obu podejść, tj. ewaluacji wpływu z ewaluacją opartą na teorii (Górnaiak, 2007; White, 2009; EVALSED, 2013). W literaturze przedmiotu (White, 2009; EVALSED, 2013) można spotkać się z określeniem takich badań jako ewaluacji wpływu opartych na teorii (*theory-based impact evaluation* – TBIE). Niektórzy autorzy zwracają w tym względzie uwagę na występowanie pewnej sekwencyjności. W szczególności Martini (2009) wskazuje, że pytanie o to czy interwencja działa, naturalnie poprzedza odpowiedź na pytanie o przyczyny tego stanu rzeczy. Połączenie obu podejść w ewaluacji – o ile jest to tylko możliwe – pozwala zatem odpowiedzieć na kompleksową listę pytań o to: **czy, dlaczego oraz – przy uwzględnieniu warunków kontekstowych – dla kogo, w jakich okolicznościach, czasie itd., dana interwencja działa?**

Wykorzystanie podejścia ewaluacji opartej na teorii oznacza w praktyce konieczność uwzględnienia teorii interwencji w całym procesie badawczym. Wybrani autorzy wskazują wręcz, że ewaluacja powinna być **wiedziona przez teorię** (Funnell i Rogers, 2011; Rogers, 2007). Rozumieć przez to należy sytuację, w której projekt badania, w tym stawiane pytania badawcze, wybierane metody zbierania i analizy danych, wskaźniki oraz narzędzia badawcze wynikają wprost z kształtu ustalonej teorii interwencji. Choć wymóg ten może z pozoru brzmieć jak dodatkowe utrudnienie, to w praktyce jest wprost przeciwnie. Znając teorię interwencji, wiemy bowiem dokładnie jakie fakty musimy ustalić oraz kogo i o co zapytać. Kompletna, wykonana krok po kroku, weryfikacja teorii

<sup>17</sup> Naturalnie jeszcze bardziej kłopotliwą jest sytuacja, w której badania wskazują na brak skuteczności interwencji przy jednoczesnym braku wiedzy, dlaczego tak się stało.

interwencji powinna prowadzić również do użytecznych rekomendacji, gdyż pozwoli nam ustalić, które elementy tej teorii (nie)zadziałały oraz dlaczego tak się stało.

## Jak ustalić (odtworzyć) teorię interwencji?

Rzadko kiedy teoria interwencji jest wyartykułowana wprost (EVALSED, 2013). Dlatego jednym z elementów ewaluacji powinno być jej odtworzenie. Proces ten typowo podzielony będzie na fazę gabinetową oraz terenową. W pierwszym przypadku konieczne będzie zapoznanie się z dokumentacją interwencji, gdzie powinny być przedstawione cele i działania planowane do zrealizowania w ramach danego przedsięwzięcia. Wśród ważnych dokumentów wymienić tu należy w szczególności te związane z procesem selekcji odbiorców wsparcia (warunki udziału, kryteria oceny i wyboru projektów, regulaminy konkursowe itp.). W praktyce to one będą często operacyjnym odzwierciedleniem oczekiwań autorów danej interwencji. Ważnym aspektem badania na tym etapie powinna być również analiza literatury przedmiotu, w tym wcześniejszych badań zrealizowanych w podobnym obszarze. Cenne będą w szczególności te pozycje, w których dokonane zostało już odtworzenie teorii lub zidentyfikowane zostały przyczyny sukcesu lub porażki podobnych instrumentów (White 2009). Faza gabinetowa powinna być następnie uzupełniona o etap konsultacji teorii z kluczowymi interesariuszami danej interwencji, włączając w to zarówno przedstawicieli administracji odpowiedzialnych za jej wdrażanie jak również podmioty, czy osoby mające bezpośredni kontakt z odbiorcami wsparcia. Perspektywa wybranych reprezentantów tej ostatniej grupy również będzie wartościowa. Co więcej, może się okazać, że odbiega ona od teorii interwencji oficjalnie reprezentowanej przez instytucje, które ją realizują. Innymi słowy, wyniki procesu odtwarzania teorii mogą wskazywać, że mamy do czynienia z kilkoma „rywalizującymi” ze sobą teoriami.

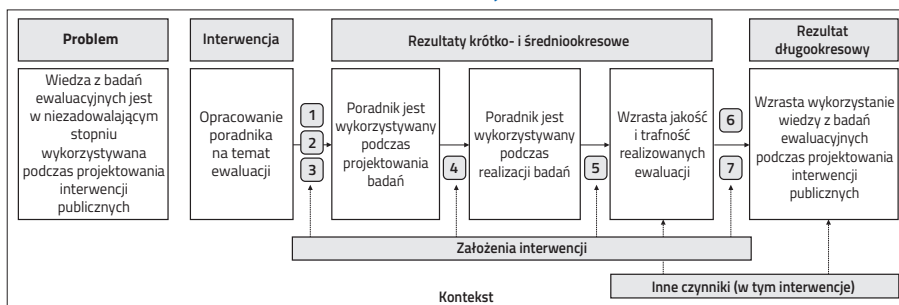
Rezultatem całego procesu powinien być kompleksowy opis teorii interwencji, zwany czasem łańcuchem przyczynowym zdarzeń, łączący działania z oczekiwanymi rezultatami – bezpośrednimi i końcowymi. Poszczególne relacje przyczynowe między działaniami i rezultatami – oznaczane najczęściej strzałkami – powinny być wyjaśnione poprzez wskazanie założeń na jakich się opierają. Całość typowo prezentowana jest w formie schematu graficznego uzupełnionego opisem, choć dopuszczalne są również inne formy prezentacji teorii interwencji (np. wyłącznie opis)<sup>18</sup>. Kryterium oceny poprawności odtworzonej teorii – poza jej kompletnością – jest jej czytelność dla potencjalnych odbiorców wyników badania.

<sup>18</sup> Przykłady prezentacji graficznych znaleźć można np. w obszernym opracowaniu, które poświęcone jest podejściom ewaluacji opartej na teorii – Funnell S.C., Rogers P. J. (2011). *Purposeful Program Theory: Effective Use of Theories of Change and Logic Models*, Jossey Bass.



Wykorzystując jako przykład interwencji opracowanie i publikację niniejszego poradnika na temat ewaluacji, pokazujemy jak w praktyce stosuje się ten typ wnioskowania. Różne analizy wskazują, że jakość ewaluacji pozostaje generalnie na niezadowalającym poziomie, co przekłada się na niskie wykorzystanie wyników badań w procesie tworzenia polityk publicznych (Pellegrin i Colnot, 2020). Choć wybrane opracowania wskazują, że związek między jakością ewaluacji, a jej wykorzystaniem nie jest jednoznaczny, to jednocześnie ustalono, że wdrażanie rekomendacji z badań ewaluacyjnych jest wspierane przez solidność zastosowanej metodyki badawczej, jakość wykonanych analiz i przejrzystość raportu końcowego (Kupiec, 2022). Niniejszy poradnik obejmuje wszystkie te trzy zagadnienia, proponując dla nich odpowiednie rozwiązania. Wierzymy, że zastosowanie ich w praktyce pozwoli zwiększyć jakość zlecanych i realizowanych badań. Długoterminowo powinno to przełożyć się również na wzrost wykorzystania wyników ewaluacji w projektowaniu polityk publicznych. Ten uproszczony ciąg zakładanych rezultatów przedstawiony został na kolejnym rysunku, który można określić mianem schematu logicznego tej interwencji.

**Rysunek 2. Schemat logiczny interwencji polegającej na opracowaniu poradnika ewaluacji**



Źródło: Opracowanie własne

W praktyce skuteczność przedsięwzięcia będzie zależała od spełnienia wielu założeń, w tym uruchomienia odpowiednich mechanizmów. Zostały one oznaczone numerami przy poszczególnych strzałkach. W szczególności materializacja przedstawionego ciągu przyczynowego zależeć będzie od następujących kwestii:

1. Odbiorcy dowiedzą się o niniejszym poradniku.
2. Odbiorcy faktycznie zapoznają się z jego treścią, a ewaluatorzy dodatkowo zapoznają się z literaturą uzupełniającą.
3. Poradnik musi posiadać walory zarówno jakościowe oraz związane z przekazywaną treścią (walor praktyczny).
4. Zewnętrzni ewaluatorzy, zakontraktowani do realizacji badania, posiadają odpowiednie zasoby i potencjał oraz są odpowiednio zmotywowani, by stosować się do zaleceń wynikających z poradnika (o ile realizujemy nasze badanie przy wsparciu zewnętrznych wykonawców).

5. Wdrożenie prezentowanych w poradniku rozwiązań, aby było możliwe, wymaga różnorodnych środków. Nawet najlepszy poradnik nie zastąpi doświadczonego ewaluatora, a bez odpowiednich, często kosztownych danych części analiz wykonać się nie da.
6. Istnieją założenia dotyczące samego mechanizmu wykorzystania wyników badań przez instytucje i osoby odpowiedzialne za projektowanie interwencji. Wiąże się z tym rozwijana od wielu lat koncepcja brokera oraz brokeryng wiedzy. Jest to złożona teoria wychodząca poza proces upowszechniania wyników z danego badania ewaluacyjnego. Obejmuje ona m.in. takie kwestie jak odpowiednio wcześnie włączenie odbiorców wyników w proces projektowania badań (identyfikację potrzeb wiedzy), budowanie relacji między producentami i odbiorcami wiedzy, czy też budowę szeroko rozumianej kultury polityk publicznych opartych na dowodach. Kluczowym elementem tej teorii są również uwarunkowania wpływające na proces podejmowania decyzji w danej organizacji.
7. W danej instytucji musi być wystarczający potencjał do wdrożenia wyników ewaluacji. Proponowane przez ewaluatorów rozwiązania muszą być więc w zasięgu możliwości prawnych, organizacyjnych i finansowych danej organizacji.

Oczywiście nie jest to kompletna lista założeń warunkujących osiągnięcie docelowego efektu. Jak wspomniano, mechanizmowi wykorzystania wyników ewaluacji poświęcony jest osobny wątek badań nad ewaluacją (np. Kupiec 2022, Olejniczak 2008). Przykład ten jednak pokazuje, jak wiele założeń może potencjalnie ograniczyć skuteczność z pozoru prostej interwencji. Czasem wystarczy niespełnienie jednego z nich, by przerwać ciąg przyczynowy zakładanych zdarzeń.

Etap odtwarzania teorii programu powinien być bezpośrednio powiązany z analizą kontekstu oraz obszarów możliwego zróżnicowania efektów interwencji (np. w wymiarze wspartych podgrup, miejsc realizacji lub czasu implementacji programu). Jest to szczególnie ważne z perspektywy kolejnego kroku, jakim jest zdefiniowanie szczegółowych parametrów badania, w tym projektowania narzędzi badawczych oraz doboru prób badawczych, które wykorzystane zostaną w fazie weryfikacji przedmiotowej teorii. Dla naszej przykładowej teorii poradnika ewaluacyjnego, ważne może być np. doświadczenie poszczególnych jednostek ewaluacyjnych. W szczególności może się okazać, że dla doświadczonych zespołów poradnik nie będzie stanowił żadnej wartości dodanej, stąd jego wpływ na podnoszenie jakości realizowanych ewaluacji będzie znikomy lub żaden. Z drugiej strony poradnik może się okazać przydatny dla osób i instytucji, które dopiero zaczynają przygodę z ewaluacją. Analogiczna sytuacja może dotyczyć wykonawców ewaluacji. Chcąc zatem ustalić skuteczność interwencji, nie możemy ograniczyć się do zbadania jednej lub drugiej grupy, gdyż ocena będzie obciążona błędem (poradnik może się okazać wielkim sukcesem lub całkowitą kląpą). Schemat badania musi zatem uwzględnić to zróżnicowanie, w szczególności podczas doboru prób badawczych.

Na etapie odtwarzania teorii powinniśmy się również zastanowić nad innymi czynnikami, w tym innymi interwencjami i działaniami, które mogą oddziaływać na oczekiwane rezultaty naszych działań. Powinny one znaleźć odzwierciedlenie w prezentowanej teorii, a ich ewentualny wpływ powinien być ustalony na etapie jej weryfikacji. Wracając ponownie do przykładu naszego poradnika, naiwnie byłoby sądzić, że będzie to jedyny, czy też kluczowy czynnik wpływający na jakość realizowanych ewaluacji, czy tym bardziej wykorzystanie wyników badań w projektowaniu interwencji. Na oczekiwany wzrost jakości ewaluacji wpłynąć może np. równoległe uruchomienie cyklu szkoleń lub studiów podyplomowych dla ewaluatorów. Jednocześnie mogą zmienić się regulacje prawne na temat prowadzenia lub wykorzystania wyników ewaluacji. Wreszcie organizacje mogą wdrażać rozwiązania rozwijane w ramach koncepcji brokerów wiedzy, które uruchomią procesy sprzyjające wykorzystaniu wyników z badań. Wszystkie te działania mogą przekładać się na poziom wykorzystania wyników ewaluacji. Oczywiście poradnik również może odegrać w tym procesie ważną rolę, jednak nie będzie on jedyną przyczyną tej zmiany. W przypadku rozbudowanych interwencji publicznych – np. programów UE – skala powiązania wdrażanych działań z innymi interwencjami, w tym krajowymi politykami publicznymi, jest jeszcze większa niż w przedstawionym przykładzie. Nie zmienia to jednak naszego zadania, czyli próby znalezienia odpowiedzi na pytanie, czy nasza interwencja działa oraz dlaczego, dla kogo i w jakich okolicznościach.



### Uwaga, Pułapka!

Wykorzystanie w praktyce podejścia TBE nie jest równoznaczne z prezentacją w raporcie schematu logicznego interwencji (serią bloków powiązanych niewyjaśnionymi strzałkami).

Ewaluacja prowadzona zgodnie z podejściem TBE musi:

- odkrywać i sprawdzać założenia i mechanizmy, które warunkują wywołanie oczekiwanej zmiany.
- być wiedzona przez teorię interwencji, co powinno znaleźć bezpośrednie odzwierciedlenie na poziomie pytań, metod, wskaźników oraz narzędzi badawczych.

## Jak zweryfikować teorię interwencji? Dowody faktyczne i kontrfaktyczne

Podejście TBE nie faworyzuje żadnej z metod badawczych, nie tworzy w tym względzie również żadnej hierarchii (EVALSED, 2013). Projektujący ewaluację powinni wybrać takie metody i dane, które umożliwią dokonanie weryfikacji ustalonej teorii. Istotą jest, by – tak

jak zostało to wspomniane wyżej – ewaluacja była faktycznie wiedzona przez teorię programu oraz była w stanie zweryfikować ją w możliwie pełnym zakresie.

Zwolennicy łączenia ewaluacji opartej na teorii z ewaluacją wpływu, zwracają jednocześnie uwagę na dwa typy danych niezbędnych do dokonania kompletnej weryfikacji teorii interwencji (zgodnie z podejściem TBIE). Są to dowody faktyczne i kontrfaktyczne (White, 2009). W pierwszym przypadku pod uwagę brane są wszelkiego rodzaju dane faktograficzne, bazujące na bezpośredniej obserwacji. Dane te wskazują, czy spodziewane w ramach interwencji zdarzenia faktycznie miały miejsce, czy np. wsparcie otrzymały grupy, do których miało ono trafić, czy wsparcie uzyskały projekty zgodne z ogólną teorią interwencji, czy jakość udzielonego wsparcia była na odpowiednim poziomie. Punktem wyjścia do analizy w tym zakresie powinny być systemy monitorowania, sprawozdania z realizacji. Analiza w tym względzie powinna również umożliwić zweryfikowanie, czy założenia warunkujące materializację teorii faktycznie zostały spełnione. Podobnie oceniony powinien być potencjalny wpływ czynników zewnętrznych względem interwencji na materializację teorii interwencji. W analizie faktycznej preferowane będzie wykorzystanie szerokiego zestawu metod zarówno ilościowych i jakościowych, w tym m.in. badań kwestionariuszowych, wywiadów, warsztatów (metody i techniki bardziej szczegółowo zostały omówione w rozdziale 7 poradnika). Ich dobór powinien być warunkowany bezpośrednio potrzebami informacyjnymi zawartymi w pytaniach badawczych oraz zdiagnozowaną wcześniej dostępnością danych zastanych (więcej na ten temat w rozdziale 3 poradnika). Pytania badawcze powinny być wprost powiązane z poszczególnymi elementami teorii interwencji.

W przypadku dowodów kontrfaktycznych chodzi o możliwie rygorystyczny (dokładny) pomiar efektów przyczynowych wsparcia, tj. porównanie zmian jakie są obserwowane w grupie beneficjentów z hipotetyczną (kontrfaktyczną) sytuacją, w której ta sama grupa podmiotów wsparcia nie otrzymuje. W tym zakresie metodyka badań ewaluacyjnych oferuje szerokie spektrum podejść, technik i metod analitycznych. Ich ogólne założenia przybliżone zostały w rozdziale 7 poradnika.

W części ewaluacji, zastosowanie podejścia oraz dowodów kontrfaktycznych nie będzie możliwe – może to wynikać zarówno z dostępności odpowiednich danych, momentu realizacji badania, czy też specyfiki analizowanej interwencji<sup>19</sup>. Część podejść rozwijanych w ramach rodziny ewaluacji opartych na teorii – w szczególności **ewaluacja realistyczna** (Pawson i Tilley, 1997) oraz **analiza kontrybucji** (Mayne, 2019) – rezygnuje ze stosowania kontrfaktycznej definicji przyczynowości, na rzecz perspektywy tzw. przyczynowości generatywnej (Pawson i Tilley, 1997). O ile przyczynowość kontrfaktyczna skupia się

---

<sup>19</sup> W przypadku części interwencji zastosowanie podejścia kontrfaktycznego jest bardzo trudne lub wręcz niemożliwe. Dotyczy to np. ewaluacji wsparcia dużych projektów infrastrukturalnych.

na alternatywnych scenariuszach i pytaniach „co by było, gdyby”, o tyle przyczynowość generatywna koncentruje się na tym, jak konkretne czynniki wpływają w danym kontekście na obserwowane rezultaty. W analizie kontrybucji zaproponowanej przez Mayne’a<sup>20</sup> (2011) wskazuje się, że „[...] można wysuwać uzasadnione wnioski o wpływie interwencji jeśli:

- Istnieje uzasadniona teoria zmiany<sup>21</sup> dotycząca interwencji: kluczowe założenia dotyczące tego, dlaczego oczekuje się, że interwencja będzie działać są sensowne, prawdopodobne, mogą być podparte dowodami oraz są zaakceptowane przez co najmniej niektórych kluczowych interesariuszy;
- Działania zaplanowane w interwencji zostały wdrożone zgodnie z teorią zmiany;
- Teoria zmiany lub jej kluczowe elementy są poparte i potwierdzone dowodami w zakresie obserwowanych rezultatów oraz warunkujących je założeń, wyrażanych przez ekspertów i fakty – zdefiniowane w łańcuchu przyczynowym rezultaty wystąpiły. Teoria zmiany nie została odrzucona;
- Inne czynniki wpływające na rezultaty zostały ocenione i albo okazało się, że nie miały znaczącego wkładu, albo ich względna rola w tym zakresie została ustalona”.

Jednocześnie trzeba mieć świadomość, że analiza kontrybucji, czy szerzej analizy oparte na generatywnej definicji przyczynowości nie pozwalają określić, w jakim stopniu wynik końcowy (obserwowaną zmianę) można przypisać interwencji (Mayne, 2019). Innymi słowy nie jest możliwe dokładne, ilościowe oszacowanie wpływu danej interwencji.

### Kiedy można zweryfikować teorię interwencji?

Nie ma ograniczeń czasowych w zakresie wykorzystania podejścia TBE. Znajdzie ono zastosowanie zarówno na etapie projektowania, realizacji i po zakończeniu interwencji (EVALSED, 2013). Naturalnie na każdym z tych etapów dostępne będą inne typy dowodów. W kolejnej części niniejszego podrozdziału przedstawiamy rolę teorii interwencji na kolejnych etapach jej wdrażania.

<sup>20</sup> Por. Mayne (2011), s.7–8.

<sup>21</sup> Wyjątkowo w tym przypadku zastosowaliśmy oryginalne tłumaczenie – teoria zmiany występuje tu jednak w znaczeniu teorii interwencji/teorii programu.



### Dobra praktyka

Teorię interwencji warto odtworzyć już na początkowym etapie realizacji danej strategii, polityki czy programu. Podczas ewaluacji prowadzonych w toku lub po jego zakończeniu warto jednocześnie każdorazowo zweryfikować, czy nie jest konieczna aktualizacja tej teorii. Może tak się stać, jeśli w ramach interwencji zaszły istotne zmiany związane np. z procesem selekcji beneficjentów do wsparcia lub zasadami udzielania pomocy. W długich interwencjach publicznych, a do takich zaliczają się np. programy UE wspierane w kolejnych okresach programowania, zmienić się mógł również kontekst realizacji interwencji (by wspomnieć tylko przykłady pandemii COVID-19 lub wojny w Ukrainie w trakcie perspektywy 2014–2020). Aspekt ten powinien znaleźć również odzwierciedlenie w odtwarzanych i weryfikowanych teoriach programu.

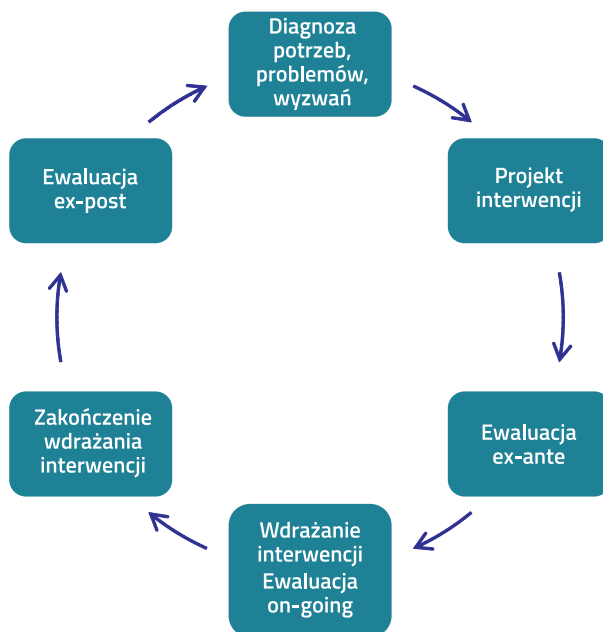
## Ewaluacja w cyklu wdrażania interwencji publicznej

Prawie każdą interwencję można przedstawić, jako cykl następujących po sobie etapów planistycznych i wdrożeniowych. Ewaluacja może, a wręcz powinna towarzyszyć interwencji na każdym etapie jej tworzenia i realizacji. W tym kontekście wyróżniamy typy badań ze względu na moment ich realizacji. Zazwyczaj wyróżnia się ewaluację *ex ante*, ewaluację *on-going* oraz ewaluację *ex post*. Można również wydzielić ewaluację *mid-term*, realizowaną w połowie okresu interwencji. Niemniej ten ostatni typ w ostatnich latach jest mniej stosowany w Polsce i rzadziej opisywany jako odrębny typ ewaluacji. Wynika to ze znacznego podobieństwa, zarówno w zakresie czasu realizacji, jak i celów, do ewaluacji *on-going*. Co do zasady ewaluację *ex ante* przeprowadzamy po zaprojektowaniu, ale przed uruchomieniem interwencji, ewaluację *on-going* w trakcie jej realizacji, zaś ewaluację *ex post* po jej zakończeniu.

W praktyce jednak moment realizacji poszczególnych rodzajów ewaluacji może odbiegać od powyższej zasady. Przykładowo ewaluacja *ex ante* może dotyczyć wybranych elementów, np. kryteriów wyboru projektów i być realizowana po formalnym zatwierdzeniu dokumentów programowych i po rzeczywistym uruchomieniu interwencji. Także w przypadku ewaluacji *ex post* badaniu podlegać może jedynie zakończona część interwencji, pomimo tego, że jako całość pozostaje ona wciąż w trakcie realizacji (np. ewaluacja *ex post* działań ukierunkowanych na wsparcie MŚP w wielosektorowych programach

regionalnych realizowanych w ramach polityki spójności). Miejsce poszczególnych rodzajów ewaluacji w cyklu wdrażania interwencji publicznej obrazuje poniższy schemat.

**Rysunek 3.** Ewaluacja w cyklu wdrażania interwencji publicznej



Źródło: Opracowanie własne na podstawie: Górniak (2007), Olejniczak (2010)

## Ewaluacja ex ante

### Co i w jakim celu badamy?

Celem ewaluacji ex ante jest weryfikacja trafności i spójności przygotowanej do wdrożenia interwencji. Zgodnie z zapisami ustawy o prowadzeniu polityki rozwoju<sup>22</sup> weryfikacja taka przeprowadzana jest w odniesieniu do spodziewanej skuteczności planowanej interwencji. Dlatego ewaluacja ex ante powinna nam dać odpowiedź na cztery podstawowe pytania:

<sup>22</sup> Patrz: Art. 10a Ustawy z dnia 6 grudnia 2006 r. o zasadach prowadzenia polityki rozwoju (Dz. U. 2006 Nr 227 poz. 1658)

1. Czy cele, które zostały założone do realizacji, odpowiadają na potrzeby odbiorców i właściwie adresują problemy i wyzwania zidentyfikowane na etapie diagnozy<sup>23</sup>?
2. Czy zaplanowane działania oraz środki finansowe są adekwatne do założonych celów<sup>24</sup>?
3. Czy zestaw działań (policy mix) przewidziany do realizacji jest spójny tj. czy zaplanowane działania uzupełniają się w sposób komplementarny w odniesieniu do założonych celów interwencji?
4. Czy zaprojektowany system wdrażania daje możliwość sprawnej i skutecznej realizacji interwencji?

W ramach odpowiedzi na pytanie pierwsze sprawdzamy, czy istnieje logiczny związek pomiędzy problemami, które chcemy rozwiązać, a celami jakie postawiliśmy sobie do osiągnięcia. Innymi słowy weryfikujemy, czy jeżeli uda nam się osiągnąć zakładane efekty, to przyczyni się to i w jakim zakresie do rozwiązania adresowanego problemu lub wyzwania. Zadaniem ewaluacji ex ante może być także weryfikacja samej diagnozy lub jej wybranych elementów, a także sposobu jej przeprowadzenia. Czy adresowane potrzeby są wciąż aktualne, czy w ramach diagnozy uwzględniliśmy najważniejsze problemy i wyzwania? Ważne jest też zweryfikowanie, czy uwzględniliśmy potrzeby kluczowych interesariuszy.

Jeżeli uznamy, że spodziewane efekty interwencji są adekwatne do potrzeb, problemów i wyzwań możemy przejść do kolejnego pytania. Na tym etapie szukamy odpowiedzi, czy działania jakie zaplanowaliśmy (instrumenty, projekty, regulacje prawne itp.) oraz środki finansowe, które planujemy przeznaczyć na ich wdrożenie, są odpowiednie dla skutecznej realizacji celów. Zadaniem ewaluacji w tym zakresie jest weryfikacja poprawności zakładanych związków przyczynowo- skutkowych pomiędzy planowanymi do realizacji działaniami, a spodziewanymi efektami tych działań. Na tym etapie oceny sprawdzamy także, czy zdefiniowane wskaźniki realizacji są właściwe dla mierzenia postępów i stopnia realizacji celów. Weryfikujemy również, czy przy zakładanych do wydatkowania środkach, istnieje możliwość osiągnięcia zaplanowanych wartości docelowych wskaźników.

Kolejnym elementem ewaluacji ex ante jest ocena wewnętrznej i zewnętrznej spójności i komplementarności zestawu działań przewidzianych do wdrożenia (policy mix). Celem badania w tym zakresie jest sprawdzenie, czy zaplanowane działania nie są ze sobą sprzeczne w kontekście zakładanych celów oraz czy wzajemnie się uzupełniają i wzmacniają (tworząc efekt synergii). Analogicznie oceniamy także spójność i komplementarność interwencji z innymi powiązanymi interwencjami. W szczególności należy

<sup>23</sup> Jest to ocena trafności interwencji, o której mowa w art. 10a ustawy o prowadzeniu polityki rozwoju.

<sup>24</sup> Jest to ocena spodziewanej skuteczności i efektywności programu, o których mowa w art. 10a ustawy o prowadzeniu polityki rozwoju.



tu przeanalizować spójność ze strategiami i programami wyższego rzędu oraz innymi działaniami (także prowadzonymi przez inne podmioty). Mogłoby się bowiem okazać, że inny podmiot realizuje interwencję, która ma niespójne (lub wręcz sprzeczne) cele z naszą interwencją.

Przedmiotem oceny powinien być także zaprojektowany dla danej interwencji system wdrażania. W ramach badania *ex ante* upewniamy, czy dostępne zasoby instytucjonalne, kadrowe i finansowe oraz sposób organizacji systemu dają możliwość sprawnej i skutecznej realizacji interwencji.

Docelowo ewaluacja *ex ante* służyć powinna udoskonaleniu interwencji. Może to polegać na lepszym dopasowaniu celów i działań do potrzeb odbiorców, czy też wyzwani w danym obszarze. Zmianie podlegać może również zakres i charakter planowanych działań, tak aby zwiększyć prawdopodobieństwo osiągnięcia zakładanych efektów. Ewaluacja *ex ante* powinna także służyć optymalnemu alokowaniu zasobów, w sposób umożliwiający ich jak najefektywniejsze wykorzystanie w odniesieniu do przyjętych celów.

## Jak badamy<sup>25</sup>?

Kluczowym elementem podlegającym ocenie w ramach ewaluacji *ex ante* jest teoria interwencji. Ewaluacja *ex ante* bazuje głównie na danych zastanych, takich jak dane diagnostyczne i dokumentacja programowa (ocena trafności i spójności interwencji). Istotnym jest również zebranie dowodów (np. z badań dot. podobnych interwencji zrealizowanych wcześniej) wskazujących na potencjalną skuteczność zaplanowanych działań. W zakresie danych pierwotnych kluczowymi źródłami informacji są instytucje i osoby bezpośrednio zaangażowane w przygotowanie interwencji (wywiady indywidualne i grupowe, warsztaty) oraz przedstawiciele grup docelowych (wywiady, ankiety). Osoby zaangażowane w przygotowanie interwencji (twórcy strategii lub programu) są również głównymi odbiorcami wyników ewaluacji. Dlatego w przypadku ewaluacji *ex ante* szczególnie użytecznym i pożądanym jest stosowanie partycypacyjnego modelu ewaluacji, opartego na dialogu i wczesnym angażowaniu głównych interesariuszy.

---

<sup>25</sup> Bardziej szczegółowe informacje dotyczące metod i technik badawczych zawiera Rozdział 7.

## Ewaluacja on-going

### Co i w jakim celu badamy?

Ewaluacja on-going, realizowana w trakcie wdrażania interwencji, obejmuje zazwyczaj węższy zakres tematyczny i skupia się na wybranych problemach i zagadnieniach. Przeprowadzana jest najczęściej w postaci szeregu badań inicjowanych w odpowiedzi na pojawiające się w toku wdrażania potrzeby informacyjne. Ewaluacja on-going odpowiada na następujące główne pytania:

1. Czy zidentyfikowane na etapie diagnozy potrzeby, problemy i wyzwania społeczno-ekonomiczne i odpowiadające im cele pozostają aktualne?
2. Jakie są główne bariery dla sprawnego wdrażania interwencji (prawne, administracyjne, organizacyjne itp.)?
3. Czy osiągnięte są zakładane cele na poszczególnych etapach realizacji i tym samym, czy zakładana teoria interwencji sprawdza się w praktyce?

Zadaniem ewaluacji on-going jest zatem bieżąca i systematyczna ocena sprawności i postępów we wdrażaniu interwencji w odniesieniu do założonych celów, aktualnych potrzeb, problemów i wyzwań społeczno-ekonomicznych w danym obszarze. Charakter badań on-going uzależniony jest od etapu na którym są one realizowane. Na wstępnych etapach wdrażania badania dotyczą głównie aktualności oraz trafności interwencji. Tym samym zbliżone są tematycznie do ewaluacji *ex ante*. Na końcowych etapach – zakres badań dotyczy w większym stopniu efektów interwencji. Zbliża je to zatem, w swojej specyfice, do ewaluacji *ex post*.

Jako specyficzny rodzaj ewaluacji on-going możemy wyróżnić ewaluację *mid-term* (nazywaną też *śródkresową*). Typ ten wyróżnia czas realizacji badania – w połowie okresu interwencji. Ewaluacja *mid-term* powinna sprawdzać czy wszystko, co zaplanowano, zmierza we właściwym kierunku, ewentualnie umożliwiając niezbędne korekty. Podobnie jak klasyczna ewaluacja on-going skupia się na bieżącej realizacji interwencji. Jej założenia i warsztat badawczy nie różnią się od ewaluacji on-going. Jedynym jej wyróżnikiem jest moment realizacji oraz zazwyczaj dość kompleksowy charakter.

Docelowo ewaluacja on-going służy dokonywaniu usprawnień i korekt w sposobie wdrażania interwencji (zatem jest ewaluacją formatywną – zobacz rozdział 1 poradnika). Zmiany takie dotyczyć mogą zakresu i charakteru realizowanych działań (np. zmiany kryteriów wyboru projektów), czy też alokacji zasobów przeznaczonych na poszczególne działania (np. przesunięcie środków w odpowiedzi na zmieniające się potrzeby

odbiorców). W niektórych przypadkach korekty dotyczyć mogą również samych celów interwencji wyrażonych w wartościach docelowych wskaźników realizacji. Ewaluacja on-going dostarcza również informacji służących projektowaniu przyszłych interwencji.

## Jak badamy?

Ewaluacja on-going w zakresie oceny aktualności potrzeb i teorii interwencji oraz barier wdrożeniowych bazuje najczęściej na analizie danych zastanych (analiza dokumentów programowych oraz aktualnych danych diagnostycznych i monitoringowych) oraz badaniach jakościowych (wywiady pogłębione, warsztaty z głównymi interesariuszami i odbiorcami wsparcia). Dla oceny postępów we wdrażaniu i pierwszych efektów interwencji wykorzystywane są dodatkowo analizy ilościowe, bazujące zarówno na danych zastanych (np. danych monitoringowych), jak i na danych pierwotnych (uzyskanych np. dzięki ankietom realizowanym wśród przedstawicieli grup docelowych). Na bardziej zaawansowanych etapach wdrażania interwencji metody stosowane w ewaluacji on-going zaczynają zbliżać się do tych wykorzystywanych w badaniach typu ex post (w tym w analizach kontrfaktycznych).

## Ewaluacja ex post

### Co i w jakim celu badamy?

Ewaluacja ex post, realizowana (co do zasady) po zakończeniu interwencji, stanowi najpełniejszą wersję ewaluacji. Służy ona podsumowaniu najważniejszych efektów i odpowiada na następujące główne pytania badawcze:

1. Jakie są krótkoterminowe i długoterminowe oraz bezpośrednie i pośrednie efekty interwencji?
2. Czy i w jakim zakresie zrealizowane zostały cele i czy przyczyniło się to do zaspokojenia potrzeb i rozwiązania problemów społeczno- ekonomicznych?
3. Jakie były główne czynniki zewnętrzne i wewnętrzne warunkujące zaistnienie zakładanych efektów interwencji?
4. Czy zaprojektowana teoria interwencji, w tym związki przyczynowo-skutkowe oraz założenia warunkujące osiągnięcie efektów sprawdziły się w praktyce?
5. Czy istnieje uzasadnienie dla kontynuacji interwencji, a jeżeli tak, to jakich ewentualnych zmian i usprawnień należy w niej dokonać?

Zadaniem ewaluacji ex post jest zatem kompleksowa weryfikacja teorii interwencji w oparciu o zebrane w toku badania dane w zakresie uzyskanych efektów. W jej ramach ustalane jest, czy interwencja osiągnęła zamierzone efekty, od czego to zależało i jakie na tej podstawie możemy wyciągnąć wnioski dla interwencji planowanych w przyszłości. Zaistnienie zakładanych efektów uzależnione jest od rodzaju realizowanej interwencji. Dla przykładu perspektywa zaistnienia docelowych efektów interwencji w przypadku programów szkoleniowych jest znacznie krótsza niż dla programów infrastrukturalnych lub badawczo- rozwojowych. Ważnym jest zatem wybranie odpowiedniego dla danej interwencji momentu realizacji ewaluacji ex post, tak aby istniała możliwość wiarygodnej oceny. Efekty muszą się bowiem zmaterializować zanim będziemy poddawać je analizie, ale też nie powinny zaniknąć, abyśmy mogli je zbadać.

Ewaluacja ex post dostarcza głównym interesariuszom, jak i ogółowi społeczeństwa informacji na temat efektów wdrażania interwencji publicznych. Jest zatem instrumentem wspierającym urzeczywistnienie zasad dobrego rządzenia, takich jak rozliczalność i przejrzystość działań władz publicznych. Wyniki ewaluacji ex post i sformułowane na tej podstawie wnioski i rekomendacje służyć powinny jednak przede wszystkim udoskonaleniu i lepszemu ukierunkowaniu innych, wdrażanych lub planowanych do wdrożenia programów, strategii i projektów.

## Jak badamy?

Zadaniem ewaluacji ex post jest praktyczna weryfikacja teorii interwencji. Polega ona na ocenie spodziewanych efektów (odpowiedź na pytanie, czy zakładane efekty zaistniały w wyniku interwencji) oraz wyjaśnieniu mechanizmów przyczynowo- skutkowych za nimi stojących (odpowiedź na pytanie dlaczego zakładane efekty wystąpiły lub nie). Ocenie zakresu i stopnia zaistnienia efektów służą przede wszystkim metody ilościowe, takie jak analizy danych statystycznych i monitoringowych, czy też badania ankietowe. Szczególnie użytecznym i pożądanym jest badanie efektów przyczynowych z wykorzystaniem metod kontrfaktycznych i grup kontrolnych. Wyjaśnieniu mechanizmów przyczynowo- skutkowych interwencji służą z kolei głównie metody jakościowe, takie jak wywiady pogłębione, badania focusowe, panele ekspertów czy studia przypadku (więcej informacji na temat metod i technik badawczych – zobacz rozdział 7 poradnika).

# Rozdział 3. Co jest potrzebne, aby przeprowadzić ewaluację i w jaki sposób możemy ją realizować?

Jacek Pokorski

## Wprowadzenie

Decyzja o przeprowadzeniu ewaluacji jest wypadkową wielu przesłanek, związanych przede wszystkim z potrzebami wiedzy w określonym obszarze działania instytucji, z gotowością na wykorzystanie wyników ewaluacji przez kierownictwo lub partnerów instytucjonalnych, ale także z potencjałem realizacyjnym instytucji (wewnętrznym lub możliwością wynajęcia ewaluatora).

Planowaniu zakresu badania od strony merytorycznej (m.in. określaniu celów i pytań badawczych) poświęcony jest Rozdział 4. W tym rozdziale skupimy się na zasobach potrzebnych do przeprowadzania ewaluacji. Trzeba jednak pamiętać, że merytoryka badania determinuje zasoby i kwestie operacyjne. Wskazówki przedstawione w tym rozdziale, co do planowania i wykorzystania zasobów, formułowane są zatem przy założeniu, że mamy już za sobą istotne ustalenia merytoryczne dotyczące potrzeb wiedzy i zakresu planowanego badania.



**Jednostka ewaluacyjna to struktura powołana i odpowiedzialna za realizację ewaluacji w ramach instytucji lub programu. Może nią być osoba, stanowisko wieloosobowe/zespół lub komórka organizacyjna w instytucji.**

## Wybór ekonomiczny w procesie planowania ewaluacji

Należy mieć na uwadze, że zakres badania oraz poszczególne typy zasobów które do jego zrealizowania angażujemy, są względem siebie w silnej zależności. Przykładowo, im szybciej chcemy otrzymać wyniki, tym większych środków finansowych potrzebujemy i/ lub lepszego zespołu. Z kolei im szerszy zakres badania, tym więcej czasu na realizację

należy zaplanować, uwzględnić wyższy budżet i zróżnicowany kompetencyjnie zespół. Tym samym planowanie badania przyjmuje charakter wyboru ekonomicznego, który powinien prowadzić do uzyskania racjonalnej równowagi pomiędzy potrzebami i zasobami, w sytuacji w której te pierwsze mogą być praktycznie nieskończone, a te drugie są ograniczone. Z tej perspektywy jednostka ewaluacyjna nie różni się od klasycznego podmiotu gospodarującego. Jako że nie jest możliwym „wyprodukowanie” odpowiedzi na tyle pytań badawczych, ażeby zaspokoić zapotrzebowanie wszystkich interesariuszy, pojawia się konieczność dokonywania wyboru tych, które decydujemy się dostarczać i tych z których dostarczenia rezygnujemy<sup>26</sup>.

W tym zakresie jednostki planujące ewaluację muszą być świadome zasobochłonności określonych oczekiwań i prowadzić w tym procesie efektywną komunikację. Komunikacja ta powinna być prowadzona z interesariuszami (inicjatorami potrzeb wiedzy, odbiorcami badania), ekspertami (własnymi pracownikami, zewnętrznymi doradcami) angażowanymi do procesu projektowania badania, i z rynkiem dostawców usług (badawczych, eksperckich), jeśli ewaluację dostarczyć ma ewaluator zewnętrzny. Z kolei gdy proces ewaluacji ma być prowadzony wewnętrznie (zarówno jego planowanie jak i wykonywanie badań), należy brać pod uwagę ryzyko nieutrzymania nakładów niefinansowych (np. czasu poświęconego na proces badawczo-analityczny, liczby angażowanych osób/osobogodzin – zarówno w zespole wykonawczym jak i innych osób angażowanych podczas spotkań/konsultacji/warsztatów).

Warto mieć także na uwadze, że na proces planowania ewaluacji może wpływać szereg czynników kontekstowych, które zniekształcają racjonalność dokonywanych wyborów. Takimi czynnikami mogą być presja czasowa na dostarczenie wyników (np. w obliczu pilnej potrzeby rozwiązania określonego problemu interwencji), ciekawość, ambicja badawcza i duży zakres oczekiwań merytorycznych (te nierzadko abstrahujące od kosztów, lub forsowane przy niewspółmiernie niskim budżecie). Czynniki kontekstowe mogą także prowokować różnego typu próby zastosowania nieprzemyślanej, czy mało wnikliwej metodyki, w oderwaniu od dorobku badań społeczno-ekonomicznych w tym zakresie i standardów ewaluacji, pod presją redukcji czasu i kosztów badania, prowadząc jednocześnie do spłykania badania lub eliminowania kluczowej problematyki. To ostatnie może być także powodowane innego rodzaju presją, polegającą na obawie części interesariuszy przed „zbyt dokładnym” zgłębieniem określonej problematyki interwencji przez zewnętrznych ewaluatorów.

Na początek zatem trzeba mieć świadomość co do możliwości występowania tego typu mechanizmów (zależności) i je odpowiednio zaadresować przy rozpisywaniu planu działań, jak również przeanalizować koszty i korzyści określonych wyborów. Trafne

<sup>26</sup> Por. [https://www.abc-ekonomii.net.pl/s/wybor\\_ekonomiczny.html](https://www.abc-ekonomii.net.pl/s/wybor_ekonomiczny.html)

zidentyfikowanie tych mechanizmów i odpowiednie wyjście im naprzeciw w procesie planowania i projektowania, uchroni nas przed realizacją badań o ograniczonej użyteczności (co za tym idzie – nieefektywnym wykorzystaniem zasobów). A co gorsze – także przed dostarczeniem błędnych przesłanek dla kierownictwa, bo opartych na potencjalnie niskiej jakości badaniach (np. anegdotycznych, fragmentarycznych, powierzchownych), dla decyzji o potencjalnie dużych kosztach finansowych (publicznych) i skutkach społecznych (co może dodatkowo zwielokrotnić straty).

## Przegląd zasobów a wybór modelu realizacji ewaluacji

Aby wybrać sposób prowadzenia ewaluacji – czy będzie to badanie wewnętrzne, zewnętrzne (zlecone wykonawcy) czy prowadzone w formule mieszanej (hybryda prac wewnętrznych i zleconych) – należy skonfrontować potrzeby z dostępnymi zasobami instytucji<sup>27</sup>. Warto podkreślić, że kluczowym kryterium w wyborze podejścia (wewnętrznego, zewnętrznego, hybrydowego) pozostaje dostępność zasobów niezbędnych do sprostania celom i zakresowi planowanego badania, a nie wymóg zapewnienia niezależności procesu<sup>28</sup>.

<sup>27</sup> Ten dylemat dotyczyć może nie tylko etapu realizacji ewaluacji – jeśli jednostka nie ma wystarczającego doświadczenia badawczego czy wiedzy eksperckiej (sektorowej, metodycznej), może sięgnąć po wsparcie zewnętrzne także w procesie projektowania (wynajęty ekspert, przygotowujący kompleksowy projekt badawczy do docelowego zlecenia przez jednostkę zewnętrznemu wykonawcy) i/lub asystę merytoryczną w trakcie procesu ewaluacji (nadzór ekspercki nad wykonawcą w imieniu jednostki, kontrola jakości procesu badawczego, wsparcie na etapie odbioru wyników/recenzji raportów itp.).

<sup>28</sup> Rozróżnienie ewaluacji wewnętrznych od zewnętrznych pod względem stopnia obiektywizmu dostarczanych wyników (rzetelności i niezależności w procesie wnioskowania), jak pokazuje praktyka badań ewaluacyjnych w Polsce, ma charakter pozorny. Przy odpowiedniej kulturze ewaluacyjnej instytucji i umieszczeniu w jej strukturze jednostki ewaluacyjnej (tj. zapewnienie jednostce nadzoru i możliwości działania niezależnego od pionów odpowiedzialnych za ocenianą interwencję), wykonanie ewaluacji jest możliwe co najmniej tak samo dobrze (zgodnie ze standardami) „siłami własnymi” instytucji, jak przy pomocy usługodawcy zewnętrznego. Należy podkreślić, że potencjalne ryzyko występowania oddziaływania instytucji na wyniki ewaluatora zewnętrznego może być nie mniejsze, o ile kultura organizacyjna instytucji czy poziom profesjonalizmu wykonawcy pozostawiają wiele do życzenia. Mowa tu – przykładowo – o takich zjawiskach jak skłonność ewaluatora do odstąpienia od formułowania krytycznych osądów wartości interwencji na rzecz spokojnego dopełnienia kontraktu bez komplikacji w procesie, które mogłyby zwiększyć koszty/pomniejszyć zyski, np. dodatkową serią konsultacji/poprawek raportu, *versus* skłonność pracownika jednostki/ewaluatora wewnętrznego do napisania raportu dyplomatycznie bądź nieporuszania problematyki, która mogłaby stać w negatywnym świetle instytucję i prowokować nieprzyjemności, np. dodatkowe wyjaśnienia dla kierownictwa, poddawanie w wątpliwość jakość wykonywanej pracy. Należy mieć na względzie, że w niektórych sytuacjach, wymóg niezależności ewaluatora (przynajmniej na poziomie funkcjonalnym) może być narzucony prawnie. Ma to miejsce np. dla badań realizowanych w ramach funduszy UE.

Prowadzenie ewaluacji wiąże się z określonymi nakładami, przeliczalnymi na jednostki pieniężne. W niektórych instytucjach (w zależności od specyfiki np. dostępności funduszy zewnętrznych, uwarunkowań strategicznych) łatwiej zagwarantować budżet na tzw. usługi obce, pozwalające zamówić badania ewaluacyjne, niż zatrudnić pracowników stanowiących personel jednostki ewaluacyjnej (stanowiska eksperckie/specjalistyczne). W innych instytucjach może być odwrotnie – nie ma problemu ze zwiększeniem zatrudnienia, a wygospodarowanie w roku budżetowym kilkuset tys. zł na zlecenie badań, pozostaje poza zasięgiem. Powyższe może być również warunkowane określonymi zadaniami/projektami realizowanymi w instytucji (np. na ewaluację programów unijnych są dostępne etaty i budżet w jednostce; na ewaluację programów krajowych, czy ocenę skutków regulacji ustawy dot. tej samej instytucji, do dyspozycji jednostki nie ma ani jednego ani drugiego). Wówczas dobrym rozwiązaniem może być podejście hybrydowe, ograniczające koszty usług zewnętrznych. Można to zrobić np. poprzez przeprowadzenie samodzielnie w jednostce części prac analitycznych na dokumentach i danych zastanych, samodzielne zaprojektowanie i wykorzystanie narzędzi badawczych (np. projektu kwestionariusza ankiety internetowej i zebranie za jego pomocą danych od uczestników programu z wykorzystaniem techniki CAWI). Z kolei zlecić można wykonawcom/ekspertom tylko bardziej złożone badania terenowe i prace analityczne (np. realizację zogniskowanych wywiadów grupowych, ankietowanie innymi technikami w obrębie niedostępnych on-line podgrup uczestników, studia przypadku z wizytami i wywiadami u beneficjentów oraz opracowanie w oczekiwanej formie wyników tego typu badań ilościowych i jakościowych). Ponadto, warto zwrócić uwagę, że aktualne możliwości prowadzenia wywiadów (indywidualnych, grupowych) i różnego typu spotkań konsultacyjnych *on-line* z wykorzystaniem komunikatorów internetowych, są korzystne dla rozwoju modelu ewaluacji wewnętrznej, realizowanej co do zasady „zza biurka”, bez konieczności podróżowania czy angażowania struktur terenowych.

Warto wspomnieć, że ewaluacja wewnętrzna, choć jest istotnym obciążeniem zasobów kadrowych (zwłaszcza w przypadku jednostek ewaluacyjnych realizujących kilka dużych badań rocznie), ma także niebagatelne znaczenie dla rozwoju kompetencji badawczo-analitycznych pracowników jednostki. W instytucjach, w których jednostka ewaluacyjna zleca na zewnątrz wszystkie lub zdecydowaną większość prac ewaluacyjnych, pracownicy odpowiedzialni za ewaluację rozwijają (umacniają) z kolei głównie kompetencje z zakresu zamawiania usług, zarządzania kontraktem z wykonawcą, nadzorowania jego prac. W dłuższej perspektywie kompetencje badawczo-analityczne jednostki mogą być zastępowane przez te administracyjne. Zatem w zakresie wyboru modelu realizacji, również warto rozważyć odpowiedni balans pomiędzy ewaluacjami wewnętrznymi a zewnętrznymi<sup>29</sup>.

<sup>29</sup> Należy pamiętać przy tym o przedstawionym potencjale badań on-line, oraz coraz większej ilości danych w zasobach publicznych (danych administracyjnych gromadzonych przez różnego typu rejestry publiczne i danych typu big data, rejestrowanych przez liczne systemy monitorowania). Dodatkowe znaczenie ma fakt, że



Należy stwierdzić, że ewaluacje o charakterze strategicznym, obejmujące złożony zakres przedmiotowy i rozbudowaną problematykę (np. ewaluacje strategii ogólnopolskich lub sektorowych; badania kwestii horyzontalnych lub zakresów dot. kilku programów), pozostają niezwykle trudne, w praktyce niewykonalne w modelu wewnętrznym. Zarówno rekonstrukcja logiki kompleksowej interwencji, przygotowanie i prowadzenie badań odnośnie do niej (w wielu obszarach, na różnych grupach docelowych interwencji), jak i formułowanie osądu wartości określonych jej wymiarów, na podstawie zebranych obserwacji i analiz, wymaga wiedzy eksperckiej szerszego zespołu wykonawczego i odwoływania się nierzadko do dorobku naukowo-badawczego w danej dziedzinie/sektorze<sup>30</sup>. Stąd dla tego typu ewaluacji, lepszym rozwiązaniem są zlecenia zewnętrzne w zakresie kompleksowej realizacji ewaluacji.



### Zapamiętaj

**Im więcej zaplanowanych badań terenowych, im trudniejsza metodyka ewaluacji, im bardziej skomplikowane podstawy teoretyczne dot. problematyki w obszarze przedmiotowym badania, tym bardziej uzasadnione jest skorzystanie ze wsparcia ekspertów zewnętrznych. Jednakże poziom skomplikowania uzasadniający model zewnętrzny, może być różny dla różnych organizacji.**

Ewaluacje wewnętrzne mogą sprostać potrzebom dostarczania wiedzy z obszaru pojedynczego projektu, w który jest zaangażowana instytucja (np. działanie o charakterze systemowym, nie-konkursowym) lub inicjatywy o mniejszej skali (np. pilotaże, wydarzenia, serie aktywności terenowych/*on-line* prowadzonych przez instytucję itp.). Wówczas należy mieć na uwadze, że pracownik jednostki (ewaluator wewnętrzny) oddelegowany do takiego zadania, powinien mieć zapewniony odpowiedni komfort czasu i uwagi na

część zasobów danych, pochodzących z tego typu systemów, rejestrujących zachowania osób (np. korzystanie z różnego typu dóbr/usług, mobilność), które pracownicy jednostek ewaluacyjnych posiadających odpowiednie kompetencje analityczne mogliby wykorzystywać w prowadzonych pracach badawczych w kontekście interwencji, mogą nie być możliwe do przekazania podmiotom trzecim (wykonawcom ewaluacji). Wrażliwość tych danych (np. korzystanie z usług zdrowotnych, publiczny monitoring ruchu itp.) lub ograniczenia wynikające ze zgód osób, których dane dotyczą mogą to uniemożliwiać.

<sup>30</sup> W tego typu badaniach, wykorzystywana bywa także metodyka, której zastosowanie pozostaje zazwyczaj poza zasięgiem potencjału wewnętrznego jednostki np. przygotowanie analiz ekonometrycznych z wykorzystaniem modelu makroekonomicznego dużej skali, analizy porównawcze/kontrfaktyczne z grupą kontrolną, badania terenowe ogólnopolskie, lub w miejscu docelowym dla projektu (tj. badanie użytkowników/pasażerów nowego taboru, zakupionego w ramach projektu dot. inwestycji w obszarze transportu publicznego), badania panelowe i delfickie angażujące szerokie grono ekspertów itp. Wówczas skala zgromadzonych w procesie danych, pozostaje trudna do opracowania przez jednostkę. Także zarządzanie (pod względem zasobochłonności, terminowości, sprawności prac) takim kompleksowym projektem badawczym, w warunkach struktur i procedur administracyjnych, pozostaje wyzwaniem.

powierzone zadanie, tj. w ograniczonym stopniu powinien być włączany w inne inicjatywy/procesy administracyjne czy koordynowanie ewaluacji zewnętrznych w jednostce.

Niezależnie od wyboru modelu realizacji ewaluacji (wewnętrzna, zewnętrzna, hybrydowa), musimy zapewnić określone zasoby, które są konieczne do wykonania ewaluacji (w odpowiedniej proporcji - przez jednostkę ewaluacyjną i wykonawcę). Tymi zasobami są:

- Kompetencje (wiedza, umiejętności, *know-how*)
- Źródła wiedzy (dane, informatorzy)
- Czas (harmonogram prac badawczych, termin wdrożenia wyników)
- Zaplecze organizacyjno-techniczne (infrastruktura IT, licencje, technologia)

Czasami wskazywany jest jeszcze jeden konieczny zasób - budżet. Jest on różnie rozumiany. W zależności od wybranego modelu ewaluacji może przyjmować w ramach jednostki ewaluacyjnej postać finansową (koszty usług obcych) lub pozafinansową (wynagrodzenia/etaty pracowników jednostki, ujmowane w ramach kosztów funkcjonowania instytucji), jest to zatem zasób wtórny, względem pozostałych (pierwotnych).

## Kompetencje (wiedza, umiejętności, *know-how*)

W zależności od uwarunkowań, kompetencje ewaluacyjne mogą oznaczać kompetencje w zakresie przygotowania zamówienia, współpracy z ewaluatorem zewnętrznym i ostatecznie upowszechnienia i wykorzystania wyników badania, lub obejmować także etapy obserwacji, analizy i oceny - w przypadku modelu wewnętrznego.



### Zapamiętaj

**Zastanów się ile osób potrzeba do wykonania danej ewaluacji. Jedna osoba nie wykona nawet względnie prostego badania w 1 miesiąc, jeśli będą tam zaplanowane np. badania terenowe. Z kolei zamawianie i współpraca z wykonawcą, to także istotnie angażująca praca dla pracownika jednostki. Jedna osoba nie powinna zajmować się 5 badaniami w tym samym czasie.**

W ostatnim czasie szereg trendów (m.in. upowszechnienia zdalnych form pracy i komunikacji z respondentami – jako uboczny skutek pandemii Covid-19, otwieranie danych publicznych, korzystanie z technologii generatywnej sztucznej inteligencji) wprowadził

rewolucyjne zmiany w zakresie kompetencji, które mogą być wykorzystane na potrzeby ewaluacji. Zmiany te z kolei istotnie wpływają na sam przebieg procesu ewaluacji – silniej obecnie wykorzystujący komunikację zdalną, wtórne wykorzystanie danych zastanych i dorobek technologii cyfrowych. Powszechne stosowanie technik badań *online*, pozyskiwanie i analiza dużych zbiorów danych budują zapotrzebowanie na nowe kompetencje. Wraz z rozpowszechnieniem pracy zdalnej pojawiło się również zapotrzebowanie na zmienione kompetencje w zakresie zarządzania zespołem oraz nowymi ryzykami. Kompetencje w tym zakresie muszą być rozwijane i podążać za nowymi możliwościami, które pozwalają przezwyciężyć z kolei nowe ograniczenia w badaniach (np. niechęć respondentów do udziału w wywiadach/ankietach, wzrost wartości prywatności i konieczność zapewnienia poufności). Osoby mające pełnić określone role w procesie ewaluacji, w świetle tych nowych trendów wynikających z transformacji cyfrowej, powinny jednocześnie intensywnie inwestować w podnoszenie kompetencji (zwłaszcza w obszarze cyfryzacji, dotychczas wykonywanych procesów/czynności lub ich przeprojektowania), aby pozostać produktywnymi.

Natomiast podstawowe profile kompetencyjne i role osób, które są potrzebne do obsługi procesu ewaluacji (o ile jednostka planuje realizację badań ewaluacyjnych siłami własnymi), pozostają względnie stałe:

- 1) badawcze (np. badacze społeczno-ekonomiczni o doświadczeniu/specjalizacji zbieżnej z przedmiotem dziedziny lub specyfice prowadzonych badań w jednostce/zespole wykonawczym – metodycy i specjaliści; badacze terenowi, moderatorzy, ankieteryzy, koordynatorzy, rekruterzy/asystenci – jeśli dotyczy);
- 2) analityczne (np. analitycy jakościowi; analitycy ilościowi – statystycy i analitycy danych; inżynierowie danych/*data scientists*, inżynierowie SI, ekonometrycy – jeśli dotyczy);
- 3) sektorowe/dziedzinowe (np. konsultanci; osoby z dorobkiem naukowo-badawczym i autorytetem w dziedzinie, w tym posiadający stopień/tytuł naukowy – jeśli dotyczy);
- 4) wspierające np. osoby odpowiedzialne za komunikację (strony www, kontakt z mediami, media społecznościowe, w tym utrzymanie uwagi/zainteresowania/aktualności); osoby odpowiadające za postępowania o udzielenie zamówienia na usługi ewaluatora zewnętrznego – jeśli dotyczy; osoby odpowiedzialne za utrzymanie infrastruktury i licencji IT; asystenci odpowiedzialni za dokumentację, redakcję techniczną, dostępność cyfrową utworów (WCAG) – jeśli dotyczy itp.;
- 5) związane z zarządzaniem (np. osoby odpowiedzialne za kierowanie jednostką i/lub zespołem wykonawczym; menadżerowie projektu odpowiedzialni za gospodarowanie zasobami, zarządzanie procesem badawczym, dostarczenie wyników klientowi/zamawiającemu – jeśli dotyczy).

Warto podkreślić, że w mniejszych jednostkach ewaluacyjnych (i/lub przy realizacji badań o mniejszej skali), ww. profile kompetencyjne i role w procesie ewaluacji nierzadko będą musiały być łączone przez jedną osobę.

## Źródła wiedzy (dane, informatorzy)

Dostępność zasobów, jakimi są odpowiednie źródła wiedzy, decyduje o tym czy ewaluacja jest w ogóle wykonalna, tj. czy jest możliwe udzielenie odpowiedzi na postawione pytania badawcze. Zaznaczyć należy, że nie chodzi tu o jakiegokolwiek źródła, ale takie które dostarczą informacji odpowiedniej jakości<sup>31</sup> i których pozyskanie będzie efektywne.



### Zapamiętaj

**Źródła wiedzy to informacje pochodzące ze źródeł zastanych i danych, które mogą być zebrane od informatorów.**

Powyższe zastrzeżenie wskazuje na konieczność weryfikacji wstępnie wytypowanej listy pytań badawczych (redukcji, zastąpienia innymi). W tym przypadku będzie to weryfikacja, czy dane źródła wiedzy umożliwią dostarczenie rzetelnej odpowiedzi i czy dotarcie do nich będzie opłacalne (wykonalne przy określonym budżecie zamówienia, lub nakładzie pracy własnej jednostki). Ponadto, na etapie planowania ewaluacji warto zweryfikować potencjalne koszty pozyskania danych także pod kątem realnej możliwości skorzystania z nich w procesie analizy i oceny. Nadmiar danych lub ich nieprzystępna forma, zwłaszcza tych zastanych (nieustrukturyzowanie, forma mieszana cyfrowo-analogowa, fragmentaryczność), może wpływać na wzrost kosztów i nakładów pracy na późniejszych etapach ewaluacji.

W badaniach wykorzystujemy dwa rodzaje danych:

1. Dane pierwotne, które są wywoływane specjalnie, aby zgromadzić wiedzę pomagającą przy udzielaniu odpowiedzi na postawione pytania badawcze. Wywoływaniu takich danych służą różnego typu metody badań społecznych. Polegają one zwykle na angażowaniu wytypowanych informatorów do udzielenia informacji w zaplanowany metodycznie sposób, albo prowadzeniu bezpośrednich obserwacji – więcej o metodach i technikach badań w rozdziale 7.

<sup>31</sup> Informacje powinny być rzetelne, istotne i aktualne – to kluczowe kryteria selekcji źródeł informacji, którymi należy się kierować w ewaluacji.

2. Dane wtórne pozyskiwane są ze źródeł zastanych. To dane tworzone w innych celach niż przedmiotowa ewaluacja, ale powiązane z przedmiotem badania. Mogą one obrazować kontekst dla realizowanej interwencji (np. odnoszą się do obszaru lub grup docelowych, jednostek niepoddanych oddziaływaniu), zachowania uczestników programu (wg określonych cech, w zależności od typu wsparcia itp.), przebieg programu na osi czasu (rzeczowy, finansowy) i tym samym stanowić wkład do odpowiedzi na część pytań ewaluacyjnych. Dane wtórne to m.in. informacje pochodzące ze statystyki publicznej, a także z rozmaitych dokumentów (programowych, regulacji różnego szczebla – np. przepisy prawa, wytyczne, regulaminy, procedury) i opracowań (raporty ewaluacyjne, monitoringowe, dorobek naukowo-badawczy dot. problematyki interwencji).

Dane pierwotne są z założenia lepiej dopasowane do celów przedmiotowego badania (sposób ich generowania jest od początku podporządkowany temu badaniu). Jednocześnie ich niewątpliwymi wadami są m.in. wysokie koszty realizacji badań społecznych, w tym wynikające z ograniczonej dostępności niektórych grup respondentów, będących docelowymi dla interwencji, takich jak osoby starsze, osoby z niepełnosprawnościami, osoby bezrobotne, bezdomne, imigranci (o ograniczonej mobilności, kompetencjach cyfrowych, z różnego typu ograniczeniami), przedsiębiorcy/eksperci/menedżerowie/liderzy itp. Wysokie koszty generują np. sondaże na odpowiednio licznych próbach reprezentatywnych lub przekrojowe studia przypadków oparte na badaniach jakościowych. Ponadto, w świetle powszechnego zalewu informacyjnego, troski o prywatność i dane osobowe, ale i obaw względem różnego typu działań akwizycyjnych i niechcianych wiadomości (spam), oraz obaw związanych z możliwością podszywania się innych osób i ich złych intencji, rekrutacja respondentów i realizacja badań społecznych staje się niezmiernie trudna. Trudności te z kolei prowadzić mogą do zniekształcania prób i nieuprawnionego wnioskowania na przypadkowych zbiorowościach (lub wręcz wyciągania wniosków ze specyficznych przypadków, które zgodziły się wziąć udział w badaniu).

Należy podkreślić, że potencjał wtórnych źródeł informacji systematycznie rośnie. Sprzyjają temu wyżej omówione ograniczenia badań pierwotnych, oraz liczne – utrzymywane, stale rozwijane i nowo tworzone – systemy monitorowania i rejestrowania aktywności obywateli, w tym systemy usług (publicznych i prywatnych). W przypadku danych wtórnych mamy do czynienia z różnymi źródłami. Ich zasoby danych mogą w niewielkim stopniu być dopasowane do przedmiotowego badania. Ich niewątpliwym walorem jest to, że nie wymagają wywoływania *ad hoc* na potrzeby badania, a jedynie odpowiedniego zidentyfikowania, wyselekcjonowania i przetworzenia – w celu dopasowania do celów przedmiotowego badania. Najczęściej dane, które pochodzą ze struktur odpowiedzialnych za realizację programu, są dobrej jakości (rzetelne, istotne i aktualne). Do takich danych należą np. rejestr aplikacji/bazy danych pochodzących z systemu monitorowania beneficjentów w trakcie ich udziału w programie, w tym rozliczeń uzyskanej pomocy/ dane sprawozdawcze, w tym raporty okresowe przygotowywane przez beneficjentów/

instytucje, zestawienia ad-hoc. Jednak nie zawsze są to wystarczające źródła do analizy skutków interwencji. Przez taką analizę rozumiemy zbadanie reakcji uczestników programu na działanie bodźca w postaci interwencji, w kontekście teoretycznej zmiany, która miała zostać nią wywołana. Prowadzenie tego typu analiz wymaga ewentualnego sięgnięcia po dane, które obrazują właśnie owe zmiany dotyczące obiektów poddanych interwencji lub grup/obszarów docelowych w szerszym kontekście teorii programu.

Często prowadzone rejestry publiczne gromadzą dane, które mogłyby być wykorzystane jako rezultat interwencji. Przykładowo mogłyby to być rejestry dot. składek na ubezpieczenia społeczne ZUS – jako dane o zmianie sytuacji osób na rynku pracy, czy też rejestr przedsiębiorców GUS i sprawozdań finansowych/podatkowych KRS/MF – jako dane o zmianie sytuacji ekonomicznej przedsiębiorców<sup>32</sup>. Wartościowym producentem danych od lat są służby statystyki publicznej (Główny Urząd Statystyczny, wraz z siecią wojewódzkich urzędów statystycznych). Gromadzą one dane w ramach wieloletnich badań spisowych i reprezentatywnych (na dużych próbach) w ramach corocznych Programów Badań Statystycznych Statystyki Publicznej. Badania te dotyczą wielu obszarów społeczno-gospodarczych, w tym nierzadko powiązanych z interwencją, dostarczając danych (statystyk) na dużym poziomie szczegółowości.

Poza wspomnianymi źródłami, do dyspozycji ewaluatorów pozostaje także szereg źródeł dostępnych na stronach internetowych i/lub u prywatnych gestorów (np. dane z systemu nawigacji GoogleMaps, statystyki stron www GoogleAnalytics, dane pochodzące od operatorów różnego typu sensorów/detektorów, systemów monitoringu ruchu itp.). Nierzadko dane te mają charakter tylko częściowo ustrukturyzowany i są zbiorami dużej objętości (por. dane „wydrapane” ze stron internetowych, ang. *web scraping*; zasoby pochodzące z różnego typu rejestratorów pracujących w trybie ciągłym). Przy odpowiednich technikach pobierania i przetwarzania takie dane mogą być wartościowym i darmowym źródłem informacji w ewaluacji. Należy jednak podkreślić, że korzystanie z tego typu danych pozostaje wyzwaniem – zarówno od strony prawnej (np. gdy dane dotyczą osób lub gdy dane choć wystawione publicznie, pozostają własnością prywatną, zastrzeżoną pod względem celów wykorzystania), technicznej (gdy dane wystawione na stronach internetowych są zabezpieczone przed pobieraniem



### Zapamiętaj

**Dane otaczają nas zewsząd. Trzeba tylko je poprawnie zidentyfikować, upewnić się co do legalności ich w korzystania w ewaluacji, pozyskać i przetworzyć.**

<sup>32</sup> Część publicznych rejestrów udostępnia gromadzone dane, z wykorzystaniem aplikacji webowej (API) do której zarejestrowany użytkownik pod pewnymi warunkami może się podłączyć i pobrać dane w wystawionym przez gestora zakresie (por. inicjatywa Otwieramy dane publiczne – <https://www.gov.pl/web/cyfryzacja/otwieramy-dane-publiczne--sejm-za>).

zautomatyzowanym), jak również od strony analitycznej (docelowe ustrukturyzowanie i przetwarzanie dużych zbiorów danych, wymaga wysokich kompetencji analitycznych typu *data science*, dużych nakładów pracy i infrastruktury IT, zapewniającej odpowiedni poziom bezpieczeństwa i szybkości procesowej).

## Czas (harmonogram prac badawczych, termin wdrożenia wyników)

Zaplanowanie czasu jako zasobu potrzebnego do realizacji badania ewaluacyjnego można rozpatrywać w trzech wymiarach. Pierwszy dotyczy zakresu czasowego interwencji, która zostaje objęta badaniem. Drugi, to wyzwanie związane z dostarczeniem wiedzy z ewaluacji w momencie kiedy jest ona faktycznie potrzebna, np. do wsparcia podejmowania decyzji dot. interwencji. I wreszcie trzeci wymiar operacyjny, który dotyczy harmonogramu prac, związanych z samym procesem ewaluacji.

1. Pierwszy wymiar przekłada się na definiowany przez jednostkę ewaluacyjną zakres czasowy badania (np. efekty programu w okresie funkcjonowania ograniczeń w kontaktach społecznych, związanych z epidemią i zagrożeniem epidemicznym, tj. w okresie 14 marca 2020 – 30 czerwca 2023 r.), co ma przełożenie na uszczegółowienie zestawu pytań badawczych i zapotrzebowania na dane, które powinny być wykorzystane w badaniu.
2. Aby sprostać potrzebom użytkowników, powinniśmy przed rozpoczęciem ewaluacji, rozpoznać dokładne uwarunkowania czasowe dla wykorzystania wyników badania, tj. na kiedy są potrzebne wyniki badania (termin optymalny *versus* termin krytyczny) a następnie skonfrontować je z etapami prac (ich harmonogramem), które wiążą się z wykonaniem ewaluacji. Jeśli wykonanie zaplanowanych prac w tych terminach jest wątpliwe, warto rozważyć ich ograniczenie (np. rezygnacja z wybranych pytań badawczych, z czasochłonnej metodyki badań), prowadzenie działań równoległe (w tym zwiększenie zasobów kadrowych) lub zastąpienie wybranych elementów badania innymi, mniej czasochłonnymi a równie satysfakcjonującymi pod względem ich alternatywnej wartości dodanej. Gdy dostarczenie wyników w terminie, nawet w okrojonym zakresie, jest niemożliwe, warto zrezygnować z przedmiotowej ewaluacji.
3. Harmonogram ewaluacji powinien uwzględniać wszystkie etapy procesu, począwszy od planowania, poprzez projektowanie i realizację po wykorzystanie wyników.
  - Jeśli mamy do czynienia z ewaluacją zewnętrzną (lub z angażowaniem na którymkolwiek z ww. etapów ekspertów zewnętrznych), należy wkalkulować także czas wymagany na zakontraktowanie wykonawcy (tj. przygotowanie

i poprowadzenie postępowania o udzielenie zamówienia, zawarcie umowy). Trzeba mieć na uwadze, że zamówienie realizowane w świetle Prawa zamówień publicznych (najczęściej w trybie przetargowym) ma swoje specyficzne uwarunkowania czasowe. Jeśli jednostka ewaluacyjna przygotowuje projekt badania (w formie Opisu przedmiotu zamówienia) i istotne postanowienia umowy oraz specyfikacji warunków zamówienia, należy dodatkowo uwzględnić czas na prace innych jednostek, odpowiedzialnych w instytucji za zamówienia i finanse, oraz osób akceptujących zakres zadań zleczanych:

- sformułowanie docelowego wzoru umowy, wraz dokumentacją przetargową; kontrola dokumentów od strony formalno-prawnej i finansowej; zatwierdzenie dokumentów (minimum 1 miesiąc),
- upublicznienie zamówienia i składanie ofert (minimum 0,5 miesiąca);
- ocena ofert, uzupełnienie dokumentów dot. spełnienia warunków zamówienia, w tym prace komisji oceniającej, biura zamówień i wykonawcy (1-2 miesiące, w zależności od liczby ofert, skomplikowania dokumentów/specyfiki zamówienia)
- rozstrzygnięcie postępowania i zawarcie umowy, w tym zatwierdzenie dokumentów, przygotowanie projektu umowy i podpisanie jej przez strony (0,5–1 miesiąc).

Uwzględniając powyższe, realnie na pozyskanie ewaluatora zewnętrznego w ramach procedur zamówień publicznych, należy zarezerwować między 3 a 4,5 miesiąca<sup>33</sup>.

- Jeśli chodzi o czas, który należy zaplanować na samą realizację badania, sprawa nie jest prosta i zależy od wielu czynników. W zależności od modelu w jakim realizowane jest badanie (wewnętrzne, zewnętrzne, hybrydowe), jego zakresu merytorycznego (założona liczba pytań badawczych) i metodycznego (zaplanowany zestaw metod/technik badawczych i analiz) oraz zasobów jakie zostały na badanie alokowane (zespół kompetentnych osób, dostępne dane, budżet zamówienia – jeśli dotyczy), intensywność prac i ich czasochłonność może się rozkładać w różny sposób. Na to nakłada się także sezonowość (okres wakacyjny – w tym niedostępność respondentów i części zespołu realizującego jak też okres kumulacji prac w ramach zleceń sektora publicznego, w IV kwartale roku), która również powinna być uwzględniona w harmonogramie prac lub ryzykach potencjalnych opóźnień. Co do zasady, duże

<sup>33</sup> Warto pamiętać, że w zapisach umowy zawartej w wyniku przetargu (krajowego trybu podstawowego lub unijnego przetargu nieograniczonego) można zaplanować możliwości jej przedłużenia, gdyby wystąpiły uzasadnione, przewidziane na etapie przetargu, okoliczności. Warto też wziąć pod uwagę fakt, że w zamówieniach publicznych od niedawna nie jest możliwe określanie terminów wg dat kalendarzowych, a jedynie w ujęciu długości maksymalnego okresu realizacji – liczby miesięcy, tygodni lub dni – od dnia zawarcia umowy. Biorąc pod uwagę uwarunkowania, związane z optymalnym lub krytycznym terminem wykorzystania wyników ewaluacji, należy bardzo rygorystycznie kontrolować przebieg tych procesów, aby nie wywołały „efektu domina” w całościowym harmonogramie procesu ewaluacji.



- badania, obejmujące szeroki zakres przedmiotowy (np. duży program ogólnopolski, cele strategiczne/priorytety, kwestie horyzontalne) i prowadzone przez ewaluatora zewnętrznego, powinny być realizowane co najmniej 6–9 miesięcy (od podpisania umowy do przedstawienia projektu raportu końcowego). Analogicznie, ten sam okres realizacji może dotyczyć średniego badania, wykonywanego w jednostce w modelu wewnętrznym lub hybrydowym. Z kolei, na realizację mniejszych ewaluacji tematycznych (w przypadku zewnętrznych – chodzi o badania dotyczące względnie wąskiej problematyki lub pojedynczego schematu wsparcia/działania; w przypadku wewnętrznych, realizowanych przez oddelegowanego ewaluatora/pracownika jednostki – mowa tu o badaniach dotyczących jednego projektu/konkursu/pilotażu) zwykle powinny wystarczyć 3–4 miesiące. Należy mieć na względzie, że z punktu widzenia zewnętrznego wykonawcy faktycznie kluczowy czas dla realizacji badania to okres między akceptacją raportu metodycznego/wstępnego, a przekazaniem projektu raportu końcowego. Okres ten musi być wystarczający na faktyczną realizację badania (co najmniej 3 miesiące w przypadku mniejszych badań i 5–7 miesięcy w przypadku dużych i kompleksowych ewaluacji).
- Na długość procesu silnie wpływa skala badań pierwotnych (w tym wielkość próby i specyfika dostępności respondentów), udział badań jakościowych, które dotychczas wymagały czasochłonnej analizy (w tym transkrypcji i kodowania), przebieg uzgodnień dot. szczegółów metodycznych, w tym służących dopracowaniu narzędzi badawczych (np. przeprowadzenie konsultacji w kilku iteracjach, pilotaż), wypracowanie decyzji analitycznych (np. wybór zmiennych kategorialnych/podgrup, w odniesieniu do których prezentowane będą wyniki ogółem; ustalenia dot. doboru ewentualnej próby kontrolnej w analizie kontrfaktycznej), czas potrzebny na uzyskanie dostępu do danych zastanych (jednostkowych, statystycznych), bieżąca interakcja z interesariuszami, w tym dodatkowe konsultacje w procesie badawczym (np. metodyki, wyników cząstkowych itp.).
  - Na etapach poprzedzającym realizację badania i następującym po niej, zwykle prowadzone są konsultacje z interesariuszami. Zazwyczaj dotyczą one całego projektu badania (opisu przedmiotu zamówienia lub koncepcji badania – w przypadku ewaluacji wewnętrznych) oraz raportu końcowego wraz z rekomendacjami. Biorąc pod uwagę potencjalną czasochłonność tego elementu procesu – w zależności oczywiście od liczby zaangażowanych w niego instytucji i interesariuszy różnych szczebli – należy również uwzględnić go w harmonogramie. Z reguły na konsultacje projektu badania wystarczy 1–1,5 miesiąca (w tym czas na uwzględnienie poprawek/uzupełnienia projektu). Z kolei konsultacje raportu mogą być procesem iteracyjnym i trwać od 1 do 3 miesięcy (w tym czas na uzgodnienie rekomendacji z ich adresatami i poprawki/uzupełnienia raportu). Jeśli współpracujemy z ewaluatorem zewnętrznym, warto zaplanować w kontrakcie jego dyspozycyjność na etapie konsultacji (np. w ramach 3 miesięcznej asysty porealizacyjnej) i gotowość na wprowadzenie udoskonaleń raportu w ich wyniku.

- Za przygotowanie i pilnowanie harmonogramu (w tym jego aktualizację) samego badania odpowiada kierownik projektu w zespole wykonawczym. W przypadku ewaluacji zewnętrznych jest on zwykle dość szczegółowy, opracowywany w formie wykresu GANTT-a i pozwala także zamawiającemu na nadzór nad realizowaną umową i postępem prac. Z kolei w przypadku ewaluacji wewnętrznych (co do zasady działań o mniejszej skali) częściej praktykuje się w jednostkach ustalanie i kontrolowanie postępu, za pomocą kamieni milowych przewidzianych prac ewaluatora (np. z wykorzystaniem narzędzi IT, wspomagających zarządzanie projektem w jednostce).

## Zaplecze organizacyjno-techniczne (infrastruktura IT, licencje, technologia)

Abstrahując od podstawowej infrastruktury, jaką stanowi biuro (które coraz częściej ma charakter wirtualnej platformy do współpracy) należy także wspomnieć o innych zasobach organizacyjno-technicznych, które są niezbędne w pracy jednostki ewaluacyjnej/ewaluatorów.

1. Infrastruktura IT. Jakkolwiek nie wydawałoby się to oczywiste, praca ewaluatora na odpowiednim sprzęcie komputerowym (z Internetem o przepustowości pozwalającej na prowadzenie wideokonferencji oraz pobieranie dużych zbiorów danych; z kilkoma dużymi ekranami – do pracy na kilku dokumentach/arkuszach danych jednocześnie; z podzespołami pozwalającymi na szybkie operacje na dużych dokumentach i zbiorach danych itp.) w bezpiecznym środowisku (zarówno lokalnie jak i podczas współpracy w chmurze i na fizycznych dyskach sieciowych) gwarantującym efektywność pracy i trwałość gromadzonych zasobów (w tym danych podlegających szczególnej ochronie i wytworzonych produktów cyfrowych), pozostaje zasobem kluczowym.
2. Narzędzia badawcze i analityczne. Dziś znaczna część procesów badawczo-ewaluacyjnych może być wykonywana „zza biurka”, stąd zapewnienie dostępu do podstawowych narzędzi zdecydowanie może usprawnić te prace (przy założeniu, że wraz z dostępem do tych rozwiązań, zatroszczymy się o dostosowanie kompetencji do korzystania z nich). Do tych narzędzi należą m.in. programy do przygotowywania i realizacji badań typu CAWI (ankietowania internetowego, ale są to także narzędzia użyteczne do udostępnienia adresatom różnego typu formularzy *online* np. rekrutacyjnych, do głosowania, raportowania wpływu/trwałości itp.), programy do statystycznej analizy danych, kompleksowe środowiska do pracy z zbiorami danych (narzędzia typu *business intelligence*), programy do analizy danych jakościowych

(analizy transkrypcji wywiadów, dokumentów i różnego typu opracowań), programy do wizualizacji danych, programy do pracy grupowej/warsztatowej (narzędzia typu „elektroniczny flipchart”). Poza badawczo-analitycznym oprogramowaniem, niezastąpiony pozostaje typowy pakiet biurowy, zapewniający ewaluatorowi narzędzia do przetwarzania tekstu, pracy z arkuszami kalkulacyjnymi, tworzenia prezentacji. Warto podkreślić, że część z tych narzędzi i wiele innych (np. wspierających zarządzanie) może być dostępnych bezpłatnie, jednakże zazwyczaj wiąże się to z określonymi ograniczeniami lub nie gwarantuje poufności dla wprowadzonych do nich treści.

Poza tradycyjnym zestawem powyższych narzędzi w warsztacie ewaluatora, warto wspomnieć o możliwościach technologii generatywnej sztucznej inteligencji ([Chat GPT4](#) lub [równoważne](#)). Może ona obecnie efektywnie wspomagać prace na każdym etapie procesu ewaluacji (od projektowania badania, poprzez proces badawczy np. *desk research* i analityczny, po opracowywanie wyników w formie syntetycznej, raportowej czy prezentacyjnej). Coraz więcej tych rozwiązań, doskonale radzi sobie z zapytaniem i źródłami w języku polskim, co sprawia, że stają się one rewolucyjnymi narzędziami w pracach ewaluatora i zmieniają sposób prowadzenia badań (przebieg procesu konceptualizacji, badań i analiz, raportowania) przy wspomaganie AI. Należy jednakże mieć na uwadze, że korzystanie z tego typu usług, powinno uwzględniać regulacje dot. poufności i ochrony danych osobowych, którym podlega instytucja i/lub stosuje w umowach z ewaluatorami zewnętrznymi. Dotychczas dostępne usługi oparte na AI są najczęściej rozwijane („uczą się”) uwzględniając m.in. treści i zapytania, które użytkownicy im przekazują („karmią”). Oznacza to, że zasoby (informacje, treści, dokumenty, dane) udostępniane przez użytkowników w ramach publicznych narzędzi AI, mogą zostać trwale przechwycone, a ich dalsze przetwarzanie wykraczać może poza cele statutowe instytucji/warunki kontraktu z ewaluatorem. Z pewnością korzystanie z tego typu narzędzi będzie coraz powszechniejsze, a ww. ograniczenia stymulować będą powstawanie dedykowanych rozwiązań opartych na AI (prywatnych botów AI), które będą mogły wspierać prace jednostek i ewaluatorów.

## Rozdział 4. Planowanie ewaluacji

Stanisław Bienias, Jakub Wróblewski

Proces ewaluacji to sekwencja względnie powtarzalnych działań i decyzji – od identyfikacji potrzeby informacyjnej, którą chcemy zaspokoić, do wykorzystania wyników badania. W niniejszym poradniku proces ten przedstawiamy jako 10 następujących po sobie kroków. W tym rozdziale prezentujemy 5 pierwszych. To w ramach tych pięciu kroków rozstrzygamy m.in., czy będziemy realizować ewaluację w formule wewnętrznej, czy zewnętrznej.

### Krok 1: Zrozumienie potrzeb i podjęcie decyzji o realizacji ewaluacji

U podstaw projektu ewaluacyjnego powinna stać konkretna potrzeba informacyjna. Zatem pierwszym pytaniem na które należy udzielić odpowiedzi jest „po co ewaluację przeprowadzamy?”. Badania ewaluacyjne są koszt- i czasochłonne, zatem decyzja o ich realizacji powinna być uzasadniona i wypływać z realnego zapotrzebowania na wyniki, a w szczególności chęci dostarczenia wiedzy, która może wesprzeć istotne decyzje w zakresie realizowanych lub przyszłych interwencji.

Przyczynkiem do ewaluacji może być też wymóg prawny (np. Ustawa o zasadach prowadzenia polityki rozwoju<sup>34</sup> lub wymóg z Rozporządzenia ogólnego<sup>35</sup>), ale ten również można interpretować, jako sformalizowaną wcześniej przez prawodawcę potrzebę wiedzy.

Potrzeby informacyjne nie istnieją w próżni, tzn. są potrzebami konkretnych odbiorców. W związku z tym identyfikacja potrzeb jest równoznaczna z koniecznością ustalenia użytkowników i odbiorców ewaluacji. Tych – jak przekonujemy w następnych rozdziałach – warto angażować w proces ewaluacji od samego początku.

<sup>34</sup> Patrz: Ustawa z dnia 6 grudnia 2006 r. o zasadach prowadzenia polityki rozwoju (Dz. U. 2006 Nr 227 poz. 1658)

<sup>35</sup> Patrz: Rozporządzenie Parlamentu Europejskiego i Rady (UE) 2021/1060 z dnia 24 czerwca 2021 r.

Dobra identyfikacja odbiorców i ich potrzeb informacyjnych jest kluczem do sukcesu ewaluacji, czyli wykorzystania jej wyników. Tym samym już w pierwszym kroku rozstrzyga się powodzenie działań z kroku 10. Ma to związek z faktem, iż różni odbiorcy mają różne preferencje co do kanałów i form komunikacji wyników ewaluacji (patrz szerzej rozdział 7).

Zrozumienie potrzeb i motywacji odbiorców ewaluacji jest kluczowe również dla jej wykonawców, zarówno wewnętrznych, jak i zewnętrznych, w zależności od przyjętego modelu ewaluacji. Dlatego rekomendujemy, by syntetyczny opis, z czego wynika potrzeba realizacji badania, był zawsze elementem opisu przedmiotu zamówienia, czy zarysu planu badawczego. O ile jest to uzasadnione, można również zawrzeć opis szerszego kontekstu ocenianych interwencji. Jest to szczególnie istotne jeżeli uległ on modyfikacji w stosunku do założeń dokumentów wdrożeniowych (przykładowo zaszły nowe okoliczności, które są przesłanką do realizacji badania).

## Krok 2: Definiujemy cel/e badania, stawiamy pytania badawcze oraz określamy zakres ewaluacji

Zdefiniowanie celu/celów i pytań badawczych jest kluczowym krokiem w procesie planowania badania ewaluacyjnego. Zazwyczaj określamy jeden cel główny oraz kilka celów szczegółowych ewaluacji.

Pytania badawcze stanowią uszczegółowienie i operacjonalizację celów i kryteriów ewaluacji. Dla zachowania spójności i jasności koncepcji ewaluacji pytania powinny jednoznacznie wynikać z celów badania i być możliwie precyzyjne. Wykonalność ewaluacji wymaga, by pytań nie formułować zbyt wiele (nie więcej niż 5–7).



### Dobra praktyka

**Najbardziej trafne pytania ewaluacyjne zadają na ogół osoby... spoza jednostki ewaluacyjnej. Ewaluacja ma dostarczyć wiedzy konkretnym osobom, które mogą ją potem wykorzystać. Dlatego tak ważne jest, by w proces formułowania pytań badawczych, na jak najwcześniejszym etapie, zaangażować przyszłych odbiorców ewaluacji.**

Praktyka realizacji ewaluacji pokazuje, że projekty zawierające zbyt dużo pytań badawczych, o zbyt szerokim zakresie, są realizowane na wyższym poziomie ogólności i finalnie nie zaspokajają potrzeb informacyjnych odbiorców. Często, proponowana w takiej

sytuacji „uniwersalna metodyka” (bo dająca „jakąś” odpowiedź na każde z pytań) nie pozwala na udzielenie odpowiedzi rzetelnych i dokładnych. W przypadku bardzo szerokich potrzeb informacyjnych warto rozważyć realizację nie jednego, ale dwóch-trzech badań ewaluacyjnych skoncentrowanych na konkretnych obszarach tematycznych. Alternatywą jest priorytetyzacja i wybór z wielu faktycznych potrzeb tych najważniejszych w danym momencie.

W zależności od momentu, w którym realizujemy nasze badanie (ex ante, ex post) zadając pytania koncentrujemy się na różnych kwestiach (więcej patrz rozdział 2). Pytania badawcze formułujemy na pewnym poziomie ogólności. Nie formułujemy ich zbyt wąsko (od tego są pytania w konkretnych narzędziach badawczych, np. ankietach), ani zbyt szeroko (na poziomie celu ewaluacji). Stawiając pytania ewaluacyjne staramy się je sformułować w sposób obiektywny/niesugerujący (np. „Czy, a jeżeli tak ...?”).

Kolejnym elementem, który musimy doprecyzować, jest zakres ewaluacji, czyli co, kogo, kiedy i w jakim układzie przestrzennym będziemy badać. Zakres badania często jest prezentowany w przekroju:

- Zakres przedmiotowy – w tym wymiarze należy wskazać co będziemy badać. Może to być cały program, strategia lub jego/jej wycinek (np. priorytet, działanie, a czasem pojedynczy projekt).
- Zakres podmiotowy – czyli z kim planujemy przeprowadzić badania. Mogą to być osoby zaangażowane w opracowanie dokumentu programowego, odbiorcy oferowanego wsparcia, inne osoby/grupy, których perspektywa powinna zostać uwzględniona w naszej ewaluacji.
- Zakres czasowy – czyli z jakiego okresu będziemy zbierać i analizować informacje. Możemy koncentrować się przykładowo na projektach zakończonych zgodnie ze stanem na wybrany dzień. Określając zakres czasowy warto wziąć pod uwagę cykliczność zbieranych danych (z systemu monitorowania, danych ze statystyk publicznych), tak by dysponować możliwie najbardziej aktualnymi informacjami.
- Zakres terytorialny – czyli z jakiego obszaru będziemy analizować informacje. Nie zawsze badanie musi obejmować cały obszar oddziaływania naszej interwencji. Uzasadnione może być skoncentrowanie na wybranych podobszarach



#### Dobra praktyka

**Zakres badania może być elementem istotnie wpływającym na czas i koszt realizacji badania. Każdorazowo należy się zatem zastanowić, czy nie możemy w jakiś sposób ograniczyć zakresu ewaluacji bez istotnej szkody dla realizacji celu badania.**

np. interwencja dotyczy całej Polski, ale badamy efekty jej realizacji w wybranych województwach, interwencja realizowana jest w całym mieście, ale ewaluacja skupia się na określonej dzielnicy. Doprecyzowanie terytorium może też oznaczać wybór obszarów miejskich lub wiejskich, przygranicznych, górskich, przybrzeżnych, albo o innych interesujących nas cechach.

Cel, zakres i pytania badawcze musimy doprecyzować niezależnie od tego, czy planujemy zlecić ewaluację zewnętrznemu wykonawcy, czy też wykonać ją siłami własnymi. W tym pierwszym przypadku opis stanowi element OPZ, w drugim – element koncepcji badania.

## Krok 3: Określamy sposób realizacji badania

O ile krok 1 i 2 mają charakter uniwersalny i powinny być przeprowadzone niezależnie od tego, czy ewaluacja będzie mieć charakter wewnętrzny czy zewnętrzny, to już sposób realizacji badania jest wypadkową przyjętej strategii wykonawczej. Instytucje podejmujące się realizacji badania ewaluacyjnego zasobami własnymi decydują się na samodzielne opracowanie metodyki badania.

Zamawiający poszukujący wykonawcy zewnętrznego mogą przyjąć jedną z dwóch strategii. Pierwszą jest precyzyjne określenie wymagań dotyczących metodyki przyszłego badania ewaluacyjnego. Często łącznie z określeniem metod, profilu respondentów i wielkości prób badawczych oraz sposobu wnioskowania. Nie mając odpowiedniego *know-how* i doświadczenia, można zdecydować się oddać odpowiedzialność za zaplanowanie metodyki potencjalnym wykonawcom ewaluacji.

Umiejętność skonstruowania adekwatnej metodyki badania jest również bardzo dobrym sposobem na zweryfikowanie potencjału wykonawcy. Dlatego wskazane jest by metodyka (ew. jej uzupełnienie, rozwinięcie) podlegało ocenie w trakcie postępowania o zamówienie publiczne. Bardzo często spotykanym rozwiązaniem przy ewaluacjach zewnętrznych jest zawarcie w OPZ wstępnej metodyki (czasem nazywanej minimum metodycznym), i oczekiwanie jej uszczegółowienia, rozwinięcia, uzupełnienia przez podmioty zainteresowane pozyskaniem zamówienia.

Metodykę należy projektować zawsze pod kątem realizacji celów badania i uzyskania rzetelnej odpowiedzi na pytania ewaluacyjne. Odpowiedź na każde pytanie, by była wiarygodna, wymaga unikalnego zestawu metod i technik badawczych. W badaniach konieczną praktyką jest stosowanie **triangulacji**, czyli zaplanowania badania w sposób zakładający udzielanie odpowiedzi na pytania badawcze w oparciu o więcej niż jedno

źródło danych, więcej niż jedną metodę ich pozyskania/analizy. Dodatkowym wymiarem wzmacniającym triangulację są zróżnicowana wiedza i kompetencje członków zespołu badawczego.



#### Dobra praktyka

**W przypadku braku *know-how* w zakresie formułowania zakresu ewaluacji i przygotowywania OPZ warto skorzystać ze wsparcia w postaci eksperta zewnętrznego. Ekspertskie wsparcie pomoże uniknąć błędów i przygotować wyższej jakości koncepcję ewaluacji. Ekspert może również wspomóc nas na późniejszych etapach realizacji badania (wyborze wykonawcy i odbiorze produktów ewaluacji).**

Samodzielne zdefiniowanie sposobu realizacji badania ewaluacyjnego jest skomplikowane. Wymaga uwzględnienia wielu czynników (zasoby, specyfika pytań badawczych, dostępne dane) i dopasowania do tych czynników odpowiednich metod i technik zbierania i analizy danych (więcej o metodach i technikach badawczych w rozdziale 7).

## Krok 4: Określamy zasoby niezbędne do realizacji badania

Po określeniu celów ewaluacji, opracowaniu listy pytań i metodyki badania należy koncepcję projektu badawczego skonfrontować z posiadanymi zasobami (więcej na ten temat patrz rozdział 3). W tym kroku skoncentrujemy się na jednym z kluczowych zasobów koniecznych do zlecenia ewaluacji na zewnątrz (bądź jej realizacji w modelu hybrydowym) – środkach finansowych. Budżet jest podstawowym zasobem, determinującym skalę i złożoność badania, gdy decydujemy się na ewaluację zewnętrzną. Poniżej prezentujemy tabelę zawierającą przybliżone rynkowe ceny usług, które składają się na koszt badania ewaluacyjnego.



**Tabela 3.** Szacunkowe koszty realizacji metod i technik badawczych. Kwoty na wrzesień 2023.

| Metoda                                        | Specyfikacja                                                                                                                | Baza do wyceny                                                                                                                                                | Szacunkowy koszt netto (bez VAT) |
|-----------------------------------------------|-----------------------------------------------------------------------------------------------------------------------------|---------------------------------------------------------------------------------------------------------------------------------------------------------------|----------------------------------|
| <b>IDI</b>                                    | Czas trwania: 60 min                                                                                                        | Koszt moderatora                                                                                                                                              | 200-800 zł                       |
| <b>TDI</b>                                    | Czas trwania: 60 min                                                                                                        | Koszt moderatora                                                                                                                                              | 150-600 zł                       |
| <b>FGI</b>                                    | Czas trwania: 90 min                                                                                                        | Koszt moderatora                                                                                                                                              | 400-5000 zł                      |
| <b>FGI-online</b>                             | Czas trwania: 90 min                                                                                                        | Koszt moderatora                                                                                                                                              | 300-3500 zł                      |
| <b>CAPI</b>                                   | Czas trwania: 15 min                                                                                                        | Koszt pozyskania jednej ankiety                                                                                                                               | 20-300 zł                        |
| <b>CAWI</b>                                   | Czas trwania: 15 min                                                                                                        | Koszt pozyskania jednej ankiety                                                                                                                               | 5-50 zł                          |
| <b>CATI</b>                                   | Czas trwania: 15 min                                                                                                        | Koszt pozyskania jednej ankiety                                                                                                                               | 15-80 zł                         |
| <b>Desk-research</b>                          | Od 26 do 50 źródeł                                                                                                          | Analiza przekraczająca 25 źródeł polskojęzycznych – źródła o charakterze tekstowym (np. raporty, strategie) i bazodanowym (np. statystyka publiczna, dane SL) | 4000-20000 zł                    |
| <b>Dashboard</b>                              | Jedna strona                                                                                                                | Opracowanie dashboardu na bazie istniejącego raportu                                                                                                          | 1000-5000 zł                     |
| <b>Opracowanie prezentacji multimedialnej</b> | Objętość: 10-20 slajdów                                                                                                     | Koszt opracowania kompletnej prezentacji                                                                                                                      | 1000-5000 zł                     |
| <b>Prezentacja wyników</b>                    | Czas trwania: 15-30 min                                                                                                     | Prezentacja w siedzibie Zamawiającego                                                                                                                         | 500-5000 zł                      |
| <b>Prezentacja wyników on-line</b>            | Czas trwania: 15-30 min                                                                                                     | Prezentacja <i>online</i> (np. via Zoom)                                                                                                                      | 500-3000 zł                      |
| <b>Studium przypadku</b>                      | Opracowanie CS na podstawie 3 IDI, analizy danych zastanych (10 źródeł) i jednodniowej wizyty studyjnej (na terenie Polski) | Realizacja CS i przygotowanie podsumowania CS                                                                                                                 | 2000-11000 zł                    |
| <b>Analiza danych jakościowych</b>            | Na podstawie zakodowanych wywiadów                                                                                          | Koszt za jedną godzinę pracy analityka                                                                                                                        | 100-500                          |
| <b>Analiza danych ilościowych</b>             | Na podstawie wyciszczonych baz danych                                                                                       | Koszt za jedną godzinę pracy analityka                                                                                                                        | 100-500                          |
| <b>Opracowanie Raportu metodycznego</b>       | Opracowanie dokumentu wraz z narzędziami: 1 kwestionariuszem CAWI, 1 scenariuszem IDI i 1 schematem studium przypadku       | Koszt opracowania kompletnego dokumentu                                                                                                                       | 3500-20000                       |
| <b>Opracowanie raportu końcowego</b>          | Objętość: 101-200 stron                                                                                                     | Koszt opracowania kompletnego dokumentu                                                                                                                       | 6000-40000                       |

Źródło: J. Wróblewski, Badanie ankietowe przeprowadzone na potrzeby przygotowania niniejszego poradnika

Przedstawione powyżej wyceny wskazują kwoty przybliżone, aktualne na wrzesień 2023 roku. Mogą być one punktem odniesienia, ale należy pamiętać, że każda wycena rynkowa jest uzależniona od wielu czynników. Planując budżet należy w szczególności wziąć pod uwagę możliwe zmiany cen wynikające z inflacji i innych procesów rynkowych.

**Tabela 4.** Czynniki wpływające na wycenę badania

| Metoda            | Czynniki wpływające na wycenę                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     |
|-------------------|-----------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|
| <b>IDI</b>        | <ul style="list-style-type: none"> <li>• złożoność tematyki wpływająca na czas konieczny do przygotowania się do rozmowy</li> <li>• liczba zagadnień będących tematem spotkania (odrębnych wątków w scenariuszu)</li> <li>• profil respondenta, tj. respondent ekspert vs. respondent nie-ekspert, co także przekłada się na czas potrzebny do przygotowania się do rozmowy</li> <li>• profil respondenta, tj. osoby o utrudnionym dostępie (np. pełniące funkcje kierownicze) vs. osoby, z którymi łatwo się umówić</li> <li>• lokalizacja wywiadu</li> <li>• komfort ewentualnego dojazdu (liczba przesiadek, szczególnie wczesny wyjazd, konieczność noclegu, etc.)</li> </ul> |
| <b>TDI</b>        | <ul style="list-style-type: none"> <li>• złożoność tematyki wpływająca na czas konieczny do przygotowania się do rozmowy</li> <li>• liczba zagadnień będących tematem spotkania (odrębnych wątków w scenariuszu)</li> <li>• profil respondenta, tj. respondent ekspert vs. respondent nie-ekspert, co także przekłada się na czas potrzebny do przygotowania się do rozmowy</li> <li>• profil respondenta, tj. osoby o utrudnionym dostępie (np. pełniące funkcje kierownicze) vs. osoby, z którymi łatwo się umówić</li> </ul>                                                                                                                                                   |
| <b>FGI</b>        | <ul style="list-style-type: none"> <li>• złożoność tematyki wpływająca na czas konieczny do przygotowania się do rozmowy</li> <li>• liczba zagadnień będących tematem spotkania (odrębnych wątków w scenariuszu)</li> <li>• profil respondentów, tj. eksperci vs. nie-eksperci, co także przekłada się na czas potrzebny do przygotowania się do rozmowy</li> <li>• ewentualna konieczność wypłaty wynagrodzenie respondentom (dodatkowy obowiązek moderatora)</li> <li>• lokalizacja FGI</li> <li>• komfort ewentualnego dojazdu na miejsce (liczba przesiadek, szczególnie wczesny wyjazd, konieczność noclegu, etc.)</li> </ul>                                                |
| <b>FGI-online</b> | <ul style="list-style-type: none"> <li>• złożoność tematyki wpływająca na czas konieczny do przygotowania się do rozmowy</li> <li>• liczba zagadnień będących tematem spotkania (odrębnych wątków w scenariuszu)</li> <li>• profil respondentów, tj. eksperci vs. nie-eksperci, co także przekłada się na czas potrzebny do przygotowania się do rozmowy</li> <li>• ewentualna konieczność wypłaty wynagrodzenie respondentom (dodatkowy obowiązek moderatora)</li> </ul>                                                                                                                                                                                                         |

| Metoda                                        | Czynniki wpływające na wycenę                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             |
|-----------------------------------------------|-----------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|
| <b>CAPI</b>                                   | <ul style="list-style-type: none"> <li>• miejsce rekrutacji respondentów</li> <li>• rozproszenie respondentów</li> <li>• prawdopodobieństwo odmów</li> <li>• profil respondentów – grupy trudne (np. niechętne, agresywne) vs. łatwe</li> <li>• profil respondentów – osoby prywatne vs. przedstawiciele instytucji</li> <li>• dostęp do miejsca ładowania urządzenia mobilnego</li> <li>• czas na realizację próby</li> <li>• wielkość próby</li> <li>• dostęp do bazy respondentów</li> <li>• wielkość bazy respondentów w stosunku do próby</li> </ul> |
| <b>CAWI</b>                                   | <ul style="list-style-type: none"> <li>• dostęp do bazy mailowej respondentów</li> <li>• dostęp do bazy telefonicznej respondentów (w celu monitowania)</li> <li>• prawdopodobieństwo odmów</li> <li>• profil respondentów – osoby prywatne vs. przedstawiciele instytucji</li> <li>• dostęp do aplikacji umożliwiającej automatyczne wysyłanie dedykowanego mailingu do dużej liczby adresatów</li> <li>• czas na realizację próby</li> <li>• wielkość próby</li> <li>• wielkość bazy respondentów w stosunku do próby</li> </ul>                        |
| <b>CATI</b>                                   | <ul style="list-style-type: none"> <li>• dostęp do bazy mailowej respondentów (w celu zapowiedzenia kontaktu przed rozpoczęciem realizacji telefonicznej)</li> <li>• dostęp do bazy telefonicznej respondentów</li> <li>• prawdopodobieństwo odmów</li> <li>• profil respondentów – osoby prywatne vs. przedstawiciele instytucji</li> <li>• konieczność nagrywania rozmów</li> <li>• czas na realizację próby</li> <li>• wielkość próby</li> <li>• wielkość bazy respondentów w stosunku do próby</li> </ul>                                             |
| <b>Desk research</b>                          | <ul style="list-style-type: none"> <li>• dostęp do źródeł</li> <li>• format źródeł (format edytowalny vs. format nieedytowalny, np. skany)</li> <li>• złożoność merytoryczna źródeł</li> <li>• dostęp do wspierającego eksperta dziedzinowego</li> <li>• czas na przygotowanie analizy</li> </ul>                                                                                                                                                                                                                                                         |
| <b>Dash-board</b>                             | <ul style="list-style-type: none"> <li>• jakość danych (dane gotowe do wizualizacji vs. wymagające przygotowania)</li> <li>• liczba analiz prezentowana na stronie</li> <li>• złożoność poszczególnych wizualizacji</li> <li>• czas na przygotowanie dash-boardu</li> </ul>                                                                                                                                                                                                                                                                               |
| <b>Opracowanie prezentacji multimedialnej</b> | <ul style="list-style-type: none"> <li>• jakość danych (dane gotowe do wizualizacji vs. wymagające przygotowania)</li> <li>• liczba analiz informacji prezentowanych na slajdach</li> <li>• złożoność poszczególnych grafik</li> <li>• konieczność dopasowania prezentacji do standardu WCAG</li> <li>• czas na przygotowanie prezentacji</li> </ul>                                                                                                                                                                                                      |
| <b>Prezentacja wyników</b>                    | <ul style="list-style-type: none"> <li>• złożoność merytoryczna prezentacji</li> <li>• wymagania dot. roli prelegenta w projekcie (np. kierownik vs. dowolny członek zespołu)</li> <li>• czas potrzebny na przygotowanie do prelekcji</li> <li>• wystąpienie kończące się po prezentacji vs. uwzględnienie w/po prezentacji dyskusji ze słuchaczami</li> <li>• komfort ewentualnego dojazdu na miejsce (liczba przesiadek, szczególnie wczesny wyjazd, konieczność noclegu, etc.)</li> </ul>                                                              |

| Metoda                                  | Czynniki wpływające na wycenę                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             |
|-----------------------------------------|---------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|
| <b>Prezentacja wyników on-line</b>      | <ul style="list-style-type: none"> <li>• złożoność merytoryczna prezentacji</li> <li>• wymagania dot. roli prelegenta w projekcie (np. kierownik vs. dowolny członek zespołu)</li> <li>• czas potrzebny na przygotowanie do prelekcji</li> <li>• wystąpienie kończące się po prezentacji vs. uwzględnienie w/po prezentacji dyskusji ze słuchaczami</li> </ul>                                                                                                                                                                                                                                                                                                            |
| <b>Studium przypadku</b>                | <ul style="list-style-type: none"> <li>• złożoność tematyki wpływająca na czas konieczny do przygotowania się do wywiadów i analiz desk-research</li> <li>• pełny/ograniczony dostęp do źródeł desk-research</li> <li>• terminarz realizacji wywiadów (w jednym dniu vs. w dłuższym okresie)</li> <li>• profil respondentów, tj. eksperci vs. nie-eksperci, co także przekłada się na czas potrzebny do przygotowania się do rozmowy</li> <li>• lokalizacje wywiadów i wizyty studyjnej</li> <li>• komfort ewentualnego dojazdu na miejsce (liczba przesiadek, szczególnie wczesny wyjazd, konieczność noclegu, etc.)</li> <li>• czas na przygotowanie studium</li> </ul> |
| <b>Analiza danych jakościowych</b>      | <ul style="list-style-type: none"> <li>• ilość materiału bazowego</li> <li>• złożoność drzewa kodowego</li> <li>• liczba pytań badawczych</li> <li>• czas na przygotowanie analizy</li> </ul>                                                                                                                                                                                                                                                                                                                                                                                                                                                                             |
| <b>Analiza danych ilościowych</b>       | <ul style="list-style-type: none"> <li>• charakter materiału bazowego (liczba i złożoność pytań, liczba pytań otwartych)</li> <li>• liczba pytań badawczych</li> <li>• zakres analiz (tabele krzyżowe, rozkłady częstości, testy statystyczne, inne np. wielowymiarowe analizy)</li> <li>• czas na przygotowanie analizy</li> </ul>                                                                                                                                                                                                                                                                                                                                       |
| <b>Opracowanie raportu metodycznego</b> | <ul style="list-style-type: none"> <li>• złożoność merytoryczna tematu badania</li> <li>• dostęp do wspierającego/cych eksperta/ów dziedzinowego/ych</li> <li>• czas na przygotowanie raportu</li> </ul>                                                                                                                                                                                                                                                                                                                                                                                                                                                                  |
| <b>Opracowanie raportu końcowego</b>    | <ul style="list-style-type: none"> <li>• złożoność merytoryczna tematu badania</li> <li>• liczba pytań badawczych</li> <li>• ilość analiz, na podstawie których powstanie raport końcowy</li> <li>• dostęp do wspierającego/cych eksperta/ów dziedzinowego/ych</li> <li>• czas na przygotowanie raportu</li> <li>• konieczność dopasowania raportu do standardu WCAG</li> </ul>                                                                                                                                                                                                                                                                                           |

Źródło: Opracowanie własne

Dysponując skromniejszymi budżetami wciąż można zaplanować i zrealizować wartościowe badanie, choć będzie to wymagało skupienia uwagi na elementach najistotniejszych. W szczególności możemy ograniczyć liczbę pytań badawczych, zawęzić zakres badania, czy liczbę wykorzystanych metod.

Planując budżet badania warto rozesłać zapytanie o szacunek cenowy, który pozwoli określić koszt ewaluacji. To rozwiązanie jest szczególnie cenne, gdy brakuje nam doświadczeń płynących z wcześniejszych podobnych projektów lub nie jesteśmy w stanie precyzyjnie określić metodyki (więcej na ten temat patrz rozdział 5).

## Krok 5: Ustalamy harmonogram badania

Plan prac pełni kluczową funkcję w każdym projekcie. Nie inaczej jest w przypadku badań ewaluacyjnych. W sytuacji, gdy zlecamy badanie na zewnątrz, w OPZ warto wskazać ramowy harmonogram prac określający nie tylko datę zakończenia badania, ale także terminy zamknięcia kluczowych jego etapów. W związku z tym, że przeprowadzenie zamówienia publicznego zawsze obarczone jest ryzykiem opóźnienia, w harmonogramie nie zaleca się wskazywania konkretnych dat, ale posługiwanie się liczbami dni, tygodni, od daty podpisania umowy. Tworząc ramowy harmonogram prac, warto uwzględnić wszelkie okoliczności, które mogą wpływać na współpracę z zewnętrznym wykonawcą ewaluacji, np. urlopy pracowników odbierających prace, czy okresy świąteczne i wakacyjne (które mogą utrudniać lub uniemożliwiać dotarcie do respondentów).



### Dobra praktyka

**Zaleca się sprecyzowanie 2–3 kluczowych produktów w badaniu będących jednocześnie kamieniami milowymi. Może to być np. akceptacja raportu metodycznego, raportu częściowego i raportu końcowego.**

Można zaplanować w harmonogramie badania ewaluacyjnego cykliczne – np. raz na dwa tygodnie – sprawozdania z realizacji badania. Sprawozdania mogą mieć formę krótkiej informacji obejmującej syntetyczne podsumowanie: 1) działań wykonanych w ostatnich dwóch tygodniach, 2) planowanych do realizacji w następnym okresie prac, 3) ewentualnych problemów w realizacji.

Raporty ewaluacyjne są często dokumentami złożonymi, a ich rzetelna ocena wymaga zaangażowania wielu osób. Dobrą praktyką jest zaplanowanie harmonogramu tak, by RK trafił do odbiorców na przynajmniej 4 tygodnie przed planowanym terminem przekazania jego ostatecznej wersji (patrz rozdział 6). Dzięki temu zamawiający może zapoznać się z treścią raportu, zgłosić uwagi i dać wykonawcy odpowiednią ilość czasu na ich uwzględnienie. Podobną praktykę, tj. stosowanie roboczych wymian dokumentów warto wdrożyć także w przypadku innych kluczowych materiałów (np. raportu metodycznego, raportów częściowych) choć czasokresy zaplanowane na ich analizę i przekazanie uwag mogą być odpowiednio krótsze.

## Podsumowanie

Proces ewaluacji składa się z szeregu etapów, których właściwe zrealizowanie zwiększa szanse na jej wysoką jakość i użyteczność. Niezależnie od sposobu realizacji badania (ewaluacja wewnętrzna, zewnętrzna i hybrydowa), każdy z pięciu przedstawionych w tym rozdziale kroków należy precyzyjnie zaplanować i zrealizować. W przypadku realizacji badania siłami zewnętrznymi możemy tylko nieco wężiej opisać metodykę realizacji i polegać tu na propozycji wykonawców zewnętrznych. To, jak nazwiemy dokument, w którym opisaliśmy powyższe kwestie, będzie zależne ostatecznie od trybu realizacji badania. W przypadku zlecenia go na zewnątrz – będzie to Opis Przedmiotu Zamówienia. W przypadku ewaluacji realizowanej wewnętrznie – opis koncepcji badawczej.

## Rozdział 5. Realizacja badania ewaluacyjnego

Stanisław Bienias, Piotr Strzęboszewski, Paulina Skórska

Zaplanowaliśmy już przeprowadzenie badania. Wiemy już czego będzie dotyczyć, znamy podstawowe założenia dotyczące celu, zakresu i pytań badawczych. Wiemy, kiedy potrzebujemy wyników badania oraz jakim zasobami dysponujemy. Podjęliśmy decyzję, czy realizujemy badanie siłami wewnętrznymi, czy też zlecamy je na zewnątrz. W Polsce najczęściej badania ewaluacyjne zlecane są podmiotom takim jak wyspecjalizowane firmy, organizacje pozarządowe, instytucje badawcze, uczelnie, oraz niezależnym ekspertom. Rzadziej ewaluacje realizowane są siłami wewnętrznymi instytucji. W niniejszym rozdziale opiszemy obydwie ścieżki. Krok 6. poświęcony został sytuacji, w której zlecamy badanie na zewnątrz. Krok 7. opisuje natomiast fazę realizacyjną ewaluacji w dwóch wariantach – tj. ewaluacji zewnętrznej oraz badania wewnętrznego. Oczywiście możliwa jest też sytuacja, że badanie realizujemy w modelu hybrydowym (część prac zlecamy na zewnątrz, a część realizujemy samodzielnie, wewnątrz instytucji).

### Krok 6: Przygotowujemy i przeprowadzamy zamówienie na ewaluację

W przypadku ewaluacji zewnętrznych szacowanym kosztem badania (to robione jest w ramach poprzednich kroków – patrz rozdział 4) oraz źródłem środków z których sfinansujemy ewaluację. Źródłu finansowania mogą towarzyszyć specyficzne reguły określające ich wydawanie. Niemniej, prawie zawsze kluczowe będą dla nas zapisy ustawy Prawo Zamówień Publicznych<sup>36</sup>. W ramach ww. ustawy mamy jedno zasadnicze rozróżnienie wynikające z kosztu planowanego działania. Art. 2 ust. 1 określa, że jeśli przewidujemy, że nasz projekt będzie kosztował więcej niż 130 tysięcy PLN netto, to przy jego realizacji musimy zastosować PZP. Kolejnym progiem, który na bazie PZP może wpływać na nasze działania jest tzw. próg unijny. Ta kwota jest zmienna w czasie i określana w odpowiednich aktach unijnych. Aktualnie dla badań ewaluacyjnych realizowanych przez sektor rządowy

<sup>36</sup> W chwili powstawania tego opracowania obowiązywała Ustawa z dnia 11 września 2019r. w wersji zmienionej, opublikowanej w sierpniu 2023r. Dziennik Ustaw RP pozycja 1605, Warszawa, dnia 14 sierpnia 2023 r. akty prawne do ISAP-u ([sejm.gov.pl](http://sejm.gov.pl)) dostęp 23.08.2023r.

Realizacja badania ewaluacyjnego  
 kwota ta wynosi 140 tysięcy EUR netto co przelicza się wg ustalonego specjalnego kursu na 623 504 PLN<sup>37</sup> (obowiązujący od 1.01.2024). Przekroczenie progu unijnego ma pewne konsekwencje dla sposobu prowadzenia postępowania, niemniej nie zmienia jego istoty, więc nie będziemy go szczegółowo opisywać w niniejszym opracowaniu.



#### Dobra praktyka

**Wszystkie aktualne dokumenty dotyczące realizacja postępowań w ramach PZP w Polsce można znaleźć na stronie Urzędu Zamówień Publicznych.**  
<https://www.gov.pl/web/uzp/>

W przypadku realizacji ewaluacji w ramach PZP wybór najkorzystniejszej oferty na jej przeprowadzenie musi zostać poprzedzony postępowaniem o udzielenie zamówienia publicznego w jednym z 4 trybów<sup>38</sup>. W praktyce w Polsce wykonawców ewaluacji najczęściej wybiera się w ramach tzw. trybu podstawowego opisanego w art. 275 ustawy. Nie znaczy to, że inne tryby nie mają zastosowania przy realizacji badań. Ich zastosowanie wiąże się jednak z koniecznością spełnienia określonych warunków (np. dysponowaniem znacznie większą ilością czasu na przeprowadzenie badania, lub uzasadnieniem dla zastosowania tzw. wolnej ręki).

W ramach działań poprzedzających wybór wykonawcy ewaluacji kluczowe jest przygotowanie następujących dokumentów:

- oszacowania wartości zamówienia,
- opisu przedmiotu zamówienia,
- specyfikacji warunków zamówienia, a w niej warunki, które musi spełnić wykonawca i kryteria oceny ofert,
- projektowanych postanowień umowy.

Oczywiście w ramach postępowania tworzonych jest więcej dokumentów niemniej te wymienione powyżej są najistotniejsze z punktu widzenia określenia przyszłej ewaluacji oraz wyboru jej wykonawcy.

Kalkulacja kosztów – o szacowaniu kosztów planowanej ewaluacji była już mowa w rozdziale 4. Na potrzeby postępowania prowadzonego w ramach PZP musimy jednak

<sup>37</sup> Próg unijny ustalono na poziomie 143 000 EUR (czyli 663 105 PLN). Zob. <https://www.portalzp.pl/nawosci/progi-unijne-w-zamowieniach-publicznych-i-sredni-kurs-euro-na-lata-2024-i-2025-25791.html>

<sup>38</sup> Lub 7 jeśli realizujemy zamówienie o wartości powyżej progu unijnego. Tryby dot. postępowań realizowanych poniżej progu zostały opisane od art. 275 do 306 Ustawy PZP, natomiast tryby dla postępowań powyżej progu unijnego zostały wymienione w art. 129 ww. ustawy. W przypadku 7 trybów będą to: przetarg nieograniczony, przetarg ograniczony, negocjacje z ogłoszeniem, dialog konkurencyjny, partnerstwo innowacyjne, negocjacje bez ogłoszenia, zamówienie z wolnej ręki. W przypadku 4 trybów będą to natomiast: tryb podstawowy bez negocjacji, tryb podstawowy z możliwością negocjacji, tryb podstawowy z obowiązkiem negocjacji; partnerstwo innowacyjne; negocjacje bez ogłoszenia; zamówienie z wolnej ręki.



stworzyć oficjalny dokument, którego opracowanie opiera się na pewnych zasadach. Proces określenia kosztów musi być transparentny i jednoznaczny oraz musi zostać przeprowadzony w odpowiednim terminie. W stosunku do badań ewaluacyjnych szacowanie może być przeprowadzone wcześniej niż 3 miesiące przed postępowaniem. Jeśli nasze szacowanie jest starsze, musimy dokonać jego aktualizacji.



### Dobra praktyka

**W zakresie szacowania kosztów zamówienia dobrą praktyką jest nawiązanie dialogu z rynkiem. Możemy to osiągnąć upubliczniając projekt OPZ. Wysłanie tego dokumentu do potencjalnych wykonawców lub umieszczenie na stronie www. przyczyni się do rzetelnego oszacowania kosztów oraz zbuduje zainteresowanie projektem. W trakcie szacowania wartości bardzo ważne jest uzyskanie uzasadnień dla wskazań wykonawców. Wykonawców możemy również prosić o uwagi do OPZ. To również może się przyczynić do lepszej jakości badania.**

Opis przedmiot zamówienia (OPZ) – to jeden z najważniejszych dokumentów tworzonych w ramach postępowania mającego wyłonić wykonawcę ewaluacji. Ujmujemy w nim opis naszego badania. OPZ powinien zawierać następujące elementy<sup>39</sup>:

- uzasadnienie realizacji ewaluacji – w tej części opisujemy przede wszystkim dlaczego chcemy przeprowadzić badanie i powinniśmy wprost i jednoznacznie wskazać, jakie przesłanki skłaniają nas do realizacji ewaluacji;
- zakres badania – określimy czy i jaki zakres przedmiotowy, podmiotowy, terytorialny i czasowy będzie istotny dla naszej ewaluacji;
- cele badania – zazwyczaj określamy jeden cel główny oraz kilka celów szczegółowych ewaluacji. Pamiętajmy, że powinny one korespondować z uzasadnieniem realizacji badania i nie powinno być ich za dużo (cele badania to nie pytania badawcze);
- kryteria ewaluacyjne – kryteria zostały opisane szerzej w rozdziale 1. Tworząc OPZ musimy dobrać odpowiednie kryteria do naszego badania. Wybierzmy tylko te naprawdę adekwatne i istotne dla naszej ewaluacji. Zbyt duża liczba kryteriów oceny przełoży się negatywnie na objętość ofert oraz ich przejrzystość i jakość;
- pytania ewaluacyjne – stanowią uszczegółowienie i operacjonalizację celów ewaluacji. Nie powinno być ich za dużo. Powinniśmy skupić się tylko na najistotniejszych kwestiach, wynikających z celów, zakresów i kryteriów planowanej ewaluacji.;
- sposób realizacji ewaluacji – często przygotowując badanie z góry wiemy już, jakie metody powinny zostać zastosowane w trakcie jego realizacji. Jeśli posiadamy taką

<sup>39</sup> Więcej informacji na temat poszczególnych elementów OPZ znajduje się w rozdziale 4.

wiedzę to powinniśmy opisać te metody w OPZ. Możemy też nie określać metodyki i pozostawić ją do opracowania wykonawcy – oceniając ten element w ramach oferty;

- harmonogram – częścią planowanego projektu, którą warto osobno przedstawić, jest harmonogram realizacji ewaluacji. Powinniśmy w nim dość szczegółowo przedstawić terminy realizacji kluczowych etapów badania;
- Wymogi względem produktów końcowych ewaluacji – na końcu OPZ możemy opisać cechy jakimi chcielibyśmy, żeby charakteryzowały się końcowe produkty ewaluacji. Zazwyczaj jest to raport końcowy oraz krótsze formy – typu prezentacje lub broszury. Opisując tę część pamiętajmy o celach projektu i odbiorcach badania.

Kolejnym z tworzonych dokumentów będzie specyfikacja warunków zamówienia (SWZ)<sup>40</sup>. W dokumencie tym skupiamy się już na potencjalnym wykonawcy ewaluacji oraz zasadach jego wyboru.



### Zapamiętaj

**Pamiętajmy o tym, żeby nasz OPZ był spójny wewnętrznie. Uzasadnienie ewaluacji powinno być powiązane bezpośrednio z celami. Cele badania powinny zostać uszczegółowione przez pytania i kryteria. Metody obowiązkowe opisane w OPZ powinny być odpowiednie do zakresu i pytań ewaluacyjnych, a harmonogram musi oferować odpowiednią ilość czasu dla rzetelnej realizacji badania.**

Decyzja o wyborze wykonawcy ma kluczowe znaczenie dla jakości produktu końcowego, czyli raportu ewaluacyjnego i dla przyszłej współpracy z wykonawcą. Dlatego w celu zapewnienia sobie oferty wysokiej jakości i odpowiednio doświadczonego wykonawcy, powinniśmy precyzyjnie określić wymagania, jakie naszym zdaniem powinny zostać spełnione w tym zakresie. Przede wszystkim należy możliwie precyzyjnie określić wymagania dotyczące składu zespołu ekspertów i ich kwalifikacji. To właśnie zatrudnieni eksperci będą stanowili w największym stopniu o rzetelności, jakości oraz ostatecznie o użyteczności naszej ewaluacji. Wymagania dotyczące tych kwestii ujmujemy w SWZ. W tworzeniu tego dokumentu jednostki ewaluacyjne zazwyczaj wspierane są przez odpowiednie wydziały zajmujące się zamówieniami publicznymi. Do obowiązków osób bezpośrednio odpowiedzialnych za realizację ewaluacji należy zazwyczaj określenie warunków udziału w postępowaniu (kryteria formalne/dostępowe) oraz określenie kryteriów oceny ofert<sup>41</sup>.

<sup>40</sup> Formalnie OPZ oraz istotne postanowienia umowy są załącznikiem do SWZ.

<sup>41</sup> Pełną listę elementów, które musi zawierać SWZ prezentuje art.281 Ustawy PZP.

Kryteria formalne weryfikują posiadaną wiedzę i doświadczenie wykonawcy. Najczęściej w ramach tych kryteriów wymagamy od potencjalnego wykonawcy przedstawienia dowodów na zrealizowanie w przeszłości określonych usług (np. badania ewaluacyjne i/lub analizy społeczno-ekonomiczne w wymaganym obszarze tematycznym) oraz dowodów na dysponowanie odpowiednim składem zespołu badawczego (liczba osób oraz ich wiedza i doświadczenie). Oferta Wykonawcy, który nie spełnia kryteriów formalnych zostaje odrzucona.



### Zapamiętaj

**Określając minimalne doświadczenie dla członków zespołu badawczego należy mieć na względzie, że eksperci zdobywają doświadczenie na ogół znacznie wolniej niż firmy. W związku z tym, należy realistycznie określać wymogi co do doświadczenia ekspertów jednocześnie możliwie wydłużając okres, w którym jest ono mierzone.**

Kryteria jakościowe (zwane też merytorycznymi) dotyczą oceny jakości przedstawionej w ofercie koncepcji badawczej, metodyki i sposobu realizacji badania.

W ramach tego typu kryteriów najczęściej przedmiotem szczegółowej oceny jest adekwatność i innowacyjność podejścia badawczego do postawionego celu badania i pytań badawczych. Oceniana jest również jakość zaproponowanych przez wykonawcę metod i technik badawczych (w różnych fazach badania), adekwatność doboru prób badawczych czy sposoby radzenia sobie z ryzykiem mogącym wystąpić w trakcie realizacji badania (np. niedostępność czy niekompletność baz danych, niedostępność kluczowych „respondentów”). Pamiętajmy, że im szerzej i bardziej szczegółowo opisaliśmy wymagane metody tworząc OPZ, tym mniejsze wymagania w tym zakresie powinniśmy formułować względem oferty. Przykładowo, jeśli w naszym odczuciu wymagane przez nas w OPZ metody są kompletne, to nie powinniśmy zachęcać wykonawcy (poprzez przyznawanie punktów) do proponowania dodatkowych metod (oczywiście dalej pozostawiając mu możliwość proponowania pewnych uzupełnień).

Wybierając kryteria jakościowe należy pamiętać, że każde z nich można zastosować w formie stopniowalnej (stopień przydatności, spójności, etc. punktowany w sposób



### Dobra praktyka

**Dobłą praktyką jest zabezpieczenie się przed wyborem słabej oferty poprzez odpowiednie zapisy SWZ. Można to osiągnąć ujmując w tym dokumencie zastrzeżenie, iż oferty które nie osiągną określonej minimalnej wartości punktów w ramach kryteriów jakościowych będą odrzucane jako niezgodne z warunkami zamówienia (na podstawie art. 226 ust. 1 pkt 5 ustawy PZP).**

Wybierając kryteria jakościowe należy pamiętać, że każde z nich można zastosować w formie stopniowalnej (stopień przydatności, spójności, etc. punktowany w sposób

„do ... pkt” lub poprzez wskazanie stopni, tj. np. „2 pkt za ..., 3 pkt za ...”) lub zerojedynkowej (trafność, przydatność, spójność, etc.) dostosowując odpowiednio punktację.

**Tabela 5.** Przykładowe kryteria jakościowe oceny ofert na badanie ewaluacyjne

|                                                                                                                                                                                                                                                                                                                                                                                             |
|---------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|
| Stopień przydatności/trafność (w kontekście realizacji celów badania) oparcia koncepcji badawczej na teorii naukowej popartej odniesieniem/ami do publikacji;                                                                                                                                                                                                                               |
| Stopień przydatności/trafność (zapewniającą realizację celu głównego oraz celów szczegółowych badania) rozszerzenia koncepcji badawczej o dodatkowe pytania badawcze;                                                                                                                                                                                                                       |
| Trafne, niepowielające pytań zaproponowanych przez Zamawiającego, niepogłębiające pytań zaproponowanych przez Zamawiającego, ale eksplorujące nowe wymiary, choć mieszczące się w zakresie badania (najlepiej wskazać liczbę ocenianych dodatkowych pytań);                                                                                                                                 |
| Stopień przydatności/trafność (zapewniającą realizację celu głównego i celów szczegółowych badania oraz odpowiedzi na wszystkie pytania ewaluacyjne) rozszerzenia metodyki badawczej o dodatkowe metody/techniki badawcze (najlepiej wskazać wprost w SWZ, których celów/obszarów badawczych powinny dotyczyć te metody, czy powinny to być metody gromadzenia czy analizy i oceny danych); |
| Stopień trafności/trafność zidentyfikowania potencjalnych obszarów ryzyka, mogących wystąpić w trakcie realizacji badania oraz środków minimalizujących i wpływających na poprawę rzetelności badania.                                                                                                                                                                                      |

Źródło: Olejniczak i in. 2012

Poza kryteriami jakościowymi w SWZ określamy również ile punktów potencjalny wykonawca będzie mógł maksymalnie otrzymać za cenę wykonania usługi. W większości zamówień publicznych stosunek kryterium cenowego do merytorycznych kryteriów oceny ofert wynosi nie więcej niż 50:50. W przypadkach jednak gdy oczekujemy od wykonawcy zaproponowania w ofercie szczegółowej metodyki badania, gdyż sami sprecyzowaliśmy tylko pewne minimalne wymagania – waga kryteriów jakościowych może być większa i sięgać przykładowo nawet 70% (analogicznie, jeśli znamy metodykę realizacji badania to będzie interesować nas w większym stopniu cena). By ograniczyć ryzyko wyboru bardzo niskiej jakościowo oferty, wskazane jest określenie pewnego minimum punktowego w ramach kryteriów jakościowych, poniżej której oferta podlega odrzuceniu. Dobrym rozwiązaniem jest też podanie już przy ogłoszeniu zamówienia, jaką kwotę maksymalnie chcemy przeznaczyć na realizację ewaluacji. W przeciwnej sytuacji potencjalny Wykonawca nie ma niezbędnych informacji pozwalających na przygotowanie trafnej i realistycznej oferty.



### Uwaga, Pułapka!

Wymaganie przedstawienia bardzo szczegółowych (i co za tym idzie pracochłonnych i kosztownych) elementów metodyki już w ofercie nie zawsze jest uzasadnione. Generuje to znaczne koszty po stronie oferenta. Takie materiały, jak np. konkretne narzędzia badawcze (kwestionariusze ankiet, scenariusze wywiadów), specjalnie dla danego badania opracowane sposoby prezentacji danych (infografiki/mapy), powinny być przygotowywane już po podpisaniu umowy z wykonawcą.

W końcowym etapie prac nad przygotowaniem zamówienia na ewaluację warto sprawdzić zrozumiałość, zasadność i spójność przyjętych elementów zamówienia (czyli zapisów OPZ i SWZ). Dobrą praktyką jest powierzenie tego zadania osobie innej niż ta, która przygotowywała dokumenty (przykładową listę sprawdzającą zawarto w Tabeli nr 6).

**Tabela 6.** Lista sprawdzająca zrozumiałość i spójność przyjętych elementów zamówienia na ewaluację

| Pytanie sprawdzające                                                                                                                                | Tak/Nie |
|-----------------------------------------------------------------------------------------------------------------------------------------------------|---------|
| Czy kryteria są zrozumiałe (dla potencjalnych wykonawców ale również dla oceniających)? Czy można jasno wytłumaczyć stosowane w kryteriach pojęcia? |         |
| Czy w wymogach wobec oferty zapisanych w SWZ zawarto wszystkie elementy, które umożliwią ocenę ofert w oparciu o kryteria?                          |         |
| Czy w OPZ zawarto elementy niezbędne do przygotowania oferty?                                                                                       |         |
| Czy kryteria zawierają czynnik oceniający (np. trafność, spójność z celami, użyteczność z punktu widzenia celów badania)?                           |         |
| Czy kryteria zawierają więcej niż 1 czynnik oceniający (np. spójność i jasność)?                                                                    |         |
| Czy kryteria powtarzają się w jakimś zakresie?                                                                                                      |         |
| Czy w przypadku stosowania kryteriów zerojedynkowych lub wg stopni, określono dokładnie pułapy (np. liczba pytań, wielkość próby, etc.)?            |         |
| Czy w OPZ i SWZ piszemy tylko o kwestiach istotnych z punktu widzenia zaspokojenia naszych potrzeb informacyjnych?                                  |         |

Źródło: Opracowanie własne na podstawie K. Olejniczak, B. Ledzion, A. Weremiuk (red.), Realizacja badań ewaluacyjnych w świetle Prawa Zamówień Publicznych, MRR, Warszawa 2012 r.

Ważnym, lecz często niedocenianym elementem SWZ (stanowiącym załącznik), są projektowane postanowienia umowy (PPU). Badania ewaluacyjne są bardzo specyficznym rodzajem zamówień publicznych i dotyczą pozyskania nowej wiedzy. Standardowe zapisy umów funkcjonujące w urzędzie nie zawsze się w tym względzie sprawdzają.

Osoby zlecające ewaluację powinny więc szczegółowo przeanalizować (z punktu widzenia specyfiki realizacji konkretnego badania ewaluacyjnego) wzór postanowień umowy. Jest to szczególnie istotne, gdy instytucja ma niewielkie doświadczenie w zlecaniu badań ewaluacyjnych.

Warto w tym względzie rozważyć pewne zapisy, których zastosowanie może wpłynąć na zwiększenie elastyczności prowadzenia badania. Przykładem takiego zapisu są kwestie terminów. Realizacja ewaluacji często jest zależna od wielu czynników, które są trudne do przewidzenia (np. zmiana harmonogramu opracowania strategii lokalnej, wynikająca ze zmian ustawowych, rzutuje istotnie na przebieg ewaluacji *ex ante*). Warto więc zawrzeć zapis w umowie, wskazujący że za obopólną zgodą stron możliwe jest wydłużenie czasu realizacji badania (lub poszczególnych etapów jego realizacji) – o ile służyć to będzie podniesieniu finalnej jakości wyników ewaluacji.

Kolejną kwestią nad którą można się zastanowić jest możliwość przekazania płatności częściowych. Jest to istotne w przypadku większych badań (szczególnie z dużym komponentem badań terenowych), bo pozwala na zapewnienie płynności finansowej niewielkim firmom badawczym funkcjonującym na rynku lokalnym.

Odbiór wyników badania, często obejmuje liczne interakcje. W umowie konieczne jest więc także dokładne opisanie tej kwestii. Umowa powinna też dawać możliwość (o ile w całości nie jest realizowana w formie elektronicznej) podpisywania różnego rodzaju dokumentów (oświadczenia, wyjaśnienia, protokoły) również elektronicznym podpisem kwalifikowalnym. Wpływa to nie tylko na czas realizacji oraz koszty dla zamawiającego i wykonawcy, lecz również jest istotne w kontekście ochrony środowiska naturalnego.

Na końcu opiszemy pokrótce w jaki sposób dokonywać oceny ofert na podstawie określonych uprzednio kryteriów. Zamawiający zobowiązany jest do przeprowadzenia oceny ofert w rzetelny sposób zapewniający zachowanie uczciwej konkurencji oraz równe traktowanie wykonawców. Punktem wyjścia do oceny przedłożonych ofert w ramach PZP są zapisy SWZ, a w szczególności kryteria formalne i merytoryczne.

Wybierając ofertę na badanie o wartości poniżej 130 tysięcy PLN również należy stosować powyższe zasady. W tym wypadku rolę SWZ-u pełni ogłoszenie zamawiającego informujące o planowanej ewaluacji i zapraszające do składania ofert. W związku z tym ogłoszenie to musi koniecznie zawierać precyzyjnie i zrozumiale opisane kryteria, na podstawie których wyłoniony zostanie Wykonawca badania.

Dla wyboru wykonawcy, czy to w ramach PZP czy w ramach postępowania prowadzonego poza ustawą, warto powołać wieloosobową komisję<sup>42</sup>. W skład tego zespołu oceniającego mogą wejść nie tylko osoby odpowiedzialne bezpośrednio za realizację badania, ale również osoby posiadające merytoryczną wiedzę na temat przedmiotu ewaluacji. Przykładowo ewaluując efekty realizacji strategii rozwoju gminy do komisji zaprosimy nie tylko przedstawicieli jednostki ewaluacyjnej, ale również wybrane osoby odpowiedzialne za opracowanie lub realizację tejże strategii. Stworzenie kolegialnego ciała oceniającego, posiadającego szeroką wiedzę i różne perspektywy pozytywnie wpłynie na rzetelność oceny.



### Zapamiętaj

**Opracowanie dobrej pod względem jakości oferty jest pracochłonne. Udzielając zamówienia należy przewidzieć odpowiedni czas na jej przygotowanie – co najmniej 2-3 tygodnie.**

## Krok 7a: Współpracujemy z wykonawcą w trakcie realizacji badania

W niniejszym kroku opisano współpracę z zewnętrznym wykonawcą badania od momentu podpisania umowy do przekazania raportu końcowego (kwestie związane z samym raportem końcowym szerzej zostaną opisane w rozdziale 6).

W przypadku ewaluacji zewnętrznych i hybrydowych będziemy mówili o współpracy osób odpowiadających za ewaluację po stronie zamawiającego z zewnętrznymi ekspertami. Mijmy świadomość, że częściowo ramy współpracy będą w tej sytuacji dodatkowo określone przez zapisy umowy, czy też Prawa Zamówień Publicznych.

Można zidentyfikować pewne ogólne zasady czy też dobre praktyki współpracy, jakimi należy się kierować w trakcie realizacji badania. Po pierwsze wszystkie role w procesie ewaluacji muszą zostać określone w sposób jednoznaczny. Podstawowe postanowienia w tym zakresie powinny zostać ujęte już w umowie i OPZ dotyczącym realizacji danej ewaluacji. Doprecyzowanie tych kwestii powinno nastąpić na kolejnych etapach realizacji badania.

<sup>42</sup> W przypadku postępowań prowadzonych powyżej progu unijnego komisja musi liczyć sobie przynajmniej 3 członków. W przypadku postępowań poniżej ww. progu powoływanie komisji de iure nie jest nawet wymagane.

Kolejne zasady to zaufanie i wzajemny szacunek. Są one powiązane z opisanym w pierwszej kolejności jasnym określeniem ról. Skoro ustalone zostały role, należy uszanować konsekwencje tego faktu. Przykładowo polega to na nie negowaniu prawa zespołu badawczego do formułowania wniosków i rekomendacji, nawet jeśli są one krytyczne. Z drugiej strony musimy pamiętać, że osoby odpowiedzialne za realizację zadań, które są ewaluowane zazwyczaj są ekspertami w ich zakresie. W związku z powyższym ich wiedza i osądy muszą zostać uwzględnione w procesie badawczym.

Bardzo istotną i dobrą praktyką jest ustanowienie grupy sterującej dla realizacji ewaluacji. Mając na uwadze jakość procesu, na jak najwcześniejszym etapie, powinniśmy zebrać wszystkie adekwatne i istotne dla realizacji ewaluacji osoby i instytucje. Najlepiej włączyć je do prac już na etapie formułowania koncepcji badania lub OPZ-u. Niemniej jeśli nie miało to miejsca wcześniej, to etap po podpisaniu umowy z zewnętrznym wykonawcą (w przypadku ewaluacji zewnętrznej) jest najpóźniejszym momentem kiedy należy to zrobić.

Ostatnią dobrą praktyką, o której warto wspomnieć jest zadbanie o odpowiednią intensywność kontaktów pomiędzy stronami procesu. Dobrze jest umówić się na regularne spotkania lub przysyłanie wiadomości sprawozdawczych, w ramach których zespół prowadzący badanie będzie na bieżąco przekazywał informacje na temat postępu prac, pojawiających się barier itp.

W przypadku, gdy pojawiają się jakiekolwiek poważniejsze problemy w realizacji procesu badawczego, zespół badawczy powinien od razu informować o nich zamawiającego. Jest to szczególnie istotne w przypadku ewaluacji zewnętrznej, ponieważ nie zrealizowanie zadania niesie ze sobą konsekwencje prawne. Informacja od zespołu powinna zawierać w miarę możliwości od razu propozycje dostępnych sposobów rozwiązania problemów. Taka bieżąca komunikacja na temat przebiegu badania pozwala na wczesną identyfikację barier, zagrożeń terminów i ewentualne modyfikacje koncepcji badawczej (np. w przypadku obiektywnych trudności w uzyskaniu danych; weryfikacji i zmian wymagań odnośnie zakresu wyników badania, w tym modyfikacji prób, grup respondentów, narzędzi).

Bieżąca komunikacja może zostać uzupełniona również bezpośrednim uczestnictwem przedstawicieli zamawiającego w realizacji badania w roli obserwatora. Jeśli dysponujemy odpowiednimi zasobami kadrowymi oraz czasem, warto rozważyć wysłanie pracowników do partycypacji w tej roli przykładowo w wywiadach fokusowych, wizytach studyjnych czy innych metodach badawczych realizowanych przez wykonawcę. Pozwoli to z jednej strony zweryfikować realnie jakość i sposób przeprowadzenia zaplanowanych metod, z drugiej zaś strony przyczyni się do zwiększenia poziomu kompetencji pracowników zamawiającego.



Wracając do chronologicznego opisu etapów badania należy wskazać, że zazwyczaj współpraca pomiędzy wykonawcą zewnętrznym a zamawiającym rozpoczyna się od spotkania uzgodnieniowego. Spotkanie takie jest szansą na doprecyzowanie oczekiwań względem ewaluacji i jej wyników. Zamawiający może przedstawić w jego trakcie wskazówki dotyczące raportu metodycznego i jednocześnie odpowiedzieć na pytania wykonawcy związane z badaniem. Spotkanie tego typu jest realizowane w formie tzw. warsztatu strukturyzującego (więcej na ten temat w rozdziale 7). W praktyce często jest dzielone na dwa osobne spotkania. Jedno w mniejszym gronie, które służy bardziej wyjaśnieniu kwestii organizacyjnych dot. realizacji umowy. Drugie, z udziałem większej liczby osób i poświęcone dyskusji merytorycznej dot. interpretacji i doprecyzowania koncepcji badania.

Na początku badania wszyscy członkowie zespołu badawczego powinni również podpisać deklarację poufności (przykładowa deklaracja poniżej).

#### DEKLARACJA POUFNOŚCI CZŁONKA ZESPOŁU EWALUACYJNEGO

w trakcie realizacji badania ewaluacyjnego „.....”

Zobowiązuję się do zachowania poufności i tajności odnośnie do jakichkolwiek informacji, danych czy dokumentów ujawnionych mi lub odkrytych przeze mnie, lub przygotowanych przeze mnie w trakcie przeprowadzania przedmiotowej oceny, i zobowiązuję się, że informacje te będą wykorzystane jedynie do celów oceny i nie zostaną ujawnione osobom trzecim.

Zobowiązuję się również nie zatrzymywać kopii jakichkolwiek dostarczonych mi pisemnych informacji i opracowań, w szczególności objętych prawem autorskim, i zwrócić wszystkie udostępnione mi kopie po ich wykorzystaniu do przeprowadzenia oceny. W przypadku dokumentów w wersji elektronicznej zobowiązuję się do ich usunięcia.

Nazwisko: .....

Podpis: .....

Data: .....

Zazwyczaj kolejnym kamieniem milowym w procesie ewaluacji jest opracowanie raportu metodycznego (RM). W przypadku zlecenia badania na zewnątrz, RM stanowi w praktyce trzeci etap strukturyzacji – doprecyzowujący cały proces (pierwszy etap to OPZ a drugi to oferta merytoryczna wykonawcy).

Dobry RM powinien stanowić precyzyjne uszczegółowienie oferty lub koncepcji opracowanej wewnętrznie i zawierać przynajmniej następujące elementy:

- opis metodyki i propozycje narzędzi badawczych,
- szczegółowe pytania badawcze,

- określenie doboru prób i sposobów dotarcia do danych,
- określenie źródeł danych.

Warto tu wskazać, że szczegółowe opisy poszczególnych narzędzi badawczych (czyli np. kwestionariusze ankiet i scenariusze wywiadów) są w wielu przypadkach opracowywane na późniejszym etapie realizacji ewaluacji. Zazwyczaj opracowywanie tych materiałów jest procesem ewolucyjnym i iteracyjnym. Przykładowo można wykorzystać wnioski z wcześniej zrealizowanych metod do optymalizacji czy wręcz opracowania narzędzi realizowanych w kolejnych krokach.

Dobłą praktyką jest przeprowadzanie szerokich konsultacji RM. W przypadku ewaluacji zewnętrznej zadaniem pracowników odpowiedzialnych po stronie zamawiającego za realizację ewaluacji jest animowanie i moderowanie dyskusji nad raportem pozostałych uczestników procesu oraz weryfikacja zasadności zgłaszanych przez nich uwag.

Opracowywanie RM jest dobrym momentem, aby ustalić dostępność danych niezbędnych do przeprowadzenia badania. Istotne jest również pozyskanie danych kontaktowych osób, które mogą wesprzeć realizację badania i dostarczyć ewaluatorom istotne dokumenty i dane teledadresowe.

W przypadku realizacji ewaluacji zewnętrznych zamawiający na początku procesu przygotowuje także dla ewaluatora list uwierzytelniający. Dobrym rozwiązaniem jest również umieszczenie informacji o badaniu na stronach instytucji zlecającej projekt. Bardzo istotne jest, by zarówno list jak i informacja na stronie zawierała wszystkie podstawowe informacje na temat celów badania oraz planowanych w jego ramach działań. Ważne jest również ujęcie w komunikacie wskazań kontaktowych umożliwiających uzyskania dodatkowych informacji na temat projektu. Takie uwierzytelnienie może znaleźć oczywiście zastosowanie również w innego typu ewaluacjach (wewnętrznych i hybrydowych).

W przypadku ewaluacji zewnętrznych formalna współpraca zamawiającego z wykonawcą często kończy się wraz z przyjęciem przez Zamawiającego finalnej wersji raportu końcowego (RK) z badania. W późniejszym terminie zazwyczaj odbywa się spotkanie na którym prezentowane są wyniki przeprowadzonej ewaluacji. Dobłą praktyką jest ujęcie w umowie również tych ww. działań, mających miejsce po przyjęciu RK (prezentacja wyników, prace nad rekomendacjami itp.).

**LIST UWIERZYTELNIĄCY**

Uprzejmie informuję, że na zlecenie [nazwa instytucji] firma [nazwa] realizuje badanie ewaluacyjne pod tytułem: [nazwa]. Powyższe badanie będzie realizowane w terminie od ..... roku do ..... roku.

Celem badania jest .....

W ramach badania ewaluowany będzie wpływ interwencji [opis/nazwa zakresu badanej interwencji/dokumentu]. Badanie powinno pozwolić na sformułowanie wniosków i rekomendacji w tym zakresie.

Pragnę zapewnić, że członkowie zespołu kierować się będą zasadą zachowania poufności informacji uzyskiwanych w trakcie badania i gromadzić wyłącznie informacje bezpośrednio związane z badaniem.

W związku z powyższym zwracam się z uprzejmą prośbą o udostępnienie dla potrzeb analizy niezbędnej dokumentacji oraz informacji przydatnych w realizacji powyższego badania, jak również współdziałanie z Wykonawcą w trakcie jego przeprowadzania.

Liczę na przychyłność z Państwa strony i z góry serdecznie dziękuję za współpracę.

Informacji na temat przedmiotowego badania udzielają:

- ze strony [nazwa instytucji zlecającej] – Pani ....., adres e-mail: ....
- ze strony firmy ..... – Pan ....., adres e-mail: ....., tel.: .....

## Krok 7b: Realizacja badania ewaluacyjnego siłami wewnętrznymi

W niniejszym kroku opiszemy pewne, specyficzne dla ewaluacji wewnętrznej, zasady którymi należy się kierować w trakcie fazy realizacyjnej badania.



### Dobra praktyka

Przystępując do opracowania koncepcji badania warto jest, jeżeli do tej pory tego nie zrobiliśmy, rozpisać logikę badanej interwencji (patrz rozdział 2). Bardzo nam to ułatwi ewentualne doprecyzowanie zakresu badania i pytań badawczych. Pozwoli na ustrukturyzowanie procesu analizy danych zastanych. Ułatwi też opracowanie konkretnych narzędzi badawczych do badań jakościowych (i ewentualnie ilościowych).

Podjęliśmy decyzję, że realizujemy ewaluację wewnętrznie. Mamy opracowaną wstępną koncepcję realizacji badania (więcej w rozdziale 4) oraz zabezpieczone odpowiednie ku temu zasoby (więcej w rozdziale 3). Kolejnym działaniem, które musimy przeprowadzić jest uszczegółowienie koncepcji. Dotyczy to zwłaszcza metodyki badania.

Jeżeli nie posiadamy zespołu doświadczonego w realizacji badań ewaluacyjnych wskazane jest przygotowanie względnie prostego schematu badawczego. Warto rozważenia jest również ograniczenie się do zastosowania relatywnie prostszych metod i technik. Dotyczy to zwłaszcza etapu analizy danych (powinniśmy ograniczyć się do metod, które już znamy, albo wiemy, że możemy szybko nabyć odpowiednie kompetencje).



#### Uwaga, Pułapka!

**Drogi czytelniku! Zanim zdecydujesz się na realizację badania siłami wewnętrznymi (szczególnie gdy nie masz wcześniejszych doświadczeń badawczych) przeczytaj do końca poradnik. W szczególności zapoznaj się z rozdziałem 7. dotyczącym metodyki badań.**

W przypadku prostszych badań często sprawdza się schemat, w którym:

1. Najpierw analizujemy dostępne dane na temat badanej interwencji (dokumenty programowe, dane z monitoringu, wcześniejsze badania itp.);
2. Przeprowadzamy kilka rozmów (wywiadów) z najważniejszymi interesariuszami (osobami odpowiedzialnymi za zarządzanie i wdrażanie interwencji, oraz wybranymi osobami reprezentującymi odbiorców wsparcia);
3. Jeżeli dostępne dane z systemu monitorowania (lub z innych rejestrów publicznych) nie są wystarczające, możemy rozważyć realizację badania ankietowego;
4. Analizujemy zgromadzone dane i przygotowujemy pierwszy zarys raportu końcowego (RK). W międzyczasie możemy uzupełnić zidentyfikowane luki w informacjach metodami jakościowymi (wywiady, studia przypadków);
5. Na ostatnim etapie warto przeprowadzić spotkanie/a podsumowujące wyniki - warsztat z ekspertami dziedzinowymi oraz kluczowymi odbiorcami wyników. Pozwoli nam to lepiej przygotować finalne wnioski i rekomendacje z badania.

Oczywiście powyższy schemat jest bardzo dużym uproszczeniem. Nie sprawdzi się w przypadku każdej ewaluacji. To, czy możemy go zastosować, zależy w szczególności od pytań badawczych, czy też dostępnych danych (np. danych kontaktowych w przypadku chęci realizacji badania ankietowego). W zależności od naszych kompetencji badawczych i analitycznych możemy go rozbudowywać i modyfikować.



### Zapamiętaj

**W przypadku realizacji badania ewaluacyjnego absolutnie kluczowa jest rzetelna analiza danych zastanych. Powinna ona być zaplanowana do realizacji na początku każdego badania. Nie warto bowiem zbierać nowych informacji, jeżeli mamy dostęp do już istniejących danych/rejestrów. Jest to szczególnie istotne w przypadku badań wewnętrznych. Dzięki temu, że pracujemy w danej instytucji, możemy mieć dostęp do unikalnych danych, do których dostępu mogą nie mieć ewaluatorzy zewnętrzni.**

Najważniejszym kryterium doboru metodyki badania są cele i pytania ewaluacyjne. Cele i pytania, nie tylko w ewaluacji, ale każdym innym badaniu, ukierunkowują cały proces badawczy. Projektując metodykę badania powinniśmy upewnić się, że każde pytanie badawcze jakie postawiliśmy ma przypisane źródła danych i metodę badawczą, najlepiej więcej niż jedną (to tzw. triangulacja metod badawczych).

Najlepsze rezultaty daje stosowanie **mieszanej metodyki**, czyli uwzględnienie w badaniu różnych metod i technik badawczych. Metody badań mają różne założenia i przeznaczenie, więc mogą się doskonale uzupełniać. Metody ilościowe zazwyczaj pozwalają uzyskać wyniki od dużych grup badanych, a odpowiednio zaimplementowane (np. z wykorzystaniem metody reprezentacyjnej) dają możliwość uogólniania wyników na całą ich populację. Metody jakościowe pozwalają z jednej strony na eksplorację tematów nowych, nieznanych, a z drugiej na pogłębienie i interpretację kluczowych zagadnień, np. poznanie oczekiwań, motywacji interesariuszy, czynników wpływających na rezultaty interwencji.



### Zapamiętaj

**Rekomenduje się stosowanie tzw. triangulacji metod badawczych. Oznacza to, że w badaniu tego samego problemu, czy odpowiedzi na to samo pytanie badawcze, użyjemy kilku różnych metod badawczych, różnych technik analizy i różnych źródeł danych. Celem takiego działania jest konfrontacja otrzymanych informacji (np. od różnych interesariuszy i pozyskanych różnymi metodami) oraz ich podsumowanie. Im szersza triangulacja, tym silniejsza moc wnioskowania w badaniach – jeśli wyniki uzyskane różnymi metodami i z różnymi respondentami/interesariuszami będą się potwierdzać, my będziemy mieć większe zaufanie, co do ich trafności.**

Warto zadbać, by w badaniu znalazła się konfiguracja metod, które możliwie wyczerpująco pozwolą odpowiedzieć na postawione pytania. Użyteczne jest przygotowanie tabeli, w której uszeregujemy ważne dla nas zagadnienia/pytania i wskażemy, które metody

badawcze zostaną wykorzystane do uzyskania odpowiedzi na nie. Rzut oka na tabelę pozwoli też ocenić, które pytania będą dogłębnie zanalizowane, poprzez użycie większej liczby metod, a więc w obrębie których pytań/tematów nasze wnioski będą najsilniejsze i najmocniej uzasadnione.

**Tabela 7.** Przykładowa tabela celów/pytań badawczych i metod

|                 | analiza danych zastanych | badanie sondażowe CAWI | wywiady pogłębione IDI | panel ekspertów |
|-----------------|--------------------------|------------------------|------------------------|-----------------|
| pytanie/temat 1 | x                        |                        | x                      | x               |
| pytanie/temat 2 | x                        | X                      |                        |                 |
| pytanie/temat 3 | x                        |                        | x                      |                 |
| pytanie/temat 4 | x                        | X                      |                        |                 |
| pytanie/temat 5 | x                        |                        |                        | x               |

Źródło: opracowanie własne

Ważnym etapem konceptualizacji badania, gdy realizujemy je siłami wewnętrznymi, jest opracowanie narzędzi badawczych. W pierwszej kolejności warto przeprowadzić wstępną analizę danych zastanych, a w kolejnym kroku opracować scenariusze wywiadów. Dopiero po przeprowadzeniu analizy danych zastanych i przeprowadzeniu kilku wywiadów przejść do opracowania narzędzia ankietowego (jeżeli te są przewidziane w badaniu). Analiza danych zastanych oraz wywiady mogą nas bowiem naprowadzić na odpowiednie dane. Być może tak uzyskane dane naprowadzą nas na trop kolejnych zagadnień, które możemy poruszyć i zgłębić w badaniach ankietowych, a tym samym przygotować lepszy kwestionariusz.

Gdy zakończymy fazę konceptualizacji badania warto jeszcze raz zweryfikować wszystkie działania, jakie do tej pory podjęliśmy. Warto powtórnie zastanowić się, czy zakres i pytania badawcze nie są zbyt ambitnie określone. Wskazane jest też jeszcze raz dokonać przeglądu dostępnych zasobów. W szczególności dotyczy to czasu, który będzie nam potrzebny na realizację badania. Upewnijmy się też, że właściwe osoby decyzyjne są świadome, że w najbliższych (na ogół kilku - kilkunastu tygodniach) ich podwładni będą w znacznym stopniu zajęci realizacją badania i ograniczy to ich możliwość wywiązywania się z innych obowiązków.

Teraz pozostaje nam przystąpić do realizacji opracowanej koncepcji. Pomocne przy tym mogą być wskazówki dot. wykorzystania poszczególnych metod przedstawione w rozdziale 7. W wielu przypadkach będziemy musieli jednak sięgnąć głębiej i uzupełnić naszą wiedzę również innymi źródłami. Nie bójmy się uczyć przez praktykę. Ważne jest tylko, by szybko wyciągać wnioski z popełnionych błędów. Gdy obserwujemy, że scenariusz wywiadu nie pozwala nam przeprowadzić udanej rozmowy – zmodyfikujmy go przed

następnym wywiadem. Zanim wyślemy ankietę – przeprowadźmy próbę (pilotaż) na kilku respondentach i ewentualnie dokonajmy poprawek<sup>43</sup>.

Na etapie opracowania wyników bardzo istotne jest odróżnienie subiektywnych opinii czy ocen prezentowanych przez respondentów od wniosków, jakie finalnie wyciągamy. Czasem negatywna ocena badanego zjawiska może pochodzić od jednego respondenta czy interesariusza. Aby dana obserwacja miała charakter wiarygodnej powinna zostać skonfrontowana z opiniami innych badanych i innymi wnioskami na temat badanej interwencji. Dlatego w każdym badaniu tak ważna jest triangulacja, a każdy wniosek, który formułujemy powinien być uwiarygodniony użyciem zróżnicowanych danych i metod, a także uwzględniać perspektywę różnych interesariuszy.

Opracowanie raportu końcowego z badania zajmuje dużo czasu. W przypadku badań wewnętrznych zamiast obszernego raportu warto rozważyć krótsze formy. Być może w zupełności wystarczy rozbudowana prezentacja (więcej informacji na temat raportu końcowego oraz innych form i kanałów prezentacji wyników ewaluacji w rozdziale 6).

Na koniec warto zwrócić uwagę na ważny aspekt realizacji badania siłami wewnętrznymi – obiektywizm i niezależność w formułowaniu wniosków i rekomendacji. Realizując ewaluację wewnętrznie często oceniamy swoją pracę lub pracę naszych kolegów i koleżanek (a nierzadko również przełożonych). Stawia nas to w trudnej sytuacji, szczególnie gdy musimy być krytyczni. Gdy zakładamy, że nasza przygoda z realizacją badań ewaluacyjnych nie będzie tylko jednorazowa - warto rozważyć wdrożenie działań systemowych zapobiegających ewentualnym konfliktom. W szczególności można określić zadania i zasady funkcjonowania takiej jednostki ewaluacyjnej (nawet jeżeli jest to bardzo niewielki zespół) w regulaminie organizacyjnymi instytucji. Za dobrą praktykę należy uznać rozwiązania organizacyjne, w ramach których jednostka ewaluacyjna podlega bezpośrednio kierownictwu instytucji (dyrektora generalnego, zarządu itp.). Z jednej strony ułatwi to formułowanie obiektywnych wniosków i rekomendacji, nawet gdy są one krytyczne względem sposobu działania pozostałych jednostek danej instytucji. Z drugiej, zwiększy to szanse, że ustalenia z ewaluacji będą traktowane „poważnie”, a zalecenia wdrażane w działania organizacji.

<sup>43</sup> Pierwszy, wstępny pilotaż można też przeprowadzić na współpracownikach. Szczególnie osoby zaznajomione z ewaluowaną interwencją mogą dostarczyć nam cennych wskazówek.

## Rozdział 6. Produkty badania oraz wykorzystanie wyników ewaluacji

Tomasz Kupiec

### Krok 8: Opracowujemy raport końcowy z badania

#### Opracowanie raportu

Raport końcowy (RK) z badania jest istotnym elementem procesu ewaluacji, wieńczącym fazę badawczą. W przypadku ewaluacji zewnętrznej jego odbiór najczęściej kończy formalną współpracę zamawiającego z wykonawcą, a więc odbiór RK jest potwierdzeniem prawidłowości wykonania zleconych prac. Również przy ewaluacjach wewnętrznych w większych organizacjach jasno przydzielone są role wykonawców i odbiorców ewaluacji, a RK potwierdza wykonanie pracy przez pierwszych na rzecz drugich. Choć w dobie rozwoju technologii informacyjno-telekomunikacyjnych, multimedialnych i mediów społecznych RK przestał być jedynym narzędziem komunikowania wyników ewaluacji, to wciąż jego rola w tym zakresie jest bardzo ważna. Po pierwsze, oczekiwania odbiorców ewaluacji różnią się i nadal dla niektórych to RK jest preferowaną formą przekazywania ustaleń ewaluacji (zobacz Krok 9). Po drugie, i być może ważniejsze, RK komunikuje wyniki ewaluacji w sposób wyczerpujący i kompleksowy, dokumentując przebieg badania, w tym przyjęte założenia i wykorzystane metody, pokazując ścieżkę dochodzenia do wniosków i w związku z tym pozwalając weryfikować ich wiarygodność. RK umożliwia porównywanie i syntezywanie wyników ewaluacji realizowanych w różnym czasie, kontekstach, przez różnych wykonawców.

Nie istnieje jeden właściwy **rozmiar i struktura RK**. Powinny one wręcz różnić się w zależności od przedmiotu i celu ewaluacji, czy podejścia badawczego. W praktyce często decydujące są po prostu preferencje odbiorców. Instytucje zlecające często ściśle określają obie te kwestie w opisie przedmiotu zamówienia, albo dysponują szczegółowymi wytycznymi (np. CDC, 2013; UNICEF, 2017). W kontekście ewaluacji programów polityki spójności znacznie częściej spotyka się określenie górnego limitu stron RK i zwykle jest to 100–130 stron bez załączników. Wymóg ten stanowi dobrą motywację dla ewaluatora, by priorytetyzować dane, fakty, wnioski, które chce przekazać, eksponując te najważniejsze. Takie podejście zwiększa szansę dotarcia z przekazem do odbiorców. Dolny limit rzadko jest narzucany, ale zdecydowana większość RK przekracza 70 stron.



RK można uznać za odmianę raportu naukowego. W obu przypadkach konieczne jest przedstawienie po co badanie zostało zrobione, jak zostało zrobione i jakie uzyskaliśmy wyniki. Niemniej naukowcy kierują swoje raporty głównie do innych naukowców i kładą relatywnie duży nacisk na uwiarygodnienie badania oraz jego metodykę. Raporty ewaluacyjne są dla praktyków, decydentów, interesariuszy danej interwencji. Tu najważniejsze jest pokazanie **co z tego wynika i co należy zrobić**. Modelowa struktura RK może wyglądać następująco:



#### Modelowa struktura raportu końcowego

- Streszczenie wykonawcze
- Wprowadzenie (cel, przedmiot, metodyka)
- Przedstawienie wyników (na ogół kilka rozdziałów)
- Kluczowe wnioski i rekomendacje

**Streszczenie wykonawcze** przeznaczone jest przede wszystkim dla decydentów i powinno syntezywać na kilku stronach najistotniejsze ustalenia badania. Główny nacisk powinien być położony na wnioski, rekomendacje i odpowiednią formę – niespecjalistyczny, zrozumiały i angażujący język. W streszczeniu powinno być wyeksponowane to, co Grob (2015) nazywa „przekazem” – najważniejsza myśl, którą autor chce przekazać czytelnikowi, jedno, dwa zdania z podstawowym wnioskiem i rekomendacją.



#### Dobra praktyka

Sposobem na sprawdzenie skuteczności przekazu zawartego w RK jest tzw. **test mamy (mom test)**. To eksperyment myślowy w ramach którego wyobrażamy sobie, że w trakcie rodzinnego spotkania mówimy mamie, że właśnie skończyliśmy pisać raport ewaluacyjny, a ta gratulując nam pyta: „i co z niego wynika?”. Nasza odpowiedź zaczyna się od „Mamo, ten raport wskazał, że...” i powinna zamknąć się w jednym, maksymalnie dwóch zdaniach. Odpowiedź ta powinna pozwolić naszej matce, niemającej żadnego specjalistycznego przygotowania, na zrozumienie istoty problemu /rozwiązania, o którym mówimy. Przekaz powinien być, prosty, ale równocześnie konkretny, inspirujący, interesujący i prowokować reakcję w stylu: „Ojej, mam nadzieję, że władze rzeczywiście coś z tym zrobią.” Przykładem przekazu zdającego test może być to stwierdzenie ze streszczenia raportu z ewaluacji wpływu funduszy europejskich na rozwój Polski Wschodniej.

*W innych częściach kraju można zidentyfikować obszary, których sytuacja społeczno-gospodarcza jest porównywalna do słabiej rozwijających się terenów w Polsce Wschodniej. Bez dedykowanej pomocy w przyszłości, obszary te mogą stać się obszarami szczególnego natężenia problemów.*

**Wprowadzenie** zawiera podstawowe założenia ewaluacji. Opis celów badania czy charakterystyka ewaluowanej interwencji nie są może najważniejszymi elementami raportu, ale ułatwiają zrozumienie prezentowanych później wyników. Metodyka badania prezentowana jest również we wprowadzeniu albo jako następujący po nim odrębny rozdział. Podejście badawcze trzeba opisać możliwie zwięźle, ale ujmując wszystkie istotne elementy, w tym potencjalne ograniczenia metodyki. Paradoksalnie odbiorcy widząc opis ograniczeń uznają badanie za bardziej wiarygodne. Warto pamiętać, że choć w większości wypadków odbiorcy skupiają się na wynikach, to w przypadku ewentualnych zastrzeżeń do wyników i rekomendacji, krytyka skupia się bardzo często na podważaniu wiarygodności metodyki.



### Dobra praktyka

**Prezentację metodyki badania ewaluacyjnego zapewniającą równocześnie zwięźłość i wyczerpujące informacje można zorganizować następująco:**

- nie więcej niż jeden akapit w streszczeniu,
- strona lub dwie w głównej części raportu,
- pełny opis wraz z narzędziami w załączniku do raportu,
- umożliwienie kontaktu z autorami na wypadek dodatkowych pytań.

**Prezentowanie wyników** zajmuje największą część raportu i stanowi spore wyzwanie. Przede wszystkim należy pamiętać, że warto ograniczyć prezentację do wyników wprost odpowiadających na przyjęte wcześniej kryteria i pytania badawcze. Wszystkie inne, nawet ciekawe w opinii wykonawcy wykresy, zestawienia, techniczne detale, będą rozpraszały uwagę odbiorcy. Zacząć warto od odpowiedzi na pytania najważniejsze dla odbiorcy<sup>44</sup> i na nich skupiać się bardziej, a inne traktować pobieżnie. Podobna selekcja powinna być przeprowadzona ze względu na kryterium 'nowości'. Odbiorcy to bardzo często specjaliści w obszarze ewaluowanej interwencji. Ich uwagę przykują spostrzeżenia nowe, ciekawe, uzupełniające to, co już wiedzą. Konieczność przedzierania się przez rozwlekłe opisy kwestii oczywistych będzie ich tylko zniechęcać. W miarę możliwości prezentacja wyników powinna układać się w jakąś historię, np. wpływu programu na grupę docelową, czy problemów które trzeba rozwiązać – to również zwiększa szansę na zaangażowanie czytelnika (Evalcommunity, 2013). Jeżeli dane wynikające bezpośrednio z badania przeplatane są komentarzami, opiniami, interpretacjami, dopowiedzeniami autorów, trzeba to każdorazowo jasno zaznaczać. Podobnie jak w przypadku metodyki, więcej szczegółowych wyników można ująć w załączniku.

<sup>44</sup> Które to, należy ustalić już na etapie planowania badania, ale z góry można przyjąć, że nie będą to np. pytania dodane w ofercie przez wykonawcę.

**Wnioski i Rekomendacje.** Wyniki ewaluacji nabierają wartości jeśli są poprawnie zinterpretowane i uzupełnione wnioskami, które następnie stanowią podstawę formułowania rekomendacji. Wnioski prawie zawsze towarzyszą już prezentacji wyników, ale typową praktyką jest też ich podsumowanie czy powtórzenie na końcu raportu by jasno pokazać, że formułowane następnie rekomendacje wynikają z wniosków z badania. Więcej o rekomendacjach i ich znaczeniu dla wykorzystania ewaluacji w Kroku 10.

**Formatowanie raportu** w oczywisty sposób kieruje uwagę na niektóre treści i odciąga od innych. Wielu odbiorców pobieżnie przeglądając RK najprawdopodobniej zwróci uwagę na wszystko poza zwykłym tekstem. Dlatego ważne przekazy (i tylko one) powinny być podkreślone stylem, nagłówkami, punktami, pogrubieniem, dopiero w następnej kolejności rozmiarem fontu, podkreśleniem, czy kolorem.

Podobne zalecenie, czyli stosowanie z umiarem, tylko do podkreślenia najważniejszego przekazu, dotyczy wszelkich **grafik, wykresów, tabel**. Odbiorcy chętnie się na nich skupiają, samodzielnie analizują i interpretują, dlatego przed przygotowaniem RK warto zapoznać się z podstawowymi zasadami wizualizacji danych. Zagadnienie to wykracza poza zakres niniejszego poradnika, toteż proponujemy poniżej kilka dodatkowych lektur.



**Proponowane lektury w zakresie wizualizacji danych w ewaluacji:**

**Knaflíc, C.N. (2015). Storytelling danych. Poradnik wizualizacji danych dla profesjonalistów. Onepress**

**Evergreen, S. D. (2019). Effective data visualization: The right chart for the right data. SAGE publications.**

**Azzam, T., Evergreen, S., Germuth, A. A., & Kistler, S. J. (2013). Data visualization and evaluation. New Directions for Evaluation, 2013(139), 7-32.**

**Język i styl** którym się posługujemy decyduje o tym, czy nasz RK jest wyrazisty, zrozumiały, a równocześnie sugestywny i przekonujący. Osiągniemy to m.in. używając krótkich wyrazów i zdań, wyrazów powszechnie znanych, a unikając terminologii naukowej. Piszząc miejmy na uwadze, że najważniejszym zdaniem akapitu jest to pierwsze i w nim powinna być zawarta najważniejsza myśl, którą potem rozwijamy. Więc przydatnych zaleceń można znaleźć w proponowanych publikacjach.



#### Proponowane lektury w zakresie wizualizacji danych w ewaluacji:

Maziarz i in. (2012). Język raportów ewaluacyjnych. Ministerstwo Rozwoju Regionalnego.

Grob, G. F. (2015). Writing for impact. w: J. S. Wholey, H. P. Hatry, K. E. Newcomer (red.), Handbook of Practical Program Evaluation, wydanie trzecie, Jossey-Bass A Wiley Imprint, s. 739-764.

Hopson, R. K. (2000). How and Why Language Matters in Evaluation. New Directions for Evaluation, 86.

## Odbiór i ocena raportu

Doświadczenia z pola ewaluacji polityki spójności pokazują, że przy ewaluacjach zewnętrznych projekt RK powinien trafić do odbiorców na przynajmniej 4 tygodnie przed planowanym terminem przekazania jego ostatecznej wersji. W tym czasie zamawiający może zapoznać się z treścią raportu, zgłosić uwagi i dać wykonawcy 2-3 tygodnie na ich uwzględnienie.

Etap oceny projektu RK, to ostatni moment w którym odbiorca może jeszcze zaingerować i zgłosić oczekiwanie dostosowania ewaluacji do swoich potrzeb. Oczywiście jest to już za późno by zmieniać np. rozwiązania z fazy zbierania danych, ale możliwe jest dopracowanie warstwy analizy, interpretacji, wnioskowania, rekomendacji, sposobu prezentacji wyników.

By uniknąć rozczarowań na tym etapie, jak najwięcej oczekiwań dotyczących RK – przede wszystkim jego struktury, rozmiaru, ale też bardziej szczegółowe preferencje dotyczące zawartości i formy – warto ustalić najpóźniej w trakcie uzgadniania RM. Tylko precyzyjnie sformułowane oczekiwania są następnie możliwe do wyegzekwowania.



### **Dobra praktyka** **przykłady jasno sformułowanych oczekiwań dotyczących RK**

**Przedstawienie wyników nie może sprowadzać się jedynie do prezentacji danych i odpowiedzi pochodzących z realizacji poszczególnych metod/technik badawczych, lecz powinno być syntezą wyników uzyskanych na różnych etapach realizacji badania (Województwo Śląskie, 2021).**

**Opis wyników musi zawierać analizę i interpretację danych z poszczególnych źródeł (...) wraz z oceną ewaluatora (Ministerstwo Zdrowia, 2018).**

**Co oceniamy?** Ocena RK, w praktyce oznacza zwykle ocenę całego procesu badawczego. Świadomy odbiorca ewaluacji po zakończeniu badania i lekturze RK po prostu wie, czuje, czy badanie się udało i jest użyteczne. Z różnych powodów warto jednak etap oceny bardziej usystematyzować i sformalizować. Wiele organizacji o rozwiniętej praktyce ewaluacji posiada szczegółowe listy sprawdzające do oceny RK, a w nich wskazuje jej następujące cele:

- ukierunkowanie dyskusji między ewaluatorami i odbiorcami o preferowanej zawartości RK;
- informacja zwrotna dla autorów raportu;
- autorefleksja zamawiających ewaluację<sup>45</sup>.

Oczywistym kryterium oceny jest zgodność RK i procesu badania z określonymi na wstępie wymogami<sup>46</sup>. Pozostałe kryteria sprowadzają się najczęściej do sprawdzenia czy RK spełnia standardy opisane w pierwszej części tego podrozdziału, m.in.:

(1) czy jest jasny, kompletny i ma logiczną strukturę; (2) czy wszystkie wnioski wynikają wprost z wyników badania, (3) czy rekomendacje są zorientowane na działanie, praktyczne i konkretne?

<sup>45</sup> Polski system ewaluacji polityki spójności jest ciekawym przykładem w którym nastąpiła ewolucja podejścia do oceny RK i procesu ewaluacji i obecnie służy on głównie autorefleksji nad aspektami, które należy poprawić przy realizacji przyszłych badań i dobrymi praktykami, które warto powtarzać i rozpowszechniać (MFiPR, 2022).

<sup>46</sup> Np. zapisanymi w OPZ, jeśli to ewaluacja zewnętrzna.



**Przykłady wytycznych, check-list do oceny raportów końcowych z ewaluacji:**  
 Karta oceny badania wraz z instrukcją do ewaluacji polityki spójności 2021-27, <https://www.ewaluacja.gov.pl/media/118105/KartaOcenyBadania.docx>  
 UNEG Quality Checklist for Evaluation Reports, <http://unevaluation.org/document/download/853>  
 Departament Stanu, USA, Checklist for Assessing Evaluation Reports, <https://www.state.gov/wp-content/uploads/2018/12/Checklist-for-Assessing-Evaluation-Reports.pdf>  
 USAID, Evaluation Report And Review Template, [https://usaidlearninglab.org/sites/default/files/2022-09/evaluation\\_report\\_checklist\\_and\\_review\\_template\\_final2022.pdf](https://usaidlearninglab.org/sites/default/files/2022-09/evaluation_report_checklist_and_review_template_final2022.pdf)

RK niespełniający przyjętych wymogów nie powinien być akceptowany, ale kierowany do poprawy. W harmonogramie badania i wykorzystania jego wyników warto przewidzieć taką możliwość.



### Uwaga, Pułapka!

Częstym niedomaganiem RK, szczególnie w badaniach, które miały odpowiedzieć na relatywnie dużo pytań, jest ich niekompletność, tzn. pomijanie albo bardzo pobieżne traktowanie odpowiedzi na pytania ważne, ale trudne do zaadresowania i maskowanie tego faktu rozwlekaniem odpowiedzi na pytania mniej istotne. Rolą odbierającego raport jest zauważenie tego problemu i wyegzekwowanie jego usunięcia.

**Kto ocenia?** W przypadku ewaluacji zewnętrznych najczęściej pierwszym recenzentem jest osoba(y) odpowiedzialna za zlecenie ewaluacji w organizacji zamawiającej. Na jej barkach spoczywa wnikliwa weryfikacja spełnienia wymogów formalnych, ale też ta osoba powinna być wystarczająco kompetentna, by ocenić jakość procesu badawczego, logikę wnioskowania, poprawność rekomendacji. W przypadku ewaluacji wewnętrznych również warto wyznaczyć osobę, która będzie pełnić funkcję niezależnego recenzenta ewaluacji i ma ku temu odpowiednie kompetencje.



### Dobra praktyka – zewnętrzny ekspert

Odbiór i ocena RK jest odpowiedzialnym i trudnym zadaniem, wymagającym doświadczenia ogólnego w zakresie ewaluacji, ale też często wiedzy specjalistycznej związanej z tematem badania. Rozwiązaniem problemu niedostatecznego potencjału organizacji zamawiającej ewaluację może być zatrudnienie eksperta zewnętrznego, który może pełnić rolę recenzenta RK, ale też szerzej wesprzeć przygotowanie koncepcji badania, zamówienie, obsługę kontaktów z wykonawcą w trakcie realizacji ewaluacji.



### Zapamiętaj!

**W przypadku ewaluacji zewnętrznych, niewypełnienie wymogów dotyczących RK i całego procesu badania powinno być obciążone sankcją redukcji wynagrodzenia. To ważny argument dla zamawiającego na etapie oceny i odbioru RK. Zapisy dotyczące kar muszą być bardzo precyzyjnie sformułowane w umowie.**

Drugą, być może ważniejszą przy ocenie RK, jest perspektywa użytkowników / odbiorców ewaluacji. To oni, nawet nie mając kompetencji w zakresie ewaluacji, powinni ocenić, czy rekomendacje z badania są dla nich użyteczne: adekwatne, znaczące, wnoszące coś nowego, wykonalne. Odbiorcy najlepiej oceniają, czy wnioski z badania dostarczają im nowej wiedzy, albo wiarygodnie potwierdzają dotychczasowe przypuszczenia. Dopiero RK uwzględniający uwagi użytkowników i spełniający ich oczekiwania powinien być akceptowany.

## Krok 9: Rozpowszechniamy wyniki badania

Jak wskazano w rozdziale 1, sensem ewaluacji jest wspieranie decyzji wynikami badań. Praktyczne doświadczenia pokazują, że wyniki ewaluacji nie docierają jednak do odbiorców „tak po prostu”, nie jest to proces samoistny, ani łatwy. Rozpowszechnianie wyników ewaluacji wymaga dobrego planowania, aktywnych działań i dobrze dobranych narzędzi.

Choć konwencja kroków, którą przyjęto w tym poradniku, każe zaprezentować rozpowszechnianie jako element następujący po zakończeniu badania i odebraniu RK, to w rzeczywistości **rozpowszechnianie zaczyna się już od pierwszego kroku, gdy identyfikujemy potrzebę badawczą**, bo przecież jest to potrzeba konkretnych odbiorców, do których będziemy chcieli dotrzeć.

Skuteczne przekazywanie wiedzy z ewaluacji wymaga identyfikacji właściwych odbiorców i ich potrzeb informacyjnych oraz dobrania do nich odpowiednich: formy przekazu, kanału komunikacji, momentu przekazu.

Konieczność **identyfikacji odbiorców** już na początku procesu ewaluacji wynika przede wszystkim z faktu, że komunikacja w ramach ewaluacji zaczyna się już przed rozpoczęciem badania, ma miejsce w jego trakcie i po zakończeniu. Sensem komunikacji z odbiorcami przed i w trakcie ewaluacji jest:

- włączenie w decyzje dotyczące celów, zakresu ewaluacji, sposobu jej realizacji;
- informowanie o nadchodzących aktywnościach w ramach badania;
- przekazywanie częściowych wyników.

Działania te zwiększają szanse, że skutecznie zrealizujemy cele rozpowszechniania wyników po zakończeniu badania (Torres i in., 2005):

- przekazanie informacji, faktów, wyników;
- ułatwienie zrozumienia i interpretacji wyników;
- wsparcie podejmowania decyzji.

O ile dwa pierwsze cele dotyczą wszystkich odbiorców, trzeci związany jest z kategorią, którą możemy nazwać **bezpośrednimi użytkownikami**.

**Tabela 9.** Typowi odbiorcy i użytkownicy ewaluacji

| Osoby lub instytucje:                                                                                                                                                                                                                                                                  |                         |
|----------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|-------------------------|
| <ul style="list-style-type: none"> <li>• podejmujące decyzje o uruchomieniu, kontynuacji, modyfikacji lub zakończeniu ewaluowanej interwencji,</li> <li>• zaangażowane w projektowanie interwencji,</li> <li>• zaangażowane we wdrażanie interwencji,</li> </ul>                       | Bezpośredni użytkownicy |
| <ul style="list-style-type: none"> <li>• dotknięte pozytywnie lub negatywnie skutkami interwencji,</li> <li>• mające obowiązek lub prawo zapoznać się z wynikami (legislatorzy, media, społeczeństwo),</li> <li>• zainteresowane interwencją (adwokaci, krytycy, naukowcy).</li> </ul> | Pozostali odbiorcy      |

Źródło: Opracowanie własne na podstawie MFIPR (2022), Olejniczak i in. (2014)

Planując rozpowszechnianie produktów badania powinniśmy mieć na uwadze w pierwszej kolejności potrzeby i preferencje bezpośrednich użytkowników badania. W ich



identyfikacji pomaga świadomość, że osoby na różnych stanowiskach angażują się zwykle w interwencję i potrzebują wiedzy o niej na różnych etapach cyklu jej funkcjonowania (zobacz rozdział 2) (Olejniczak i in., 2014):

- decydenci polityczni – etapy wyboru opcji rozwiązania problemu i oceny efektów;
- dyrektorzy, urzędnicy wyższego szczebla – etap projektowania interwencji i oceny efektów;
- menedżerowie, osoby na niższych stanowiskach wdrażające i zarządzające interwencją – etap realizacji interwencji.

**Kanały i formy komunikacji** to użyteczna koncepcja pozwalająca dobrać sposób rozpowszechniania do poszczególnych odbiorców. Zgodnie z nią dla każdego odbiorcy potrzebujemy dobrać i preferowaną formę i kanał, dopiero razem zapewniają one skuteczne przekazanie wiedzy z ewaluacji (Olejniczak i in., 2016).

**Tabela 10.** Wybrane formy i kanały rozpowszechniania wyników badania z sugerowanymi odbiorcami

| Formy                                                              | Odbiorca                                   |
|--------------------------------------------------------------------|--------------------------------------------|
| Streszczenie wykonawcze, podsumowanie dowodów, tabela rekomendacji | Dyrektorzy, menedżerowie                   |
| Broszury, ikonografiki, wideo prezentacja                          | Decydenci polityczni, media, społeczeństwo |
| Raport końcowy*                                                    | Menedżerowie, naukowcy                     |

| Kanały                                                               | Odbiorca                                              |
|----------------------------------------------------------------------|-------------------------------------------------------|
| Małe spotkanie dyskusyjne                                            | Dyrektorzy, menedżerowie                              |
| Duże spotkanie lub konferencja                                       | Decydenci polityczni, dyrektorzy, media, beneficjenci |
| Kontakt poprzez doradców, osoby opiniotwórcze                        | Decydenci polityczni                                  |
| Indywidualny kontakt z odbiorcą                                      | Dyrektorzy, menedżerowie                              |
| Media społecznościowe, blog, podcast, strona internetowa, newsletter | Beneficjenci, społeczeństwo, media                    |
| Mapa argumentów, policy brief <sup>47</sup>                          | Decydenci polityczni, dyrektorzy                      |

\*w tabeli ujęto zarówno raport końcowy jak i formy, które stanowią często jego element, ale równocześnie są oddzielnymi kompletnymi produktami – streszczenie wykonawcze, tabela rekomendacji, mapa argumentów.

Źródło: Opracowanie własne na podstawie Olejniczak i in. (2016), Torres i in. (2005)

<sup>47</sup> Zwarte, krótkie opracowanie, wskazujące problem, możliwe opcje działania i rekomendujące preferowaną.



### Jednostki ewaluacyjne w mediach społecznościowych

Bank Światowy: [www.facebook.com/IndependentEvaluationGroup](https://www.facebook.com/IndependentEvaluationGroup)

Międzynarodowy Fundusz Walutowy:

[www.youtube.com/@independentevaluationoffic5780](https://www.youtube.com/@independentevaluationoffic5780)

Bank Światowy: [ieg.worldbankgroup.org/blogs](https://ieg.worldbankgroup.org/blogs)

Organizacja Narodów Zjednoczonych: [twitter.com/un\\_evaluation](https://twitter.com/un_evaluation)

Krajowa Jednostka Ewaluacji w Ministerstwie Funduszy i Polityki Regionalnej

[www.facebook.com/news.eval](https://www.facebook.com/news.eval)



### Dobra praktyka – grupy sterujące ewaluacją jako narzędzie rozpowszechniania

Od wielu lat instytucje prowadzące ewaluacje programów polityki spójności tworzą grupy lub zespoły ewaluacyjne, które kształtują proces ewaluacji w perspektywie wieloletniej, ale także mają rolę doradczą w przygotowaniu, realizacji i odbiorze poszczególnych ewaluacji. Członkowie grup wywodzą się z różnorodnych komórek organizacyjnych, więc w naturalny sposób pełnią też funkcję kanałów komunikacji rozprowadzających wiedzę o wynikach ewaluacji po organizacji.

Ciekawe spojrzenie na prezentowanie wniosków z ewaluacji proponuje też Grob (2015), wskazując że jest 6 poziomów szczegółowości, na których powinniśmy przygotować swój przekaz:

- **Podsumowanie do testu mamy**<sup>48</sup> – wykorzystywane najczęściej w ustnych wypowiedziach;
- **„Killer paragraph”** – jeden akapit z najważniejszym wnioskiem i koniecznym działaniem, wykorzystywany w komunikacji pisemnej, np. gdy przekazujemy komuś raport, albo jako abstrakt artykułu;
- **Outline** – jednostronicowe ustrukturyzowane opracowanie, w którym każdy ważny element wyrażony jest jednym równoważnikiem zdania. Strukturyzuje wystąpienia ustne i może być podsumowaniem dla słuchaczy;

<sup>48</sup> Zobacz krok 8.

- Dwustronicowe **streszczenie wykonawcze** – tworzone na podobnej zasadzie jak „killer paragraph” ale pozwalające nieco rozwinąć temat. Najważniejszy format dla większości odbiorców, którzy postanowią zapoznać się z badaniem;
- **Dziesięciostronicowy raport**<sup>49</sup> – rozwinięcie streszczenia, uzupełnienie o wyjaśnienia, kontekst, podstawowe grafiki. Podstawowy materiał dla osób opiniotwórczych;
- **Raport techniczny** – kompletny RK, głównie dla osób zaangażowanych w zarządzanie interwencją, analityków, naukowców.

Tak przygotowany materiał stanowi potem wkład do form i kanałów wymienionych wcześniej, np. treści broszur, wypowiedzi w podcaście, tekstu na blogu itp.



#### Proponowane lektury w zakresie komunikowania wyników ewaluacji:

Grob, G. F. (2015). *Writing for impact*. w: J. S. Wholey, H. P. Hatry, K. E. Newcomer (red.), *Handbook of Practical Program Evaluation*, wydanie trzecie, Jossey-Bass A Wiley Imprint, s. 739–764.

Torres, R. T., Preskill, H., & Piontek, M. E. (2005). *Evaluation strategies for communicating and reporting: Enhancing learning in organizations*. Sage.

Olejniczak K., Kupiec T., Raimondo E. (2014). *Brokerzy wiedzy. Nowe spojrzenie na rolę jednostek ewaluacyjnych*. W: A. Haber, K. Olejniczak (red.). *(R)ewaluacja 2. Wiedza w działaniu*. Warszawa: Polska Agencja Rozwoju Przedsiębiorczości, s. 67–114.

Rozwijając uwagi z kroku 8 dotyczące formatowania, stylu i języka RK warto zwrócić uwagę na ogólniejsze zagadnienie dostępności produktów ewaluacji dla osób ze szczególnymi potrzebami. Minimalne wymogi w zakresie dostępności informacyjno-komunikacyjnej zawiera ustawa o zapewnianiu dostępności osobom ze szczególnymi potrzebami<sup>50</sup>, a ich użyteczne rozwinięcie w zakresie materiałów pisanych, elektronicznych czy kampanii medialnych można znaleźć w wytycznych MFIPR (2022).

Zgodnie z nimi zakładanie, że osoby ze szczególnymi potrzebami nie są odbiorcami naszej ewaluacji, jest formą dyskryminacji. Poza najbardziej oczywistą kategorią osób z niepełnosprawnościami, warto rozważyć czy wynikami naszego badania nie będą zainteresowane np. osoby starsze, przedstawiciele mniejszości narodowej, czy osoby o ograniczonym dostępie do niektórych środków przekazu ze względu na niskie dochody.

<sup>49</sup> W polskiej praktyce ewaluacji polityki spójności streszczenie wykonawcze dołączane do raportów ma często rozmiar 10 stron, lub nieco więcej.

<sup>50</sup> Ustawa z dnia 19 lipca 2019 r. o zapewnianiu dostępności osobom ze szczególnymi potrzebami.



### Materiały zawierające wytyczne w zakresie przygotowania dostępnych tekstów:

Ośrodek Rozwoju Edukacji (2020) Standardy przygotowywania treści zgodnie z wytycznymi WCAG 2.1 na poziomie AA.

MFIPR (2022) Wytyczne dotyczące realizacji zasad równościowych w ramach funduszy unijnych na lata 2021–2027. Załącznik nr 2 Standardy dostępności dla polityki spójności 2021–2027.

Dobierając formy i kanały komunikacji warto zatem potrzeby takich grup uwzględnić. Zgodnie ze standardami dostępności materiały powinny być łatwe do przeczytania i zrozumienia, co może wiązać się z koniecznością tworzenia wersji w druku powiększonym, o wysokim kontraście, w języku łatwym do czytania, w systemie Braille'a, czy nagrania tłumaczenia na język migowy na nośniku elektronicznym. Przy wyborze sposobów komunikacji, dobrą praktyką jest uwzględnienie co najmniej dwóch kanałów sensorycznych (wzrok, słuch). Tworząc jakiegokolwiek dokumenty warto pamiętać o wypełnieniu metadanych oraz zapewnieniu dostępu do warstwy tekstowej. Produkty cyfrowe powinny być zgodne z wymaganiami WCAG 2.1.

## Krok 10: Wykorzystujemy wnioski i rekomendacje z ewaluacji

**Wykorzystanie wyników** jest kluczowym zagadnieniem dla dziedziny ewaluacji. Większość ewaluatorów wewnętrznych i zewnętrznych oczekuje i dąży do tego, by ich badania były wykorzystywane. Większość akceptuje fakt, iż właśnie to ostatecznie świadczy o jakości ich pracy. Wykorzystanie ewaluacji może przyjmować różne formy. Podstawowym jest podejmowanie decyzji modyfikujących interwencję (w tym sposób jej wdrażania), lub jej przyszłe edycje w oparciu o wnioski i zalecenia z badania. To tzw. wykorzystanie instrumentalne i jemu głównie poświęcony jest ten podrozdział.

Wykorzystanie zależy od wielu czynników, z których sporo jest poza kontrolą ewaluatora (tak zewnętrznego, jak i osób obsługujących ewaluację wewnątrz organizacji). Jest jednak też szereg aspektów, o które ewaluator może zadbać w krokach opisanych w tym poradniku, np. kompetencje ewaluatora (krok 6), jakość badania (kroki 2–7), odpowiedni czas na badanie (kroki 1–5). Ten podrozdział skupia się szczególnie na elementach trafności i wiarygodności wniosków z ewaluacji.

**Wnioski i rekomendacje** wieńczą RK, są skonkretyzowaną propozycją wykorzystania ustaleń ewaluacji. Żeby były użyteczne powinny spełniać co najmniej kilka kryteriów (Kupiec, 2014):

- adekwatność – rekomendowane działania odpowiadają na faktyczne potrzeby i problemy odbiorców zidentyfikowane przed albo w trakcie ewaluacji;
- wykonalność – proponowane działania są opisane w sposób zrozumiały, wystarczająco precyzyjny i są możliwe do wdrożenia;
- wiarygodność – rekomendacje jednoznacznie wynikają z ustaleń badania;
- istotność – rekomendacje dostarczają odbiorcom wiedzę dla nich nową oraz dotyczącą ważnych z ich perspektywy kwestii, związanych z zarządzaniem i wdrażaniem programu<sup>51</sup>;
- odpowiedni czas – rekomendacje trafiają do odbiorców na niedługo przed momentem podejmowania istotnych decyzji, które mogą zostać nimi wsparte.

Dobrą praktyką znaną w Polsce z pola ewaluacji programów polityki spójności jest przedstawianie rekomendacji w formie tabelarycznej. Rozwiązanie takie jest wskazane przede wszystkim dla dużych organizacji o rozbudowanej strukturze, lub sieci organizacji. Tabela, poza samą rekomendacją, zawierać może m.in. wniosek, z którego ona wynika, podmiot/komórkę organizacyjną odpowiedzialną za wdrożenie, termin i proponowany sposób wdrożenia rekomendacji.

**Proces ustalania rekomendacji** powinien mieć charakter interakcyjny. To dialog pomiędzy wykonawcą badania i użytkownikami, którzy mają rekomendacje wdrożyć. Przy ewaluacjach zewnętrznych osoba odpowiedzialna za badanie po stronie zamawiającego może pełnić rolę pośrednika, animatora tego procesu. Dialog zaczyna się w momencie przedstawienia wstępnego RK, który zawiera rekomendacje będące na tym etapie tylko propozycją wykonawcy. Ta podlega ocenie odbiorców, adresatów, osób i podmiotów, od których oczekuje się wdrożenia rekomendacji. Proces ten może, a w zasadzie powinien mieć kilka iteracji i trwać też po przyjęciu RK. Zaleca się, by nie były to tylko konsultacje korespondencyjne, ale faktyczny dialog w trakcie spotkań wykonawców i odbiorców. Bezpośrednia dyskusja o ustaleniach ewaluacji sprawia, że mają one szansę przestać być zewnętrznym wnioskiem, a stać się zasobem wiedzy decydenta. Istotą procesu konsultacji rekomendacji jest to, że stopniowo przekształcają się one z propozycji ewaluatora w plan i zobowiązanie adresatów do ich wdrożenia<sup>52</sup>. Nie oznacza to oczywiście,

<sup>51</sup> Czyli nie wskazują rozwiązań oczywistych, nieodkrywczych (np. powszechnie stosowanych i z dużym prawdopodobieństwem już znanych odbiorcom), niedotykających istoty poruszanego problemu, a koncentrujących się na działaniach drobnych, technicznych, które nie wpłyną zauważalnie na jakość realizacji i efekty interwencji.

<sup>52</sup> W nazewnictwie angielskim to zobowiązanie i plan nazywa się *management response* (zobacz np.: <https://erc.undp.org/>).

że pod wpływem sugestii odbiorców ewaluator ma wycofywać się z rekomendacji, które w oparciu o wyniki badania uważa za słuszne. Zupełnie normalną jest sytuacja, w której ewaluator formułuje rekomendacje w RK a odbiorca rezygnuje z ich wdrożenia przedstawiając uzasadnienie tej decyzji. Takie rozwiązanie jest pożądane z perspektywy przejrzystości procesu wykorzystania ewaluacji.



### Dobra praktyka

Sposobem na odpowiednie zaangażowanie wykonawcy ewaluacji w proces konsultacji rekomendacji może być dodanie do procesu ewaluacji etapu tzw. asysty. To dodatkowy element zamówienia, w ramach którego wykonawca jest zobowiązany do współpracy – uczestnictwa w spotkaniach konsultacyjnych, prezentacji wyników badania – również po odebraniu RK. Rozwiązanie to stosowane jest np. przez Polską Agencję Rozwoju Przedsiębiorczości.

Doprecyzowując powyższe stwierdzenia warto zauważyć, że wykorzystanie wniosków z ewaluacji zależy nie tylko od ich faktycznej trafności i wiarygodności, ale **oceny wiarygodności w oczach odbiorców** ewaluacji, a ta jest subiektywna i ograniczona. Donaldson i in. (2014) wskazują np. na duże znaczenie tzw. sygnałów peryferyjnych związanych m.in. z reputacją źródła informacji (poważana instytucja, osoba, wykonawca z którym mamy dobre doświadczenia), lub formą przekazu (poprawny język, profesjonalny wygląd). Dodatkowo, mamy mniejszą skłonność do zapamiętywania i akceptowania informacji, które kłócą się z naszym dotychczasowym stanem wiedzy (Weiss & Bucuvalas, 1980).



### Uwaga, Pułapka!

W procesie konsultacji kształtu rekomendacji obie strony negocjacji mogą kierować się motywacjami, które prowadzą do nieoptymalnych rozwiązań. Odbiorcy ewaluacji, adresaci rekomendacji mogą bardziej kwestionować adekwatność i wiarygodność rekomendacji, które uważają za trudne do wdrożenia, czasochłonne, kosztowne. Ewaluatorzy, szczególnie zewnątrzni, mogą natomiast mieć interes w promowaniu rekomendacji łatwych, drobnych, jeśli to przyspieszy akceptację RK. Konsekwencją są mało użyteczne ewaluacje, w których np. odpowiedzią na strategiczne problemy są nieadekwatne operacyjne rekomendacje.

To wyjaśnia dlaczego tak ważny jest styl i język RK, docieranie z przekazem do odbiorców przez poważanych ekspertów, zlecenie ewaluacji osobom i zespołom o dobrej reputacji, z którymi nasi odbiorcy mają dobre doświadczenia. To też pokazuje, dlaczego wnioski

z ewaluacji warto prezentować w kontekście innych wiarygodnych raportów, badań, artykułów, które potwierdzają nasze ustalenia.

Ewaluacja jest narzędziem uczenia się i rozliczania z realizowanych działań. Sama ewaluacja też jest specyficznym działaniem, interwencją, która angażuje zasoby i ma przynosić określone efekty. Dlatego pożądane jest by także ewaluacja, w tym **wykorzystanie ewaluacji, było monitorowane**. W najprostszej formie może to być po prostu wywiad, rozmowa z odbiorcami ewaluacji zmierzająca do ustalenia, które rekomendacje i które badania były dla nich użyteczne. Bardziej rozwiniętym rozwiązaniem jest baza rekomendacji ze wszystkich realizowanych w organizacji badań, w której dokumentowany jest status ich wdrożenia (np. wdrożona w całości, częściowo, niewdrożona) z ewentualnymi wyjaśnieniami odstąpienia od wdrożenia. Taki materiał monitoringowy może następnie stanowić wstęp do okresowej refleksji – ewaluacji procesu wykorzystania ewaluacji poświęconej temu, jakie rekomendacje są pomocne i wdrażane, a jakie nie powinny być formułowane, jakie ewaluacje warto realizować, a jakie okazują się nieużyteczne.



#### Dobra praktyka

Najbardziej rozbudowany w kraju system monitorowania wykorzystania ewaluacji, który może stanowić inspirację przy tworzeniu własnych rozwiązań prowadzi Krajowa Jednostka Ewaluacji w Ministerstwie Funduszy i Polityki Regionalnej (MFiPR, 2022)

## Rozdział 7. Metodyka badań

Paulina Skórska, Szymon Kajma

### Słowo wstępu

Na wstępie należy odróżnić metodologię badań (czyli naukę o metodach badawczych) od metodyki, czyli zbioru zasad dotyczących sposobów wykonywania jakiegś pracy.

Metodyka badania, to całościowa strategia czy też podejście do rozwiązania problemu badawczego. Metodyka to więc również zestaw metod. Na potrzeby odpowiedzi na interesujące nas pytania projektujemy więc określoną konfigurację metod badawczych. Są to to zestawy procedur dedykowane konkretnym typom pytań badawczych. Każda metoda charakteryzuje się specyficznym podejściem do zbierania danych, a następnie ich analizy. Określenia technika najczęściej używa się w kontekście analizy danych, przykładem technik są: analiza korelacji czy regresji.

Choć część autorów próbuje dość restrykcyjnie kategoryzować metody badawcze i techniki analizy danych, to z praktycznego punktu widzenia nie ma to znaczenia – ostatecznie przecież chodzi o pozyskanie informacji, która nas interesuje (Jasiński i M. Kowalski, 2007). Często zresztą pojęcia w obszarze badań są przez praktyków stosowane wymienianie i definiowane dość płynnie. Przykładowo najczęściej poprzez ankiety zbiera się dane ilościowe, ale nic nie stoi na przeszkodzie, by w ankietach uwzględniać pytania otwarte i analizować je jakościowo. W tym rozdziale przedstawimy i opiszemy najczęściej wykorzystywane metody wraz z informacją do czego można je stosować. Wskażemy również jakie typy danych i narzędzi są najczęściej wykorzystywane w obrębie danej metody. Należy jednak pamiętać, że przedstawione sugestie nie są sztywnymi regułami, które wykluczają zastosowanie danej metody w inny niż najczęstszy sposób.



#### Zapamiętaj

**Zazwyczaj rekomenduje się stosowanie mieszanej metodyki, tj. korzystanie w jednym badaniu z różnych metod i technik analizy danych (ilościowych i jakościowych). Zakładamy, że dzięki temu pozyskiwana przez nas informacja będzie możliwie pewna. Wtedy możemy mieć większe zaufanie do wyników, mając pewność, że zostały zweryfikowane z pomocą różnych metod badawczych i technik analizy danych.**



## Najpopularniejsze metody i techniki badań wykorzystywanych w ewaluacji

### Warsztat strukturyzujący/warsztat z interesariuszami

Na początku każdego badania warto zrealizować spotkanie/warsztat, który pozwoli uporządkować wszystkie oczekiwania wobec badania. Jeśli wykorzystujemy tę metodę w ramach ewaluacji zewnętrznej, to najczęściej **realizuje się** ją w formie spotkania z Zamawiającym, na którym doprecyzowuje się jego oczekiwania oraz uzgadnia inne priorytetowe kwestie. W warsztacie takim powinny wziąć udział osoby, które będą realizowały badanie, osoby które je zlecają oraz inni, ważni aktorzy zainteresowani wynikami. W ewaluacji wewnętrznej taki warsztat może przyjąć formę spotkania osób, które będą faktycznie realizować badanie, osób mających wiedzę o przedmiocie badania i wszystkich zainteresowanych wynikami. Warsztat **pozwala** wyjaśnić kwestie merytoryczne i techniczne (takie jak sposób organizacji badania, harmonogram). Należy zwrócić uwagę na wszelkie pułapki lub problematyczne kwestie, które mogą skomplikować realizację badania. Kiedy ewaluujemy określoną interwencję publiczną, warto na spotkaniu uzgodnić jej logikę. Na ogół wstępną koncepcję logiki, czy też teorii interwencji, w oparciu o dokumenty, opracowuje ewaluator i jest ona przedmiotem dyskusji z kluczowymi interesariuszami (o odtworzeniu teorii interwencji więcej w rozdziale 2). **Zaletą** metody jest możliwość pozyskania ważnych informacji już na początku badania i uwzględnienie ich we wszystkich dalszych etapach.

### Desk research

Metoda desk research, czyli tzw. badań gabinetowych jest jedną z podstawowych metod badawczych. Analizę **realizuje się** w oparciu o dane zastane (Babbie, 2007; Hox i Boeije, 2005). Polega ona na przeglądzie i analizie istniejących: raportów, opracowań naukowych, czy też innych istotnych (w kontekście przedmiotu badania) dokumentów i danych.

Dane zastane, czyli wszelkie materiały istniejące przed rozpoczęciem badania, są bardzo wartościowym źródłem informacji. Zapoznanie się z nimi **pozwoli** nam zaoszczędzić sporo czasu w badaniu, wiele interesujących



#### Zapamiętaj

**W praktyce każde badanie powinno zaczynać się od szczegółowej analizy danych zastanych.**

nas kwestii prawdopodobnie zostało już wcześniej zbadanych, czy zanalizowanych w pewnym zakresie. Nawet jeśli istniejące materiały nie dotyczą ściśle badanego przez nas zagadnienia, to pozwolą na rozpoznanie szerszego kontekstu (społecznego, gospodarczego, kulturowego), w którym można umieścić nasze badanie.

Analizie można poddać:

- dane jakościowe: w postaci istniejących dokumentów (np. programy operacyjne, dokumenty strategiczne), innych badań<sup>53</sup>, opracowań, literatury przedmiotu;
- dane ilościowe: bazy danych zagregowanych upubliczniane przez instytucje zajmujące się statystyką publiczną (np. GUS, Eurostat), bazy danych administracyjnych z rejestrów (dane jednostkowe, np. z KRS), bazy wewnętrzne tworzone przez różne instytucje (np. dane monitoringowe, SL2014 itp.).

Najbardziej efektywny desk research przyjmuje formę przeglądu systematycznego (Pet-  
ticrew i Roberts, 2006). W takiej analizie precyzyjne określamy kryteria włączenia danych materiałów do przeglądu, gromadzimy dane, a następnie je kategoryzujemy w ustrukturyzowanej formie, np. w formie tabeli.

Najważniejszą **zaletą** metody jest niski koszt i krótki czas realizacji – korzystamy z już istniejących danych (choć dostęp do części baz może być odpłatny, to gros materiałów jest upublicznianych). Dobra analiza danych zastanych dostarczy nam wiedzy, którą wykorzystamy w formułowaniu wniosków z badania, ale także pomoże nam stworzyć szczegółowe, wyczerpujące i dostosowane do odbiorców narzędzia badawcze, tj. kwestionariusze ankiet czy scenariusze wywiadu. Dzięki wcześniejszemu analizowaniu danych istniejących koncentrujemy zbieranie danych pierwotnych, na tych kwestiach, które rzeczywiście stanowią lukę informacyjną, przez co bardziej optymalnie wykorzystujemy dostępne w badaniu zasoby (np. czas, zaangażowani eksperci itp.).

## Indywidualny wywiad pogłębiony (IDI)

IDI – (*Individual In-Depth Interview*) to indywidualny wywiad pogłębiony. **Polega na** rozmowie, którą z respondentem prowadzi ankieter. Indywidualny wywiad pogłębiony (Babbie, 2007; Morris, 2015) może być prowadzony poprzez fizyczne spotkanie rozmówców, za pośrednictwem telefonu (TDI, *Telephone In-Depth Interview*) lub Internetu (poprzez narzędzia umożliwiające wirtualne spotkania *online*, telekonferencje).

---

<sup>53</sup> Należy mieć na względzie, że wyniki badań mają też często charakter ilościowy i syntezuje się je wówczas ilościowo.

Ankieter w trakcie rozmowy stara się zgłębić dany temat poprzez dopytywanie respondenta o daną interwencję czy sytuację. Badacza interesują szczególnie oceny i opinie respondenta oraz ich uzasadnienie, motywy działania, indywidualna perspektywa. Wywiad **pozwala** zgłębić ważne, w kontekście analizowanego problemu badawczego, kwestie. Zakres i sposób zadawania pytań powinien być każdorazowo dostosowany do wiedzy i kompetencji rozmówcy. Dlatego ważne jest doświadczenie ankietera, który powinien obserwować reakcje respondenta i odpowiednio na nie reagować. Często poruszonymi wątkami są:

- oczekiwania respondenta, potrzeby;
- rola respondenta lub reprezentowanej przez niego instytucji w całym procesie;
- doświadczenia dotyczące realizacji interwencji, w tym trudności w realizacji;
- efekty, skutki interwencji (czy się udało osiągnąć cel);
- czynniki wpływające na efekty interwencji (co sprzyjało, a co przeszkadzało w skutecznej realizacji);
- pomysły i rekomendacje respondenta dot. przyszłości (jak w przyszłości działać, jakie są dalsze potrzeby do zaspokojenia).

Najczęściej scenariusz rozmowy przygotowuje się, a samą rozmowę przeprowadza się, zgodnie z zasadą od ogółu do szczegółu. Oznacza to, że najpierw poruszane są bardziej ogólne kwestie, które następnie są uszczegóławiane i pogłębiane (stąd nazwa metody). Scenariusz, czyli spis pytań i problemów warto podzielić tematycznie na części i przypisać im orientacyjny czas trwania rozmowy. Inaczej pierwszym w kolejności tematom możemy poświęcić zbyt wiele czasu, co sprawi, że zabraknie go do dokończenia wywiadu. Czas trwania wywiadu zależy od złożoności problemu i zakresu informacji, jaki może posiadać respondent (np. ze względu na swoją rolę). Ważne, by wywiad nie był zbyt długi. Na ogół trwa od 30 do 60 minut. Gdy planujemy dłuższy wywiad, powinniśmy zawczasu uprzedzić o tym respondenta. Przed właściwą realizacją wywiadów, z docelową grupą respondentów, warto zrealizować pilotaż scenariusza wywiadu. Polega on na przeprowadzeniu „próbego” wywiadu, podczas którego sprawdzimy, czy pytania są jasne i zrozumiałe oraz czy mieścimy się w zaplanowanym czasie. Na początku rozmowy warto zebrać podstawowe informacje o respondencie, w zakresie zależnym od tematu badania (czasami będą interesować nas kwestie dot. badanej osoby, takie jak wiek, wykształcenie, miejsce zamieszkania, a czasem także instytucji, którą reprezentuje).



#### Dobra praktyka

**Wcześniejsze zebranie tematów ważnych do poruszenia w trakcie wywiadu pomaga upewnić się, że żadne ważne wątki nie zostaną pominięte. Dlatego większość wywiadów jest prowadzona w oparciu o narzędzie, tzw. scenariusz wywiadu. Jest to spis najważniejszych pytań i tematów do rozwinięcia w trakcie wywiadu.**

Dobór próby w wywiadach ma zazwyczaj charakter celowy, najczęściej zależy nam na przeprowadzeniu rozmowy z kluczowymi interesariuszami. Ważne by pamiętać, że celowy dobór próby nie umożliwia późniejszego uogólniania (generalizowania) wyników na inne osoby z badanej populacji, tak jak ma to miejsce w próbach losowych. Uzyskane w ten sposób wyniki możemy więc odnosić tylko i wyłącznie do przebadanej osoby/osób.

Najlepsze efekty badania można osiągnąć, gdy ankieter zostanie przeszkolony w zakresie scenariusza wywiadu. Ważne, by osoba prowadząca wywiad była świadoma celów badania i dobrze zaznajomiona z przedmiotem ewaluacji. Doświadczony prowadzący często ma większą swobodę w decydowaniu o kolejności poruszanych tematów, pomijaniu ich i rozwijaniu. Rozmowy często nagrywa się, a następnie przygotowuje z nich notatki lub transkrypcje (tekstowy zapis rozmowy, obecnie często przygotowany automatycznie za pośrednictwem narzędzi sztucznej inteligencji przetwarzającej mowę).

Szczególną formą tej metody są diady i triady – tj. wywiady pogłębione prowadzone odpowiednio z dwoma lub trzema osobami. W takich sytuacjach zazwyczaj badamy jedną rolę/grupę, reprezentowaną przez więcej osób. Dla przykładu badamy kierownictwo firmy, a więc np. osobę zarządzającą i zastępcę lub badamy rodzinę (albo osobę niepełnoletnią, w towarzystwie rodzica lub opiekuna prawnego). W praktyce diady lub triady często realizuje się w sytuacji, gdy pracownicy jednej instytucji, czy firmy uzupełniają wzajemnie swoje odpowiedzi (posiadając różną wiedzę z racji na różne pełnione funkcje).

Z wywiadów pozyskuje się głównie dane jakościowe. Najczęściej analiza polega na „wydobyciu” z treści wywiadów pewnych kategorii, prawidłowości, uogólnień czy typologii. W praktyce pozyskane dane, najczęściej tekst, dzieli się i zaznacza tzw. etykietami (kodami), które później można pogrupować tematycznie w szersze kategorie. Następnie, w ramach tych kategorii podsumowujemy, a nawet zliczamy wystąpienie poszczególnych zdarzeń (np. sprawdzamy, jak często respondenci podawali daną przyczynę, czy wymieniali dany problem).

Wywiady mają tę **zaletę**, że dają możliwość zgłębienia szczegółowych i złożonych kwestii. Pozwalają zrozumieć przyczyny różnych zjawisk i poznać złożone motywy, z perspektywy różnych interesariuszy. Jednocześnie metoda ta jest dość kosztowna, w danym



#### Dobra praktyka

Warto zaplanować kilka pierwszych wywiadów już na początku badania, co pozwoli nam lepiej dopasować inne narzędzia, które planujemy wykorzystać na większą skalę, np. kwestionariusz ankiety planowany do zrealizowania na większej próbie respondentów. To pozwoli nam określić, czy zadawane typy pytań są zrozumiałe dla interesariuszy i dobrze zaadresowane.

czasie jesteśmy w stanie przebadać ograniczoną liczbę osób. Stąd, zazwyczaj w badaniu prowadzi się kilka, kilkanaście, rzadziej kilkadziesiąt wywiadów.

## Zogniskowany wywiad grupowy (FGI)

Zogniskowany wywiad grupowy (FGI, *Focus Group Interview*) jest pewnym rozszerzeniem koncepcji IDI. FGI **realizowany jest** jako rozmowa z większą liczbą respondentów, czasem przyjmuje formę warsztatową (Odimegwu, 2000). Osobą prowadzącą wywiad jest moderator, ponieważ osoba ta nie tylko zadaje pytania, ale musi moderować relacje i dynamikę między uczestnikami biorącymi udział w dyskusji grupowej. Oznacza to, że jej zadaniem, poza pozyskaniem informacji jest dbałość o to, by nikt z respondentów nadmiernie nie zdominował rozmowy i/lub narzucił własny punkt widzenia oraz by każdy uczestnik wywiadu czuł się swobodnie i komfortowo w grupie.



### Dobra praktyka

**Liczba osób badanych w trakcie jednego spotkania nie powinna przekraczać 4-8 uczestników, tak by rozmowa nadal była komfortowa i pogłębiona oraz by każdy respondent mógł zabrać głos.**

Poza tym sposób badania jest analogiczny, jak w przypadku IDI, a podstawowym narzędziem badawczym jest również scenariusz wywiadu. Ze względu na większą liczbę badanych, zogniskowany wywiad grupowy trwa zazwyczaj nieco dłużej niż IDI – nawet około dwóch do trzech godzin.

FGI **pozwała** zebrać przydatne informacje w sytuacjach, gdy chcemy poznać opinie, co do których wiemy, że są niejednoznaczne i kompleksowe. W przypadku problemów badawczych, w których zależy nam na skonfrontowaniu opinii i różnych poglądów osób badanych. **Zaletą** metody jest jej relatywnie krótszy czas realizacji niż w przypadku wywiadów indywidualnych (szybciej możemy pozyskać informacje od większej liczby osób, choć wyzwaniem jest kwestia organizacji samego wywiadu). Mamy też możliwość skonfrontowania informacji. Ograniczeniem zaś proces grupowy, który może utrudniać dyskusję, stąd duże znaczenie ma doświadczenie moderatora.

## Badania sondażowe

Badania sondażowe **polegają na** pozyskiwaniu informacji od określonej (zazwyczaj dużej) liczby osób (tzw. respondentów), poprzez wypełnienie kwestionariusza ankiety (Babbie, 2007; Bradburn i in., 2004). O ile opisane powyżej wywiady pogłębione w większym

stopniu pozwalają zrozumieć badane zjawisko, o tyle badania sondażowe **pozwalają** oszacować skalę pewnych zjawisk. W sytuacji, kiedy ze względu na cel badania zależy nam na pozyskaniu informacji od dużej grupy respondentów, sondaż będzie skutecznym i efektywnym kosztowo wyborem. Zastosowanie badań sondażowych umożliwia pozyskanie danych w wystandaryzowanej formie, co oznacza, że wszyscy badani mają zadawane pytania w tej samej postaci. Dzięki temu w analizie danych można uchwycić wymiar porównawczy. To znaczy, że przy odpowiednio dużej liczbie badanych, możemy zestawić oceny i opinie różnych grup, np. porównać odpowiedzi na to samo pytanie w grupie kobiet i mężczyzn, młodszych i starszych respondentów i ze względu na inne, interesujące nas cechy.

Kwestionariusz ankiety może być zrealizowany na różne sposoby:

- drogą internetową: CAWI (*Computer Assisted Web Interview*).

Link do ankiety udostępniany jest respondentom drogą elektroniczną (Sills i Song, 2002; Dillman, 2000). Po kliknięciu w link respondent jest przekierowany do pytań ankiety. Link może zostać szeroko udostępniony, np. w mediach społecznościowych, zostać wysłany na wskazany adres e-mail respondenta lub przyjsć smsem (jeśli dysponujemy numerem telefonu respondenta). Ankiety realizowane internetowo mają wiele przewag nad pozostałymi sposobami dotarcia do respondenta, dlatego są bardzo popularną, obecnie podstawową techniką realizacji badań sondażowych. Udział badań CAWI w rynku usług agencji badawczych wynosił w 2021 r. 28%, wskazując, że jest to najpopularniejsza technika badawcza<sup>54</sup>. Popularność metody wynika też z faktu, że ponad 80% Polaków korzysta z Internetu. Do **zalet** CAWI należą:

- niski koszt realizacji badania – nie wymaga zatrudnienia ankierów, respondent sam wypełnia ankietę. Koszt realizacji sprowadza się do przygotowania elektronicznej wersji ankiety oraz ewentualnego serwisu wsparcia, tzw. helpdesku (obsługa e-mail lub infolinii, na które respondent może zgłosić problemy z wypełnieniem ankiety);
- wygoda respondenta – respondent może wypełnić ankietę w najbardziej dla siebie dogodnym czasie i miejscu. Większość ankiet umożliwia zapamiętywanie dotychczasowych odpowiedzi, tzn. respondent może wypełnić część ankiety, a potem wrócić do linka i dokończyć wypełnianie, a poprzednie odpowiedzi będą już zapamiętane;
- wygoda badacza – badacze mogą na bieżąco monitorować realizację badania, tzw. poziom zwrotu (*response rate*) oraz na bieżąco wprowadzać modyfikacje (jeśli wykryty zostanie błąd w ankiecie, np. literówka). Do takiej ankiety można

---

<sup>54</sup> <https://insightmap.pl/ranking/>

załączyć materiały kontekstowe, np. zdjęcie czy nagranie. Bieżący monitoring pozwala stymulować zwrot – możemy dodatkowo zachęcać osoby, które jeszcze nie ukończyły badania;

- niskie ryzyko pomyłek – odpowiedzi respondenta zapisywane są automatycznie w bazie danych, eliminowany jest czynnik ludzki i pomyłki, gdy odpowiedzi respondenta zapisuje ankieter (co ma miejsce w technikach realizacji badania z ankieterem);
- wysoka anonimowość – ze względu na brak wpływu ankietera, respondent w największym stopniu może mieć poczucie anonimowości, co pozwala zbierać odpowiedzi także na pytania wrażliwe, np. takie o stan zdrowia albo stan finansów.

Ograniczenia badania realizowanego drogą internetową wynikają przede wszystkim z braku obecności ankietera. Ryzykiem w tym zakresie jest możliwość wypełnienia ankiety przez osoby, do których nie była ostatecznie adresowana. Istnieje szereg narzędzi minimalizujących to ograniczenie, takich jak: ograniczenie liczby wypełnionych ankiet z jednego urządzenia/adresu IP do jednej, zabezpieczenie ankiety hasłem itp.

- telefonicznie: CATI (*Computer Assisted Telephone Interview*). To badanie **realizowane** z respondentem telefonicznie, a wspomagane komputerowo (Dabalen i in., 2016). **Polega na tym**, że podczas kontaktu telefonicznego ankieter przez telefon zadaje pytania respondentowi i wpisuje lub zaznacza jego odpowiedzi w komputerze. Badania realizowane drogą telefoniczną są nieco mniej popularne na rynku badawczym, w 2021 r. stanowiły ok. 12% wszystkich metod badawczych stosowanych przez agencje badawcze. Badania CATI są droższe niż CAWI, ze względu na konieczność zaangażowania ankietera, sam wywiad telefoniczny zazwyczaj też trwa dłużej niż samodzielne wypełnienie ankiety przez respondenta drogą internetową.
- **Zalety** metody są analogiczne, co w wypadku CAWI, choć koszt badania jest wyższy. Badania sondażowe realizowane techniką CATI są preferowane w sytuacji, gdy istnieje prawdopodobieństwo, że osoby badane mają ograniczony dostęp i umiejętności korzystania z Internetu. Sprawdzają się więc w badaniach, w których uwzględniamy przykładowo osoby starsze, czy migrantów posługujących się tylko językiem polskim mówionym.
- osobiście, twarzą-w-twarz (Fowler, 2002). To badanie, które **polega na tym**, że ankieter fizycznie spotyka się z respondentem. Może być realizowane technikami CAPI (*Computer Assisted Personal Interview*) oraz PAPI (*Paper And Pen Personal Interview*). W tym pierwszym przypadku ankieter zadaje pytania respondentowi, a odpowiedzi zapisuje w komputerze. W drugim przypadku, ankietę ma formę papierową. Ankieter lub respondent samodzielnie zaznacza odpowiedzi na pytania na papierowym kwestionariuszu. Jest to relatywnie najdroższa technika zbierania danych, a wersje papierowe są obecnie stosowane bardzo rzadko. Badania twarzą-w-twarz mają tę **zaletę**, że sprawdzają się w sytuacji, gdy

respondent w stopniu ograniczonym korzysta z telefonu i/lub internetu. W przypadku uzupełniania ankiety przez respondenta, dzięki obecności ankietera, uzyskane odpowiedzi cechuje zazwyczaj bardzo wysoka jakość. Jeśli respondent ma wątpliwości, ankieter może na bieżąco udzielać wyjaśnień co do treści i sposobu wypełnienia ankiety. Technikę warto zastosować w sytuacji, kiedy zadawane pytania są stosunkowo skomplikowane i mogą wymagać wyjaśnienia. Badania twarzą-w-twarz są drogie, gdyż wymagają osobistego dotarcia ankietera do respondenta. Stanowią około 10% usług badawczych.

Podstawą sukcesu badań sondażowych jest każdorazowo dobry kwestionariusz, który spełnia następujące warunki:

- jest możliwie krótki i zwięzły – im dłuższa ankieta, tym mniejsze prawdopodobieństwo, że respondent weźmie udział w badaniu i/lub dokończy ankietę (odpowie na wszystkie pytania);
- pytania są zrozumiałe i precyzyjne – kwestionariusz jest trafny, jeśli respondent nie ma wątpliwości, o co jest pytany;
- ankieta jest spójna logicznie i dostosowana do respondenta – każdy respondent odpowiada tylko na te pytania, które go dotyczą. M.in. w tym celu stosowane są pytania filtrujące. Pytania te pozwalają ominąć część pytań, które nie pasują do respondenta. Np. badając ocenę działania pomocy społecznej w gminie, zadamy pytanie o to, czy respondent korzystał z usług pomocy społecznej. Jeśli odpowiedź będzie przecząca część pytań o ocenę działalności instytucji zostanie pominięta, co dodatkowo skróci czas wypełnienia kwestionariusza.

Zasadniczo w ankietach możemy umieszczać zarówno pytania zamknięte, jak i otwarte. Pytania zamknięte to takie, w których respondent wybiera jedną lub więcej odpowiedzi z ograniczonej puli możliwości (tzw. kafeterii). Pytania otwarte umożliwiają wpisanie odpowiedzi samodzielnie przez respondenta. Służą pogłębieniu tematu i lepszemu zrozumieniu ocen i motywów respondenta. Często zaczynają się od sformułowań: „dlaczego”, „jak/w jaki sposób”, „uzasadnij”, „wymień”, „opisz”. Co do zasady w ankietach przeważają pytania zamknięte, co skraca czas wypełnienia badania, zaś do badań pogłębionych służą przede wszystkim inne metody. Dlatego badania sondażowe w przeważającym stopniu służą do gromadzenia przede wszystkim danych ilościowych, w standaryzowanej formie, które można poddawać statystycznej analizie danych.

W przypadku niewielkiej populacji można robić badania na wszystkich osobach.



#### Dobra praktyka

**W badaniach sondażowych zaleca się wykonywanie pilotażu narzędzia, tj. przeprowadzenie ankiety dla kilku respondentów i sprawdzenie, jak narzędzie jest odbierane i oceniane przez badanych.**



Zazwyczaj badaniami sondażowymi obejmuje się tylko część interesującej nas populacji, tj. próbę. Przykładowo, nie badamy wszystkich mieszkańców województwa, by poznać ich opinie o różnych aspektach życia w regionie, a do badań wybieramy tylko część mieszkańców. Jeśli próba jest dobrana poprawnie, losowa oraz reprezentatywna, wyniki uzyskane na próbie mieszkańców, mogą być, z dużym prawdopodobieństwem, generalizowane na całą populację. Oznacza to, że na podstawie wyników ankiety zrealizowanej na próbie mieszkańców województwa, możemy z dużą dozą pewności, wypowiadać się o wszystkich mieszkańcach regionu. Próba losowa to taka, w której każda osoba ma równe prawdopodobieństwo znaleźć się w badaniu. Próba reprezentatywna, to taka, w której proporcje ważnych cech są takie same, jak w całej populacji. Próba reprezentatywna dla Polski ze względu na płeć powinna mieć 51,7% kobiet i 48,3% mężczyzn, bo taki jest rozkład płci w populacji. Należy pamiętać, że reprezentatywność próby jest pewnym uproszczeniem, a próba nigdy nie będzie odpowiadać wszystkim cechom populacji jednocześnie, a jedynie wybranym. Dlatego w każdym badaniu należy uważnie zdefiniować, jakie cechy kontekstowe mają dla nas znaczenie.

## Studium przypadku

Studium przypadku (*case study*) to metoda **polegająca na** dogłębnym opisie przypadku, czyli opisie osoby, grupy, sytuacji czy działania. Przypadkiem może być jakaś instytucja, przedsiębiorstwo czy interwencja publiczna. Opis przypadku przybiera formę studium, ponieważ stanowi podstawę dogłębnego poznania i zrozumienia przypadku/sytuacji (Stake, 1997, Yin, 2009).

Są różne strategie doboru przypadków. Często staramy się wskazać najbardziej typowe przypadki. Inną strategią jest pokazanie najlepszych praktyk lub najgorszych praktyk. Przykładowo, badając dobre szkoły w naszej gminie będziemy chcieli zgłębić przypadki najlepszych uczniów i nauczycieli, a przez to przeanalizować cechy osób i czynniki zewnętrzne wpływające na najlepsze wyniki edukacyjne. Dla kontrastu możemy analizować przypadki przedsiębiorstw, które w naszym regionie miały najgorsze wyniki finansowe, a przez to zidentyfikować bariery rozwoju przedsiębiorczości.



### Zapamiętaj

**Ważne, by każdorazowo tak dobierać przypadki do opisu, by stanowiły najlepszą ilustrację, czy reprezentację badanego zjawiska.**

Studium przypadku **pozwała** zasadniczo pozyskać dane jakościowe. **Ograniczeniem** tej metody jest to, że jest czasochłonna i kosztochłonna. Wymaga poświęcenia większej ilości czasu, by dogłębnie poznać analizowany przypadek, zebrać szczegółowe

materiały. Dlatego zazwyczaj liczba studiów przypadku w pojedynczym badaniu jest ograniczona. Wyniki praktycznie nie dają się uogólniać na inne sytuacje/przypadki, ale **zaletą** jest to, że stanowią doskonałą ilustrację badanego zjawiska i pozwala poznać mechanizm zmiany. Często w studium wykorzystuje się wyniki innych zrealizowanych w badaniu metod. Dla przykładu badając przypadek osoby

z niepełnosprawnością, którą wsparto w znalezieniu pracy, w studium wykorzystamy wywiady indywidualne z tą osobą, ale także z rodziną, doradcami zawodowymi, rehabilitantami itp. Badając przypadek szkoły odnoszącej sukcesy edukacyjne, możemy wykorzystać wyniki sondażu wśród uczniów i wywiadów z nauczycielami oraz kadrą zarządzającą, a nawet przedstawicielami organu prowadzącego szkołę. W praktyce więc do przygotowania studium wykorzystuje się dane zastane i pierwotne pozyskane w toku całego badania, a następnie integruje w jeden opis przypadku.



#### Dobra praktyka

Jeśli to możliwe, studia przypadku często opatruje się zdjęciami, np. by pokazać jak zmieniło się miejsce zamieszkania, jakie inwestycje zbudowano itd.



#### Dobra praktyka

Ważne, by dobierając przypadki do analizy porównawczej kierować się zasadą podobieństwa do przedmiotu naszej ewaluacji i jego kontekstu. Przykładowo, jeśli będą nas interesowały skuteczne rozwiązania z zakresu ochrony środowiska w innych gminach, to powinniśmy, kierując się zasadą realizmu wdrożenia, wybierać takie do porównania, które są podobne do naszej gminy pod kątem ważnych cech, np. dysponują porównywalnymi zasobami finansowymi, uwarunkowaniami przestrzennymi itp.

## Benchmarking

Benchmarking stanowi jakościową metodę porównawczą. Zasadniczo **polega na** przeglądzie rozwiązań i efektów innych działań, porównywalnych do tego, które aktualnie ewaluujemy (Hey, 2017; Keehley i Abercrombie, 2008). Jeśli oceniamy działanie naszej firmy lub instytucji, inspirujące będzie przyjrzenie się rozwiązaniom w innych, porównywalnych podmiotach. Często do benchmarkingu wybierane są takie jednostki porównawcze, które stanowią jednocześnie dobrą praktykę, co **pozwała** sprawdzić w badaniu, dlaczego działają skutecznie i efektywnie. W ewaluacji często stosowany jest benchmarking międzynarodowy, gdzie porównujemy rozwiązania, czy modele stosowane w Polsce do tych za granicą. Podobnie możemy porównywać działanie naszego miasta czy gminy, do innych,

stosujących innowacyjne rozwiązania. W tym wymiarze metoda pomaga w implementacji rozwiązań, które sprawdziły się wcześniej i/lub gdzieś indziej.

**Zaletą** benchmarkingu jest relatywnie niski koszt, wynikający z tego, że głównie posługujemy się danymi zastanymi. Jest krótko mówiąc metodą uczenia się na błędach i sukcesach innych. Porównywane podmioty powinny być dogłębnie przeanalizowane, a każdy będzie stanowił swego rodzaju pogłębione studium przypadku. Stąd metodzie poddawana jest zazwyczaj w ramach jednej ewaluacji niewielka pula porównywanych podmiotów/działań.

## Panel/warsztat ekspercki

Jednym z ważnych elementów badania ewaluacyjnego jest zasięgnięcie opinii ekspertów. Metoda **polega na** nawiązaniu współpracy z ekspertami i najczęściej przybiera formę ustrukturyzowanej dyskusji, w ramach której konfrontowane są różne stanowiska, kumulowana wiedza i wypracowywane rozwiązania oraz rekomendacje (Waltz i in., 2015). Panel ekspertów **pozwała** osiągnąć różne cele, m.in. pomóc w weryfikacji i interpretacji ustaleń badania. O wsparciu eksperckim często mówi się także w kontekście podejścia tzw. *foresight*. Oznacza to, że eksperci wspierają ewaluatorów i decydentów w podejmowaniu decyzji, uwzględniając przyszłe ewentualności, scenariusze i efekty. W tym sensie eksperci bazując na swojej wiedzy i doświadczeniu, mogą prognozować, jak wypracowane rozwiązania i działania mogą sprawdzić się w przyszłości. W tym zakresie warto zadbać, by dobrane grono ekspertów było wyspecjalizowane w danej dziedzinie, ale również możliwie interdyscyplinarne. **Zaletą** metody jest możliwość skorzystania z wiedzy ludzi o specjalistycznym i zróżnicowanym doświadczeniu, relatywnie niewielkim kosztem – zazwyczaj w panelu realizowanym w jednym miejscu i czasie uczestniczy kilku ekspertów.

## Warsztat podsumowujący wyniki

Tak, jak często badanie rozpoczyna się warsztatem z interesariuszami, tak też powinno się kończyć. Dobrą praktyką jest spotkanie w gronie osób zaangażowanych w badanie celem przedstawienia najważniejszych wyników, wniosków i rekomendacji. Na spotkaniu można je przedyskutować i sprawdzić na ile są realistyczne i wykonalne, a następnie dostosować tak, by dało się je bezpośrednio wdrożyć.

## Podjęcia analityczne, które wzbogacą badanie ewaluacyjne

W poniższym podrozdziale chcemy przybliżyć czytelnikowi zestaw pewnych całościowych podejść do badań, które wzbogacą wnioskowanie w każdej ewaluacji. Nie są to pojedyncze metody badawcze, ale całe schematy analityczne, w skład których wchodzi zestaw różnych metod badawczych i technik analizy danych, skupionych na realizacji specyficznych funkcji.

### Podjęcie quasi-eksperymentalne (kontrfaktyczne)

W ewaluacji skuteczności jakiegokolwiek interwencji zasadnicze pytanie dotyczy tego, czy udało się osiągnąć założone efekty, a jeśli tak, to jaka jest ich skala. Jeśli badamy jakiś projekt, np. projekt oferujący darmowe badania mammograficzne dla mieszkanki gminy, po upływie jakiegoś czasu może nas interesować, jak to wpłynęło na zachorowalność na raka piersi. Odpowiedź na pytanie, czy pomoc przyniosła oczekiwane zmiany u osób, które z niej skorzystały, wymaga określenia przyczynowej zmiany obserwowanej po wdrożeniu interwencji. To z kolei wymaga znalezienia i zdefiniowania przybliżenia sytuacji, która wystąpiłaby, gdyby osoby nie skorzystały ze wsparcia (tzw. sytuacja kontrfaktyczna) i odniesienia jej do tego, co rzeczywiście miało miejsce (Rosenbaum i Rubin 1983; Rosenbaum 2002). Na naszym przykładzie, chodzi więc o wyobrażenie sobie, jaka byłaby sytuacja kobiet, które skorzystały z badań przesiewowych, gdyby nie wzięły w nich udziału.

Ze względu na oparcie całego podejścia o sytuację kontrfaktyczną, podejście to jest nazywane kontrfaktyczną ewaluacją wpływu interwencji (ang. *counterfactual impact evaluation – CIE*). Sytuacja kontrfaktyczna jest jednak czysto hipotetyczna i nie może być zaobserwowana (kobiety korzystające z badań już raz je zrobiły), a co za tym idzie nie może zostać zmierzona. Trzeba więc zdefiniować ją hipotetycznie, czyli znaleźć jakieś przybliżenie sytuacji kontrfaktycznej. Najczęściej wykorzystywanym sposobem jest znalezienie podobnej grupy, która nie wzięła udziału w projekcie, tzw. grupy kontrolnej. Następnie można porównać grupę biorącą udział w projekcie (eksperymentalna) i nie biorącą w nim udziału (kontrolna) pod kątem efektu, tj. zamierzonego celu projektu. Zgodnie z naszym przykładem, chodziłoby o porównanie sytuacji kobiet, które wzięły udział w badaniach przesiewowych oraz takich, które nie wzięły udziału w projekcie pod kątem późniejszej zachorowalności i wyleczalności raka piersi. Przy zachowaniu odpowiednich założeń (w tym przy prawidłowym dobraniu grupy kontrolnej, o czym więcej poniżej) różnica zachorowalności i wyleczalności w porównywanych grupach jest szacunkiem efektu netto interwencji (o szacowaniu efektu netto więcej w dalszej części rozdziału).

Przyjęcie, że porównanie sytuacji obu grup da odpowiedź na to, jaki jest efekt interwencji, jest obwarowane szeregiem założeń metodologicznych. Jednym z najważniejszych jest maksymalne podobieństwo porównywanych grup pod kątem wszystkich ważnych cech. Jedyne, co powinno je różnić to fakt uczestnictwa (grupa eksperymentalna), albo nieuczestnictwa (kontrolna) w interwencji. Dodatkowo obie grupy muszą być wystarczająco liczne, by różnica między nimi, np. w poziomie zachorowalności, spełniła warunek tzw. istotności statystycznej.

Wracając do przykładu, zależy nam, aby porównać zachorowalność w podobnych grupach kobiet: by odnosić do siebie kobiety w podobnym wieku, o podobnej sytuacji społecznej, z podobnego miejsca zamieszkania (wieś vs. duże ośrodki miejskie), z podobnym dostępem do służby zdrowia itp. Im więcej takich ważnych cech weźmiemy pod uwagę, tym większa szansa, że późniejszą lepszą wyleczalność w grupie kobiet korzystających z badań przesiewowych będzie można przypisać właśnie udziałowi w projekcie (efekt netto, w przeciwieństwie do efektu brutto, który ma charakter globalny i nie jest skorygowany o inne, niezależne od projektu czynniki).

Na poziomie analizy danych, by dopasować do siebie jak najbardziej podobne grupy, tzw. statystycznych bliźniaków, najczęściej stosuje się metody **statystycznego dopasowania** (*statistical matching*). Jej zastosowanie może w przybliżeniu wyglądać tak, że wszystkie interesujące nas cechy, dla których chcemy uzyskać podobieństwo (np. wiek, miejsce zamieszkania, dochody, dostęp do lekarza itp.), zliczymy do jednego indeksu (tzw. *propensity score*, PS). Dlatego ta technika jest nazywana *propensity score matching* (PSM) i jest jedną z najpopularniejszych technik dopasowania (Morgan i Winship 2014). Następnie będziemy porównywać efekty naszego projektu (np. zachorowalność, wyleczalność itp.), w grupach osób (grupa korzystająca i niekorzystająca z projektu) podobnych pod kątem wartości tego indeksu. Przy dobrym dopasowaniu grupy korzystających i niekorzystających z projektu (tj. takim, które uwzględni wszystkie ważne kontekstowo cechy tych osób), różnice w efekcie możemy definiować jako faktyczny efekt interwencji, który zwykło nazywać się efektem netto. Jest on „oczyszczony” ze zmian, które tylko pozornie były efektem interwencji.

Zazwyczaj stosuje się różne ulepszenia schematu kontrfaktycznego, po to, by oszacowanie efektu netto było jeszcze lepsze, tj. bardziej trafne i rzetelne. Dlatego często w analizach uwzględnia się też wielkość efektu **przed udziałem w projekcie**, a nie tylko po nim. Dzięki temu w analizach jesteśmy jeszcze bardziej precyzyjni. Wtedy w naszym badaniu będziemy kontrolować też fakt, czy obie porównywane grupy miały podobny efekt przed wzięciem udziału w projekcie. Jeśli np. grupa kobiet biorąca udział w badaniach przesiewowych miałaby niższy poziom zachorowalności już przed projektem, to nieuwzględnienie sytuacji przed spowodowałoby, że finalnie przypiszemy wyższą skuteczność projektowi, niż miałyby to miejsce, gdybyśmy tę kwestię kontrolowali.

W statystycznej analizie danych efekty netto często oblicza się z wykorzystaniem metod regresji, bardzo popularna jest metoda **podwójnej różnicy** lub różnicy w różnicach (*difference-in-differences*, DID). Jej nazwa wywodzi się z tego, że uwzględnia dwa typy porównań, o których wspomnieliśmy powyżej:

- między osobami, które wzięły udział w projekcie a grupą kontrolną (czyli taką, która nie wzięła udziału w projekcie);
- między sytuacją ww. grup przed i po projekcie.

Obliczenie efektów dzięki podejściu kontrfaktycznemu jest sposobem precyzyjnej oceny naszych projektów czy działań. Często okazuje się, że gdybyśmy przyjrzeni się efektom działań różnych przedsięwzięć to na pierwszy rzut oka wydaje się, że są one pozytywne i wysokie. Nierzadko, po zastosowaniu modelu kontrfaktycznego okazuje się, że efekt działania jest zerowy. Oznacza to, że nawet bez udziału w projekcie efekt wystąpiłby u osób korzystających z projektu, gdyż w porównywalnym stopniu pojawił się w grupie kontrolnej. W takiej sytuacji możemy bardziej realistycznie i krytycznie ocenić ewaluowane przez nas działania. Jeśli efekt netto danego projektu jest zerowy, oznacza to, że w przyszłości należałoby lepiej zaplanować nasze działania, gdyż bez projektu efekt byłby dokładnie taki sam.

## Ewaluacja oparta na teorii (TBE)

Ewaluacja oparta na teorii została szczegółowo opisana w rozdziale 2 poradnika. W tym miejscu warto tylko nadmienić, że częstym efektem realizacji takiego podejścia badawczego jest odtworzona logika/teoria danej interwencji. Efektem realizacji ewaluacji opartej na teorii powinny być więc zidentyfikowane i dokładnie zdefiniowane planowane efekty naszego przedsięwzięcia (bezpośrednie i pośrednie, krótko i długoterminowe, planowane i nieplanowane) oraz wszystkie działania, które miały z założenia prowadzić do tych skutków, a także wszystkie czynniki kontekstowe, które mogłyby mieć wpływ na uzyskane rezultaty. Finalnie pozwoli nam to odtworzyć mechanizm działania interwencji, tzn. dokładnie opisać, w jaki sposób (za pomocą jakich działań i narzędzi) interwencja może doprowadzić do osiągnięcia konkretnych efektów i jakie czynniki mogą mieć na to wpływ. Dzięki takiemu zobrazowaniu mechanizmu działania interwencji, możemy łatwiej weryfikować poszczególne jej założenia. Szczególnie, gdy efekty są niezadowalające pomaga wyjaśnić, który element ciągu przyczynowo skutkowego zawiódł, gdzie pojawił się problem, które założenia były błędne. Taki model interwencji czy działania, które ewaluujemy powinien towarzyszyć nam w toku całego badania, stanowić podstawę przygotowania narzędzi badawczych i punkt odniesienia na ostatnim etapie analizy i wnioskowania. Zidentyfikowana lista efektów może zaś zostać zoperacjonalizowana i oszacowana w sposób ilościowy, za pomocą m.in. opisanego powyżej podejścia kontrfaktycznego.

Jak wspomnieliśmy powyżej, podejście TBE przywołujemy tutaj i opisujemy w skrótowej formie. Niemniej wspomnienie o nim jest ważne, dlatego, że często jest ono traktowane jako uzupełniające się z podejściem kontrfaktycznym. Razem uznawane są za filary badań przyczynowych, tj. takich, które oceniają efekty i przyczyny danych interwencji. Bywają więc łączone i obejmowane jedną nazwą: ewaluacji wpływu opartej na teorii, ang. *theory-based impact evaluation* – *TBIE*. Podejście kontrfaktyczne służy bardziej do określenia, czy efekt interwencji wystąpił, czy jest pozytywny oraz jaka jest jego wielkość/skala. Z kolei podejście TBE, w większym stopniu odpowiada na pytanie dlaczego i w jaki sposób interwencja doprowadziła do tych skutków, co stanowi wkład w wyjaśnienie mechanizmu interwencji.

## Design thinking

Design thinking (DT) jest całościowym podejściem do projektowania produktów, usług czy rozwiązań, które opiera się na dogłębnym zrozumieniu problemów i potrzeb użytkownika/odbiorcy. W tym sensie jest podejściem podobnym do ewaluacji, jeśli przyjmiemy szeroką jej definicję, jako ostatecznie zorientowanej na projektowaniu coraz lepszych (tj. skutecznych i efektywnych) działań/rozwiązań. Założeniem projektowania w metodyce DT jest partycypacja, czyli takie projektowanie konkretnego rozwiązania czy usługi, w które od samego początku będą zaangażowani wszyscy ważni odbiorcy usługi/interesariusze. W tym kontekście podejście DT jest bardzo podobne do metodyki ewaluacji *ex ante* (Skowrońska, 2020). Ten typ ewaluacji jest skupiony na potrzebach osób, dla których projektowane jest dane działanie/program, a ostatecznym celem jest zaprojektowanie finalnej wersji projektu, tak by był maksymalnie dopasowany do oczekiwań odbiorców.

Jednym z podstawowych etapów DT jest tzw. empatyzacja, czyli możliwie szczegółowe rozpoznanie potrzeb, „wczucie się” w sytuację odbiorców rozwiązania. Dlatego elementem DT jest zespołowa praca projektowa, w interdyscyplinarnym składzie, pozwalającym uwzględnić perspektywę różnych interesariuszy.

Kolejny etap stanowi projektowanie danej usługi czy rozwiązania, najlepiej w interdyscyplinarnym zespole. Generowanie ostatecznego rozwiązania ma charakter iteracyjny, powstaje wiele pomysłów czy scenariuszy działania, często wzbogaconych innowacyjnymi elementami. Proces ten zazwyczaj jest kreatywny i przyjmuje formę burzy mózgów. Rekomenduje się tworzenie prototypów lub modeli działania w mniejszej skali, a następnie przetestowanie jak się sprawdzają, jak reagują na nie odbiorcy, a następnie wdrożenie potrzebnych zmian. W tym sensie DT przypomina ciągłą ewaluację *on-going*, jeśli możliwe jest wdrożenie zmian w trakcie realizacji danego działania czy sposobu funkcjonowania.

Metodyka DT w wielu aspektach przeplata się więc z ewaluacją. Ma w mniejszym stopniu charakter podsumowujący, a bardziej formatywny (zobacz rozdział 1), czyli sprawdza się tam, gdzie wypracowuje się rozwiązania i można wprowadzać zmiany.

## Analizy makroekonomiczne

Analizy makroekonomiczne pozwalają na ogólne spojrzenie na dany problem, w szczególności na całą gospodarkę. W odróżnieniu od pozostałych metod nie bada się oddziaływania poszczególnych projektów. Zamiast tego działa się na ekonomicznych agregatach – wskaźnikach makroekonomicznych. To podejście ma sporo ograniczeń, jednak może dostarczać bardzo istotnych informacji, które byłyby ciężkie do uzyskania w inny sposób. Do wad należy zaliczyć wysoki poziom abstrakcji i możliwość uchwycenia raczej trendów niż konkretnych danych – na większości etapów korzysta się z pewnych założeń i szacunków. Jednocześnie jest to zadanie dość skomplikowane i może wymagać dużych nakładów czasowych (jeżeli dane zastane nie są dostępne od ręki) i finansowych. Niemniej trudno wskazać inną metodę umożliwiającą otrzymanie podobnych informacji. Szczególnym narzędziem wykorzystywanym do analiz makroekonomicznych są modele makroekonomiczne.

Model makroekonomiczny jest uproszczonym zapisem funkcjonowania gospodarki, opisanym za pomocą równań matematycznych. Po wprowadzeniu danych ze źródeł zewnętrznych (np. GUS) za jego pomocą można generować prognozy gospodarcze i szacować wpływ określonych czynników na gałęzie gospodarki lub zmianę wskaźników, takich jak np. PKB. Jest to na swój sposób metoda pokrewna do analizy kontrfaktycznej. Modele bazują najczęściej na porównaniu dwóch scenariuszy – wpływu na przedmiot ewaluacji przy wykorzystaniu zaplanowanych środków finansowych oraz alternatywnego, bez ich wsparcia. Pozwala to oszacować bardzo pożądaną w ewaluacji efekt netto. Modele nie dają jednak możliwości zmierzenia efektów, a jedynie ich oszacowania.

Modele wykorzystywane są także w celach prognostycznych, w tym także do prognozowania przyszłego wpływu. Wymaga to jednak przyjęcia pewnych założeń i jest zwykle obarczone pewnym błędem, chociaż może stanowić ważne narzędzie wspierające podejmowanie decyzji.

Należy pamiętać, że interpretacja danych otrzymanych z modelu wymaga dużej wiedzy eksperckiej, ponieważ stanowią one wypadkową założeń autorów, ich poglądów na ekonomię, a także jakości danych zastanych czy specyfiki operacji matematycznych, na których opiera się model. Zatem same szacunki otrzymane z modelu nie są jeszcze wynikiem badania ewaluacyjnego. Dopiero po opatrzeniu ich stosownym komentarzem eksperckim można wyciągać dalsze wnioski.



## Analiza kosztów-korzyści

Analiza kosztów-korzyści (*Cost-benefit analysis, CBA*) jest zbiorem technik ilościowych służących ocenie efektywności kosztowej naszych działań czy publicznych interwencji (Adler i Posner, 2001). Porównuje więc **koszty**, jakie ponieśliśmy na realizację interwencji z **rezultatami**, przez co pozwala ocenić, czy nasze działanie było opłacalne, z punktu widzenia tego, co pozwoliło osiągnąć. Czasami w rachunku uwzględnia również przyszłe korzyści, które jeszcze nie ujawniły się.

Częstym sposobem analizy danych jest oszacowanie średniego kosztu jednostkowego efektu. Przykładowo, założmy że badamy skuteczność programu aktywizacji osób bezrobotnych, który wspierał osoby bezrobotne poprzez szkolenia, usługi doradców zawodowych, pomoc w znalezieniu stażu itp., a którego celem było zatrudnienie osoby bezrobotnej. Uwzględniając całościowy koszt takiego projektu i dzieląc go przez liczbę osób, które uzyskały zatrudnienie możemy pokazać, ile wyniósł średni koszt zatrudnienia jednej osoby w naszym projekcie. W tej analizie odnosimy koszty do uzyskanych korzyści, np. wskazujemy koszt aktywizacji osoby bezrobotnej i odnosimy go do uzyskanych korzyści, np. z tytułu odprowadzanych przez nią podatków czy oszczędności z racji tego, że nie musimy takiej osobie płacić zasiłku. Wyniki takich analiz, często wykorzystywane są w analizach porównawczych, gdzie możemy porównać efektywność kosztową naszego projektu do innych, podobnych działań, relatywnie do kosztów realizacji.

## Wybór i realizacja adekwatnej metodyki: kryteria i czynniki kontekstowe

Zaprojektowanie odpowiedniej metodyki jest kluczowym elementem każdego badania ewaluacyjnego. Poniżej wskazujemy, jakimi kryteriami powinniśmy się kierować dobierając poszczególne metody i techniki analizy danych oraz o czym pomyśleć realizując zaplanowaną metodykę.

### Typy potrzebnych danych i ich dostępność

Typ wymaganych danych często determinuje dobór technik ich analizy. To, jakie dane zbierzemy będzie wpływać na wybór sposobów analizy i charakter wyciąganych wniosków. **Dane ilościowe** to dane liczbowe, które można poddawać **statystycznej analizie**

**danych.** Technik ilościowej analizy danych jest bardzo dużo, poczynając od najprostszych i najpowszechniej stosowanych w ewaluacji takich jak opis statystyczny (np. rozkład częstości i procentowy, średnia, mediana, minimum, maksimum itp.) po bardziej złożone, np. analiza regresji, analiza skupień itp. **Dane jakościowe** mają postać nie liczbową, występują najczęściej w postaci tekstu, słów, fraz, wypowiedzi, ale także obrazów, materiałów audiowizualnych, czy obserwacji. Popularną formą analizy danych jakościowych jest ich kategoryzacja. Jeśli przykładowo badamy wpływ projektu aktywizacji zawodowej dla bezrobotnych na znalezienie pracy i w ramach badania przeprowadziliśmy serię wywiadów z osobami, którym udało się znaleźć zatrudnienie, możemy najpierw spisać czynniki sukcesu, a potem podzielić je tematycznie na kategorie. Czasami na tak zebranych materiałach jakościowych wykonuje się analizy ilościowe, tj. można zliczyć jak często pojawiły się dane czynniki sukcesu, sprzyjające znalezieniu pracy po projekcie i wyłonić te elementy, które w największym stopniu stymulują efekt.

Każdorazowo, realizując badanie, należy zadać sobie pytanie, czy dane potrzebne do realizacji określonej techniki są lub będą dostępne. Jeśli dane zostały już wcześniej przez kogoś pozyskane, mówimy o tzw. **danych zastanych** (inaczej wtórnych). Przykładem są zasoby Banku Danych Lokalnych w GUS, czy Eurostatu. Jeśli w danym projekcie występuje jakiś wewnętrzny monitoring, można skorzystać z takich baz lub poprosić o nie administratora. Jeśli jakichś ważnych danych brakuje, możemy je pozyskać w badaniu, za pomocą różnych metod, np. realizując badanie sondażowe czy wywiady indywidualne. Wtedy mówimy o **danych pierwotnych**, czyli takich, które na bieżąco wytwarzamy (wywołujemy) w badaniu terenowym. Warto pamiętać, że niektóre metody i techniki analizy danych są bardzo wymagające i restrykcyjne w zakresie potrzebnych danych. Np. analizy kontrfaktyczne wymagają zastosowania dużych prób, danych od osób korzystających z danego rozwiązania i grupy kontrolnej (nie uczestniczących w działaniu), danych sprzed i po wzięciu udziału w danym działaniu.

Oczywiście dostępność danych nie powinna być czynnikiem decydującym przy doborze metodyki, jednak należy kierować się zasadą realizmu i upewnić się, że potrzebne dane:

- już istnieją i/lub
- zostaną nam udostępnione przez odpowiednie instytucje i/lub
- możemy je samodzielnie uzupełnić.

## Dostępne zasoby

Projektując badanie musimy rozważyć, czy dysponujemy właściwymi zasobami: czasu, budżetu oraz wsparcia eksperckiego. W przypadku ograniczeń finansowych, warto rozważyć skorzystanie z mniej kosztownych rozwiązań, np. wykonać badania sondażowe

techniką CAWI, zamiast CATI czy CAPI, co jak wskazano powyżej jest częstą praktyką agencji badawczych. Badania sondażowe CAWI będą tańsze i szybsze, będą wymagać mniej zasobów (brak ankietera) niż badania sondażowe realizowane innymi sposobami. Wywiady "twarzą w twarz" można zamienić na wywiady telefoniczne lub wywiad w formie spotkania on-line itp. Tu również ograniczamy koszty badania i czas pracy badacza. Takie wywiady zyskały na popularności w czasie zwiększonych obostrzeń i zaleceń sanitarnych wynikających z pojawienia się pandemii COVID-19 i ugruntowały się w polskiej praktyce badawczej. Niektóre badania wymagają zabezpieczenia w harmonogramie większej ilości czasu, np. badania sondażowe. Po przygotowaniu i przeprowadzeniu pilotażu ankiety, a następnie jej uruchomieniu, np. w trybie CAWI powinniśmy dać wszystkim respondentom możliwość komfortowego udziału w badaniu, np. dwa-trzy tygodnie.

## Dostęp do osób badanych

Dostępność osób podlegających ewaluacji jest jednym z ważnych czynników wpływających na realizację metodyki w badaniu. Jeśli respondenci są rozproszeni geograficznie, np. rozrzućeni w różnych miejscowościach województwa, czy kraju, metody takie jak sondaż realizowany internetowo lub wywiad realizowany w formie wideokonferencji mogą być bardziej właściwym wyborem niż badania „twarzą w twarz”. Z kolei w przypadku ograniczonej dostępności lub umiejętności obsługi internetu i telefonu lepiej rozważyć osobiste spotkanie z respondentem, np. gdy badamy seniorów. Należy też pamiętać, by unikać realizowania badań w okresach wakacji lub świąt, kiedy respondenci są nieobecni lub zaabsorbowani innymi zajęciami (choć to w większym stopniu determinuje harmonogram niż metodykę badania).

## Etyka w realizacji metodyki



### Uwaga, Pułapka!

Przy zapewnieniu anonimowości wyników trzeba zachować czujność. Czasem mając małą grupę badanych, kontekstowe informacje umożliwiają identyfikację respondenta. Założmy, że pokazujemy wyniki dotyczące osiągnięć szkolnych jakiejś klasy w szkole. Jeśli w klasie jest jeden uczeń posiadający orzeczenie o niepełnosprawności, będzie można tę osobę jednoznacznie zidentyfikować.

Przy tworzeniu i realizacji metodyki badania powinniśmy kierować się zasadą poszanowania praw i godności osób badanych. Jednym z aspektów tej zasady jest zapewnienie poufności wyników i anonimowości uczestników badania. Tematy intymne, wstydliwe i drażliwe lepiej badać za pomocą metod zapewniających maksymalny komfort i anonimowość, np. w ramach badań sondażowych wybierać technikę CAWI zamiast CATI czy CAPI, a w badaniach pogłębionych wywiady telefoniczne lub realizowane on-line zamiast wywiadów „twarzą-w-twarz”. Zapewnienie anonimowości wyników oznacza, że osoba czytająca wyniki nie będzie w żaden sposób w stanie zidentyfikować, jak odpowiadała konkretna osoba. Zazwyczaj wyjściem jest prezentowanie wyników zagregowanych, tj. dla większych grup i pokazywanie średniej czy rozkładu procentowego.

Każdorazowo powinniśmy upewnić się, że także informacje kontekstowe podawane o respondentach nie naruszają poufności informacji. Podobnie, gdy przytaczamy wypowiedzi czy opinie osób badanych jakościowo, np. w wywiadzie. Uważajmy, by w przytaczanej odpowiedzi czy opinii nie znalazły się informacje czy wskazówki pozwalające jednoznacznie stwierdzić kto ją wypowiedział. Także opisując skąd pochodzi wypowiedź, unikajmy podpisów o roli czy stanowisku respondenta, jeśli jest jedna lub niewiele osób spełniających te kryteria. Przykładowo zamiast podpisu opinii „prezes firmy XXX” możemy użyć sformułowania „osoba z kadry zarządzającej dużym przedsiębiorstwem” (bez wskazywania konkretnej pozycji i konkretnej instytucji).

# Załącznik 1. Jakimi ogólnymi zasadami należy się kierować przy realizacji ewaluacji? Słów kilka o standardach ewaluacji

Monika Bartosiewicz-Niziołek

## Wprowadzenie

Prowadzenie ewaluacji wymaga zapewnienia odpowiedniej jakości na każdym etapie jej procesu. Bez spełnienia tego warunku rośnie ryzyko gromadzenia nierzetelnych danych oraz formułowania błędnych wniosków i nietrafnych rekomendacji sprawiających, że ewaluacja będzie nieprzydatna a zasoby na nią przeznaczone zostaną zmarnowane. Aby zapobiec takim przypadkom tworzone są różnego rodzaju standardy i zalecenia. W Polsce jedynym dokumentem tego typu, ukierunkowanym bezpośrednio na ewaluację są „Standardy Ewaluacji” (dalej Standardy) opracowane przez Polskie Towarzystwo Ewaluacyjne w 2008 r.<sup>55</sup> Podobne dokumenty są opracowywane również przez inne kraje członkowskie UE jak i organizacje międzynarodowe<sup>56</sup>. Niniejszy rozdział pokrótce opisuje wskazania ww. Standardów opracowanych przez PTE.

Standardy mają na celu wspieranie osób biorących udział w procesie ewaluacji (szczególnie zamawiających oraz jej wykonawców, ale także osób realizujących ewaluacje wewnętrzne) w zapewnieniu pożądanej jakości prowadzonych prac. Mogą stanowić punkt odniesienia przy przygotowywaniu, zamawianiu i realizacji ewaluacji oraz wykorzystywaniu jej wyników. Standardy służą również ujednoliceniu sposobów rozumienia różnych pojęć mających zastosowanie w procesie ewaluacji oraz przyjęciu uzgodnień w zakresie respektowania podstawowych zasad, norm i reguł, zarówno o charakterze metodologicznym, etycznym, jak i prawnym.

---

<sup>55</sup> Dokument ten powstał w ramach projektu „Wspólnie tworzymy standardy ewaluacji w Polsce” sfinansowanego w ramach Środków Przejściowych (Transition Facility) i jest efektem współpracy około 60-ciu osób reprezentujących różne środowiska.

<sup>56</sup> Przykładowe dokumenty tego typu można znaleźć na stronach Europejskiego Towarzystwa Ewaluacyjnego – <https://europeanevaluation.org/resources/evaluation-standards/>.

Choć Standardy zostały opisane w odniesieniu do każdego z etapów procesu ewaluacji, należy także pamiętać o nadrzędnych zasadach, bez których działania ewaluacyjne tracą sens. We wstępie do Standardów podkreślono, że ewaluacja powinna się opierać na wartościach, które są nierozdzielnie związane z ładem demokratycznym, takich jak: niezależność, rzetelność, uczciwość, bezstronność oraz obiektywizm. Bez zadbania o etykę i urzeczywistnienie tych wartości istnieje duże ryzyko, że ewaluacja stanie się jedynie narzędziem kontroli lub zostanie instrumentalnie wykorzystana w celu potwierdzenia z góry założonych tez (np. „wybielania” bądź „oczerniania” ewaluowanego przedsięwzięcia w zależności od chęci jego kontynuacji lub zaprzestania realizacji).

Kluczowe dla ewaluacji są jej **niezależność i bezstronność**. Autonomia oznacza brak nacisków na etapie planowania i realizacji badania ewaluacyjnego, a szczególnie formułowania wniosków i rekomendacji. Na etapie konceptualizacji należy się wystrzegać pokusy formułowania pytań badawczych oraz przyjmowania kryteriów, które pokażą przedmiot ewaluacji jedynie w dobrym świetle, pomagając ukryć jego słabsze strony wymagające udoskonalenia. Istotne jest tu także ujawnienie i jasne zdefiniowanie celów planowanej ewaluacji. Niedopuszczalne jest posiadanie ukrytych celów, szczególnie takich, które mają służyć manipulowaniu interesariuszami, a także pozyskiwanymi oraz analizowanymi informacjami, aby uzyskać pożądane wnioski i rekomendacje. Wykorzystywanie ewaluacji w taki sposób jest niedopuszczalnym nadużyciem, które należy piętnować. Ewaluatorzy/ki angażujący się w tego rodzaju działania postępują nieetycznie, łamiąc jedną z podstawowych reguł swojej profesji.

Podobne znaczenie dla ewaluacji ma zachowanie bezstronności oraz neutralności światopoglądowej, które gwarantują demokratyczny charakter tego procesu. Bezstronność oznacza, że ewaluator/ka nie opowiada się po żadnej ze stron reprezentowanych przez poszczególnych interesariuszy, nikogo nie faworyzuje ani nie lekceważy niczyjego zdania. Dotyczy to zwłaszcza tych osób, które mają słabszą pozycję i mogą być mniej słyszalne, czy wręcz pomijane w procesie ewaluacji. Neutralność ewaluatora/ki polega na nie uwzględnianiu w procesie badawczym swoich przekonań politycznych, religijnych i rasowych, a także wszelkich innych, które mogą zakłócać bezstronność. W procesie ewaluacji kluczowe jest zapewnienie równości wszystkim interesariuszom oraz szacunek dla cechujących ich różnic kulturowych, etnicznych, światopoglądowych, religijnych, politycznych i rasowych. Ewaluacja powinna gwarantować równość wszystkim osobom, niezależnie od ich płci, wieku, niepełnosprawności czy orientacji seksualnej.

## Jak Standardy Ewaluacji powinny wspierać proces ewaluacji?

### Planowanie ewaluacji

Wstępem do ewaluacji jest projektowanie jej koncepcji oraz planowanie poszczególnych działań (szerzej opisane w rozdziale 4 poradnika). Jednak jeszcze przed uruchomieniem procesu ewaluacyjnego bardzo ważna jest **ocena zasadności jego prowadzenia**. Obecnie realizacja działań ewaluacyjnych częściej wynika z zewnętrznego wymogu narzuconego przez donora/grantodawcę niż z wewnętrznej potrzeby. Ta pojawia się na ogół dopiero wtedy, gdy realizator danego przedsięwzięcia doceni korzyści płynące z ewaluacji. Ewaluacja motywowana obowiązkiem nie musi jednak być bezużyteczna, jeśli realizująca ją organizacja ma wolę doskonalenia siebie i swoich działań. Takie podejście do ewaluacji świadczy o odpowiedzialności oraz dbaniu o dobro publiczne, zamiast złe pojętego pielęgnowania własnego wizerunku czy chęci usatysfakcjonowania donora/grantodawcy laurką ewaluacyjną. Każda ewaluacja powinna posiadać wysokie walory poznawcze oraz praktyczne, które przełożą się na chęć i możliwość wykorzystania zebranych wyników.



#### Uwaga, Pułapka!

Ewaluacja, za którą nie stoi potrzeba uczenia się poprzez zdobywanie informacji na temat przebiegu i/lub efektów danego projektu, programu, czy strategii, może mieć jedynie fasadowy charakter i stać się fikcją. Takie pseudoewaluacje na ogół służą innym celom niż rozwijanie i udoskonalanie danego przedsięwzięcia, np. chęci pokazania się z jak najlepszej strony, przemilczenia napotkanych problemów oraz przedstawieniu dowodów na to, że wszystko działa idealnie. Takie podejście wywołuje jednak brak zaufania i odbiera danej instytucji nie tylko wiarygodność, ale także możliwość uczenia się na błędach, co prowadzi do ich powielania.

Jednym z kluczowych zadań na etapie planowania ewaluacji jest identyfikacja potrzeb informacyjnych. Praktyka pokazuje, że największy sens mają **ewaluacje społecznione**, w które zaangażowanych jest wielu interesariuszy. Mogą to być osoby reprezentujące uczestników/czki ewaluowanej inicjatywy, a także ich bliscy i opiekunowie (np. rodziny osób z niepełnosprawnościami, po kryzysie psychicznym czy długotrwale bezrobotnych), kadra prowadząca kierowane do nich działania, partnerzy projektowi, grantodawcy, eksperci specjalizujący się w danej dziedzinie, jak również społeczność lokalna (w tym

aktywiści społeczni z organizacji pozarządowych świadczących wsparcie dla danej grupy odbiorców).



#### Uwaga, Pułapka!

Angażowanie interesariuszy jedynie na etapie konsultowania wyników badania ewaluacyjnego czy formułowania wniosków i rekomendacji jest znacznie mniej owocne niż włączanie tych osób w proces ewaluacji od samego początku. Jeśli w ewaluację zaangażowane są tylko dwie strony, tj. zamawiający i wykonawca, rośnie ryzyko niewdrożenia rekomendacji ze względu na brak motywacji oraz wsparcia innych interesariuszy.

Interesariusze powinni być angażowani w działania ewaluacyjne od samego początku np. poprzez zapytanie ich o oczekiwania w stosunku do ewaluacji, czyli o to czego chcieliby się dzięki niej dowiedzieć oraz czemu ma ona służyć i jak chcą wykorzystać jej wyniki. W ten sposób interesariusze pomagają zamawiającemu **precyzyjnie określić cel, przedmiot, zakres i rodzaj ewaluacji**, jak również spójne z nimi **kryteria i pytania badawcze**. Dzięki temu rosną szanse na otrzymanie wsparcia ze strony tych osób, nie tylko na etapie gromadzenia informacji, konsultowania zebranych wyników, tworzenia wniosków i rekomendacji oraz ich rozpowszechniania, ale także wprowadzania w życie uzgodnionych zmian. Współpraca z osobami, które identyfikują się z przeprowadzoną ewaluacją i czują się jej właścicielami znacznie ułatwia wdrażanie wspólnie sformułowanych zaleceń, które może być złożonym procesem wymagającym wielu zasobów, jak również rodzącym opór przed wprowadzaniem nowych rozwiązań. Interesariusze, którzy rozumieją ich sens oraz mają silną motywację do usprawniania i rozwijania danego programu czy strategii, mogą mieć kluczowe znaczenie dla powodzenia wysiłków podejmowanych w tym celu.

Prawidłowe zaplanowanie ewaluacji wymaga również zapewnienia odpowiednich nakładów kadrowych, finansowych, czasowych, informacyjnych oraz organizacyjno-technicznych (więcej na ten temat w rozdziale 3 poradnika). Zakres ewaluacji często jest zdezeterminowany zasobami, którymi dysponuje podmiot ją realizujący. **Realistyczne określenie sposobu przeprowadzania ewaluacji**, adekwatne do jej przedmiotu, celu i zakresu, opiera się na racjonalnym oraz precyzyjnym oszacowaniu niezbędnych nakładów i znacząco zwiększa szanse na powodzenie całego procesu.





### Uwaga, Pułapka!

**Jeśli nie jesteśmy w stanie zapewnić stosownych zasobów na ewaluację, lepiej jej w danym momencie nie podejmować. Niedostateczna ilość czasu i środków finansowych, a także brak zaangażowania i kompetencji zespołu odbiją się negatywnie na jakości badania ewaluacyjnego, prowadząc do błędnych wniosków i bezużytecznych rekomendacji.**

Warto pamiętać o tym, że niezmiernie ważnym zasobem są kompetencje ewaluacyjne, które powinny zostać zapewnione nie tylko po stronie wykonawców, ale także zamawiających ewaluację. Jeśli w danej instytucji nie ma osób, które wiedzą czego należy wymagać od wykonawcy, to może się to negatywnie odbić na jakości ewaluacji. Zamawiający powinien umieć wyłapać niespójności i błędy w jej koncepcji (w tym w narzędziach badawczych), dopilnować prawidłowej realizacji badania ewaluacyjnego, rzetelności pozyskanych informacji, a także zweryfikować poprawność przeprowadzonych analiz oraz formułowanych wniosków i rekomendacji. W przeciwnym razie zlecający naraża się na wystąpienie błędów, które uczynią ewaluację mało użyteczną, a tym samym podważą zasadność wydatkowanych na nią zasobów. Zamawiający nie posiadający kompetencji badawczych nie będzie w stanie trafnie ocenić poprawności zastosowanych metod i technik oraz sposobów doboru badanych prób, poziomu responsywności (np. stopnia zwrotu wypełnionych ankiet), a także możliwości generalizowania uzyskanych wyników.

Podczas planowania ewaluacji należy także zwrócić uwagę na tzw. zasoby informacyjne poprzez upewnienie się, że możliwe będzie zebranie niezbędnych danych oraz zadbanie o dostęp do wymaganych źródeł informacji. W przypadku ewaluacji zewnętrznej Zamawiający powinien jasno wskazać które dane zostaną przez niego dostarczone ewaluator(k)om, a także udzielić im wszelkiego wsparcia w pozyskiwaniu potrzebnych informacji. Zamawiający powinien także zadbać o gotowość do współpracy przy zbieraniu danych przez różnych interesariuszy, którzy mogą w istotny sposób wspierać ten proces (np. poprzez udostępnianie unikalnych informacji czy nakłanianie odbiorców ewaluowanych przedsięwzięć do udziału w prowadzonych badaniach).



### Dobra praktyka

**Informowanie o wysokości środków jakie zamawiający zamierza przeznaczyć na sfinansowanie ewaluacji pozwala uniknąć niepotrzebnego przygotowania, a następnie oceny ofert, które znacznie wykraczają poza założony budżet, oszczędzając wysiłek i czas zarówno oferentów jak i komisji wyłaniającej wykonawców ewaluacji.**

Planowanie ewaluacji w przypadku ewaluacji zewnętrznej obejmuje także określenie **przejrzystych kryteriów wyboru jej wykonawcy**. Kryteria te, podobnie jak cele ewaluacji, powinny być transparentne, dające równy dostęp do realizacji danego zamówienia różnym osobom i podmiotom, zgodnie z zasadą uczciwej konkurencji. Absolutnie niedopuszczalne jest formułowanie tych wymagań pod kątem określonego wykonawcy. Proces jego wyboru musi się opierać na jawnych i czytelnych zasadach umożliwiających wyłonienie osoby/podmiotu oferującego najlepszą jakość. Wybór wykonawcy ze względu na niską cenę czy nadzwyczaj szybki termin realizacji ewaluacji pociąga za sobą ryzyko dostarczenia wyników i/lub raportu nie spełniających podstawowych standardów metodologicznych i ewaluacyjnych.

W procesie przygotowania ewaluacji bardzo ważne jest także **zapewnienie niezależności**, bez względu na to czy jest ona zlecana zewnętrznemu wykonawcy czy prowadzona własnymi siłami. W przypadku ewaluacji wewnętrznej może to być paradoksalnie nawet trudniejsze niż w przypadku badania zewnętrznego. Dobrym rozwiązaniem jest bezpośrednia podległość komórki ewaluacyjnej, która została powołana w danej instytucji jej zarządowi. Takie ułożenie zapewnia jednostce ewaluacyjnej silną pozycję w strukturze tego podmiotu zwiększając jej niezależność<sup>57</sup>. W przypadku konfliktu interesów, ewaluator powinien otwarcie o tym poinformować i o ile nie uda się rozwiązać problemu, zrezygnować z prowadzenia ewaluacji.



#### Uwaga, Pułapka!

Zamawiający nie mogą wymagać dostarczenia materiałów źródłowych np. transkrypcji wywiadów czy surowych wyników badań ankietowych, które umożliwiłyby identyfikację osób badanych. Zapewnienie im poufności oraz anonimowości w celu zapobiegania ryzyku poniesienia konsekwencji z powodu udzielenia pewnych informacji jest jednym z podstawowych obowiązków ewaluatora/ki. Nie neguje to możliwości przekazywania danych źródłowych. Czasami jest to wręcz konieczne dla weryfikacji jakości procesu. Niemniej musi to zostać przeprowadzone w odpowiedni sposób, gwarantujący zachowanie zasady poufności.

Jednym z koniecznych warunków ewaluacji jest także jej **zgodność z zasadami etycznymi**. Choć wymóg ten obejmuje cały proces ewaluacji, nie można o nim zapominać na etapie jej planowania oraz zamawiania. Warunki umowy nie mogą na przykład wymuszać na wykonawcach naruszania poufności (np. poprzez dostarczanie niezanonimizowanych wyników badań – wywiadów czy sondaży).

<sup>57</sup> Tego rodzaju rozwiązanie jest stosowane np. przez ONZ czy Bank Światowy.

## Realizacja ewaluacji

Odpowiedni potencjał merytoryczny i organizacyjny jest istotnym gwarantem udanej ewaluacji. Za **zapewnienie wysokiej jej jakości** odpowiadają obie strony tego procesu, zarówno wykonawca, jak i zamawiający. Ich sprawna współpraca umożliwia skuteczne zarządzanie działaniami ewaluacyjnymi oraz zachowanie zgodności z założonym harmonogramem, budżetem i zakresem prowadzonych prac. Ewaluator/ka musi zachować należytą staranność w gromadzeniu i analizowaniu danych, a także formułowaniu na ich podstawie uzasadnionych wniosków. Z kolei do zamawiającego należy wspieranie tych działań oraz weryfikowanie ich poprawności (więcej informacji na temat realizacji badań w rozdziałach 5 i 6). **Zapewnienie trafności i rzetelności badań ewaluacyjnych** warunkuje możliwość wykorzystania ich wyników, czyli zwiększa użyteczność ewaluacji.

Warto podkreślić, że zamawiający nie powinien traktować ewaluatorów/ek jedynie jako realizatorów stworzonej przez siebie, kompletnej koncepcji ewaluacji, ale powinien być otwarty na ich wkład merytoryczny opierający się na doświadczeniu w prowadzeniu badań ewaluacyjnych w danym obszarze. Oczywiście ostateczne decyzje podejmowane są przez osobę lub podmiot zlecający badanie, niemniej jego realizacja powinna być oparta na zasadzie partnerstwa oraz wzajemnego szacunku i zaufania. Ponadto obie strony powinny zidentyfikować potencjalne ryzyka i trudności, które mogą wpływać na przebieg i wyniki ewaluacji, a także komunikować je w trakcie trwania tego procesu.

Wspomniana wcześniej **partycypacyjność, a także jakość współpracy** pomiędzy wykonawcą oraz zamawiającym mają istotne znaczenie dla użyteczności ewaluacji, dlatego ich kooperacja powinna się opierać na otwartej i płynnej komunikacji oraz stałym dialogu poprzez wyznaczone osoby. Obie strony powinno cechować elastyczne podejście i otwartość na modyfikacje przyjętej koncepcji ewaluacji, ponieważ nie wszystkie trudności są możliwe do przewidzenia na wczesnych etapach jej procesu.

Bardzo ważna jest również **dbałość o jakość relacji z osobami badanymi**. Ewaluatorzy



### Zapamiętaj

W procesie ewaluacji najlepiej sprawdza się podejście partnerskie, oparte na wzajemnym zaufaniu i szacunku. W przypadku zaistnienia konfliktu, np. na tle pozyskanych informacji czy sposobu ich interpretacji, należy poszukiwać kompromisu i polubownych rozwiązań.



### Zapamiętaj

W uzasadnionych przypadkach (np. wystąpienia nieprzewidzianych wcześniej okoliczności) umowa pomiędzy zamawiającym a ewaluatorem/ką powinna zapewniać możliwość jej re-negocjowania czy aneksowania. Takie rozwiązania warunkują wyższą jakość oraz użyteczność ewaluacji.

powinni okazywać im szacunek i empatię, zachowywać neutralność i obiektywizm, a także zapewnić odpowiednie do potrzeb możliwości wypowiedzenia się oraz chronić ich anonimowość. Zawsze należy pamiętać o zapewnieniu równego dostępu do udziału w badaniach, czyli możliwości przedstawienia swoich opinii i perspektyw przez osoby, które często są z działań ewaluacyjnych wykluczane ze względu na wiek czy niepełnosprawność. Równy dostęp oznacza np. dostosowanie metod i narzędzi badawczych do potrzeb tych osób, likwidację barier architektonicznych, zapewnienie tłumacza (w tym języka migowego) czy asystenta/ki osoby z niepełnosprawnością.



### Dobra praktyka

**W umowie powinny się znaleźć m.in. zapisy dotyczące: niezależności ewaluatora/ki oraz praw autorskich, ochrony interesów osób, których dotyczy ewaluacja, w tym poufności przekazywanych informacji, sposobów rozstrzygania ewentualnych sporów, możliwości modyfikacji sposobu realizacji badania ewaluacyjnego, egzekwowania rzetelnego wykonania zleconych prac, a także zasad i terminów odbioru produktów z ewaluacji (zabezpieczenia przed opóźnieniami).**

Realizację ewaluacji znacznie ułatwia **przestrzeganie warunków zawartej umowy** pomiędzy zamawiającym i wykonawcą, która powinna w równym stopniu zabezpieczać interesy obu stron poprzez wskazanie ich praw, obowiązków oraz zakresu odpowiedzialności. Ważne by zapisy umowy były jasne i precyzyjne, aby nie pozostawiać miejsca na domysły i niedomówienia.



### Zapamiętaj

**Raport z ewaluacji może przybierać różne formy, nie tylko rozbudowanego dokumentu tekstowego, ale także prezentacji multimedialnej, tabeli wniosków i rekomendacji, filmu czy infografiki.**

## Komunikowanie i wykorzystanie wyników ewaluacji

Raportowanie i rozpowszechnianie wyników ewaluacji jest bardzo ważnym elementem, który z jednej strony stanowi zwieńczenie etapu badawczego, a z drugiej strony, poprzez wdrażanie rekomendacji. Jedną z najważniejszych zasad raportowania jest dostosowanie formy i sposobu prezentacji danych do potrzeb interesariuszy oraz celu ewaluacji (więcej na temat form i kanałów komunikacji oraz raportu z badania w rozdziale 6). Ważne jest przy tym zachowanie **przejrzystości struktury i języka raportu ewaluacyjnego, który powinien zawierać rzetelną i możliwie wyczerpującą odpowiedź na pytania badawcze.**

Niezwykle ważnym elementem procesu ewaluacyjnego jest **konsultowanie sformułowanych przez wykonawcę wniosków i rekomendacji**. Takie działania umożliwia ich uzgodnienie z różnymi interesariuszami nadając ewaluacji szerszy kontekst i zapewniając wielość perspektyw. Dyskusja w gronie interesariuszy nad trafnością wniosków i rekomendacji zwiększa prawdopodobieństwo ich wdrożenia, także dzięki doprecyzowaniu oraz ustaleniu właściwych adresatów i terminów realizacji poszczególnych zaleceń. Warto pamiętać o tym, że rekomendacje mogą mieć zarówno operacyjny (doraźny), jak i strategiczny czy horyzontalny charakter, odnosząc się do szerszych kwestii, dłuższej perspektywy czasowej i bardziej ogólnego poziomu niż dany projekt, program czy strategia.

### Formułowanie precyzyjnych wniosków i rekomendacji

w wydzielonym miejscu raportu ma istotne znaczenie dla ich wykorzystania. Rekomendacje powinny się odnosić nie tylko do negatywnych aspektów przedmiotu ewaluacji, które wymagają korekty czy udoskonalenia, ale także do wątków pozytywnych, wskazując na potrzebę ich kontynuowania w dotychczasowej postaci. Skonsultowane wnioski i rekomendacje powinny być wnikliwe, dokładne, jasne i precyzyjne. Źle sformułowane, zbyt ogólne, nietrafne i niemożliwe do wdrożenia rekomendacje mogą podważyć sens i zasadność całej przeprowadzonej ewaluacji.



#### Zapamiętaj

**Zamawiający może (i powinien) zgłosić uwagi oraz uzupełnienia do raportu ewaluacyjnego, jednak nie może ingerować w wyniki przeprowadzonego badania. Wprowadzanie wszelkich zmian w raporcie wymaga zgody ewaluatora/ki i powinno być z nim/nią negocjowane.**

Ostatnim etapem procesu ewaluacyjnego jest **wykorzystanie rekomendacji**. W ramach każdej ewaluacji należy zadbać o wypracowanie klarownych procedur wdrożenia przygotowanych zaleceń. Powinny one obejmować adresatów (tj. osoby które będą zaangażowane w ten proces), opis sposobów wdrożenia, a także wstępny harmonogram działań, które umożliwią zrealizowanie rekomendacji. Wszystkie te elementy istotnie zwiększają szanse na wykorzystanie wyników ewaluacji w praktyce. Celowi temu służy także organizowanie roboczych spotkań z adresatami zaleceń w celu wzmocnienia akceptacji i motywacji do zmian, np. poprzez przedstawienie korzyści i zysków wynikających z planowanych ulepszeń. Spotkania te mogą również wnieść dogłębne zrozumienie potrzeby wprowadzanych modyfikacji, umożliwić przedyskutowanie ich sensu i znaczenia, także w bardziej odległej perspektywie. Proces wdrażania rekomendacji powinien być monitorowany, ponieważ zwiększa to szanse na realizację uzgodnionych zaleceń.

## Załącznik 2. Przykładowy Opis Przedmiotu Zamówienia (OPZ) dla ewaluacji ex ante strategii lokalnej

Tomasz Gapski

### 1. Uzasadnienie i cele badania

Zgodnie z zapisami Ustawy o Zasadach Prowadzenia Polityki Rozwoju z dn. 6.12.2006 z późn.zm. władze samorządowe odpowiedzialne za przygotowanie lokalnych dokumentów strategicznych zobowiązane są do przeprowadzenia ewaluacji ex ante (ewaluacji uprzedniej) strategii rozwoju przed jej formalnym przyjęciem. Głównym odbiorcą badania jest rada gminy odpowiedzialna za przyjęcie strategii rozwoju.

**Celem głównym** badania jest ocena trafności oraz przewidywanej skuteczności i efektywności realizacji strategii rozwoju gminy.

**Celami szczegółowymi** badania są:

1. Ocena kompletności, użyteczności i aktualności diagnozy społeczno-ekonomicznej gminy przygotowanej na potrzeby strategii rozwoju.
2. Ocena trafności celów strategicznych i operacyjnych oraz zaplanowanych w strategii kierunków działań z punktu widzenia potrzeb rozwojowych gminy zidentyfikowanych w diagnozie.
3. Ocena adekwatności oraz spodziewanej skuteczności i efektywności działań zaplanowanych w strategii w odniesieniu do założonych celów strategicznych i operacyjnych.
4. Ocena spójności wewnętrznej i zewnętrznej strategii.
5. Ocena systemu realizacji strategii.
6. Sformułowanie rekomendacji w zakresie ewentualnych zmian w projekcie strategii rozwoju gminy.

### 2. Zakres badania i pytania badawcze

Przedmiotem badania jest projekt strategii rozwoju gminy x.

W ramach badania powinny zostać udzielone odpowiedzi na następujące pytania badawcze:

1. Czy zawarta w strategii rozwoju gminy diagnoza sytuacji społeczno-ekonomicznej jest kompletna, użyteczna i aktualna?
  - a. Czy diagnoza dostarcza kompletnych, precyzyjnych i poprawnych informacji w zakresie najważniejszych uwarunkowań zewnętrznych i wewnętrznych oraz barier, szans i potencjałów rozwojowych gminy?
  - b. Czy zakres przedmiotowy oraz wyniki diagnozy (wnioski) są użyteczne dla podjęcia decyzji strategicznych i operacyjnych w zakresie planowania długofalowej polityki rozwoju gminy?
  - c. Czy zakres czasowy i stopień aktualności danych i wyników diagnozy są wystarczające dla podjęcia trafnych decyzji strategicznych i operacyjnych w zakresie planowania długofalowej polityki rozwoju gminy?
2. Czy wizja rozwoju gminy, cele strategiczne i operacyjne oraz kierunki działań (zadania) trafnie odpowiadają na potrzeby rozwojowe gminy? Czy właściwie adresują problemy i wyzwania zidentyfikowane na etapie diagnozy?
  - a. Czy osiągnięcie celów i rezultatów przewidzianych w strategii przyczyni się do rozwiązania najważniejszych problemów społeczno-ekonomicznych gminy zidentyfikowanych w diagnozie?
  - b. Czy osiągnięcie celów i rezultatów przewidzianych w strategii pozwoli na wykorzystanie potencjałów i szans rozwojowych gminy zidentyfikowanych w diagnozie?
  - c. Czy cele i kierunki działań (zadania) przewidziane w strategii trafnie odpowiadają na bariery i potencjały rozwojowe lokalnych i regionalnych Obszarów Strategicznej Interwencji?
  - d. Czy cele i kierunki działań (zadania) przewidziane w strategii są trafne z punktu widzenia gminnych uwarunkowań przestrzennych (ocena trafności i spójności modelu funkcjonalno-przestrzennego)?
3. Czy zaplanowane w strategii rozwoju gminy kierunki działań (zadania) oraz środki finansowe są adekwatne dla osiągnięcia zakładanych rezultatów oraz celów operacyjnych i strategicznych?
  - a. Czy w strategii poprawnie sformułowane zostały związki przyczynowe pomiędzy kierunkami działań, rezultatami oraz celami operacyjnymi i strategicznymi?
  - b. Czy trafnie sformułowano założenia (wewnętrzne i zewnętrzne) warunkujące uzyskanie przewidywanych rezultatów i osiągnięcie celów operacyjnych i strategicznych?
  - c. Czy przewidziane do wydatkowania środki finansowe dają możliwość osiągnięcia zakładanych celów, rezultatów i wartości docelowych wskaźników? Czy przewidziane do wydatkowania środki finansowe są proporcjonalne do zakładanych

- rezultatów (ocena spodziewanej efektywności ekonomicznej planowanych działań)?
- d. Czy zdefiniowane w strategii cele, rezultaty i wskaźniki dają możliwość mierzenia postępów i stopnia jej realizacji? Czy wskaźniki są spójne z rezultatami oraz celami operacyjnymi i strategicznymi?
4. Czy strategia rozwoju gminy jest spójna wewnętrznie i zewnętrznie?
- a. Czy zaplanowane w ramach poszczególnych celów operacyjnych kierunki działań (zadania) uzupełniają się w sposób komplementarny?
- b. Czy cele operacyjne i strategiczne strategii są komplementarne tj. uzupełniają się w odniesieniu do ogólnej wizji rozwoju gminy?
- c. Czy cele i kierunki działań przewidziane do realizacji w ramach strategii są spójne ze strategiami na poziomie regionalnym (strategia rozwoju województwa) i krajowym?
5. Czy zaprojektowany system realizacji daje możliwość sprawnego i skutecznego wdrażania strategii rozwoju gminy?
- a. Czy dokonano jasnego i zgodnego z kompetencjami podziału zadań pomiędzy lokalne podmioty i instytucje zaangażowane w realizację strategii? Czy wskazane podmioty i instytucje dysponują odpowiednim potencjałem niezbędnym do realizacji przydzielonych im zadań?
- b. Czy zaproponowane mechanizmy i instrumenty wdrażania (w tym niezbędne dokumenty wykonawcze) są adekwatne dla realizacji kierunków działań (zadań) i celów sformułowanych w strategii rozwoju gminy?
- c. Czy zaproponowany system monitorowania i ewaluacji daje możliwość mierzenia postępów i stopnia realizacji strategii?

### 3. Założenia metodyczne badania

Szczegółowa metodyka badania zostanie zaproponowana przez Wykonawcę. Metodyka ewaluacji ex ante powinna uwzględniać perspektywę najważniejszych interesariuszy strategii, tj. instytucji publicznych, w tym w szczególności twórców strategii, lokalnych organizacji pozarządowych, przedsiębiorców oraz mieszkańców gminy.

W ramach badania powinny zostać zastosowane co najmniej następujące metody badawcze:

- **Analiza danych zastanych**, w tym szczególności:
  - analiza projektu strategii rozwoju gminy oraz nadrzędnych dokumentów strategicznych na poziomie regionalnym i krajowym,



- analiza dostępnych danych statystycznych i opracowań dot. sytuacji społeczno-ekonomicznej gminy.
- **Wywiady pogłębione** z głównymi interesariuszami strategii,
- **Warsztaty/badania focusowe** z głównymi interesariuszami strategii (w tym warsztat rekomendacyjny służący przedyskutowaniu sposobów wdrożenia rekomendacji),
- **Badanie ankietowe** wśród mieszkańców gminy.

#### 4. Produkty badania

W wyniku realizacji badania Wykonawca przygotuje i przekaze Zamawiającemu następujące produkty:

- **Raport metodyczny**, zawierający opis metodyki badawczej, planowanej do zastosowania w ramach ewaluacji, projekty narzędzi badawczych oraz harmonogram badania.
- **Raport końcowy**, prezentujący wyniki ewaluacji ex ante strategii rozwoju gminy i posiadający następującą strukturę:
  - spis treści,
  - streszczenie wykonawcze (syntetyczne ujęcie wyników badania),
  - wprowadzenie (zakres, cele i metodyka badania),
  - wyniki badania wraz z ich analizą i interpretacją,
  - wnioski i rekomendacje (w tym w szczególności propozycje zmian i/lub uzupełnień w strategii rozwoju gminy wraz z uzasadnieniem).
- **Pozostałe produkty:**
  - **Prezentacja multimedialna**, przedstawiająca w sposób atrakcyjny wizualnie najważniejsze informacje i ustalenia badania wraz z wnioskami i rekomendacjami.
  - **Bazy danych** z przeprowadzonego badania.

#### 5. Harmonogram realizacji badania

Badanie zrealizowane zostanie w terminie ... dni kalendarzowych w od dnia podpisania umowy.

Badanie realizowane będzie w etapach zgodnie z następującym harmonogramem:

- **Raport metodyczny:**
  - **Projekt** zostanie przekazany Zamawiającemu w terminie ... dni kalendarzowych od dnia podpisania umowy.
  - **Wersja końcowa** zostanie przekazana Zamawiającemu w terminie ... dni kalendarzowych od dnia podpisania umowy.

- **Raport końcowy**, prezentujący wyniki badania:
  - **Projekt** zostanie przekazany Zamawiającemu w terminie ... dni kalendarzowych od dnia podpisania umowy.
  - **Wersja końcowa** zostanie przekazana Zamawiającemu w terminie ... dni kalendarzowych od dnia podpisania umowy.
- Pozostałe produkty (prezentacja i bazy danych) zostaną przekazane Zamawiającemu w terminie ... dni kalendarzowych od dnia podpisania umowy. Dodatkowo Wykonawca zaprezentuje wyniki badania na dwóch spotkaniach zorganizowanych przez Zamawiającego w terminie ... dni kalendarzowych od dnia podpisania umowy.

Wersje końcowe dokumentów uwzględniać powinny wszystkie uwagi przekazane Wykonawcy przez Zamawiającego.

# Bibliografia

## Rozdział 1:

- Batterbury S.C.E. (2006), "Principles and purposes of European Union Cohesion policy evaluation", *Regional Studies*, 40(2), 179–188.
- Bienias, S., Herman-Pawłowska K. (2015) Ocena wpływu regulacji. Poradnik dla administracji publicznej.
- Cooksy L.J., Caracelli V.J. (2005). Quality, Context, and Use: Issues in Achieving the Goals of Metaevaluation. *American Journal of Evaluation*, vol. 26, 1: pp. 31–42.
- Cooper, H. & Hedges L.V. (1994), *The handbook of research synthesis* (pp. 97–110). New York: Russell Sage Foundation.
- Derlien H.U. (1990), „Genesis and Structure of Evaluation Efforts in Comparative Perspective”, w: R.C. Rist, *Program Evaluation and the Management of Government. Patterns and Prospects across Eight Nations*, New Brunswick, New Jersey, Transaction Publishers.
- DG Regio (2006) *The New Programming Period 2007–2013. Indicative guidelines on evaluation methods: monitoring and evaluation indicators*. Working Document No. 2.
- EVALSED (2013) The resource for the evaluation of Socio-Economic Development*. [https://ec.europa.eu/regional\\_policy/sources/evaluation/guide/guide\\_evalsed.pdf](https://ec.europa.eu/regional_policy/sources/evaluation/guide/guide_evalsed.pdf), dostęp: 2023.07.23.
- Henry G.T., Mark M.M. (2003), „Beyond Use: Understanding Evaluation's Influence on Attitudes and Actions”, *American Journal of Evaluation*, nr 24(3), 293–314.
- Kierzkowski T. (2002), *Ocena (ewaluacja) programów i projektów o charakterze społeczno-gospodarczym w kontekście przystąpienia Polski do Unii Europejskiej*, Warszawa, Ministerstwo Finansów.
- Korporowicz L. (red.) (1997), *Ewaluacja w edukacji*, Warszawa, Oficyna Naukowa.
- Mark, M.M., Greene, J.C., and Shaw, I. (2006) *The sage handbook of evaluation*. London: SAGE.
- MFIPR (2022) *Wytyczne dotyczące ewaluacji polityki spójności na lata 2021–2027*. Warszawa, Minister Funduszy i Polityki Regionalnej.
- OECD (2021), *Applying Evaluation Criteria Thoughtfully*, OECD Publishing, Paris, <https://doi.org/10.1787/543e84ed-en>.
- Olejniczak K. (2007), „Ewaluacja jako narzędzie zarządzania w sektorze publicznym”, w: Gorzelak G., Tucholska A. (red.), *Rozwój, region, przestrzeń*, Warszawa, Ministerstwo Rozwoju Regionalnego.
- Olejniczak K. (2008) „Wprowadzenie do zagadnień ewaluacji”, w: Olejniczak K., Kozak M., Ledzion B. (red.), *Teoria i praktyka ewaluacji interwencji publicznych. Podręcznik akademicki*. Warszawa, Wydawnictwa Akademickie i Profesjonalne.
- Patton M.Q. (1997), *Utilization-focused evaluation: The new century text* (3rd ed.), Thousand Oaks, CA, Sage.
- Patton, M.Q. (2021). Evaluation criteria for evaluating transformation: Implications for the coronavirus pandemic and the global climate emergency. *American Journal of Evaluation*, 42(1), 53–89.
- Picciotto, R. (2016). Evaluation and bureaucracy: The tricky rectangle. *Evaluation*, 22(4), 424–434.
- Potluka, O. (2023). Why and how to use the quality of life as an evaluation criterion?. *Evaluation and Program Planning*, 102342.

- Rossi P.H., Freeman H.E., Lipsey M.W. (2004) *Evaluation: A Systematic Approach*, 7<sup>th</sup> edition Thousand Oaks, London, Sage.
- Scriven M. (1991), *Evaluation Thesaurus*, Newbury Park, London, Sage.
- Scriven, M. (1991). Beyond formative and summative evaluation. W: G. W. McLaughlin & D. C. Phillips (Eds.), *Evaluation and education: At quarter century*. Chicago, IL: University of Chicago Press, pp. 19–64.
- Stufflebeam D.L. (2001) „Evaluation models”, *New Directions for Evaluation*, nr 89, 7–92.
- Stufflebeam, D. L., & Shinkfield, A. J. (2012). *Systematic evaluation: A self-instructional guide to theory and practice* (Vol. 8). Springer Science & Business Media.

## Rozdział 2:

- Astbury B., Leeuw F.L. (2010). Unpacking black boxes: Mechanisms and theory building in evaluation. *American Journal of Evaluation*, 31(3).
- Coryn C.L.S. i in. (2011). A Systematic Review of Theory-Driven Evaluation Practice From 1990 to 2009, *American Journal of Evaluation*, 32(2).
- Funnell S.C., Rogers P. J. (2011). *Purposeful Program Theory: Effective Use of Theories of Change and Logic Models*, Jossey Bass.
- Górniak J. (2007). „Ewaluacja w cyklu polityk publicznych” w: Mazur S. (red.), *Ewaluacja funduszy strukturalnych – perspektywa regionalna*. Kraków. Uniwersytet Ekonomiczny w Krakowie, Małopolska Szkoła Administracji Publicznej.
- Hermann-Pawłowska K., Skórska P. (2017). „Przegląd systematyczny koncepcji ewaluacji opartej na teorii” w: Hermann-Pawłowska K., Pokorski J., Popis Z., Wyszyńska T., (red.), *Ewaluacja oparta na teorii w złożonym otoczeniu społeczno-ekonomicznym*. Warszawa, Polska Agencja Rozwoju Przedsiębiorczości.
- Olejniczak K., Kupiec T., Raimondo E. (2014). „Brokerzy wiedzy. Nowe spojrzenie na rolę jednostek ewaluacyjnych” w: Haber A., Olejniczak K. (red.), *(R)ewaluacja 2. Wiedza w działaniu*. Warszawa, Polska Agencja Rozwoju Przedsiębiorczości.
- Kupiec, T., Celińska-Janowicz, D., & Pattyn, V. (2023). Understanding evaluation use from an organisational perspective: A review of the literature and a research agenda. *Evaluation*, 29(3), 338–355.
- Kupiec T. (2022). Does evaluation quality matter? Quantitative analysis of the use of evaluation findings in the field of cohesion policy in Poland. *Evaluation and Program Planning*. Volume 93.
- Leeuw F.L. (2012). Linking theory-based evaluation and contribution analysis: Three problems and a few solutions. *Evaluation*, 18(3).
- Martini A. (2009) Counterfactual impact evaluation: what it can (and cannot) do for cohesion policy, 6th European Conference on Evaluation of Cohesion Policy, Warszawa.
- Martini A. (2013). „Różne oblicza randomizowanych prób kontrolnych” w: Haber A., Popis Z. (red.), *(R)ewaluacja. Poszukiwanie nowych metod oceny efektów*. Warszawa, Polska Agencja Rozwoju Przedsiębiorczości.
- Mayne J. (2011). „Contribution analysis: addressing cause and effect” w: Schwartz R., Forss K., Marra M. (red.) *Evaluating the Complex*. New Brunswick, NJ: Transaction Publishers.
- Mayne J. (2019). Revisiting Contribution Analysis. *Canadian Journal of Program Evaluation*.

- Olejniczak K. (2008) Mechanizmy Wykorzystania ewaluacji. Studium ewaluacji śródkresowych INTERREG III. Centrum Europejskich Studiów Regionalnych i Lokalnych UW. Wydawnictwo Naukowe SCHOLAR.
- Pawson R., Tilley N. (1997). *Realistic Evaluation*. London, Sage.
- Pellegrin J., Colnot L. (2020). Research for REGI Committee – The Role of Evaluation In Cohesion Policy, European Parliament, Policy Department for Structural and Cohesion Policies, Brussels.
- Rogers P.J., Weiss C.H. (2007), Theory-based evaluation: Reflections ten years on: Theory-based evaluation: Past, present, and future, *New Directions for Evaluation*. Nr 114.
- Stern E. i in. (2012). Broadening The Range of Designs and Methods for Impact Evaluations, Report of a study commissioned by the Department for International Development, DFID.
- Weiss C.H. (1995). „Nothing as Practical as Good Theory: Exploring Theory-Based Evaluation for Comprehensive Community Initiatives for Children and Families.” w: Connell J., Kubisch A., Schorr L. B., Weiss C.H. (red.), *New Approaches to Evaluating Community Initiatives*. Washington, D.C., Aspen Institute.
- Weiss C.H. (1998). *Evaluation: Methods for Studying Programs and Policies*, 2<sup>nd</sup> Edition, Upper Saddle River, New Jersey, Prentice Hall.
- White H. (2009). „Theory-based impact evaluation: principles and practice”, *Journal of Development Effectiveness*, 1(3).
- EVALSED (2013). – The resource for the evaluation of Socio-Economic Development : Sourcebook – Method and techniques [https://ec.europa.eu/regional\\_policy/sources/evaluation/guide/evaluation\\_sourcebook.pdf](https://ec.europa.eu/regional_policy/sources/evaluation/guide/evaluation_sourcebook.pdf), dostęp: 2023.08.24.

### Rozdział 3:

- Bienias S., Strzęboszewski P., Opałka E. (red.), Ewaluacja. Poradnik dla pracowników administracji publicznej, Ministerstwo Rozwoju Regionalnego, Warszawa, 2012.
- Ciężka B., Rybińska A. (red.), Evaluation in Educational Practice, Fundacja Rozwoju Systemu Edukacji, Warszawa 2020.
- Chłoń-Domińczak A., Pokorski J. i in., Public Intelligence. Wykorzystanie danych administracyjnych do monitorowania i ewaluacji polityk publicznych, Polska Agencja Rozwoju Przedsiębiorczości, Warszawa 2022.
- Hermann-Pawłowska K., Pokorski J. i in. (red.), Theory-based evaluation in complex environments, Polska Agencja Rozwoju Przedsiębiorczości, Warszawa, 2017.
- Lampart M., Rozwiązania generatywnej sztucznej inteligencji – zagrożenia i aspekty prawne, Polska Agencja Rozwoju Przedsiębiorczości, Warszawa 2023.
- Musidłowska M., Wawrzyniak B., Zygmuntowski J.J., Uwolnić potencjał danych. Zarządzanie danymi jako zasobem współdzielonym, Raport opracowany na zlecenie Cyfryzacji KPRM, Warszawa 2022.
- Olejniczak K., Kozak M., Ledzion B. (red.), Teoria i praktyka ewaluacji interwencji publicznych. Podręcznik Akademicki, Akademia Leona Koźmińskiego, Warszawa 2008.
- Rogalińska D., Wtórne wykorzystanie danych administracyjnych w statystyce publicznej – obecnie i w przyszłości., Ekspertyza opracowana na zlecenie PARP, Warszawa 2022

## Rozdział 4:

- Evalsed – the resource for the evaluation of socio-economic development, Komisja Europejska, 2009 ([http://ec.europa.eu/regional\\_policy/sources/docgener/evaluation/evalsed/index\\_en.htm](http://ec.europa.eu/regional_policy/sources/docgener/evaluation/evalsed/index_en.htm))
- Olejniczak, K. Ewaluacja jako proces badawczy, [w:] K. Olejniczak, M. Kozak, B. Ledzion (red.), *Teoria i praktyka ewaluacji interwencji publicznych*. Podręcznik Akademicki, Warszawa, Wydawnictwa Akademickie i Profesjonalne, 2008.
- Haber, A. Zarządzanie badaniem ewaluacyjnym – perspektywa zlecającego, [w:] K. Olejniczak, M. Kozak, B. Ledzion (red.), *Teoria i praktyka ewaluacji interwencji publicznych*. Podręcznik Akademicki, Warszawa, Wydawnictwa Akademickie i Profesjonalne, 2008.

## Rozdział 5:

- Olejniczak Karol, Mazur Stanisław, Weremiuk Alicja, Miedziński Michał, Ledzion Bartosz, Kupiec Tomasz (2012) *Realizacja badań ewaluacyjnych w świetle Prawa Zamówień Publicznych*. Warszawa: Ministerstwo Rozwoju Regionalnego, oraz na stronie internetowej Ministerstwa Funduszy i Polityki Regionalnej w bazie dokumentacji zamówień publicznych (<https://zamowienia.mfipr.gov.pl/>)

## Rozdział 6:

- MFIPR (2022) Wytyczne dotyczące realizacji zasad równościowych w ramach funduszy unijnych na lata 2021-2027. Załącznik nr 2 Standardy dostępności dla polityki spójności 2021–2027.
- CDC (2013) *Evaluation Reporting: A Guide to Help Ensure Use of Evaluation Findings*. Atlanta, GA: US Dept of Health and Human Services. <https://www.cdc.gov/dhds/docs/Evaluation-Reporting-Guide.pdf>, dostęp: 2023-08-01.
- Donaldson, S. I., Christie, C. A., & Mark, M. M. (Eds.). (2014). *Credible and actionable evidence: The foundation for rigorous and influential evaluations*. Sage Publications.
- Evalcommunity (2013). *How to Write Evaluation Reports: Purpose, Structure, Content, Challenges, Tips, and Examples*. <https://www.evalcommunity.com/career-center/structure-of-the-evaluation-report/>, dostęp: 2023-08-01.
- Grob, G. F. (2015). Writing for impact. w: J. S. Wholey, H. P. Hatry, K. E. Newcomer (red.), *Handbook of Practical Program Evaluation*, wydanie czwarte, Jossey-Bass A Wiley Imprint, s. 739–764.
- Kupiec, T. (2014). Użyteczność ewaluacji jako narzędzia zarządzania regionalnymi programami operacyjnymi. *Studia Regionalne i Lokalne*, 15(56), 52–67.
- MFIPR (2022) *Wytyczne dotyczące ewaluacji polityki spójności na lata 2021–2027*. Warszawa, Minister Funduszy i Polityki Regionalnej.
- Ministerstwo Zdrowia (2018). *Ewaluacja on-going realizacji celów V osi priorytetowej PO WER*.
- Olejniczak K., Kupiec T., Raimondo E. (2014). Brokerzy wiedzy. Nowe spojrzenie na rolę jednostek ewaluacyjnych. W: A. Haber, K. Olejniczak (red.). *(R)ewaluacja 2. Wiedza w działaniu*. Warszawa: Polska Agencja Rozwoju Przedsiębiorczości, s. 67–114.
- Olejniczak, K., Raimondo, E., & Kupiec, T. (2016). Evaluation units as knowledge brokers: Testing and calibrating an innovative framework. *Evaluation*, 22(2), 168–189.
- Torres, R. T., Preskill, H., & Piontek, M. E. (2005). *Evaluation strategies for communicating and reporting: Enhancing learning in organizations*. Sage.

- UNICEF (2017) Adapted UNEG Evaluation Reports Standards. <https://www.unicef.org/evaluation/media/816/file/UNICEF-Adapted-UNEG-Evaluation-Report-Standards.pdf>, dostęp: 2023-08-01.
- Weiss, C. H., & Bucuvalas, M. J. (1980). *Social science research and decision-making*. Columbia University Press.
- Województwo Śląskie (2021). Ewaluacja dotycząca sposobu, w jaki wsparcie w ramach RPO WSL na lata 2014–2020 przyczyniło się do osiągnięcia celów w ramach Osi Priorytetowej V Ochrona środowiska i efektywne wykorzystanie zasobów. [https://rpo.slaskie.pl/czytaj/ewaluacja\\_7](https://rpo.slaskie.pl/czytaj/ewaluacja_7), dostęp: 2023-08-01.

## Rozdział 7:

- Adler, M.D. i Posner, E.A. (red.) (2001). *Cost-benefit analysis. Legal, economic and philosophical perspectives*. University of Chicago Press, Chicago.
- Babbie, E. (2007). *Badania społeczne w praktyce*, PWN, Warszawa.
- Bradburn, N.M., Sudman, S. i Wansink, B. (2004). *Asking questions: the definitive guide to questionnaire design-for market research, political polls, and social and health questionnaires*. John Wiley & Sons.
- Dabalen, A., Etang, A., Hoogeveen, J., Mushi, E., Schipper, Y. i von Engelhardt, J. (2016). *Mobile Phone Panel Surveys in Developing Countries: A Practical Guide for Microdata Collection. Directions in Development*. Washington, DC: World Bank.
- Dillman, D. A. (2000). *Mail and Internet Surveys: The Tailored Design Method*, 2nd ed. New York: Wiley.
- Fowler, F.J., Jr. (2002). *Survey Research Methods*, 3rd ed. Thousand Oaks, CA: Sage.
- Hey, R.B. (2017). *Chapter 27 – Using Benchmarking Results*, (w:) Hey, R.B. *Performance Management for the Oil, Gas, and Process Industries*, Gulf Professional Publishing (s. 419–428), ISBN 9780128104460
- Hox, J.J. i Boeijs, H.R. (2005). *Data collection, primary vs. secondary*. Encyclopedia of Social Measurement, 593–599.
- Jasiński, M. i Kowalski, M. (2007). *Fałszywa sprzeczność: metodologia jakościowa czy ilościowa?*, [w:] A. Haber (red.), *Ewaluacja ex-post. Teoria i praktyka badawcza*, Warszawa, Polska Agencja Rozwoju Przedsiębiorczości.
- Keehley, P. i Abercrombie, N. (2008). *Benchmarking in the Public and Nonprofit Sectors: Best Practices for Achieving Performance Breakthroughs*. Jossey-Bass.
- Morgan, S. i Winship, C. (2014). *Counterfactuals and Causal Inference: Methods and Principles for Social Research (Analytical Methods for Social Research)*, 2 edycja. Cambridge: Cambridge University Press.
- Morris, A. (2015). *A practical introduction to in-depth interviewing*. London: Sage.
- Odimegwu, C.O. (2000). *Methodological Issues in the Use of Focus Group Discussion as a Data Collection Tool*. Journal of Social Sciences, 4(2–3), 207–212.
- Petticrew, M. i Roberts, H. (2006). *Systematic Reviews in the Social Sciences: A Practical Guide*. Oxford: Blackwell.
- Rosenbaum, P. i Rubin, D. (1983). *The central role of the propensity score in observational studies for causal effects*, Biometrika, 70(1), 41–55.
- Rosenbaum, P. (2002). *Observational Studies*, 2 edycja. New York: Springer-Verlag.

- Sills, S. J. i Song, C. Y. (2002). *Innovations in survey research: an application of web-based surveys*. Soc Sci Comput Rev, 20(1), 22–30.
- Skowrońska, A. (2020). *Desing Thinking a ewaluacja – różnice, podobieństwa, przykłady zastosowania w edukacji*, w: Cięża, B. i Rybińska, A. (2020) Ewaluacja w działaniach edukacyjnych, Fundacja Rozwoju Systemu Edukacji, Warszawa, s. 118–129, ISBN 978-83-66515-04-8.
- Stake, R.E. (1995). *The art of case study research*. London: Sage Publications Ltd.
- Yin, R.K. (2009). *Case study research, design and method*. London: Sage Publications Ltd.
- Waltz, T.J., Powell, B.J., Matthieu, M.M. i in. (2015). *Innovative methods for using expert panels in identifying implementation strategies and obtaining recommendations for their use*. Implementation Science, 10, Article number: A44, <https://doi.org/10.1186/1748-5908-10-S1-A44>



## Spis tabel

|            |                                                                                         |    |
|------------|-----------------------------------------------------------------------------------------|----|
| Tabela 1.  | Zestawienie różnic między ewaluacją, monitoringiem i kontrolą/audytem                   | 11 |
| Tabela 2.  | Kryteria ewaluacyjne                                                                    | 15 |
| Tabela 3.  | Szacunkowe koszty realizacji metod i technik badawczych.<br>Kwoty na wrzesień 2023      | 57 |
| Tabela 4.  | Czynniki wpływające na wycenę badania                                                   | 58 |
| Tabela 5.  | Przykładowe kryteria jakościowe oceny ofert na badanie ewaluacyjne                      | 68 |
| Tabela 6.  | Lista sprawdzająca zrozumiałość i spójność przyjętych elementów zamówienia na ewaluację | 69 |
| Tabela 7.  | Przykładowa tabela celów/pytań badawczych i metod                                       | 78 |
| Tabela 8.  | Modelowa struktura raportu końcowego                                                    | 83 |
| Tabela 9.  | Typowi odbiorcy i użytkownicy ewaluacji                                                 | 88 |
| Tabela 10. | Wybrane formy i kanały rozpowszechniania wyników badania z sugerowanymi odbiorcami      | 89 |

## Spis ilustracji

|            |                                                                             |    |
|------------|-----------------------------------------------------------------------------|----|
| Rysunek 1. | Kryteria ewaluacyjne na tle logiki interwencji                              | 16 |
| Rysunek 2. | Schemat logiczny interwencji polegającej na opracowaniu poradnika ewaluacji | 25 |
| Rysunek 3. | Ewaluacja w cyklu wdrażania interwencji publicznej                          | 31 |