

Dariusz Zalewski

Możliwości i ograniczenia wykorzystania wyników ewaluacji przez instytucje publiczne

„Diaboliczne ćwiczenie”

Czym jest ewaluacja i jakie są jej funkcje wyjaśniano na wiele sposobów. Nie ma więc powodu, aby po raz kolejny przytaczać znane podręcznikowe definicje. Kariera, jaką zrobiło to pojęcie, wydaje się być nierozzerwalnie związana z dwoma zjawiskami. Po pierwsze – z poszerzaniem się sfery publicznej i wzrostem znaczenia administracji odpowiedzialnej za realizację określonych zadań. Po drugie – z rozczarowaniem efektywnością realizacji tychże zadań przez administrację zorganizowaną wedle klasycznych zasad sformułowanych przez M. Webera. Oczywiście, musimy pamiętać, że sama ewaluacja ma znacznie dłuższą historię – sięga ona początków XX wieku, a w jej rozwoju miał swój udział m.in. Florian Znaniecki – niż krytyka weberowskiej biurokracji i stosownie do niej poszukiwanie nowych form organizacyjnych administracji publicznej dla skuteczniejszego i efektywniejszego realizowania zadań. Należałoby też dla porządku dodać, że badania ewaluacyjne, jako narzędzie oceny realizowanych programów interwencji społecznej, stosowane są nie tylko w obrębie działalności administracji publicznej, ale to głównie jej specyfice czy też niedoskonałości zawdzięczamy w znacznej mierze ich rozwój.

Wzrost zainteresowania badaniami ewaluacyjnymi w Polsce jest ściśle związany z przystąpieniem Polski do Unii Europejskiej. Przyczyna jest prozaiczna i ma charakter obligatoryjnego wymogu ewaluowania programów publicznych finansowanych przez UE¹. Czarujący powab ewaluacji widoczny jest niemalże w każdym oficjalnym dokumencie UE poświęconym tejże problematyce. W pierwszym zdaniu jednego z dokumentów Komisji Europejskiej możemy przeczytać: „Ewaluacja może być traktowana przez kogoś jako diaboliczne ćwiczenie. Jednakże, jeśli ewaluacja jest dobrze przeprowadzona

¹ *Inicjatywa Komisji Europejskiej Sound and Efficient Management 2000* (znana jako SEM 2000) włącza użycie ewaluacji jako kluczowego elementu poprawy kultury zarządzania samej Komisji. Kluczową innowacją SEM będzie wymaganie, aby systematyczną ewaluacją objąć wszystkie programy UE; zob. „Evaluating UE Expenditure Programmes: A Guide”, European Commission, 01.1997, s. 7.

i jeśli rezultaty ewaluacji są wykorzystane przez decydentów, może ona przyczynić się do poprawy publicznych programów, jak również zwiększyć ich przejrzystość, odpowiedzialność i efektywność”².

Chociaż ewaluacja pełni wiele funkcji, a jej wartość powinna być oceniana przez pryzmat złożoności funkcji i celów, jakie przed nią się stawia, to jednak z uwagi na fakt, że największym odbiorcą jej wyników jest administracja publiczna, najważniejsza wydaje się być ocena racjonalności poniesionych kosztów dla osiągnięcia określonych celów. Myśl tę dość dobrze wyraża następujące zdanie: „Ewaluacja ma dostarczać kryteriów, metod i środków do oceny racjonalności działań publicznych, spełniając w sektorze publicznym rolę analogiczną do mechanizmu rynkowego w sektorze prywatnym”³. Można by oczywiście argumentować, że taka redukcja funkcji badań ewaluacyjnych jest nieuprawniona, a porównanie ich roli do mechanizmu rynkowego jest w tej mierze nietrafne, w jakiej państwo różni się od rynku, a logika działań administracji publicznej jest odmienna niż w przypadku prywatnego przedsiębiorstwa. Jednak krytyka administracji publicznej, inspirowana głównie przez przedstawicieli neoliberalnej ekonomii klasycznej oraz teorii wyboru publicznego, sprawiła, że to właśnie rynek ze swoimi zasadami konkurencji i dążeniem do ekonomicznej efektywności stał się płaszczyzną odniesienia dla proponowanych reform w zarządzaniu sektorem publicznym. Natomiast rolą ewaluacji miałyby być odpowiedź na pytanie, czy realizowane programy publiczne są efektywne w ekonomicznym tego słowa znaczeniu.

Można sądzić, że oczekiwania formułowane pod adresem ewaluacji są zbyt duże. Jeśli tak, to jakie są, z jednej strony, możliwości, a z drugiej strony ograniczenia wykorzystania wyników ewaluacji w praktyce administracyjnej? Odpowiedzi na te pytania wydają się być interesujące w kontekście poszukiwanej efektywności działania sektora publicznego, owego „Świętego Grała” wszystkich reformatorów, jak zręcznie ujął to V. Writgh⁴. Analiza będzie koncentrowała się na zewnętrznych ograniczeniach możliwości wykorzystania wyników ewaluacji pomimo świadomości, że jej aplikacyjna wartość zależy w równym, jeśli nie większym, stopniu od czynników wewnętrznych procesu ewaluacyjnego (poprawności założeń metodologicznych, przestrzegania standardów badawczych, kompetencji badacza itp.). Taki wybór podyktowany jest powszechnie formułowanymi pod adresem badań ewaluacyjnych oczekiwaniami (niezwykle sprzecznymi) ze strony różnych środowisk, aby ich wyniki przyczyniały się do wzrostu efektywności realizowanych programów publicznych.

² *Evaluating UE Expenditure Programmes...*, op. cit.

³ Shadish W.R., Cook T.D., Leviton L.C., *Foundations of Program Evaluation: Theories of Practice*, Sage Publications, Newbury Park, London 1991, s. 19; cyt. za K. Olejniczak, „Teoretyczne podstawy ewaluacji *ex-post*”, w: A. Haber (red.), *Ewaluacja ex-post. Teoria i praktyka*, PARP, Warszawa 2007, s. 19.

⁴ Wright V., „Paradoxes of Administrative Reform”, w: W.J.M. Kickert, *Public Management and Administrative Reform in Western Europe*, Edward Elgar, Cheltenham 1997, s. 11.

Rozdział ten jest tak pomyślany, aby w klarowny sposób przedstawić, jakie są oczekiwania wobec ewaluacji oraz jej możliwości i ograniczenia. Ponieważ znaczenie ewaluacji może być różne w różnych koncepcjach realizacji zadań publicznych, pokażemy w skrótowny sposób propozycje zmian w administracji publicznej i ewentualne konsekwencje dla ewaluacji programów interwencji społecznych. Byłoby wskazane, aby ów tekst inspirował dyskusję o możliwościach i ograniczeniach ewaluacji nie tylko wśród przedstawicieli środowiska naukowego, gdyż ci zawsze znajdą czas, by spierać się o coś, ale także wśród przedstawicieli administracji, odpowiedzialnej za realizację zadań o charakterze publicznym.

Słów kilka o reformach administracji i ich konsekwencjach dla ewaluacji

To, jaką rolę ewaluacja może odegrać w procesie efektywnego wydawania środków publicznych, w jakim stopniu przyczynić się do racjonalnego podejmowania decyzji alokacyjnych, zależy w dużej mierze od tego, w jaki sposób i kto realizuje usługi o charakterze publicznym.

Przed prawie trzydziestu laty M. Friedman wspólnie ze swoją małżonką Rose opublikowali książkę *Free to Choose*, w której przystępnie przedstawili zasady rządzące racjonalnością alokacji zasobów finansowych, ze względu na dwie pary zmiennych: czyje pieniądze wydajemy (własne bądź cudze) oraz na kogo wydajemy (na siebie bądź na kogoś innego). W wyniku połączenia tych dwóch par alternatyw uzyskujemy macierz odpowiedzi, na którą składają się cztery warianty: własne pieniądze na własne potrzeby (1), własne pieniądze na potrzeby innych (2), cudze pieniądze na własne potrzeby (3) oraz cudze pieniądze na potrzeby innych (4)⁵. Krytykując nieracjonalność alokacyjną administracji publicznej, Friedmanowie skoncentrowali się na wariantach 3 i 4, dowodząc, że wydawanie publicznych pieniędzy stanowi przesłankę do marnotrawstwa tych środków, gdyż nie ma wystarczających bodźców do uruchomienia racjonalnych strategii działania tak, aby optymalizować uzyskiwane przez beneficjentów korzyści do poniesionych publicznych kosztów. W podobnym duchu wypowiadają się przedstawiciele teorii wyboru publicznego, którzy swoje analizy opierają na mikroekonomicznych założeniach neoliberalnej ekonomii. *Nobody spends somebody else's money as carefully as he spends his own* – pisał przed laty W. Mitchell, analizując funkcjonowanie sektora publicznego w ramach teorii wyboru publicznego⁶.

⁵ Friedman M., Friedman R., *Wolny wybór*, Wyd. Panta, Sosnowiec 1994, s. 111.

⁶ Cyt. za: Dunleavy P., *Democracy, Bureaucracy and Public Choice*, Harvester Wheatsheaf, 1991, s. 173.

Uwagi poczynione przez Friedmanów raczej nie przyczyniły się do rozwoju badań ewaluacyjnych, a ich intencją było pokazanie, że administracja publiczna głównie dba o własne, a nie publiczne interesy. Jeden z możliwych wniosków, jaki moglibyśmy wyprowadzić z neoliberalnej krytyki funkcjonowania sektora publicznego, byłby taki, że wyniki badań ewaluacyjnych mogłyby być pomocne, by nie powiedzieć rozstrzygające, w dyskusji nad efektywnością programów realizowanych przez instytucje publiczne. Potwierdza to przekaz płynący z rozmaitych wypowiedzi przedstawicieli politycznych establishmentów, szczególnie dobrze widoczny dla polityków UE, którzy wyniki ewaluacji traktują jako czynnik racjonalizacji procesu decyzyjnego w sektorze publicznym. „Ewaluacja jest niezbędną częścią nowoczesnej praktyki zarządzania sektorem publicznym” – pisał przed kilku laty E. Liikanen, członek Komisji Europejskiej odpowiedzialny za budżet i administrację⁷.

Oczywiście musimy pamiętać, że ewaluacja jest tylko narzędziem służącym ocenie rozmaitych programów społecznej interwencji, podejmowanych nie tylko przez administrację publiczną, i w żadnym przypadku nie można jej traktować jako funkcjonalnej recepty na rozmaite kłopoty, których nastręcza funkcjonowanie sektora publicznego. Jeśli jednak chcemy odpowiedzieć na takie czy inne zarzuty kierowane pod adresem instytucji publicznych, często z resztą uzasadnione, lub też dostarczyć samej administracji informacji o kosztach i efektach jej funkcjonowania, to wydaje się, że powinniśmy odwołać się do wyników oceny tejże działalności, które możemy uzyskać w ramach prowadzonych badań ewaluacyjnych. Należy jednak zdawać sobie sprawę nie tylko z rozmaitych ograniczeń wykorzystania wyników tych badań, ale też i z tego, że znaczenie samej ewaluacji może być różne ze względu na to, kto i w jaki sposób będzie realizował określone zadania tradycyjnie przynależne do sektora publicznego w ramach wzajemnych powiązań państwa i rynku.

„Urynkowienie państwa” a ewaluacja

Propozycje radykalnych zmian w administracji publicznej, jakie pojawiły się m.in. w latach 80. ubiegłego wieku, głównie za sprawą ekonomistów ze Szkoły Chicagowskiej oraz zwolenników teorii wyboru publicznego, nawiązywały do zasad funkcjonowania wolnego rynku z jego logiką maksymalizowania korzyści. Doświadczenie było takie, że administracja publiczna nie radzi sobie z rozwiązywaniem rozmaitych problemów, do czego została powołana, a jej efektywność alokacyjna jest zdecydowanie mniejsza niż podmiotów prywatnych działających na rynku. Uzasadnieniu

⁷ *Evaluating UE Expenditure Programmes...*, op. cit.

utrzymywania sektora publicznego i zarządzającej nim administracji zawartemu w hasle *market failure* przeciwstawiono hasło *government failure*. Porównując zatem działanie administracji publicznej do przedsiębiorstw rynkowych (czy słusznie to zupełnie inna sprawa) neoliberalni ekonomiści zaproponowali, po pierwsze, deregulację sektora publicznego, po drugie, jego prywatyzację⁸.

Deregulacja sektora publicznego miała na celu przede wszystkim zredukowanie odpowiedzialności państwa za rozwiązywanie rozmaitych problemów tak, aby zwiększyć przestrzeń dla działań rynkowych. „Jeśli publiczna regulacja – argumentował George Stigler – nie jest konieczna, zatem wiele programów publicznego sektora regulujących gospodarkę powinno zostać usuniętych, włączając w to agencje, które wprowadzają te publiczne regulacje”⁹. Krótko mówiąc, deregulacja, a w ślad za nią prywatyzacja agencji sektora publicznego, miały zapewnić większą konkurencyjność między jednostkami odpowiedzialnymi za realizację funkcji publicznych państwa, a funkcję ewaluacyjną ich działalności miały pełnić przede wszystkim reguły gry rynkowej ze swoją zdolnością do skutecznego rozwiązywania problemów alokacji zasobów. Kłopot tylko w tym, że rynkowe czy quasi-rynkowe instytucje alokują swoje zasoby tam, gdzie istnieje możliwość uzyskania wysokiej stopy zwrotu zainwestowanych środków. Poza tym nie ma żadnych gwarancji, że konkurencja między podmiotami rywalizującymi o realizację określonych usług o charakterze publicznym będzie prowadzona wedle zasad bezstronności. Abstrahując jednak od wielu problemów i krytyk propozycji ekonomistów Szkoły Chicagowskiej czy przedstawicieli teorii wyboru publicznego, jak na przykład takich, że deregulacja wymaga nowych form regulacji¹⁰, a prywatyzacja zwiększa koszty transakcyjne, należy zauważyć, że jedną z konsekwencji tych propozycji powinno być zmniejszenie roli badań ewaluacyjnych. Tę rolę powinien w znacznym stopniu przejąć rynek ze swoją czytelną zasadą ekonomicznej efektywności. Prywatyzacja bowiem stanowi odwrócenie procesu poszerzania państwowego interwencjonizmu, który to proces był jedną z kluczowych przesłanek rozwijania badań ewaluacyjnych czy szerzej, implementacji dorobku nauk społecznych do rozwiązywania konkretnych problemów¹¹.

⁸ Lane J.E., „Public Sector Reform: Only Deregulation, Privatization and Marketization?”, w: J.E. Lane (red.), *Public Sector Reform. Rationale, Trends and Problems*, Sage Publications, London 1997, s. 3.

⁹ Ibidem.

¹⁰ Majone G., *The rise of the regulatory state in Europe*, West European Politics, 17, s. 77–101.

¹¹ K.W. Frieske np. podaje, że początki zaangażowania przedstawicieli nauk społecznych w „praktykę społeczną” w Stanach Zjednoczonych miały miejsce w okresie New Dealu, kiedy państwo zwiększyło obszar swojego interwencjonizmu; zob. Frieske K.W., *Socjologia w działaniu. Nadzieje i rozczarowania*, Wyd. Uniwersytetu Warszawskiego, Warszawa 1990, s. 78–79.

„Urynkowienie władzy” a ewaluacja

Propozycje zmian w sektorze publicznym złożone przez neoliberalnych ekonomistów z pewnością stanowiły poważne zagrożenie dla interesów ewaluatorów, a w jakiejś mierze także i dla samych badań ewaluacyjnych, gdyż rychło mogłoby się okazać, że nie ma czego ewaluować albo nie byłoby dla kogo ewaluować. Takiego zagrożenia z pewnością nie ma ze strony reformatorskiej koncepcji „publicznej przedsiębiorczości” (*public entrepreneurship*), której bodaj najbardziej znanymi orędownikami stali się w latach 90. ubiegłego wieku D. Osborne i T. Gaebler, autorzy bardzo głośnej książki *Reinventing Government*. Sama propozycja, wpisująca się w szerszą strategię „nowego zarządzania publicznego” (*New Public Management*) nie była ani specjalnie nowa, ani oryginalna, a czerpała zarówno z założeń teorii wyboru publicznego (opartej na neoliberalnym modelu racjonalności jednostki), jak i neotaylorowskiej teorii organizacji i zarządzania. To, co ewidentnie różni koncepcję „publicznej przedsiębiorczości” od koncepcji urynkowienia usług publicznych, pochodzącej ze Szkoły Chicagowskiej, zawiera się w różnym podejściu do sił rynkowych, które miały zreformować sektor publiczny. O ile neoliberalni ekonomiści proponowali ograniczyć działalność państwa, aby zwiększyć przestrzeń dla sektora prywatnego, o tyle „menedżeryści” zaproponowali włączenie mechanizmów rynkowych do działalności administracji tak, aby stworzyć tzw. wewnętrzny rynek w ramach sektora publicznego. Nawiązując do neotaylorowskiej koncepcji efektywnego zarządzania, zaproponowali oni model administracji zorientowanej na osiąganie celów, a w szczególności położyli nacisk na wykorzystywanie wskaźników wydajności i skuteczności, ewaluację realizowanych programów, wynagradzanie wedle zasady uzyskiwanych rezultatów i podnoszenie jakości świadczonych usług¹².

Osborne i Gaebler, zachęcając do systematycznej ewaluacji realizowanych przez administrację publiczną programów, wskazali jednocześnie na szereg problemów, jakie zwykle towarzyszą pomiarowi oczekiwanych wyników. Ich uwagi, sformułowane na bazie doświadczenia administracyjnej praktyki, miały przestrzegać, jakich błędów przy dokonywaniu pomiaru osiągniętych wyników należy unikać. Warto do nich odnieść się chociażby po to, aby pokazać, nie kwestionując ich merytorycznej zasadności, że ani nie wyczerpują one katalogu możliwych do popełnienia błędów, na jakie narażone jest wnioskowanie na podstawie wyników ewaluacji, ani zastrzeżeń, jakie pod adresem ewaluacji kierują rozmaici badacze.

¹² Keraudren Ph., van Mierlo H., „Theories of Public Management Reform and their Practical Implications”, w: T. Verheijen, D. Coombes, *Innovations in Public Management. Perspectives from East and West Europe*, Edward Elgar, Cheltenham 1998, s. 40.

- Po pierwsze, „między mierzeniem procesu działania a mierzeniem jego wyników jest wielka różnica”¹³. Uwaga ta brzmi dość banalnie dla każdego, kto przynajmniej w stopniu podstawowym jest wtajemniczony w arkania ewaluacyjnego rzemiosła, choć niekoniecznie musi być równie oczywista dla przedstawicieli administracji. W gruncie rzeczy autorom chodziło o to, że administracja często koncentruje się na pomiarze nie tego, co powinno być przedmiotem oceny. „Skłonność do koncentrowania się na jakimś procesie jest naturalna – menedżerowie mierzą to, co robią powierzone im organizacje, a w organizacjach kierujących się regulaminami i przepisami ludzie traktują swoją pracę jako pewien proces, którego przebieg wyznaczają te przepisy. Jeśli zatem przestrzegają dokładnie przepisów i wytwarzają tyle, ile się od nich oczekuje, to wywiązują się ze swoich zadań. Rzadko myślą o wynikach swojej pracy, o tym, jaki wpływ ma ona na tych, którym dana agencja ma służyć. Jednakże taka, nawet doskonale wykonana praca, jest stratą czasu i pieniędzy, jeśli nie prowadzi do osiągnięcia pożądaných wyników”¹⁴.
- Po drugie, „istnieje wielka różnica między mierzeniem wydajności a mierzeniem skuteczności”¹⁵. Dla każdej jednostki organizacyjnej, powiadają autorzy, ważne powinno być skuteczne osiąganie zakładanego celu. Wydajność nie jest i nie powinna być celem, do jakiego zmierza organizacja, a jedynie środkiem do jego osiągnięcia. Koncentracja uwagi na wydajności (i jego pomiarze) w oderwaniu od skuteczności działania prowadzi do paradoksalnej sytuacji marnotrawstwa środków. „Nie ma nic głupszego niż wydajniejsze robienie czegoś, czego nie powinno się w ogóle robić”¹⁶. Co więcej, Osborne i Gaebler twierdzą, że: „koncentrowanie się na wydajności prowadzi również do wytworzenia u pracowników instytucji publicznych poczucia alienacji. Kiedy władze podkreślają koszt każdej jednostki pracy, to często osiągają tylko tyle, że pracownicy widzą jedynie swój odcinek, co ogranicza ich inteligencję i umiejętności. (...) Jeśli jednak ich przełożeni koncentrują się wyłącznie na ich wydajności – na tym, jak szybko wykonują każdą jednostkę pracy – to zaczynają się oni czuć, jakby byli zatrudnieni przy taśmie montażowej”¹⁷. Na marginesie warto dodać, iż wyraźnie widać, że autorzy przestrzegając przed negatywnymi konsekwencjami pomiaru tylko wydajności pracy, wychodzą poza tradycję neotaylorizmu, a nawiązują do tradycji badań z kręgu tzw. *Organizational humanism* – wyznaczanej przez takie nazwiska, jak John Dewey, Elton Mayo, Chester Bernard

¹³ Osborne D., Gaebler T., *Rządzić inaczej. Jak duch przedsiębiorczości przenika i przekształca administrację publiczną*, Media Rodzina, Poznań 1992, s. 467.

¹⁴ Ibidem.

¹⁵ Ibidem, s. 468.

¹⁶ Ibidem, s. 469.

¹⁷ Ibidem.

czy Kurt Levin – zorientowanych na eksplorację relacji międzyludzkich wewnątrz organizacji, m.in. po to, aby testować związek między poziomem satysfakcji oraz stopniem zaspokojenia społecznych i emocjonalnych potrzeb członków organizacji a osiąganymi przez nią rezultatami¹⁸.

- Po trzecie, „istnieje też ważna różnica między «wynikami programu» a obejmującymi szerszy zakres «wynikami polityki»”¹⁹. Według Gaeblera i Osborna to, co jest ważne ze społecznego punktu widzenia, odnosi się do rezultatów szerzej rozumianej „polityki”, której poszczególne programy są tylko częścią. Wymowny jest przykład, jaki podają autorzy. Otóż instytucje realizujące programy szukania miejsc pracy dla klientów systemu pomocy społecznej często obliczały liczbę osób, dla których udało się ją znaleźć, oraz wysokość ich zarobków, jednakże nie uwzględniały liczby osób przechodzących na zasiłek dla bezrobotnych, długości okresu pobierania zasiłku itd. Jakże zatem były tego konsekwencje? „Osoby bezpośrednio nadzorujące program mogą odrzec, że są odpowiedzialne za liczbę znalezionych miejsc pracy, ale nie za liczbę osób przechodzących na zasiłek i będą miały rację. Na wielkość tej drugiej liczby wpływa bowiem ogólna polityka w tej dziedzinie, to znaczy zasady ustalające, kto kwalifikuje się do uzyskania zasiłku, w jakiej relacji pozostaje zasiłek do wynagrodzenia za wymagającą małych kwalifikacji i nisko płatną pracę i ile można znaleźć takich miejsc pracy. Jeśli jednak wydział opieki nie mierzy obu zbiorów liczb, to jego pracownikom może się wydawać, że wykonują znakomitą robotę, znajdując pracę dla swoich klientów i tym samym pozbywając się ich, gdy tymczasem szeregi podopiecznych w rzeczywistości rosną. Tak właśnie dzieje się od dziesięciu lat w wielu stanach. (...) Takie statystyki podkreślają wagę mierzenia zarówno wyników programu, jak też polityki”²⁰.

Gaebler i Osborne formułują też kilka praktycznych zaleceń dla administracji publicznej, która powinna być zainteresowana pozyskaniem wartościowych informacji pochodzących z badań ewaluacyjnych.

- Pierwsze zalecenie, które brzmi *przeprowadzaj zarówno analizę ilościową, jak i jakościową*, ma swoją bardzo obszerną literaturę i nie ma powodu aby cokolwiek w tej sprawie dodatkowo wyjaśniać²¹.
- Po drugie, *uważaj na amatorów łatwych kasków*. Chodzi o to, aby nie pozwolić zwieść się pozorom łatwych rozwiązań, które miałyby świadczyć o skuteczności takiego czy innego programu. Odwołując się do praktyki

¹⁸ Szerzej na ten temat m.in. w McKinney J.B., Howard L.C., *Public Administration. Balancing Power and Accountability*, Praeger Publishers, Westport 1998, s. 159–165.

¹⁹ Osborne D., Gaebler T., op. cit., s. 470.

²⁰ Ibidem, s. 471.

²¹ Jasiński M., Kowalski M., „Fałszywa sprzeczność: metodologia jakościowa czy ilościowa?”, w: A. Haber (red.), *Ewaluacja ex-post. Teoria i praktyka*, PARP, Warszawa 2007.

amerykańskiego sektora publicznego, autorzy piszą tak: „Usługodawcy zazwyczaj podają takie liczby, jakich się od nich oczekuje, nawet jeśli muszą w tym celu uprościć sobie pracę albo obejść przepisy. Jeśli muszą znaleźć pracę dla tysiąca osób rocznie, to wybiorą tysiąc osób, dla których najłatwiej znaleźć pracę i przeszkolą je. (...) Tak właśnie postępowały ośrodki pomocy społecznej w pierwszych latach realizacji ustawy o współudziale w kształceniu zawodowym”²².

- Po trzecie, *bądź przygotowany na silny opór*, co oznacza mniej więcej tyle, że rzetelne informacje o wydajności i skuteczności mogą stanowić duże zagrożenie dla bytu danej jednostki organizacyjnej. Autorzy podają kilka przykładów takiej sytuacji, np. Stanu Floryda, który zlikwidował program nagradzania szkół za poprawę jakości pracy czy Stanu Arizona, który zaprzestał publikowania danych dotyczących odsetka absolwentów szkół zawodowych i pomaturalnych, którzy uzyskali pracę. Powód był prozaiczny – „college i szkoły zawodowe uznały bowiem, że informacje takie stanowią dla nich zagrożenie, jako że ujawniają, jak naprawdę przygotowują młodych ludzi do podjęcia pracy”²³.
- Po czwarte, *włączaj w proces tworzenia metod pomiaru usługodawców i podwładnych*. To zalecenie odwołuje się do znanych z teorii organizacji i zarządzania modeli, jak chociażby przywoływany wcześniej *Organizational humanism*, zalecających kooperację dla wypracowania wspólnych rozwiązań i uniknięcia konfliktów wewnątrz-organizacyjnych. „Najlepszym – jak piszą autorzy – sposobem przełamania oporu przed wprowadzeniem metod mierzenia wykonania i jakości pracy jest włączenie usługodawców i wszystkich pracowników w proces ich tworzenia”²⁴.
- Po piąte, *poddawaj metody pomiaru corocznej rewizji i modyfikacji*.
- Po szóste, *nie stosuj zbyt dużo ani zbyt mało miar*. To iście arystotelesowskie zalecenie umiaru opiera się na przekonaniu, że jest on możliwy do wypracowania w ramach kooperacyjnej strategii tworzenia miar. „Jeśli (...) pracownicy i usługodawcy biorą czynny udział w tworzeniu metod pomiaru i jeśli pozwala się im na okresowe ich korygowanie, to zazwyczaj potrafią zachować umiar w tworzeniu miar”²⁵.
- Po siódme, *strzeż się antybodźców*. Chodzi o to, aby unikać rozwiązań, które prowadzą do nieefektywnych, z punktu widzenia „całościowej polityki”, rezultatów. „Metody pomiaru wykonania pracy przyjęte początkowo w ustawie o współudziale w kształceniu zawodowym zachęcały usługodawców do wyszukiwania łatwych kąsków. Takie antybodźce mogą obrócić w niwecz wszystkie wysiłki zmierzające do stworzenia miar wy-

²² Osborne D., Gaebler T., op. cit., s. 476.

²³ Ibidem, s. 477.

²⁴ Ibidem.

²⁵ Ibidem, s. 478.

konania pracy. Dla ich uniknięcia organizacje, przed wprowadzeniem nowych miar, powinny je przetestować, starając się znaleźć wszelkie możliwe sposoby robienia uników, które stosowałiby sprytni usługodawcy”²⁶.

- W kolejnym, ósmym zaleceniu, autorzy namawiają, aby „skupić się na maksymalnym wykorzystaniu danych o wykonaniu pracy. Samo stworzenie miar nie daje jeszcze gwarancji, że menedżerowie wykorzystają je do zmiany sposobu działania kierowanych przez siebie instytucji ani że ciała ustawodawcze skorzystają z nich, zmieniając kierunki i sposoby finansowania programów i instytucji”²⁷. Krótko mówiąc, nie należy zwlekać z wykorzystaniem dostępnym danych i stosować „politykę faktów dokonanych” w oparciu o zgromadzone informacje dotyczące efektywności i skuteczności działania.
- Po dziewiąte, „zlecaj mierzenie politycznie niezależnej, obiektywnej instytucji. Jeśli ludzie mają polegać na danych, to muszą wierzyć, że są one obiektywne. Dlatego dobrym rozwiązaniem jest korzystanie z usług niezależnej instytucji”²⁸. Realizacja tego zalecenia napotyka jednak na pewne trudności. Kłopot nie polega tylko na tym, czy „wiara” w obiektywność dostarczonych informacji, jak sugerują Gaebler i Osborne, czyni je rzeczywiście obiektywnymi. To fundamentalne dla nauki pytanie o „obiektywność” rzeczywistości i możliwość jej poznania możemy w tym miejscu uchylić, gdyż intencje autorów są bardzo klarowne. Chodziło im bowiem o to, aby pomiaru wydajności i skuteczności realizowanych programów dokonywały zewnętrzne w stosunku do wykonawców, równocześnie niezależne od nich, organizacje. Obiektywność oceny osiągniętych wyników należałoby zatem traktować jako funkcję niezależności instytucji oceniającej. I tu pojawia się problem z tzw. niezależnością instytucji oceniającej, która mogłaby dostarczyć administracji obiektywnych informacji. Można by oczywiście powiedzieć, zgodnie z sugestią autorów, że niezależność oznacza tyle, co brak personalnych i kapitałowych powiązań instytucji oceniającej z organizacją realizującą dany program. Ale taka redukcjonistyczna wykładnia „instytucjonalnej niezależności” jest zawodna w tej mierze, w jakiej nie uwzględnia złożoności rozmaitych punktów widzenia i interesów w sferze publicznej. Nie widać specjalnie dobrych uzasadnień, dla których tzw. niezależna instytucja (wszystko jedno czy prywatna czy publiczna), dokonująca ewaluacji programu, nie miałaby uwzględniać w swojej ocenie oczekiwań zleceniodawcy. Niewątpliwie uwzględnienie punktu widzenia czy interesów zleceniodawcy powinno mieć wpływ na ostateczny rezultat ewaluacji, nawet jeśli jej wyniki oparte byłyby na dobrze przygotowanych narzędziach badawczych. Ostateczny wynik

²⁶ Ibidem.

²⁷ Ibidem, s. 479.

²⁸ Ibidem, s. 478–479.

ewaluacji jest nie tylko, a może nawet nie przede wszystkim, rezultatem zastosowania określonych standardów badania, co ich ustalenia w fazie przedbadawczej.

Co więcej, można by dowodzić, że decyzje w sferze publicznej najczęściej kształtowane są nie wedle empirycznych ustaleń wolnych od osobistych wartości i poglądów decydentów, ale zgodnie z nimi. Wyniki rozmaitych badań ewaluacyjnych wykorzystywane w dyskusjach o polityce penalizacyjnej, edukacji seksualnej, aborcji itd. głównie służą jako instrument wzmacniający wyznawane przez strony poglądy czy wartości i zwykle kwestionuje się obiektywność wyników niezgodnych z wyrażanymi poglądami, niż zmienia poglądy. Dotychczas brakuje jakichkolwiek przesłanek (i empirycznych dowodów), by sądzić, że dobrze przeprowadzone przez niezależne instytucje badania ewaluacyjne, nad np. programami resocjalizacyjnymi groźnych przestępców, skłonią zwolenników kary śmierci do porzucenia swoich poglądów „penalizacyjnych” czy też odwrotnie – z zagorzałych zwolenników resocjalizacji uczynią równie gorliwych zwolenników kary śmierci. Krótko mówiąc, problem niezależności i obiektywności wyników badań ewaluacyjnych uwikłany jest w strukturę interesów i wartości wyznaczających dynamikę procesów decyzyjnych w sferze publicznej.

Ewaluacyjne nadzieje i rozczarowania

G. Majone, znany ze swojego sceptycyzmu wobec możliwości wyznaczania racjonalności procesu decyzyjnego w sferze publicznej wedle czysto empirycznych ustaleń badań ewaluacyjnych, podkreśla, że „profesjonalna ewaluacja stanowi zaledwie niewielką część ogólnego procesu krytyki i oceny polityk publicznych, w którym wszyscy politycznie aktywni członkowie demokratycznej wspólnoty biorą udział na różne, lecz równie użyteczne sposoby”²⁹. Dostrzegając znaczenie ewaluacji, jako rozległej i rozwijającej się dziedziny analizy polityki poświęconej „gromadzeniu, sprawdzaniu i interpretowaniu informacji na temat wdrażania i skuteczności istniejących polityk i programów publicznych”, Majone twierdzi, że dla efektywnego świadczenia usług publicznych nie wystarczy przygotowanie jakiegoś teoretycznie optymalnego programu działania. „Jeszcze ważniejsze jest poznanie, jak faktycznie przebiega wdrażanie programu, kto z niego korzysta, a kto na nim traci, czy program realizuje zamierzone cele, a jeśli tak się nie dzieje – to jak można go poprawić lub przerwać”³⁰. Przykładowo, ocena szczególnie popularnych w krajach UE

²⁹ Majone G., *Dowody, argumenty i perswazja w procesie politycznym*, Wyd. Naukowe Scholar, Warszawa 2004, s. 249.

³⁰ Ibidem, s. 248.

programów aktywizacji bezrobotnych nie zawiera zwykle informacji o tym, kto na nich zyskuje, a kto traci, gdyż milcząco zakłada się, że zwiększenie zatrudnienia leży w interesie wszystkich, a zatem uzasadnione jest, aby koszt takich szkoleń wpisać w koszty opodatkowania obywateli.

Można by argumentować, że przedstawiona kontrowersja jest raczej pozorna niż rzeczywista, gdyż ustalenie „wygranych” i „przegranych” w wyniku realizacji określonych programów czy polityk publicznych, stopnia realizacji zamierzonych celów czy sposobów wdrażania rozmaitych programów, należy do kanonu zadań ewaluacji. Trudno jednak nie zgodzić się z Majone, który twierdzi, że: „Wydaje się, że wielu ewaluatorów zakłada, iż są to czysto empiryczne ustalenia, które nie wiążą się ani z wyborami wartości, ani z osobistymi poglądami. W rzeczywistości wartości i poglądy bardzo się liczą w ewaluacji, nie tylko ze względu na niejednoznaczność wyników w praktyce – trudności z przypisaniem określonych przyczyn konkretnym rezultatom, z pomiarem wyników i oceną niezamierzonych skutków, z rozróżnieniem błędnych koncepcji i niepowodzeń we wdrażaniu – lecz jeszcze w większym stopniu wskutek nieuniknionych różnic poglądów na temat rodzaju kryteriów ewaluacyjnych, które są sensowne, uzasadnione lub politycznie dopuszczalne w danej sytuacji. Takich niejasności i różnic zdań nie można w pełni rozwiązać za pomocą ulepszonych metod pomiarowych i sprawdzających, lecz można je przedstawić i wyjaśnić poprzez dyskusję i wzajemną perswazję”³¹.

Argumenty Majone bardzo wyraźnie wskazują na ograniczone możliwości perswazyjne wyników badań ewaluacyjnych i na niewiele, jak sugeruje autor, zda się poszukiwanie coraz bardziej wyrafinowanych metod pomiaru. Kłopot polega na tym m.in., że „wyjściowy” rezultat ewaluacji zależy od tego, jakim materiałem „wejściowym” dysponują ewaluatorzy, a jego zebranie to nie tylko kwestia zastosowania takich czy innych technik badawczych, ale także pewnych uzgodnień punktów widzenia, wartości czy interesów między zainteresowanymi realizacją programu i jego ewaluacją stronami. Stosowanie określonych standardów badawczych jest kompletnie czym innym niż ustalanie tychże i aby zdawać sobie sprawę z możliwości perswazyjnych osiągniętych rezultatów ewaluacji, konieczne jest, jak twierdzi Majone, „rozróżnienie ustalania standardów i stosowania standardów”³².

Dyskusowanie określonych standardów ewaluacyjnych, np. uzgodnienie wśród zainteresowanych stron podstawowych kryteriów oceny programu ze wszelkimi tego konsekwencjami dla ostatecznych rezultatów ewaluacji może uchronić ewaluatorów przed rozmaitymi zarzutami, np. że nie mierzą tego, co powinni, albo, że nie uwzględnili istotnych aspektów ewaluowanego programu bądź też problemu, którego rozwiązanie tenże program miał przynieść. Takie zarzuty mogą być szczególnie często podnoszone, kiedy wyniki ewa-

³¹ Ibidem, s. 249.

³² Ibidem, s. 249.

luacji zagrażają mniej lub bardziej wyartykułowanym interesom grupowym zaangażowanych w realizację programu czy też dotkniętych skutkami jego implementacji³³. Powinni szczególnie dobrze pamiętać o tych zagrożeniach wszyscy ci, którzy realizują programy naruszające określony układ interesów, także wówczas, gdy dysponują empirycznymi dowodami, że utrzymywanie *status quo* jest niekorzystne dla wszystkich, a proponowana zmiana spełnia kryteria postępu wedle Pareto. Odwoływanie się jednak do argumentu, że „liczby nie kłamią”, może być równie irytujące i niezrozumiałe dla takiej czy innej grupy, jak podejmowane przed laty próby przekonywania „opornych” robotników przez żarliwych marksistów, że są oni nosicielami „fałszywej świadomości klasowej”, ale dla ich dobra trzeba to zmienić w ramach „zobiektywizowanych” procesów dziejowych.

Przyjęcie określonych kryteriów ewaluacyjnych, np. efektywności i skuteczności działania, w oczywisty sposób warunkuje ostateczny rezultat oceny programu. To, jakie kryteria oceny zostają przyjęte, to w dużej mierze konsekwencja ról społecznych, jakie odgrywają aktorzy zaangażowani w proces ewaluacji. Ten sam program inaczej będzie oceniany przez ekonomistę, dla którego najważniejsze jest poszukiwanie efektywności i skuteczności realizowanego programu, a inaczej przez prawnika, dla którego ważna jest np. ocena jego zgodności z obowiązującym prawem. Nie ma niczego złego w tym, że programy oceniane są z różnych punktów widzenia. Ważne jest, aby zachować odrębność różnych punktów widzenia, bowiem, „dopóki poglądy wyrażane z perspektywy jednej konkretnej roli nie są przedstawiane ani mylnie interpretowane jako poglądy odnoszące się do wszystkich możliwych ról lub wygłaszane w ich imieniu, mamy do czynienia ze zdrowym stanem wielokrotnej lub pluralistycznej ewaluacji”³⁴. Kłopot tylko w tym, że w praktyce ewaluacyjnej różne punkty widzenia nakładają się na siebie, a błędem jest zakładanie pewnej uniwersalności wniosków wypływających z ewaluacji w tym sensie, że powinny być one równie istotne dla aktorów odgrywających różne społeczne role. „Być może najczęstszy problem występuje, kiedy zakłada się, że wnioski z ewaluacji dokonanej na użytek określonej roli są równie istotne z perspektywy innych ról stosujących odmienne kryteria ewaluacyjne. Ponieważ role i kryteria do siebie nie pasują, jest niemal nieuniknione, że wnioski z ewaluacji okażą się nieprzekonujące”³⁵.

³³ „Kluczowy dylemat reformatorów można określić bardzo prosto: większość badań z zakresu public policy podkreśla fakt, że skuteczne reformy muszą uzyskać akceptację grup, których reformy najbardziej dotyczą. Jednakże wielu z tych, których reformy dotkną, są albo niechętni, albo wrogo nastawieni do reform i to nie jest zaskoczeniem, ponieważ ich miejsca pracy mogą być zagrożone; tylko najbardziej ekscentryczne indyki wyglądają świat”; zob. Wrigth V., „The Paradoxes of Administrative Reform”, w: W. Kickert (red.), *Public Management and Administrative Reform in Western Europe*, Edward Elgar, Cheltenham 1997, s. 9.

³⁴ Majone G., op. cit., s. 251.

³⁵ Ibidem.

Szczególny kłopot wydaje się występować wówczas, kiedy nacisk kładzie się na nieco „zmitologizowaną” efektywność i skuteczność osiągania zakładanych celów, przy czym prawie zawsze w takich przypadkach chodzi o tzw. racjonalizację publicznych wydatków, a w zasadzie o ograniczenie kosztów. Szczególny dlatego, że kryteriom efektywności i skuteczności najczęściej przypisuje się najważniejszą rolę w procesie oceny publicznych programów, często zapominając, że ich podnoszenie, zgodnie z ekonomicznym standardem racjonalności, może naruszać wiele żywotnych interesów czy podzielanых wartości głównych aktorów, których program dotyczy. Z drugiej zaś strony, nieuwzględnienie tychże interesów może sprawić, że dążność do racjonalizacji decyzji może prowadzić do skutków odwrotnych od zamierzonych i znacznego marnotrawstwa środków (*casus* obwodnicy Augustowa).

Zalecane w raportach ewaluacyjnych sposoby podniesienia efektywności i skuteczności mogą być też sprzeczne z panującą w organizacji, realizującą dany program, kulturą pracy czy obowiązującymi w niej standardami profesjonalizmu. Co więcej, jak wskazuje V. Wright, analizując paradoksalne konsekwencje podejmowanych reform administracji publicznej, ewaluacja może przynosić zgoła nieoczekiwane rezultaty z punktu widzenia podstawowego celu, jakim jest tzw. racjonalizacja wydatków publicznych. „Czwartym paradoksem jest to, że podczas gdy zamierzeniem wielu reform jest cięcie kosztów, ostateczny rezultat niektórych z nich może okazać się całkiem odmienny. Weźmy na przykład «ewaluację», to współczesne szaleństwo, która ignoruje koszty siły roboczej. Ewaluacja może identyfikować obszary marnotrawstwa środków, ale odkrywa także różnice nie tylko w wydajności (*performance*), lecz również osiągniętych ogólnych wyników (*outcomes*): pewni ludzie mogą być jednak relatywnie gorzej traktowani. Te różnice były wcześniej ukrywane. Rezultatem będzie wywieranie politycznej presji, aby takie różnice usuwać. Zgodnie z zamierzeniem, przekształcanie obywateli w klientów, z jasno określonymi przez cele prawami, może mieć wpływ na poprawę jakości realizowanych programów, lecz to także tworzy oczekiwanie i presję zadośćuczynienia (ujawnionych nierówności – przyp. D.Z.), często na drodze sądowej”³⁶. Jeśli uwaga Wrighta jest trafna, to nadzieje, że dzięki stałej ewaluacji programów publicznych czy działań administracji uda się zwiększyć ich ekonomiczną efektywność, wydają się być płonne w tej mierze, w jakiej nie uwzględniają one złożoności relacji międzyludzkich i powielają błędne przekonania o prostej relacji bodźca i efektu, zawarte w konwencjonalnych mechanistycznych teoriach organizacji i zarządzania.

G. Majone utrzymuje, że oprócz uwzględnienia ról, pozwalających na różnicowanie kryteriów ewaluacji, użyteczne jest też odniesienie się do trzech

³⁶ Wright V., „Paradoxes of Administrative Reform”, w: W.J.M. Kickert, *Public Management and Administrative Reform in Western Europe*, Edward Elgar, Cheltenham 1997, s. 10.

różnych trybów ewaluacji: trybu wynikowego, nakładowego i procesowego³⁷. Szczególnie atrakcyjny wydaje się być tryb wynikowy, bowiem ocena dokonywana na podstawie osiągniętych wyników najbardziej apeluje do wyobraźni decydentów. Należy przypomnieć, że wszelkie próby racjonalizowania wydatków publicznych i optymalizowania wyników podejmowanych działań w sferze publicznej opierają się na dążeniu do pomiaru rezultatów i ich kwantyfikowaniu. Nie bez powodu płaszczyzną odniesienia dla krytyków działalności administracji publicznej jest rynek oraz prywatne przedsiębiorstwo z przypisywaną im zdolnością do pomiaru osiągniętych wyników. Założenie jest logiczne, jak też wbudowane w metodologiczny kanon ewaluacji, a mianowicie, aby móc możliwie najlepiej ocenić realizowany program, trzeba dokonać pomiaru jego efektów. Odwołanie się do czysto ekonomicznych zasad oceny programów społecznej interwencji jest jednak kłopotliwie i to nie tylko dlatego, że logiki funkcjonowania rynku i państwa oraz funkcje, jakie każde z nich pełni, są odmienne. Kłopot polega m.in. na tym, że wyniki nie zawsze poddają się kwantyfikacji, cele nie zawsze są jednoznacznie określone, a narzędzia pomiaru często zawodzą. Jeszcze większą trudność jednak stanowi to, że związek między zarejestrowanym wynikiem a jego przyczyną często jest niepewny i bardziej polegamy na intuicji niż naukowych dowodach.

E. Babbie słusznie zauważa, że: „potencjalnie jednym z najbardziej uciążliwych aspektów badań ewaluacyjnych jest ustalenie, czy badany program odniósł sukces, czy poniósł porażkę. Celem programu nauczania języka obcego może być pomoc studentom w lepszym nauczaniu się tego języka, lecz jak duży postęp jest tu wystarczający? Celem programu odwiedzin małżonków w więzieniu może być poprawa nastrojów wśród więźniów, lecz jak bardzo mają być one poprawione, by móc uzasadnić ten program? Jak można się spodziewać, nie ma ostatecznych odpowiedzi na tego typu pytania. Dylemat ten w oczywisty sposób stał się podstawą analiz kosztów i korzyści (*cost – benefit analysis*). Niestety sytuacje, z jakimi zwykle ma się do czynienia podczas badań ewaluacyjnych, rzadko kiedy poddają się prostemu rachunkowi ekonomicznemu”³⁸.

Wydaje się jednak, że także oczekiwania, jakie ewaluatorzy mają wobec analizy kosztowo-korzyściowej, są nieco przesadzone w tym sensie, że jej wyniki bywają podważane jako nieobiektywne, gdyż sama analiza opiera się raczej na subiektywnych założeniach składających się na jej ekonomiczny model. Pamiętając o tym, że analiza kosztów i korzyści jest tylko jednym z narzędzi wykorzystywanych w badaniach ewaluacyjnych, musimy mieć

³⁷ W trybie wynikowym ewaluacja koncentruje się na rezultatach konkretnego działania. W trybie nakładowym ewaluacja kładzie nacisk na zasoby, umiejętności i ludzi, a w trybie procesowym przenosi się na metody stosowane do przekształcania nakładów w rezultaty; zob. Majone G., *Dowody, argumenty...*, op. cit., s. 253.

³⁸ Babbie E., *Badania społeczne w praktyce*, PWN, Warszawa 2004, s. 376.

świadomość ograniczonej wartości wyników. R. Formaini, w swojej książce o bardzo sugestywnym tytule *The Myth of Scientific Public Policy*, przekonuje, że nie ma czegoś takiego, jak „obiektywna” analiza kosztów i korzyści, a zatem wyniki, jakie dzięki niej uzyskujemy, nie są „obiektywne” w tym sensie, że nie dostarczają nam pewnej, możliwej do przyjęcia przez każdego, wiedzy potrzebnej do oceny realizowanego programu³⁹. Zarzuty, jakie stawia Formaini, mają charakter ontologiczno-epistemologiczny, a dotyczą błędnych, jego zdaniem, założeń neoklasycznej ekonomii co do natury porządku społecznego, na których oparta jest analiza kosztów i korzyści. „Analiza kosztów i korzyści i neoklasyczna ekonomia są symbiotycznie powiązane ze sobą przez fakt, że analiza kosztów i korzyści nie mogłaby być dokonana bez założeń dostarczonych przez neoklasyczne metody. Kiedy analitycy mierzą, definiują i agregują «koszty» i «korzyści», robią to w ramach neoklasycznej struktury poczynionych założeń. Jest zatem uprawnione, aby argumentować, że analiza kosztów i korzyści jest «dzieckiem», a neoklasyczna teoria ekonomiczna «ojcem». (...) Koniecznym zatem jest, aby zrozumieć analizę kosztów i korzyści, trzeba najpierw poznać podstawowe założenia neoklasycznej ekonomii”⁴⁰.

Dyskusja o ontologicznych podstawach modeli analizy, przeciwstawianiu modelowi „człowieka ekonomicznego” modelu „człowieka socjologicznego”, i uzyskiwanych w ich ramach wynikach stanowi kanon naukowych dociekań. Nie ma powodu, aby w tym miejscu tę dyskusję rekonstruować. Wystarczy przypomnieć tylko, że to, co dla jednych uchodzi za zobiektywizowany w ramach przyjętych teoretycznych założeń rezultat, w oparciu o który można podejmować racjonalne decyzje, dla innych jest niczym więcej niż tylko projekcją pewnych zakładanych motywów czy postaw, które nie mogą być podstawą do czynienia uogólnionych wyjaśnień w sferze teorii czy podejmowanych działań w sferze praktyki. Zarzut, jaki czyni Formaini tzw. unaukowionej *Public Policy*, jest niczym innym jak próbą obnażenia niedostatków ekonomicznego modelu „człowieka ekonomicznego” czy „kalkulującego” (*homo economicus*). Można by go uznać za zasadny także dla badań ewaluacyjnych, których orędownicy często przekonują, że o ile przeprowadzone są one wedle starannie przygotowanej metodologii, to ich wyniki stanowią podstawę do podejmowania racjonalnych decyzji. Konkluzja Formainiego brzmi, że „oparta na naukowych podstawach polityka publiczna (*Public Policy*), to marzenie, które wzrastało coraz bardziej od czasów Oświecenia i być może osiągnęło swoje apogeum pod koniec tego wieku, jest mitem, teoretyczną iluzją. Ona istnieje w naszych umysłach, analizach i metodach tylko dlatego, że szukamy by ją odkryć, a zmierzamy do odkrycia tego, czego szukamy”⁴¹.

³⁹ Formaini R., *The Myth of Scientific Public Policy*, Transaction Publishers, 1990, s. 39–65.

⁴⁰ Ibidem, s. 39.

⁴¹ Ibidem, s. 1.

Zakończenie

Wyciągnięcie wniosku, że wszelkie ograniczenia, na jakie narażona jest ewaluacja, czyni jej wyniki mało przydatnymi w procesach podejmowania decyzji, w szczególności dotyczących racjonalizowania wydatków publicznych, wydaje się być tak samo nieuprawnione, jak twierdzenie, że wyniki starannie przeprowadzonej ewaluacji z zachowaniem wszelkich prawideł sztuki są pewnym oparciem dla racjonalizowania decyzji w sferze publicznej. Powiedzenie jednak, że „prawda leży gdzieś pośrodku” jest obarczone takim samym prawdopodobieństwem błędu, jak wskazanie na którąś ze skrajnych opinii. Ponadto spór nie powinien przebiegać wedle logiki „kto ma rację” i wedle jakich argumentów tę rację jesteśmy skłonni przyznać jednej ze stron sporu. Raczej powinniśmy poszukiwać konsensusu tak, aby wyniki badań ewaluacji, nawet jeśli zgodzimy się co do ich rozmaitych niedoskonałości, stanowiły określoną wartość dla administracji publicznej odpowiedzialnej za podejmowanie takich czy innych decyzji w sektorze publicznym.

Dla „decyzjonistów”, by posłużyć się terminem używanym przez G. Majone, problem w zasadzie nie istnieje, gdyż racjonalne decyzje powinny być podejmowane w oparciu o informacje, jakie przynosi starannie przeprowadzona analiza kosztów i korzyści. Ponieważ jednak oponenci wskazują na wiele niedoskonałości takiego podejścia, głównie o charakterze ontologiczno-epistemologicznym, to pojawia się pytanie, co zrobić z ogromną ilością „wyprodukowanych” informacji w ramach prowadzonych ewaluacji programów interwencji społecznych. Gdyby bez jakichkolwiek zastrzeżeń uznać rację „antydecyzjonistów” kwestionujących przydatność analiz ewaluacyjnych dla procesu racjonalizowania decyzji publicznych, to musielibyśmy w jakiejś mierze przyznać, że dążność do alokacyjnej racjonalności w oparciu o ewaluację wyników przeradza się w swoje przeciwieństwo w tym sensie, że chcąc racjonalnie wydawać publiczne pieniądze, nie uzyskujemy żadnej pewnej wiedzy, za którą musimy jeszcze zapłacić.

Propozycja krytyków badań ewaluacyjnych, prowadzonych wedle założeń modelu ekonomicznego, jest taka, aby włączyć wysiłki ewaluatorów w szerszą demokratyczną debatę poświęconą ustalaniu kryteriów i standardów realizacji programów publicznych i ich oceny. Jedną ze stron takiej debaty powinna być administracja publiczna, która nie bezkrytycznie wykorzystywałaby dostarczane jej informacje pochodzące z przedsięwzięć ewaluacyjnych. Innymi słowy, wartość ewaluacji nie polegałaby na tym, że można by na jej podstawie podjąć racjonalną decyzję maksymalizującą korzyści przy określonych kosztach czy też minimalizującą koszty przy określonych korzyściach, ale na tym, że stanowiłaby ona wartościowy wkład do dyskusji między stronami reprezentującymi różne punkty widzenia czy interesy. „Obecnie – pisze Majone – konieczne jest nie tyle opracowanie «obiektywnych» wskaźników wyników

– tradycyjny cel analiz ewaluacyjnych – ile ułatwienie szerokiego dialogu wśród zwolenników różnych kryteriów⁴².

Gdyby udało się wyniki ewaluacji wykorzystać w szerszej debacie o zasadności prowadzenia interwencji społecznych, ich kosztach i korzyściach oraz kryteriach, wedle których oceniane byłyby rezultaty interwencji, to z pewnością moglibyśmy powiedzieć, że ewaluacja może wносить swój wkład w demokratyczny proces uzgadniania rozmaitych racji czy punktów widzenia. We współczesnych demokracjach, cierpiących na deficyt legitymizacji, taka funkcja ewaluacji wydaje się być bardzo atrakcyjna przynajmniej z dwóch powodów. Pierwszy jest taki, że prowokując debatę, w jakiejś mierze zwiększa się legitymizację procesu decyzyjnego w administracji publicznej, drugi zaś taki, że debata ta, oparta jednak na rezultatach uzyskanych zgodnie z naukowymi standardami badań empirycznych, wnosi określony potencjał wiedzy, który może przyczynić się do bardziej wysublimowanego, jak mówi Majone, postrzegania procesu kształtowania polityki publicznej. Z pewnością nie oznacza to, że wyniki badań ewaluacyjnych, z których może i powinna korzystać administracja publiczna, uwolnią nas od rozmaitych manipulacji zasobem informacyjnym, gdyż sama nauka nie jest wolna od tego typu pokus, ale to jest temat na zupełnie inną dyskusję.

Ważne jednak jest, aby możliwości i ograniczenia wykorzystania wyników badań ewaluacyjnych były szeroko dyskutowane wśród przedstawicieli administracji publicznej, gdyż to od niej zależy jakość realizowanych programów społecznej interwencji. Gdyby miało się okazać, że decydenci na poziomie systemu administracyjnego przeceniają możliwości ewaluacji, to wyniki, które dzięki niej uzyskujemy, niekoniecznie musiałyby odzwierciedlać rzeczywistą wartość realizowanych programów. Ta ostatnia, jak mówią „antydecyzyjniści”, nie zawsze jest możliwa do policzenia. Sceptycyzm wobec ewaluacji nie powinien jednak przesłaniać jej możliwości, gdyż wiele programów społecznej interwencji z powodzeniem można poddać procesom ewaluacji, dostarczając tym samym wielu cennych informacji o rozmaitych aspektach realizowanych programów. Dyskusja wokół ewaluacji powinna toczyć się zatem zarówno wśród przedstawicieli środowiska akademickiego w ramach międzydyscyplinarnej czy paradygmatycznej wymiany opinii, jak i między nimi a przedstawicielami administracji. Oni z racji swojego doświadczenia mogą znacznie więcej wnieść do procesów doskonalenia badań ewaluacyjnych, niż to wydaje się z perspektywy akademickiej katedry. Cóż to oznacza dla procesu podejmowania decyzji przez administrację? Z pewnością ewaluacja nie jest w stanie zastąpić zdrowego rozsądku i doświadczenia, które generowane jest w toku codziennej praktyki administracyjnej, a do którego często odwołują się inkrementaliści. Musimy jednak pamiętać, że postępowanie wedle określonych schematów grozi petryfikacją struktur decyzyjnych, a wyniki ewaluacji mogą

⁴² Majone G., op. cit., s. 264.

nas przed takim niebezpieczeństwem uchronić. Arystotelesowska pochwała umiaru wydaje się być rozsądną propozycją, z której powinna korzystać także administracja publiczna.

Dariusz Zalewski – doktor, adiunkt w Instytucie Socjologii WFiS Uniwersytetu Warszawskiego (Zakład Socjologii Problemów Społecznych i Planowania Społecznego) oraz w Instytucie Pracy i Spraw Socjalnych (Zakład Zbiorowych Stosunków Pracy). Od lat zajmuje się problematyką instytucji państwa opiekuńczego, administracji publicznej i dialogu społecznego. Jest autorem i współautorem m.in. takich publikacji, jak: *Opieka i pomoc społeczna. Dynamika instytucji, Utopie inkluzji. Sukcesy i porażki programów reintegracji społecznej, Dialog społeczny na poziomie regionalnym. Ocena szans rozwoju*. Wielokrotnie uczestniczył w projektach badawczych poświęconych m.in.: formom reprezentacji interesów grupowych, sposobom realizacji zadań przez administrację publiczną, dialogowi społecznemu na poziomie centralnym i regionalnym. W ramach swoich zawodowych zainteresowań brał udział także w badaniach ewaluacyjnych, np. w ewaluacji programu wymiany młodzieży „Leonardo da Vinci” współfinansowanego przez Unię Europejską.

Bibliografia

- Evaluating UE Expenditure Programmes: A Guide*, European Commission, 01.1997.
Babbie E., *Badania społeczne w praktyce*, PWN, Warszawa 2004.
Dunleavy P., *Democracy, Bureaucracy and Public Choice*, Harvester Wheatsheaf, 1991.
Formaini R., *The Myth of Scientific Public Policy*, Transaction Publishers, 1990.
Friedman M., Friedman R., *Wolny wybór*, Wyd. Panta, Sosnowiec 1994.
Frieske K.W., *Socjologia w działaniu. Nadzieje i rozczarowania*, Wyd. Uniwersytetu Warszawskiego, Warszawa 1990.
Haber A. (red.), *Ewaluacja ex-post. Teoria i praktyka*, PARP, Warszawa 2007.
Kickert W.J.M., *Public Management and Administrative Reform in Western Europe*, Edward Elgar, Cheltenham 1997.
Lane J.E. (red.), *Public Sector Reform. Rationale, Trends and Problems*, Sage Publications, London 1997.
Majone G., *Dowody, argumenty i perswazja w procesie politycznym*, Wyd. Naukowe Scholar, Warszawa 2004.
Majone G., *The rise of the regulatory state in Europe*, West European Politics, 17.
McKinney J.B., Howard L.C., *Public Administration. Balancing Power and Accountability*, Praeger Publishers, Westport 1998.
Osborne D., Gaebler T., *Rządzić inaczej. Jak duch przedsiębiorczości przenika i przekształca administrację publiczną*, Media Rodzina, Poznań 1992.
Verheijen T., Coombes D., *Innovations in Public Management. Perspectives from East and West Europe*, Edward Elgar, Cheltenham 1998.