Teoria zmiany jako architektura sprawczości: od mechanizmów przyczynowych do rozwiązywania węzłów systemowych w ujęciu Setha Godina

AUTOR

Fundacja Dobre Państwo

STRESZCZENIE / ABSTRACT

Artykuł analizuje teorię zmiany nie jako prosty schemat projektowy, lecz jako zaawansowaną architekturę sprawczości. Autor podkreśla kluczową różnicę między wykonaniem zaplanowanych działań a rzeczywistym wywarciem wpływu na system. Skuteczność interwencji zależy od zrozumienia mechanizmów przyczynowych i uniknięcia błędów w przypisywaniu rezultatów samym działaniom, bez uwzględnienia czynników zewnętrznych. Tekst wskazuje na konieczność oddzielenia intencji normatywnych (wizji stanu pożądanego) od empirycznej wiedzy o świecie społecznym. Teoria zmiany staje się zatem modelem pośredniczącym, który pozwala weryfikować hipotezy dotyczące transformacji systemowej i uniknąć pułapki pozornej efektywności operacyjnej, która nie rozwiązuje pierwotnych przyczyn problemu.

The article analyzes the Theory of Change not as a simple design scheme, but as an advanced architecture of agency. The author emphasizes the key difference between executing planned actions and exerting actual influence on a system. The effectiveness of an intervention depends on understanding causal mechanisms and avoiding errors in attributing results to actions alone without considering external factors. The text points to the necessity of separating normative intentions (the vision of a desired state) from empirical knowledge of the social world. Theory of Change thus becomes an intermediary model that allows for the verification of hypotheses regarding systemic transformation and avoids the trap of apparent operational efficiency, which fails to address the root causes of the problem.

Artykuł stanowi krytyczną analizę sprawczości w rozwiązywaniu problemów, stawiając tezę, że rzeczywista skuteczność nie wynika z samej realizacji planu czy zaobserwowania korzystnej zmiany, lecz z udowodnienia bezpośredniego związku przyczynowego między interwencją a rezultatem. Autor argumentuje, że współczesne organizacje często wpadają w pułapkę utożsamiania intencji normatywnych z mechanizmami sprawczymi, co prowadzi do błędnych wniosków o sukcesie projektów. Tekst proponuje przejście od powierzchownych modeli logicznych ku rygorystycznej Teorii Zmiany (Theory of Change), która mapuje konkretne mechanizmy społeczne i warunki systemowe. Poprzez syntezę metodologii kontrfaktycznej, analizy wkładu oraz ewaluacji realistycznej, autor pokazuje, jak odróżnić faktyczny wpływ od przypadkowych korelacji czy efektów zewnętrznych. W efekcie sprawczość zostaje zdefiniowana nie jako umiejętność wdrażania działań, lecz jako epistemiczna odpowiedzialność za zrozumienie architektury przyczynowej systemu, co pozwala na przejście od doraźnego łagodzenia skutków do trwałego przekształcania struktur generujących problemy.

SŁOWA KLUCZOWE

teoria zmiany | architektura sprawczości | mechanizmy przyczynowe | węzły systemowe | ewaluacja kontrfaktyczna | warunki INUS | interwencje złożone | model logiczny | sprawczość systemowa | potencjalne rezultaty | refleksyjność systemu społecznego | analiza instytucjonalna | metodologia ewaluacji | zależność od ścieżki

Skuteczność to dowód na przyczynowość zmiany

Teoria zmiany jako nauka o sprawczości – przyczynowość, ewaluacja kontrfaktyczna, mechanizmy społeczne i granice dowodzenia rzeczywistego wpływu

Zdolność do skutecznego działania nie może być utożsamiana z samą umiejętnością uruchamiania przedsięwzięć, wykonywania zaplanowanych czynności czy osiągania bezpośrednich rezultatów. Zasadniczym kryterium rozwiązywania problemów pozostaje bowiem uzasadniona możliwość przypisania konkretnej interwencji rzeczywistej zmiany stanu rzeczy, której przezwyciężenie stanowiło pierwotny cel wysiłku.

Współczesne organizacje dysponują rozwiniętymi narzędziami planowania, zarządzania projektowego, statystycznego monitorowania działalności i cyfrowego przetwarzania informacji. Żadna z tych kompetencji nie gwarantuje jednak, że przedsięwzięcia realizowane zgodnie z harmonogramem doprowadzą do społecznie pożądanych konsekwencji. Można poprawnie wykonać tysiące czynności, rozdysponować całość przewidzianych środków i uzyskać korzystne wskaźniki operacyjne, a jednocześnie nie ograniczyć mechanizmu odpowiedzialnego za istnienie problemu.

Możliwe jest również doprowadzenie do korzystnej zmiany, której znaczną część wyjaśniają czynniki niezależne od podjętej interwencji. Szczególnej doniosłości nabiera zatem konieczność rozdzielenia trzech rodzajów twierdzeń: że określone działanie zostało wykonane; że po jego wykonaniu zaobserwowano zmianę; oraz że zmiana nastąpiła właśnie dzięki temu działaniu. Każde kolejne stwierdzenie wymaga silniejszego uzasadnienia przyczynowego, a przechodzenie między nimi bez odpowiednich podstaw metodologicznych jest jednym z głównych źródeł błędów w ocenie skuteczności polityk publicznych, programów społecznych, reform organizacyjnych i przedsięwzięć gospodarczych.

Teoria zmiany (Theory of Change) nie powinna być zatem rozumiana wyłącznie jako atrakcyjny schemat przedstawiający kolejne etapy projektu czy deklaracja przekonania, że podjęcie określonego działania przyniesie korzystny rezultat. Jej zasadniczą funkcją jest ujawnienie założeń dotyczących tego, w jaki sposób konkretne zasoby, czynności, mechanizmy społeczne i warunki środowiskowe mają doprowadzić do osiągnięcia zamierzonego celu. Autor książki kładzie szczególny nacisk na wymóg przedstawienia teorii zmiany jako świadomego zobowiązania do wskazania, dlaczego określone przedsięwzięcie miałoby rzeczywiście przekształcić istniejący system. W ujęciu inspirowanym Sethem Godinem nie wystarcza sformułowanie wartościowej intencji – konieczne jest przedstawienie drogi łączącej rozpoznany problem z możliwym do zaobserwowania rezultatem.

Należy jednak odróżnić tę projektową intuicję od szczegółowych metod współczesnej ewaluacji, których naukowe podstawy rozwijały wiele niezależnych tradycji badawczych. Teoria zmiany stanowi miejsce spotkania praktycznej refleksji nad sprawczością z metodologią nauk społecznych, ekonomią dobrobytu, analizą instytucjonalną i filozofią wyjaśniania przyczynowego. Podstawowym wymaganiem pozostaje tu rozdzielenie logicznej spójności teorii od empirycznej prawdziwości jej założeń. Można bowiem skonstruować model, w którym wszystkie etapy są uporządkowane, a ich następstwo wydaje się intuicyjnie przekonujące, choć rzeczywiste zależności przyczynowe są słabe, niestabilne lub nie występują w warunkach planowanego wdrożenia.

Program edukacyjny może zakładać, że zwiększenie liczby godzin szkoleniowych prowadzi do wzrostu kompetencji, poprawy sytuacji zawodowej i trwałego wzrostu dochodów uczestników.

Każda z tych relacji posiada pewne uzasadnienie teoretyczne, lecz żadna nie jest bezwarunkowa. Nabycie kompetencji zależy od jakości nauczania, wcześniejszej wiedzy i warunków uczenia się; ich wykorzystanie – od rzeczywistego zapotrzebowania na rynku pracy; wzrost wynagrodzeń zaś – od struktury zatrudnienia, produktywności i siły negocjacyjnej pracowników.

Rozdzielenie intencji normatywnych od mechanizmów przyczynowych

W rezultacie model może być poprawny jako zbiór hipotez, lecz nie musi trafnie przewidywać konsekwencji konkretnego programu. Właściwa metodologia wymaga zatem przedstawiania poszczególnych ogniw jako zależności o określonym zakresie potwierdzenia empirycznego, konkretnych warunkach obowiązywania i możliwych przyczynach niepowodzenia. Z perspektywy epistemologicznej teoria zmiany stanowi model pośredniczący między normatywną wizją stanu pożądanego a empiryczną wiedzą o rzeczywistych mechanizmach świata społecznego.

Wizja może zakładać, że instytucja powinna zwiększać dostępność sprawiedliwej ochrony prawnej, ograniczać wykluczenie ekonomiczne albo poprawiać jakość edukacji, jednak sama wartość celu nie mówi nam, jakie działania będą skuteczne. Wiedza empiryczna z kolei ujawnia zależności między konkretnymi czynnikami, nie rozstrzygając automatycznie, które rezultaty powinny mieć pierwszeństwo w hierarchii wartości. Teoria zmiany wymaga więc połączenia dwóch rodzajów rozumowania: normatywnego, który uzasadnia kierunek interwencji, oraz przyczynowego, pozwalającego ocenić, czy proponowane działania rzeczywiście mogą doprowadzić do pożądanego rezultatu. Ich pomieszanie prowadzi do charakterystycznego błędu, w którym moralna doniosłość problemu zostaje potraktowana jako wystarczający dowód skuteczności proponowanego rozwiązania. Tymczasem nawet przedsięwzięcie realizujące szczególnie ważne wartości wymaga niezależnej oceny zastosowanego mechanizmu; niepowodzenie konkretnej interwencji nie podważa bowiem automatycznie słuszności samego celu. Historyczne podstawy tak rozumianej ewaluacji odnajdziemy w pracach Carol Hirschon Weiss, jednej z najważniejszych badaczek rozwijających podejście oparte na teorii programu.

W jej ujęciu ocena interwencji społecznej nie powinna ograniczać się do ustalenia, czy po wykonaniu określonych czynności zaobserwowano korzystną zmianę. Konieczne jest również odtworzenie przyjętych przez projektantów założeń dotyczących mechanizmów, przez które program miał wywołać rezultat, oraz sprawdzenie, czy mechanizmy te rzeczywiście zostały uruchomione. Kluczowe jest tu odróżnienie sytuacji, w której program pozostaje nieskuteczny z powodu błędnej teorii działania, od przypadku, gdy właściwa koncepcja nie została odpowiednio

zrealizowana. Szkolenie może nie poprawiać sytuacji zawodowej uczestników dlatego, że nie przekazuje kompetencji potrzebnych na rynku pracy, ale może też posiadać prawidłowy program merytoryczny, którego po prostu nie wykonano z wymaganą starannością. W pierwszym przypadku konieczne staje się zakwestionowanie przyjętego mechanizmu, w drugim – analiza jakości wdrożenia. Pomieszanie tych interpretacji grozi porzuceniem rozwiązania wartościowego albo kontynuowaniem przedsięwzięcia, którego podstawowe założenia pozostają nieuzasadnione.

Znaczący wkład do metodologicznego uporządkowania tego problemu wniósł również Huey-Tsyh Chen, który rozwijał ewaluację sterowaną teorią w książce Theory-Driven Evaluations z 1990 roku. Chen wskazywał na potrzebę rozróżnienia normatywnych i przyczynowych aspektów teorii programu. Pierwsze odnoszą się do pożądanego sposobu organizacji interwencji, określając warunki, zasoby i procedury uznane za właściwe. Drugie dotyczą mechanizmów wyjaśniających, dlaczego konkretne działania powinny prowadzić do oczekiwanych rezultatów. Różnica ta jest istotna, gdyż dobrze opracowany standard postępowania nie stanowi samodzielnego dowodu istnienia zakładanej relacji przyczynowej.

Można bowiem poprawnie opisać przebieg szkolenia, nie wykazując jednocześnie, że zdobywane w jego trakcie umiejętności znajdują zastosowanie w rzeczywistej pracy. Można wskazać czynności, które powinny zostać wykonane przez administrację, nie wyjaśniając dostatecznie, dlaczego ich realizacja miałaby zmienić zachowanie obywateli. W takim układzie teoria normatywna organizuje działanie, natomiast teoria przyczynowa umożliwia ocenę jego rzeczywistego znaczenia. Rozróżnienie to pozwala wyjaśnić ograniczenia modeli logicznych, często wykorzystywanych w administracji i sektorze grantowym. Model logiczny przedstawia zazwyczaj relację między nakładami, czynnościami, produktami, rezultatami pośrednimi a długofalowym oddziaływaniem, lecz sam fakt uporządkowania tych elementów nie dowodzi, że kolejne przejścia są przyczynowo uzasadnione. Strzałka łącząca szkolenie z poprawą sytuacji zawodowej uczestnika nie określa jeszcze, czy poprawa wynika ze wzrostu kompetencji, uzyskania uznawanego przez pracodawców certyfikatu, zwiększenia liczby kontaktów zawodowych czy jedynie z selekcji osób wcześniej posiadających wysoką motywację.

Podobnie związek pomiędzy publikacją raportu eksperckiego a zmianą regulacji prawnej może obejmować wiele mechanizmów pośrednich: zainteresowanie decydentów, wykorzystanie argumentów przez określone instytucje, mobilizację organizacji społecznych czy powstanie politycznych warunków sprzyjających reformie.

Rozróżnienie między skomplikowaniem a złożonością interwencji

Model logiczny pozostaje przydatnym narzędziem porządkowania przedsięwzięcia, jednak jego wartość naukowa zależy od tego, jak rzetelnie wyjaśniono relacje między poszczególnymi elementami. Ważne metodologiczne rozwinięcie tego problemu zaproponowała Patricia J. Rogers w artykule Using Programme Theory to Evaluate Complicated and Complex Aspects of Interventions z 2008 roku. Autorka odróżnia w nim interwencje skomplikowane, obejmujące liczne komponenty i wymagające rozbudowanej koordynacji, od interwencji złożonych, których rezultaty mogą wynikać z wyłaniających się interakcji, adaptacji uczestników i zmieniających się warunków otoczenia. Rozróżnienie to pozwala zrozumieć, dlaczego szczegółowość harmonogramu nie gwarantuje zdolności przewidywania konsekwencji przedsięwzięcia. Program może składać się z wielu czynności, a jednocześnie funkcjonować w środowisku na tyle stabilnym, że umożliwia ono odpowiednie modelowanie procesów. Z kolei inna interwencja może obejmować stosunkowo niewielką liczbę działań, lecz wywoływać złożone reakcje społeczne, których charakter zależy od interpretacji uczestników i wzajemnego oddziaływania instytucji. Rogers wskazywała na możliwość wykorzystania teorii programu do opisu obu rodzajów złożoności, pod warunkiem dopasowania metod analitycznych do charakteru badanego mechanizmu. Jest to kluczowe w przypadku oceny programów społecznych, których rezultaty zależą od wzajemnego oddziaływania jednostek i instytucji.

Reforma edukacji może obejmować zmianę programu nauczania, rozwój kompetencji nauczycieli, wprowadzenie nowych narzędzi dydaktycznych oraz modyfikację sposobu oceniania, co czyni ją przedsięwzięciem skomplikowanym organizacyjnie. Jednocześnie jej rzeczywiste konsekwencje mogą zależeć od zmian oczekiwań rodziców, wzajemnego uczenia się nauczycieli, relacji między uczniami i sposobu reagowania szkół na nowe kryteria oceny.

Podobnie program ograniczania wykluczenia ekonomicznego może obejmować świadczenia finansowe, wsparcie zawodowe, doradztwo i pomoc mieszkaniową, lecz jego skuteczność często kształtują czynniki zewnętrzne: lokalny rynek pracy, dostępność transportu, relacje rodzinne oraz zmiany zasad przyznawania innych świadczeń.

Teoria zmiany jako mapa mechanizmów i warunków systemowych

W takich przypadkach prosta sekwencja działań może okazać się niewystarczająca, gdyż interwencja zmienia nie tylko wybrane parametry systemu, lecz również zachowania uczestników, co wpływa na warunki jej dalszego funkcjonowania. Teoria zmiany powinna zatem obejmować zarówno zakładane rezultaty bezpośrednie, jak i możliwe sprzężenia zwrotne oraz reakcje adaptacyjne. Podobne stanowisko rozwinięto w zaktualizowanych w 2021 roku wytycznych brytyjskiego Medical Research Council dotyczących projektowania i oceny interwencji złożonych, przedstawionych przez Kathryn Skivington i współautorów na łamach The BMJ.

Dokument ten podkreśla znaczenie badań wykraczających poza samo ustalenie średniej skuteczności interwencji. Dla odpowiedzialnego podejmowania decyzji kluczowe pozostają bowiem jej wykonalność, koszty, warunki wdrożenia, możliwość zwiększania skali oraz przenoszenia rezultatów między różnymi środowiskami. Autorzy wytycznych wskazują cztery perspektywy badawcze: skuteczność w warunkach kontrolowanych, skuteczność praktyczną, wyjaśnianie mechanizmów i analizę systemową. Kluczowe jest uznanie, że odpowiedź na pytanie o skuteczność przeciętną nie wyczerpuje wiedzy niezbędnej do odpowiedzialnego zaprojektowania programu. Interwencja może bowiem przynosić korzystne rezultaty w warunkach eksperymentalnych, a jednocześnie pozostawać trudna do wdrożenia na większą skalę lub nadmiernie kosztowna w stosunku do dostępnych alternatyw.

Wytyczne MRC zwracają również uwagę na znaczenie niepewności badawczej, której charakter powinien decydować o doborze odpowiednich metod. Jeśli zasadnicza trudność polega na braku wiedzy o potencjalnej skuteczności określonego działania, uzasadnione jest skonstruowanie badania pozwalającego oszacować jego efekt. Jeżeli skuteczność jest stosunkowo dobrze potwierdzona, lecz problemem pozostaje niska zdolność organizacji do prawidłowego wdrożenia, większą wartość ma badanie procesów implementacyjnych i lokalnych ograniczeń. Gdy interwencja przynosi odmienne rezultaty w różnych środowiskach, konieczna staje się analiza mechanizmów oraz warunków ich uruchamiania. W przypadku działań przekształcających całe systemy szczególnego znaczenia nabierają relacje między uczestnikami, sprzężenia zwrotne i konsekwencje niezamierzone.

Zatem wybór metody badania zostaje podporządkowany rzeczywistemu problemowi poznawczemu, zamiast przyjmowania, że każda interwencja powinna być oceniana identycznym narzędziem. Autorzy podkreślają potrzebę uwzględniania interesariuszy, kontekstu, teorii programu, zasadniczych niepewności, możliwości doskonalenia interwencji oraz ekonomicznych konsekwencji

jej zastosowania. Teoria zmiany stanowi więc znacznie więcej niż formalną deklarację planowanych rezultatów; wymaga wyodrębnienia mechanizmów, które można poddawać niezależnej analizie. W przypadku programu edukacyjnego pierwszym ogniwem może być rzeczywista dostępność szkolenia, drugim udział odbiorcy, trzecim nabycie określonej kompetencji, czwartym możliwość jej wykorzystania, a piątym osiągnięcie rezultatu zawodowego. Każdy z tych etapów może jednak cechować się odmiennym poziomem niepewności.

Dostępność szkolenia można relatywnie łatwo potwierdzić analizą organizacji zajęć, lecz rzeczywiste nabycie kompetencji wymaga odpowiednich narzędzi pomiarowych. Wykorzystanie umiejętności może zależeć od możliwości zatrudnienia i organizacji pracy, a trwałość poprawy sytuacji zawodowej od koniunktury i struktury rynku. Możliwe jest bowiem, że program podniesie określone kompetencje, nie prowadząc do proporcjonalnego wzrostu zatrudnienia. Taki wynik nie dowodzi automatycznie całkowitej nieskuteczności przedsięwzięcia; może raczej wskazywać, że mechanizm uczenia się zadziałał, podczas gdy nie zostały spełnione warunki konieczne do przełożenia nabytej wiedzy na rezultat ekonomiczny. Właściwa ewaluacja powinna wówczas odróżniać niepowodzenie całej interwencji od niepowodzenia jej określonego ogniwa. Jeśli uczestnicy nie nabyli zakładanych kompetencji, należy zbadać metody dydaktyczne, czas zajęć, jakość pracy prowadzących i adekwatność materiałów. Jeśli kompetencje wzrosły, lecz nie zostały wykorzystane, znaczenia nabiera analiza barier rynku pracy, dostępności ofert i sposobu potwierdzania kwalifikacji. Jeśli uczestnicy podjęli pracę, ale nie uzyskali trwałej poprawy dochodowej, konieczne może okazać się rozpoznanie jakości zatrudnienia, liczby przepracowanych godzin i stabilności relacji kontraktowych.

Rozróżnienie warunków koniecznych, wystarczających i INUS

Każda z tych interpretacji prowadzi do odmiennych decyzji projektowych, mimo że wszystkie mogą zostać objęte wspólnym wskaźnikiem końcowym. Właśnie dlatego teoria zmiany pozwala podnieść jakość decyzji nie tylko w przypadku sukcesu programu, ale również wtedy, gdy wyniki odbiegają od oczekiwań. Warunkiem jest jednak wcześniejsze określenie mechanizmów, których funkcjonowanie można rzeczywiście zbadać. Szczególnego znaczenia nabiera tutaj rozróżnienie warunków koniecznych, wystarczających i sprzyjających.

Warunek konieczny to element, bez którego określony rezultat nie może wystąpić w granicach przyjętego modelu przyczynowego. Warunek wystarczający stanowi natomiast konfigurację czynników, której zaistnienie prowadzi do rezultatu w ramach określonych założeń. W systemach

społecznych wiele interwencji nie ma charakteru ani koniecznego, ani samodzielnie wystarczającego, gdyż ich skuteczność zależy od współwystępowania dodatkowych okoliczności.

Szkolenie zawodowe może zwiększać prawdopodobieństwo zatrudnienia, lecz nie jest jedyną drogą uzyskania kompetencji. Sama obecność szkolenia nie gwarantuje również poprawy sytuacji uczestnika. Program może więc stanowić jeden z czynników przyczyniających się do rezultatu, którego wystąpienie zależy od innych działań i warunków. Teoria zmiany powinna rozpoznawać takie zależności; w przeciwnym razie ryzykuje przypisanie pojedynczej interwencji zdolności wywoływania rezultatów, które wymagają znacznie szerszej konfiguracji przyczyn.

W filozofii przyczynowości strukturę tę wiąże się z koncepcją warunków INUS, zaproponowaną przez Johna Lesliego Mackiego. Zgodnie z tym ujęciem określony czynnik może stanowić niewystarczającą, lecz niezbędną część większej konfiguracji warunków wystarczających, przy czym sama konfiguracja pozostaje jedną z kilku alternatywnych dróg prowadzących do rezultatu. Koncepcja ta wyjaśnia, dlaczego pojedyncza interwencja może mieć istotne znaczenie przyczynowe, nie będąc samodzielnie wystarczającą ani bezwzględnie konieczną do wystąpienia zmiany. Przedsięwzięcie edukacyjne może przyczynić się do sukcesu zawodowego, ale jego znaczenie ujawnia się dopiero w połączeniu z dostępnymi ofertami pracy, możliwością dojazdu i uznawaniem kwalifikacji przez pracodawców. Podobnie program ochrony zdrowia może zwiększać skuteczność leczenia, lecz rezultat zależy również od biologicznych właściwości choroby, współpracy pacjenta i dostępności dalszej opieki.

Analogicznie inicjatywa legislacyjna może wnieść rzeczywisty wkład w reformę, choć jej powodzenie wymaga jednocześnie społecznej akceptacji, odpowiedniego układu politycznego oraz zdolności instytucji do przygotowania projektu.

Weryfikacja wpływu poprzez analizę kontrfaktyczną

W tym kontekście analiza przyczynowa nie powinna wymagać od każdego uczestnika procesu udowodnienia wyłącznej odpowiedzialności za rezultat, lecz raczej umożliwiać badanie jego rzeczywistego wkładu w wieloprzyczynowy mechanizm zmiany. Szczególnie doniosłą próbę uporządkowania kwestii przyczynowości w naukach społecznych stanowi tradycja potencjalnych rezultatów, rozwijana przez Jerzego Neymana i Donalda Rubina. W klasycznej pracy Estimating Causal Effects of Treatments in Randomized and Nonrandomized Studies z 1974 roku, Rubin analizował warunki, w których możliwe staje się szacowanie efektów interwencji na podstawie danych eksperymentalnych i obserwacyjnych.

W tym podejściu dla konkretnej jednostki rozważa się konceptualnie dwa rezultaty: ten, który wystąpiłby po zastosowaniu określonego działania, oraz ten, który pojawiłby się w porównywalnych warunkach bez tego działania. Efekt przyczynowy stanowi różnica między tymi potencjalnymi rezultatami, jednak zasadnicza trudność polega na niemożliwości ich jednoczesnego zaobserwowania u tej samej osoby w tym samym czasie. Badacz dostrzega bowiem tylko jeden z możliwych przebiegów zdarzeń, podczas gdy drugi posiada charakter kontrfaktyczny.

Właśnie dlatego sama obserwacja poprawy po interwencji nie jest wystarczającym dowodem jej skuteczności. Konieczne pozostaje odtworzenie wiarygodnej sytuacji, która wystąpiłaby w razie niepodjęcia działania. Problem przyczynowości różni się zatem fundamentalnie od prostego porównywania rezultatów sprzed i po zakończeniu interwencji.

Jeżeli uczestnik programu zatrudnieniowego znajduje pracę po odbyciu szkolenia, nie można automatycznie przypisać tego sukcesu samemu programowi, gdyż osoba ta mogła zostać zatrudniona również bez jego wsparcia. Jednocześnie nie można wyciągać odwrotnego wniosku, jakoby program był nieistotny – rzeczywiste doświadczenie mogło bowiem zwiększyć prawdopodobieństwo sukcesu lub przyspieszyć jego osiągnięcie.

Przedmiotem oceny pozostaje różnica między rzeczywistym przebiegiem zdarzeń a najbardziej wiarygodną alternatywą. W praktyce badawczej nie da się bezpośrednio zaobserwować obu historii tej samej osoby, ale można tworzyć grupy porównawcze, wykorzystywać naturalne zróżnicowanie warunków i stosować modele szacujące skutki interwencji przy przyjęciu określonych założeń. Istotne jest tu odróżnienie wyniku jednostkowego od przeciętnego efektu w badanej zbiorowości; interwencja może bowiem zwiększać prawdopodobieństwo rezultatu, nie gwarantując go każdemu uczestnikowi.

Z perspektywy metodologicznej konieczne jest rozpoznanie trzech grup założeń leżących u podstaw wielu analiz. Pierwsza dotyczy spójności definicji interwencji – czyli możliwości określenia, jakie działanie faktycznie zastosowano i jak odpowiada ono potencjalnemu rezultatowi w modelu. Druga obejmuje porównywalność jednostek poddanych interwencji i tych, które jej nie doświadczyły, z uwzględnieniem czynników wpływających zarówno na udział w programie, jak i na wynik końcowy. Trzecia odnosi się do istnienia dostatecznego nakładania się warunków, w których te same jednostki mogłyby znaleźć się w różnych wariantach postępowania.

W literaturze odpowiadają temu pojęcia consistency, exchangeability i positivity. Ich znaczenie nie jest wyłącznie matematyczne; wskazują one na realne ograniczenia w formułowaniu twierdzeń o skutkach działania. Jeśli określony rodzaj interwencji nigdy nie był stosowany wobec osób o szczególnych właściwościach, dane obserwacyjne mogą nie wystarczyć do oszacowania jej

skuteczności w tej grupie bez dodatkowych założeń. Podobnie nie można uznawać za identyczne działań, które formalnie noszą wspólną nazwę, lecz różnią się intensywnością, jakością lub sposobem wykonania.

Szczególnie ważnym problemem pozostaje selekcja uczestników, która może prowadzić do poważnego zniekształcenia rezultatów. Osoby zgłaszające się dobrowolnie do programu edukacyjnego mogą różnić się od nieuczestniczących pod względem motywacji, zasobów rodzinnych, stanu zdrowia, kompetencji czy gotowości do zmiany. Jeśli porównanie ogranicza się jedynie do późniejszych wyników obu grup, uzyskana różnica może odzwierciedlać te wcześniejsze cechy, a nie efekt samej interwencji. W ekonomii ekonometrycznej problem ten określa się jako obciążenie wynikające z selekcji; jego ograniczanie wymaga projektu badawczego odpowiadającego rzeczywistemu mechanizmowi przypisywania uczestników do programu.

Analiza regresyjna pomaga uwzględnić obserwowalne różnice, lecz nie eliminuje wpływu czynników nieobserwowanych. Sama rozbudowa modelu statystycznego nie gwarantuje więc identyfikacji przyczynowej, zwłaszcza gdy udział w programie zależy od motywacji lub kompetencji niemierzalnych dla badacza. Losowy przydział do interwencji jest jedną z najskuteczniejszych metod ograniczania tego problemu, gdyż przy odpowiednim projekcie eksperymentu umożliwia niezależne od wcześniejszych właściwości uczestników przypisanie wariantów postępowania. Tworzy to podstawę do szacowania przeciętnego efektu przyczynowego poprzez porównanie wyników dobranych grup. Nie oznacza to jednak, że randomizacja gwarantuje identyczny skład obu grup w każdym przypadku – przypadkowe różnice mogą nadal występować, szczególnie przy niewielkiej liczebności próby.

Metodologiczna poprawność nie gwarantuje skuteczności wdrożenia

Kluczowe pozostaje zachowanie prawidłowego sposobu realizacji przydziału, ograniczenie utraty uczestników w trakcie badania oraz odpowiednie uwzględnienie przypadków nieprzestrzegania przypisanej interwencji. W praktyce odmienne informacje dostarcza efekt przypisania do programu (intention-to-treat), a odmienne szacowanie efektu faktycznego skorzystania z interwencji, którego identyfikacja wymaga dodatkowych założeń. Nie można zatem utożsamiać poprawnie przeprowadzonego losowania z automatyczną możliwością uzasadnienia każdego twierdzenia o skuteczności programu.

Szczególnego znaczenia nabiera problem etycznej dopuszczalności eksperymentowania. W polityce społecznej, ochronie zdrowia czy administracji publicznej losowe różnicowanie dostępu do świadczeń może bowiem uderzać w podstawowe interesy uczestników. Nie każdy eksperyment jest moralnie dopuszczalny, a sam potencjał pozyskania wartościowej wiedzy nie stanowi wystarczającego uzasadnienia dla ograniczania praw jednostek. Jeśli określona interwencja jest już uznanym standardem ochronnym, arbitralne pozbawianie części osób dostępu do niej może być niedopuszczalne.

W sytuacjach, gdy zasoby pozwalają objąć programem jedynie część potencjalnych uczestników, przejrzysty i uzasadniony sposób losowego przydziału może pełnić podwójną funkcję: sprawiedliwej procedury podziału ograniczonych możliwości oraz instrumentu pozyskiwania wiedzy. Ocena zależy jednak od rodzaju świadczenia, wymogów prawnych, charakteru ryzyka i dostępności mniej ingerencyjnych metod badawczych. Właściwa teoria zmiany musi zatem integrować metodologię przyczynową z etyką badań i aksjologią instytucjonalną, nie traktując maksymalizacji wartości informacyjnej jako celu posiadającego bezwarunkowe pierwszeństwo.

Krytyczne ujęcie ograniczeń badań randomizowanych przedstawili Angus Deaton i Nancy Cartwright w artykule Understanding and Misunderstanding Randomized Controlled Trials (Social Science & Medicine, 2018). Autorzy wskazują, że choć randomizacja jest niezwykle użytecznym narzędziem wnioskowania przyczynowego, nie gwarantuje ona automatycznie trafności wszystkich wniosków dotyczących mechanizmów działania, precyzji oszacowań ani możliwości przenoszenia rezultatów na odmienne populacje. Kluczowy jest tu problem heterogeniczności efektów oraz różnic między warunkami eksperymentalnymi a rzeczywistym środowiskiem wdrożenia. Wynik uzyskany dla konkretnej grupy w określonych warunkach nie musi zachować identycznej wartości w innym systemie instytucjonalnym, nawet przy poprawnie przeprowadzonym eksperymencie.

Krytyka Deatona i Cartwright nie służy odrzuceniu badań randomizowanych, lecz wskazuje na konieczność ich interpretowania w ramach szerszej teorii przyczynowej, uwzględniającej mechanizmy, kontekst i granice uogólniania. Znaczenie tego zastrzeżenia obrazuje przykład programu edukacyjnego, który wykazuje korzystny efekt w grupie dobrowolnych uczestników objętych intensywnym wsparciem specjalistycznym. Zwiększenie skali przedsięwzięcia może prowadzić do objęcia nim osób o odmiennych potrzebach, spadku średniej intensywności pracy z uczestnikiem oraz nadmiernego obciążenia infrastruktury organizacyjnej.

Skuteczność interwencji zależy od warunków systemowych i rygoru analizy kontrfaktycznej

Rzeczywisty efekt programu może ulec zmianie, choć podstawowa konstrukcja interwencji formalnie pozostaje taka sama. Reforma skuteczna w gminie posiadającej rozwinięte zaplecze instytucjonalne może nie przynieść porównywalnych rezultatów w środowisku, w którym dostępność specjalistów i infrastruktury jest ograniczona.

W takich przypadkach problem nie polega na wadliwości pierwotnego eksperymentu, lecz na nieuzasadnionym rozszerzeniu zakresu twierdzenia przyczynowego poza warunki, w których uzyskano odpowiednie potwierdzenie. Właściwa teoria zmiany powinna więc obejmować nie tylko hipotezę o skuteczności interwencji, ale także wyjaśnienie, jakie warunki są niezbędne do zachowania jej efektywności. Badania obserwacyjne pozostają koniecznym elementem oceny wielu reform, ponieważ losowy przydział nie zawsze jest możliwy, wykonalny lub etycznie dopuszczalny. W takich sytuacjach kluczowe staje się konstruowanie wiarygodnych scenariuszy kontrfaktycznych na podstawie zmian zachodzących w porównywalnych jednostkach. Jedną z najważniejszych metod jest analiza różnic w różnicach (difference-in-differences), w której zmianę wyniku w grupie objętej interwencją porównuje się z odpowiednią zmianą w grupie kontrolnej.

Metoda ta opiera się na określonych założeniach, zwłaszcza dotyczących przebiegu trendów w warunkach braku interwencji. Jeśli obie grupy podlegałyby podobnym zmianom niezależnie od programu, różnica ich późniejszych trajektorii może dostarczać informacji o jego efekcie. Jednak wcześniejsze podobieństwo poziomów wskaźnika nie stanowi wystarczającego uzasadnienia dla takiej interpretacji, ponieważ znaczenie posiada przede wszystkim hipotetyczna porównywalność trendów. Należy dodatkowo uwzględniać możliwość zmian składu populacji, występowania innych interwencji oraz oddziaływań pomiędzy grupami. Współczesny rozwój metod różnic w różnicach ujawnił problemy związane z analizą programów wdrażanych w różnych okresach i wobec odmiennych grup. Brantly Callaway i Pedro H. C. Sant'Anna w artykule Difference-in-Differences with Multiple Time Periods (Journal of Econometrics, 2021) przedstawili procedury identyfikacji i estymacji efektów przyczynowych w sytuacjach zróżnicowanego czasu objęcia jednostek interwencją. Autorzy uwzględnili możliwość występowania odmiennych efektów między grupami i okresami oraz znaczenie założeń dotyczących trendów warunkowych. Proste uśrednienie wszystkich obserwowanych różnic może prowadzić do wyników trudnych do interpretacji, jeśli efekty programu zależą od czasu od wdrożenia lub charakterystyki uczestników. W takich warunkach właściwa metoda powinna umożliwiać odrębne oszacowanie rezultatów dla grup rozpoczynających udział w różnych okresach, a następnie ich odpowiednie agregowanie.

Zastosowanie tej metodologii do oceny reform administracyjnych pozwala zilustrować różnicę między korelacją czasową a wiarygodnym efektem przyczynowym. Jeżeli określone rozwiązanie zostaje wprowadzone początkowo tylko w części samorządów, badacz może porównywać zmiany zachodzące w jednostkach objętych reformą z odpowiednimi zmianami w pozostałych. Jednak sam fakt etapowego wdrażania nie gwarantuje poprawności wnioskowania, ponieważ wcześniejszy wybór samorządów mógł zależeć od ich kompetencji, sytuacji finansowej albo gotowości organizacyjnej.

Projekt nieciągłości regresyjnej jako metoda izolacji lokalnych efektów

Jednostki, które jako pierwsze wdrażają zmianę, mogą dysponować większymi zdolnościami adaptacyjnymi, co bezpośrednio wpływa na osiągane rezultaty. Konieczne jest zatem zbadanie procesu wyboru uczestników, wcześniejszych trendów, występowania równoległych reform oraz możliwych różnic w czasie ujawniania się efektów. W przeciwnym razie korzystna trajektoria jednostek objętych programem może zostać błędnie przypisana jego działaniu, podczas gdy w rzeczywistości odzwierciedla ona wcześniej istniejące przewagi organizacyjne. Inne możliwości identyfikacji przyczynowej stwarza projekt nieciągłości regresyjnej, określany jako regression discontinuity design. Jego podstawą jest sytuacja, w której dostęp do programu zależy od przekroczenia ustalonego progu punktowego, dochodowego lub innego parametru.

Jeśli jednostki znajdujące się bezpośrednio po obu stronach progu są dostatecznie porównywalne, a przypisanie do programu zmienia się w jego pobliżu, możliwe staje się oszacowanie lokalnego efektu interwencji. David Lee i Thomas Lemieux w opublikowanym w 2010 roku w Journal of Economic Literature przeglądzie metodologicznym przedstawili warunki poprawnego stosowania tej konstrukcji. Zwrócili uwagę między innymi na znaczenie procesu przydziału, ryzyko strategicznego wpływania na zmienną progową oraz ograniczenia w interpretacji uzyskiwanych oszacowań. Należy podkreślić, że efekt zidentyfikowany w pobliżu określonego progu nie musi odpowiadać skuteczności programu w całej populacji. W polityce społecznej podobne rozwiązanie może być przydatne do oceny świadczenia przyznawanego na podstawie konkretnego kryterium dochodowego, pod warunkiem dostępności odpowiednich danych i spełnienia wymagań metodologicznych. Osoby znajdujące się nieznacznie poniżej i powyżej progu mogą być podobne pod względem wielu istotnych właściwości, a różnica w dostępie do programu pozwala badać jego lokalne konsekwencje.

Interpretacja przyczynowa staje się jednak znacznie trudniejsza, jeśli jednostki mogą manipulować deklarowanym dochodem w celu uzyskania świadczenia lub gdy przekroczenie progu uruchamia jednocześnie szereg innych zmian instytucjonalnych. Nie można również bez dodatkowego uzasadnienia zakładać, że rezultat uzyskany dla osób znajdujących się blisko granicy kwalifikacyjnej odpowiada efektowi programu wobec osób o znacznie niższych dochodach. Metoda ta może zatem dostarczać wiarygodnej odpowiedzi na stosunkowo wąskie pytanie, nie rozstrzygając automatycznie o społecznej wartości całej polityki. Ograniczenie zakresu wniosku nie stanowi jednak wady, o ile zostaje prawidłowo przedstawione i odpowiednio wykorzystane w szerszej analizie.

Pluralizm metodologiczny jako warunek precyzyjnej teorii zmiany

Szczególnie interesującą metodą w przypadku przedsięwzięć realizowanych na poziomie regionów lub państw jest konstrukcja kontroli syntetycznej. Alberto Abadie, Alexis Diamond i Jens Hainmueller w artykule z 2010 roku pt. **Synthetic Control Methods for Comparative Case Studies: Estimating the Effect of California's Tobacco Control Program** przedstawili sposób tworzenia jednostki porównawczej poprzez odpowiednio dobraną kombinację innych jednostek. Celem jest jak najwierniejsze odtworzenie trajektorii regionu poddanego interwencji przed momentem jej wprowadzenia. Autorzy zastosowali tę metodę do analizy skutków kalifornijskiego programu ograniczania palenia tytoniu, wdrożonego w 1988 roku na mocy Proposition 99.

Oszacowali oni, że do 2000 roku roczna sprzedaż papierosów na mieszkańca w Kalifornii była o około 26 paczek niższa, niż wynikałoby to z przebiegu syntetycznego regionu porównawczego. Należy jednak pamiętać, że wynik ten jest efektem konkretnej procedury estymacyjnej i jej założeń, a nie bezpośrednio zaobserwowanym faktem dotyczącym alternatywnej historii Kalifornii. Wartość tej metody tkwi w możliwości prowadzenia uporządkowanej analizy interwencji, dla których brakuje naturalnej, pojedynczej jednostki porównawczej. Reformę regionalną można wówczas oceniać w odniesieniu do kombinacji jednostek o podobnym wcześniejszym przebiegu kluczowych wskaźników. Wiarygodność takiego oszacowania zależy jednak od jakości dopasowania przed interwencją, właściwego doboru grupy porównawczej oraz braku innych zdarzeń, które mogłyby oddziaływać na badaną jednostkę w sposób odmienny od reszty grupy. Konieczne jest również sprawdzenie, czy program nie wpłynął na obszary wykorzystane do budowy kontroli, gdyż podważyłoby to założenie o niezależności porównania. Kluczowe stają się tu analizy wrażliwości –

obejmujące m.in. alternatywne sposoby doboru jednostek oraz weryfikację, czy podobne efekty nie występują w obszarach, gdzie interwencja nie miała miejsca. Metoda kontroli syntetycznej nie zastępuje zatem przyczynowego rozumowania, lecz nadaje mu przejrzystszą i łatwiejszą do skontrolowania strukturę.

W badaniach społecznych zdarzają się sytuacje, w których uczestnicy nie stosują się do przypisanej interwencji, a ich faktyczny udział zależy od osobistych preferencji i zasobów. W takich warunkach można wykorzystać zmienne instrumentalne, o ile istnieje czynnik wpływający na prawdopodobieństwo udziału w programie, który nie oddziałuje na wynik końcowy inaczej niż poprzez określoną interwencję. Wymaga to jednak przyjęcia konkretnych założeń dotyczących zachowania uczestników, charakteru działania instrumentu oraz braku nieuzasadnionych ścieżek wpływu.

W tradycji Joshuy Angrista i Guido Imbensa kluczowe stało się pojęcie lokalnego średniego efektu oddziaływania (Local Average Treatment Effect – LATE), odnoszące się do osób, których rzeczywisty udział w programie zmienił się pod wpływem zastosowanego instrumentu. Takiego oszacowania nie wolno utożsamiać z efektem przeciętnym dla całej populacji, gdyż może ono dotyczyć specyficznej grupy uczestników. W praktyce pokazuje to, że nawet zaawansowana metoda przyczynowa dostarcza informacji o konkretnym parametrze, którego znaczenie dla decyzji politycznej wymaga odrębnego uzasadnienia. Rzetelna teoria zmiany powinna zatem pytać nie tylko o to, czy interwencja przyniosła rezultat, ale o rodzaj efektu, który faktycznie zidentyfikowano. Efekt zamiaru objęcia programem może różnić się od efektu faktycznego skorzystania ze świadczenia. Wynik oszacowany dla osób w pobliżu progu kwalifikacyjnego nie musi być reprezentatywny dla wszystkich uprawnionych, a średni rezultat grupy uczestników nie rozstrzyga automatycznie o korzyściach dla osób w najtrudniejszym położeniu. Z kolei zmiana zaobserwowana bezpośrednio po zakończeniu programu nie gwarantuje jej trwałości w dłuższym horyzoncie czasowym.

W każdym z tych przypadków niezbędne jest jednoznaczne określenie jednostki analizy, populacji docelowej, rodzaju interwencji, punktu odniesienia, horyzontu pomiaru oraz właściwości rezultatu. Bez takiej precyzji nawet statystycznie poprawna liczba może prowadzić do nieuzasadnionych uogólnień. Z perspektywy filozofii nauki metody te są interesujące, ponieważ pozwalają analizować przyczynowość bez konieczności przyjmowania jednej, uniwersalnej procedury badawczej. Eksperyment losowy daje silne podstawy do szacowania efektu, ale nie musi wyjaśniać wszystkich mechanizmów jego powstawania. Analiza obserwacyjna może ujawniać ważne zależności i pozwalać na odpowiedzialne wnioskowanie przyczynowe, jeśli spełnione są warunki identyfikacji. Badanie jakościowe dostarcza szczegółowej wiedzy o tym, jak uczestnicy interpretują interwencję, lecz nie

uprawnia do szacowania przeciętnego efektu dla populacji. Z kolei analiza instytucjonalna może wyjaśnić, dlaczego rozwiązanie o potwierdzonej skuteczności nie zostaje wdrożone w innych warunkach.

Właściwa teoria zmiany wymaga zatem pluralizmu metodologicznego podporządkowanego precyzyjnie sformułowanym pytaniom. Pluralizm ten nie oznacza relatywizmu poznawczego, gdyż różne metody mają odmienne wymagania dowodowe, ograniczenia i zakres uprawnionych twierdzeń.

Konfiguracja kontekstu i mechanizmów warunkuje rezultat

Szczególną propozycję integracji analizy mechanizmów i kontekstu przedstawili Ray Pawson i Nick Tilley w książce Realistic Evaluation z 1997 roku. W rozwiniętej przez nich ewaluacji realistycznej centralne znaczenie nabiera pytanie o to, jakie rozwiązanie działa, dla kogo, w jakich warunkach oraz dzięki jakim mechanizmom. Interwencja nie jest tu traktowana jako automatyczny generator określonych rezultatów, ponieważ jej zastosowanie może uruchamiać odmienne reakcje uczestników zależnie od kontekstu społecznego, organizacyjnego i instytucjonalnego. Podstawową konstrukcję analityczną tworzą konfiguracje kontekstu, mechanizmu i wyniku, określane jako Context-Mechanism-Outcome (CMO). Kontekst obejmuje właściwości sytuacji wpływające na możliwość uruchomienia określonego mechanizmu, mechanizm odnosi się do procesów generujących zmianę, natomiast rezultat stanowi obserwowalną konsekwencję ich oddziaływania.

Kluczowe jest to, że mechanizmu nie należy utożsamiać z samą czynnością wykonywaną przez organizację. Szkolenie jest działaniem programowym, lecz mechanizmem prowadzącym do zmiany może być na przykład wzrost poczucia skuteczności, nabycie konkretnej umiejętności albo uzyskanie dostępu do nowych zasobów. Rozróżnienie to pozwala uniknąć jednego z najczęstszych błędów w konstruowaniu teorii zmiany: przedstawiania kolejnych czynności jako mechanizmów przyczynowych.

Organizacja może prowadzić konsultacje, ale sam fakt zorganizowania spotkania nie wyjaśnia, dlaczego miałoby ono doprowadzić do zmiany decyzji publicznej. Znaczenie może mieć dostarczenie decydentom nowej wiedzy, ujawnienie wcześniej niewidocznych interesów, zmniejszenie niepewności co do skutków reformy lub zwiększenie społecznej legitymizacji rozwiązania. Każdy z tych mechanizmów wymaga odmiennego sposobu badania. Podobnie program profilaktyczny może wykorzystywać kampanię informacyjną, lecz jego skuteczność zależy

od tego, czy doszło do zwiększenia wiedzy, zmiany norm społecznych, ograniczenia barier czy poprawy dostępności usług. Sam zasięg kampanii nie pozwala rozstrzygnąć, który mechanizm został rzeczywiście uruchomiony. Właściwa ewaluacja realistyczna wymaga więc konstruowania wyjaśnień na poziomie procesów zachodzących w relacji między zasobami interwencji a sposobem rozumowania i działania uczestników.

Obrazuje to przykład odmiennej skuteczności programu zatrudnieniowego w dwóch społecznościach o porównywalnych charakterystykach demograficznych. W pierwszej istnieje rozwinięty rynek pracy, sprawny transport i sieć przedsiębiorstw poszukujących kompetencji. W drugiej liczba miejsc pracy jest ograniczona, a uczestnicy napotykają bariery komunikacyjne i organizacyjne. Ten sam program szkoleniowy może zwiększać wiedzę w obu środowiskach, ale tylko w pierwszym prowadzić do wzrostu zatrudnienia. Różnica rezultatów nie musi oznaczać braku staranności w realizacji programu.

Może ona wynikać z tego, że w jednym środowisku nabyte kompetencje otwierają realną drogę na rynek pracy, podczas gdy w drugim brak warunków ogranicza znaczenie tej zmiany. Skuteczność nie jest zatem wyłącznie właściwością samego programu, lecz rezultatem relacji między konstrukcją interwencji a otoczeniem, w którym jej mechanizmy mają zostać uruchomione. Ewaluacja realistyczna wymaga jednak rygoru metodologicznego, gdyż odwołanie do kontekstu nie może stać się uniwersalnym wyjaśnieniem każdego nieoczekiwanego rezultatu. Jeżeli badacz po zakończeniu interwencji arbitralnie wskazuje właściwości środowiska uzasadniające sukces albo porażkę, tworzy narrację przekonującą, lecz niedostatecznie sprawdzalną. Dlatego kluczowe jest wcześniejsze formułowanie hipotez dotyczących mechanizmów, gromadzenie danych do porównywania przypadków i systematyczna ocena wyjaśnień konkurencyjnych. Wyrazem rozwoju tych standardów stały się wytyczne RAMESES II, przedstawione przez Geoffa Wonga i współautorów w BMC Medicine w 2016 roku.

Autorzy zwracali uwagę na potrzebę przejrzystego przedstawiania celu ewaluacji, sposobu konstruowania teorii programu, doboru metod, charakterystyki uczestników oraz procesu analizy hipotez. Podejście realistyczne nie stanowi rezygnacji z empirycznej kontroli, lecz proponuje odmienny sposób organizowania rozumowania przyczynowego, szczególnie użyteczny w badaniach złożonych interwencji społecznych. W nauce o ewaluacji istotne jest również odróżnienie przypisania skutku konkretnemu działaniu od wykazania, że działanie to było jednym z czynników wytwarzających rezultat. John Mayne w pracy Addressing Attribution Through Contribution Analysis: Using Performance Measures Sensibly z 2001 roku zaproponował analizę wkładu (contribution analysis), której celem jest konstruowanie wiarygodnego wyjaśnienia roli programu w procesie zmiany. W późniejszym artykule Contribution Analysis: Coming of Age? z 2012 roku

Mayne podkreślał znaczenie przyczyn współtworzących rezultat, które samodzielnie nie muszą być ani konieczne, ani wystarczające.

Podejście to ma szczególną wartość w środowiskach obejmujących liczne równoległe działania, gdzie wyizolowanie jednego efektu jest trudne lub wymagałoby nierealnych założeń. Analiza wkładu nie eliminuje wymogu dowodzenia, lecz zmienia charakter pytania: przedmiotem badania staje się nie tylko wielkość wyłącznego efektu programu, ale przede wszystkim wiarygodność twierdzenia, że określone mechanizmy interwencji rzeczywiście uczestniczyły w wytwarzaniu rezultatu.

Weryfikacja wkładu poprzez rygorystyczną analizę procesową

Kluczowe znaczenie ma tutaj konstrukcja opowieści przyczynowej, której nie wolno utożsamiać z dowolną narracją przedstawiającą korzystny przebieg zdarzeń. Wiarygodne wyjaśnienie wkładu wymaga sformułowania teorii zmiany, zgromadzenia informacji o poszczególnych mechanizmach, rozpoznania założeń oraz systematycznego badania alternatywnych wyjaśnień rezultatu.

Jeśli organizacja obywatelska publikuje opracowanie, po którym zostaje przyjęta reforma prawna, nie wystarczy stwierdzić, że publikacja poprzedzała zmianę. Konieczne jest ustalenie, czy argumenty zawarte w tekście rzeczywiście dotarły do osób decyzyjnych, czy zostały wykorzystane w procesie przygotowywania reformy oraz jakie inne czynniki mogły wpłynąć na jej przyjęcie. Znaczenie mogą mieć wcześniejsze inicjatywy innych organizacji, zmiany polityczne, stanowiska instytucji międzynarodowych lub niezależne analizy ekspertów. W takim przypadku udowodnienie wyłącznego autorstwa zmiany byłoby nieuzasadnione, lecz rzetelna analiza może dostarczyć przekonujących przesłanek dotyczących rzeczywistego wkładu organizacji w konkretny etap procesu. Autor książki podkreśla, że szczególnie ważne jest rozróżnienie między wpływem bezpośrednim a pośrednim. Organizacja może nie uczestniczyć formalnie w przygotowaniu rozwiązania, ale jej wcześniejsze działania mogą zwiększać społeczną widoczność problemu, upowszechniać określony aparat pojęciowy lub tworzyć przestrzeń do współpracy między podmiotami, które wcześniej funkcjonowały odrębnie.

Podobne procesy kształtują późniejsze decyzje, choć ich znaczenia trudno wyrazić za pomocą jednego wskaźnika. Nie oznacza to jednak, że każda zmiana zgodna z misją organizacji stanowi dowód jej sprawczości. Właściwa analiza wymaga ustalenia, czy istnieją rzeczywiste ścieżki przekazywania informacji, mobilizacji zasobów i oddziaływania na zachowania uczestników. Związek przyczynowy może być odroczony i wieloetapowy, lecz musi pozostać możliwy do

uzasadnienia poprzez dane, dokumenty, relacje świadków i analizę konkurencyjnych scenariuszy. Szczególnie użytecznym narzędziem badania takich zależności jest process tracing, czyli analiza śledzenia procesu przyczynowego w konkretnym przypadku. Jej zasadnicza funkcja polega na rekonstrukcji sekwencji zdarzeń i mechanizmów łączących przyczynę z rezultatem przy wykorzystaniu dowodów dotyczących poszczególnych etapów. W przeciwieństwie do prostego stwierdzenia współwystępowania zjawisk, metoda ta dąży do wykazania, w jaki sposób konkretne działanie uruchomiło proces prowadzący do obserwowanego skutku.

W praktyce może to obejmować analizę dokumentów, kolejnych wersji projektów, przebiegu negocjacji oraz stanowisk decydentów. Kluczowa jest tutaj różna wartość dowodowa poszczególnych obserwacji. Ogólne podobieństwo języka dwóch dokumentów może być zgodne z hipotezą wpływu, ale nie stanowi jej silnego potwierdzenia, jeśli obie instytucje mogły korzystać z tego samego źródła.

Znacznie większą wagę ma udokumentowane wykorzystanie konkretnej propozycji w procesie decyzyjnym, zwłaszcza gdy znana jest sekwencja jej przekazania i późniejszej modyfikacji projektu. Nawet takie ustalenie nie musi jednak dowodzić wyłącznego wpływu, gdyż decyzja mogła być współtworzona przez kilka podmiotów. Istotne staje się zatem rozróżnienie między dowodami potwierdzającymi możliwość istnienia mechanizmu a takimi, które pozwalają wykluczyć konkurencyjne wyjaśnienia. Obserwacja zgodna z hipotezą nie musi mieć wysokiej wartości diagnostycznej, jeśli podobny rezultat byłby równie prawdopodobny przy innych mechanizmach. Z drugiej strony zdarzenie mało prawdopodobne w przypadku alternatywnych hipotez istotnie zwiększa wiarygodność konkretnej interpretacji przyczynowej. Analiza procesu może być zatem ujmowana przez pryzmat bayesowskiej aktualizacji przekonań, zgodnie z którą znaczenie nowej informacji zależy od relacji jej prawdopodobieństwa przy różnych konkurencyjnych wyjaśnieniach.

Nie oznacza to konieczności przedstawiania każdej analizy historycznej czy instytucjonalnej w formie liczbowego modelu probabilistycznego. Wymaga natomiast krytycznej oceny, czy przedstawione dowody rzeczywiście odróżniają hipotezę preferowaną od alternatywnych, czy jedynie pozostają z nią zgodne.

Efekty pośrednie a pułapka równowagi ogólnej

Szczególnie trudna okazuje się ocena rezultatów przedsięwzięć, w których występują efekty pośrednie i wzajemne oddziaływania uczestników. Program skierowany do wybranej grupy może bowiem wpływać na osoby formalnie nieobjęte interwencją – dzieje się tak, gdy uczestnicy przekazują sobie informacje, zmieniają normy społeczne lub oddziałują na wspólne zasoby. W

edukacji przejawia się to w przenoszeniu kompetencji poprzez relacje rówieśnicze, w ochronie zdrowia w ograniczaniu ryzyka dla osób nieuczestniczących bezpośrednio w programie, a na rynku pracy w zmianach konkurencji o konkretne stanowiska.

W takich warunkach klasyczne założenie braku interferencji między jednostkami bywa nieuzasadnione. Kluczowe jest przy tym, że efekty pośrednie mogą mieć charakter zarówno korzystny, jak i niekorzystny. Zwiększenie kompetencji części uczestników rynku pracy może poprawić ogólną produktywność, lecz w określonych okolicznościach zmienić względną sytuację osób nieobjętych szkoleniem. Właściwa teoria zmiany powinna zatem uwzględniać, czy interwencja tworzy dodatkową wartość, czy przekształca jedynie sposób jej podziału. Problem ten wiąże się bezpośrednio z różnicą między efektami częściowymi a konsekwencjami równowagi ogólnej. Interwencja może poprawiać sytuację określonej grupy przy zachowaniu stałych warunków otoczenia, jednak jej rozszerzenie na całą populację może wpłynąć na ceny, wynagrodzenia, zachowania przedsiębiorców czy dostępność dóbr.

Przykładowo: program szkoleniowy zwiększa szanse zatrudnienia uczestników, jeśli posiadają oni kompetencje rzadko występujące na lokalnym rynku. Jednak powszechne upowszechnienie tych samych umiejętności nie musi prowadzić do proporcjonalnego wzrostu zatrudnienia, o ile liczba odpowiednich miejsc pracy pozostaje ograniczona. Podobnie dopłata do zakupu określonego dobra zwiększa dostępność dla części odbiorców, lecz jej szerokie zastosowanie może wpłynąć na cenę rynkową, zależnie od elastyczności podaży i innych uwarunkowań.

Skuteczność rozwiązania to nie to samo co skuteczność wdrożenia

Efekty interwencji nie mogą być zawsze obliczane przez prostą multiplikację rezultatu uzyskanego w niewielkiej grupie przez liczbę przyszłych uczestników. Właściwe projektowanie skali wymaga analizy zmian systemowych, które powstają wskutek samego rozszerzenia programu. Z perspektywy badań nad wdrażaniem kluczowe jest odróżnienie skuteczności rozwiązania od skuteczności strategii jego upowszechniania. Enola Proctor i współautorzy w artykule Outcomes for Implementation Research: Conceptual Distinctions, Measurement Challenges, and Research Agenda (2011) zaproponowali rozróżnienie ośmiu kategorii rezultatów implementacyjnych.

Obejmują one akceptowalność, faktyczne przyjęcie interwencji, ocenę jej odpowiedniości, wykonalność, wierność wykonania, koszt wdrożenia, stopień rozpowszechnienia w organizacji lub systemie oraz trwałość stosowania. Kategorie te stanowią ramę analityczną do badania procesów

wdrożeniowych i nie powinny być utożsamiane z samymi wynikami zdrowotnymi czy społecznymi. Interwencja może bowiem wykazywać wysoką skuteczność w warunkach kontrolowanych, a przynosić ograniczone korzyści w rzeczywistości, jeśli nie zostanie odpowiednio przyjęta przez organizacje lub nie dotrze do właściwych odbiorców. Szczególna wartość tego rozróżnienia ujawnia się, gdy organizacja stwierdza niepowodzenie programu i próbuje określić rodzaj niezbędnej korekty. Jeśli problemem jest niska akceptowalność, konieczna może być analiza doświadczeń odbiorców i sposobu komunikowania znaczenia interwencji. Przy niedostatecznej wierności wykonania należy zbadać kompetencje pracowników, jakość instrukcji i dostępność zasobów.

Jeżeli rozwiązanie pozostaje zbyt kosztowne, potrzebna może być zmiana modelu finansowania lub konstrukcji procesu. Jeśli natomiast program został poprawnie wdrożony, lecz nie przyniósł oczekiwanych rezultatów społecznych, trudność może dotyczyć samej teorii przyczynowej. Ewaluacja implementacyjna i ewaluacja oddziaływania dostarczają zatem informacji komplementarnych; wzajemne zastępowanie tych dwóch perspektyw prowadziłoby do błędów diagnostycznych. Sukces wdrożeniowy nie gwarantuje sukcesu społecznego, a brak oczekiwanego oddziaływania nie zawsze oznacza, że interwencja oparta była na błędnych założeniach.

Znaczenie tego problemu rozwinęli Graham Moore i współautorzy w wytycznych Medical Research Council z 2015 roku dotyczących ewaluacji procesu. Podkreślali oni potrzebę równoczesnego badania sposobu wdrożenia, mechanizmów oddziaływania i kontekstu, gdyż sama informacja o wyniku końcowym bywa niewystarczająca do wyjaśnienia przyczyn sukcesu lub porażki. Szczególnie ważne jest rozpoznawanie wierności wykonania, intensywności interwencji, jej zasięgu, lokalnych adaptacji oraz reakcji uczestników.

Nie każda zmiana procedury wdrożeniowej stanowi naruszenie pierwotnej konstrukcji programu – odpowiednie dostosowanie do środowiska bywa konieczne dla zachowania mechanizmu skuteczności. Jednocześnie nie każda adaptacja jest uzasadniona, zwłaszcza gdy prowadzi do usunięcia elementu odgrywającego zasadniczą rolę przyczynową. Właściwa teoria programu powinna wskazywać, które składniki są konstytutywne dla mechanizmu zmiany, a które można modyfikować bez utraty jego funkcji. Zagadnienie wierności wykonania prowadzi zatem bezpośrednio do teorii standaryzacji i zróżnicowania instytucjonalnego. Nadmiernie sztywny program może nie odpowiadać lokalnym potrzebom, podczas gdy rozwiązanie całkowicie pozostawione uznaniu wykonawców może utracić spójność i możliwość rzetelnej ewaluacji. Odpowiedzialne projektowanie wymaga odróżnienia rdzenia przyczynowego interwencji od jej zmiennych form organizacyjnych.

Program aktywizacji zawodowej może wymagać utrzymania określonej jakości diagnozy kompetencji, indywidualnego dopasowania wsparcia i rzeczywistego kontaktu z pracodawcami, ale

sposób prowadzenia spotkań może zależeć od lokalnej dostępności transportu, kompetencji cyfrownych uczestników i struktury rynku pracy.

Wielowymiarowa weryfikacja skuteczności: od kosztów i etyki po rygor falsyfikacji

W ochronie zdrowia określony standard bezpieczeństwa może pozostawać niezmienny, podczas gdy organizacja procesu jego realizacji powinna być dostosowana do wielkości placówki i dostępnych zasobów. W administracji publicznej gwarancje prawne mogą wymagać zachowania wspólnego standardu, lecz sposób komunikowania informacji obywatelom może podlegać racjonalnej adaptacji. Właściwa teoria zmiany powinna zatem chronić mechanizm i wartości nadrzędne, nie utożsamiając ich z każdym szczegółem pierwotnej procedury.

Na poziomie ekonomii dobrobytu kluczowym wymaganiem pozostaje uwzględnianie kosztów interwencji oraz dostępnych alternatyw. Program może przynosić rzeczywisty, korzystny rezultat, jednak jego społeczna wartość wymaga oceny zasobów zużytych do jego uzyskania. Jeśli dwa rozwiązania prowadzą do porównywalnych efektów, ale różnią się kosztami i zakresem niezamierzonych konsekwencji, racjonalny wybór wymaga rzetelnego zestawienia tych właściwości. Analiza efektywności kosztowej pozwala określić koszt uzyskania jednostki danego rezultatu, podczas gdy analiza kosztów i korzyści zmierza do szerszego porównania ich wartości ekonomicznej.

Nie każda wartość społeczna może jednak zostać bezspornie wyrażona w pieniądzu, a zastosowanie konkretnej metody wyceny wymaga ujawnienia przyjętych założeń. Szczególnie istotne jest więc rozróżnienie między sprawnością ekonomiczną a moralną dopuszczalnością interwencji. Rozwiązanie o korzystnym bilansie ekonomicznym nie zyskuje automatycznie uzasadnienia, jeśli jego realizacja narusza podstawowe prawa lub prowadzi do nieproporcjonalnego obciążenia określonej grupy. Zagadnienie to można pogłębić, odwołując się do teorii sprawiedliwości dystrybucyjnej i problemu agregowania indywidualnych korzyści; średni pozytywny rezultat programu może bowiem współistnieć z istotnym pogorszeniem sytuacji części uczestników.

Reforma może zwiększać przeciętną dostępność usług, a jednocześnie ograniczać ją osobom wymagającym szczególnego wsparcia. Przedsięwzięcie gospodarcze może prowadzić do wzrostu całkowitej produktywności, lecz przekształcać podział ryzyka w sposób obciążający słabszych uczestników rynku. Właściwa ewaluacja powinna zatem obejmować nie tylko wielkość przeciętnego efektu, lecz także jego rozkład oraz znaczenie dla osób o odmiennych zasobach i możliwościach

działania. Z perspektywy Rawlsowskiej teorii sprawiedliwości szczególne znaczenie ma wpływ na położenie grup najmniej uprzywilejowanych, podczas gdy podejście oparte na możliwościach Sena i Nussbaum kieruje uwagę ku rzeczywistemu poszerzaniu autonomii i zdolności urzeczywistniania wartościowych form życia.

Nie oznacza to, że wszystkie programy powinny być oceniane przez jedno, uniwersalne kryterium moralne. Wymaga jednak jawnego określenia, jakie wartości uzasadniają wybór interwencji i w jaki sposób mają być uwzględniane w ocenie jej konsekwencji. Doniosła pozostaje również potrzeba rozróżniania skutków zamierzonych, niezamierzonych i wtórnych. Interwencja może osiągać bezpośredni cel, a jednocześnie generować nowe trudności w innych częściach systemu. Program dofinansowania określonego dobra może zwiększać jego dostępność dla części odbiorców, lecz oddziaływać również na strukturę cen i zachowanie dostawców. Reforma edukacyjna może poprawiać wyniki w obszarze objętym oceną, ale zmniejszać czas przeznaczany na rozwijanie innych kompetencji. System oceny pracowników może zwiększać produktywność konkretnej czynności, a jednocześnie ograniczać skłonność do współpracy i dzielenia się wiedzą.

W każdym przypadku właściwa teoria zmiany powinna uwzględniać możliwość wystąpienia negatywnych efektów pośrednich oraz określać narzędzia pozwalające je rozpoznawać. Nie wystarcza bowiem wykazanie, że określony parametr uległ poprawie, jeżeli konsekwencje tej zmiany prowadzą do nieuzasadnionego pogorszenia innych właściwości systemu. Kluczowe jest, aby badacz nie ograniczał pomiaru wyłącznie do rezultatów zgodnych z pierwotnymi oczekiwaniami projektantów. Tutaj teoria zmiany spotyka się z problematyką falsyfikowalności oraz metodologią krytycznego sprawdzania hipotez. Karl Popper wskazywał na znaczenie poddawania twierdzeń próbom empirycznym, których wyniki mogłyby podważyć ich prawdziwość.

W badaniach społecznych zastosowanie tej intuicji wymaga uwzględnienia złożoności modeli przyczynowych, ponieważ niezgodność między przewidywaniem a rezultatem może wynikać z różnych przyczyn: błędów teorii, niedostatecznego wykonania programu lub niewłaściwego pomiaru. Nie oznacza to jednak, że teorię zmiany można chronić przed krytyką poprzez nieustanne dodawanie nowych warunków wyjaśniających brak rezultatu. Właściwy model powinien określać, jakie obserwacje stanowiłyby istotne wyzwanie dla przyjętej hipotezy oraz jakie wyniki uzasadniałyby jej modyfikację albo odrzucenie. Szczególne znaczenie posiada więc wcześniejsze dokumentowanie przewidywań, ponieważ ogranicza to możliwość późniejszego dostosowywania opowieści przyczynowej do dowolnego przebiegu zdarzeń. Problem ten ma bezpośredni związek z kryzysem replikowalności i przejrzystości badań naukowych, który zwrócił uwagę na ryzyko selektywnego przedstawiania wyników, wielokrotnego testowania hipotez i niedostatecznego rozróżniania analiz zaplanowanych od eksploracyjnych.

W ewaluacji społecznej analogiczne ryzyko pojawia się wtedy, gdy organizacja dysponuje znaczną liczbą wskaźników, spośród których po zakończeniu przedsięwzięcia wybiera te przedstawiające najbardziej korzystny rezultat. Jeżeli pierwotnym celem programu była poprawa sytuacji zawodowej, a po jego niepowodzeniu sukces zostaje zdefiniowany wyłącznie przez liczbę przeprowadzonych spotkań, dochodzi do zmiany kryterium oceny, która powinna zostać jawnie ujawniona. Nie oznacza to, że nieoczekiwane rezultaty są poznawczo bezwartościowe – mogą one bowiem prowadzić do odkrycia nowych mechanizmów.

Teoria zmiany jako architektura warunkowej sprawczości

Wymagają jednak wyraźnego rozróżnienia między pierwotnie zakładanym efektem a rezultatem odkrytym w toku analizy. Właściwa teoria zmiany powinna umożliwiać elastyczne rozwijanie wiedzy, lecz nie może stać się narzędziem służącym do przedstawiania każdej możliwej konsekwencji jako realizacji wcześniej określonego celu.

Kluczowa pozostaje tutaj kwestia niezależności procesu ewaluacyjnego. Organizacja realizująca przedsięwzięcie ma naturalny interes w wykazaniu jego użyteczności, zwłaszcza gdy od oceny zależą finansowanie, reputacja i kontynuacja działalności. Sam fakt istnienia takiego interesu nie dowodzi manipulacji, lecz tworzy warunki wymagające odpowiednich zabezpieczeń metodologicznych. Znaczenie może mieć tu wcześniejsze ustalenie kryteriów sukcesu, rozdzielenie funkcji wykonawczych od ewaluacyjnych, przejrzystość dokumentacji oraz możliwość niezależnej weryfikacji podstawowych ustaleń.

Szczególnie niebezpieczne jest połączenie silnych bodźców finansowych z możliwością jednostronnego definiowania rezultatów; system może wówczas premiować nie tyle rzeczywistą skuteczność, co zdolność konstruowania korzystnej narracji sprawozdawczej. Właściwe zarządzanie powinno zatem odróżniać uzasadnioną potrzebę wykazywania wyników od obowiązku przedstawiania pełnej informacji o ich ograniczeniach i możliwych przyczynach niepowodzenia. Teoria zmiany nie powinna być pojmowana wyłącznie jako instrument kontroli instytucjonalnej, gdyż jej funkcją jest również organizowanie odpowiedzialnego uczenia się. Program może nie osiągnąć zamierzonego efektu, a mimo to dostarczyć wartościowej wiedzy na temat nieskuteczności określonej hipotezy i warunków niezbędnych dla przyszłych interwencji. Podobnie przedsięwzięcie może przynieść korzystny rezultat, jednocześnie ujawniając, że jego koszty lub wymagania organizacyjne ograniczają możliwość zwiększenia skali.

W każdym przypadku rzetelna ocena powinna pozwalać na rozróżnienie wartości uzyskanego rezultatu od wartości poznawczej samego procesu. Nie oznacza to jednak, że każde niepowodzenie należy usprawiedliwiać jako sukces edukacyjny – działania podejmowane z wykorzystaniem zasobów publicznych lub społecznych wymagają odpowiedzialności za ich wydatkowanie. Wiedza uzyskana z przedsięwzięcia może stanowić rzeczywistą korzyść, lecz jej wartość należy oceniać w odniesieniu do kosztów, pierwotnych celów i możliwości przyszłego zastosowania. Z perspektywy filozofii praktycznej szczególnego znaczenia nabiera rozróżnienie pomiędzy odpowiedzialnością za intencję, odpowiedzialnością za sposób działania i odpowiedzialnością za konsekwencje.

Człowiek może działać w dobrej wierze, stosować uzasadnione metody i mimo to nie osiągnąć oczekiwanego rezultatu z powodu niezależnych okoliczności. Może również przypadkowo uzyskać korzystny wynik, choć decyzja została podjęta bez odpowiedniej analizy ryzyka lub z naruszeniem standardów. Dlatego ocena moralna i zawodowa nie może sprowadzać się wyłącznie do końcowego rezultatu. Kluczowa jest jakość procesu decyzyjnego, dostępność informacji w chwili podejmowania działania, przewidywalność konsekwencji oraz gotowość do korygowania błędów. Jednocześnie dobre intencje nie są wystarczającym usprawiedliwieniem dla kontynuowania nieskutecznej interwencji, jeśli dostępne dane wskazują na brak wartościowego rezultatu lub narastanie szkód. Odpowiedzialność wymaga zatem zdolności aktualizowania przekonań i modyfikacji sposobu działania w świetle wiarygodnych obserwacji. Przedstawione rozważania pozwalają autorowi książki zaproponować syntezę teorii zmiany jako architektury warunkowej sprawczości.

Jej podstawą jest określenie wartości nadrzędnej, której realizacja stanowi normatywne uzasadnienie działania. Następnie konieczne jest rozpoznanie mechanizmu odpowiedzialnego za powstawanie problemu oraz warunków podtrzymujących jego funkcjonowanie. Kolejny element obejmuje konstrukcję interwencji posiadającej uzasadnioną zdolność oddziaływania na określone ogniwo mechanizmu, wraz z określeniem wymaganych zasobów, kompetencji i form współpracy.

Teoria musi następnie definiować możliwe reakcje uczestników, skutki bezpośrednie, rezultaty pośrednie i długofalowe oddziaływanie, uwzględniając ryzyko wystąpienia konsekwencji niezamierzonych. Szczególnie ważne jest wyznaczenie alternatywnego przebiegu zdarzeń, co umożliwia ocenę rzeczywistego wkładu interwencji, oraz wskazanie danych zdolnych potwierdzać lub podważać poszczególne hipotezy. Wreszcie niezbędne jest powiązanie wiedzy uzyskanej z procesu ewaluacji z możliwością przekształcania działania, zmiany skali bądź zakończenia przedsięwzięcia.

Rozróżnienie między formalną zmianą a realnym rezultatem systemowym

Konstrukcja ta nie jest gotową teorią empiryczną, lecz modelem integrującym normatywne, przyczynowe i instytucjonalne warunki odpowiedzialnego rozwiązywania problemów. Szczególne znaczenie tego modelu ujawnia się w przypadku przedsięwzięć legislacyjnych, gdzie łatwo pomylić formalne przyjęcie regulacji z rzeczywistą zmianą zachowań. Jeśli celem jest poprawa ochrony praw określonej grupy, teoria zmiany musi obejmować nie tylko projekt i uchwalenie ustawy, ale także zdolność adresatów do rozpoznawania własnych uprawnień, dostępność instytucji odpowiedzialnych za ich realizację, sposób reagowania podmiotów zobowiązanych oraz skuteczność mechanizmów kontroli. Przyznanie prawa nie gwarantuje bowiem automatycznie jego realizacji, jeśli osoba uprawniona nie dysponuje odpowiednimi informacjami, zasobami lub możliwością uzyskania ochrony. Podobnie wprowadzenie sankcji może nie przynieść oczekiwanej zmiany zachowań, gdy prawdopodobieństwo wykrycia naruszenia jest niewielkie lub gdy ekonomiczne korzyści z łamania prawa przewyższają przewidywane koszty.

Właściwa teoria legislacyjna wymaga zatem rozpoznania relacji między normą, mechanizmem egzekwowania a rzeczywistym zachowaniem uczestników rynku. Dopiero na tym poziomie można zasadnie twierdzić, że reforma oddziałuje na mechanizm problemu, a nie jedynie zmienia formalną konstrukcję porządku prawnego. Analogiczna trudność występuje w działaniach na rzecz ochrony środowiska, gdzie rezultat społeczny zależy od wzajemnego oddziaływania regulacji, technologii i zachowań gospodarczych. Program ograniczania emisji może prowadzić do wzrostu wykorzystania konkretnych urządzeń, co nie musi oznaczać proporcjonalnego spadku całkowitego obciążenia środowiska.

Kluczowe mogą być tu sposób wytwarzania energii, intensywność użytkowania technologii czy zmiana cen i związana z nią reakcja konsumentów. Szczególnie istotne jest zjawisko efektu odbicia, w którym zwiększenie efektywności wykorzystania zasobu może prowadzić do częściowego wzrostu jego zużycia wskutek obniżenia kosztu jednostkowego usługi. Skala tego zjawiska zależy od rodzaju technologii, zachowań odbiorców i warunków gospodarczych, zatem nie można go traktować jako automatycznej konsekwencji każdej poprawy efektywności. Przykład ten uwypukla konieczność odróżnienia poprawy parametru technicznego od zmiany rezultatu systemowego, której skala zależy również od procesów adaptacyjnych zachodzących poza bezpośrednim zakresem interwencji.

Teoria zmiany jako odpowiedzialność za sprawczość

W ochronie zdrowia teoria zmiany wymaga szczególnego rozróżnienia między skutecznością biologiczną, organizacyjną dostępnością świadczenia a rzeczywistym wynikiem zdrowotnym. Interwencja może posiadać potwierdzoną skuteczność kliniczną, jednak jej praktyczne zastosowanie zależy od terminowości diagnozy, przeszkolenia personelu, infrastruktury oraz zdolności pacjenta do kontynuowania zaleceń. Program profilaktyczny może zwiększać wiedzę uczestników, nie zmieniając ich zachowań, jeśli w otoczeniu brakuje możliwości praktycznego wykorzystania tych informacji. System może również osiągać korzystne rezultaty średnie, nie poprawiając jednocześnie sytuacji osób napotykających największe bariery dostępu.

W każdym przypadku konieczne jest połączenie dowodów na skuteczność działania z analizą warunków jego realnego wykonania. Dlatego współczesne badania nad interwencjami złożonymi coraz częściej pytają nie tylko o efekt końcowy, lecz także o możliwość wdrożenia, trwałość, koszty i zmienność rezultatów między populacjami. Z perspektywy teorii systemów kluczowe jest to, że sama interwencja może przekształcać mechanizm, którego skuteczność wykazano wcześniej. Jeśli program zwiększa dostępność informacji, uczestnicy mogą zmienić sposób podejmowania decyzji, a instytucje zmodyfikować procedury. Reforma systemu finansowania może skłonić organizacje do reorganizacji działań w odpowiedzi na nowe bodźce, podobnie jak nowa technologia zwiększająca produktywność może wymusić zmianę struktury zatrudnienia i zakresu usług.

W rezultacie warunki będące podstawą wcześniejszej prognozy mogą ulec zmianie po wdrożeniu interwencji. Powstaje wtedy problem refleksyjności systemu społecznego, w którym uczestnicy reagują na reguły i przewidywania dotyczące własnego zachowania. Właściwa teoria zmiany nie może być zatem zamkniętym modelem; powinna uwzględniać nowe sprzężenia zwrotne i aktualizować założenia, zachowując rygor dokumentowania hipotez. W tym świetle kluczowe staje się rozróżnienie między rozwiązaniem problemu a zmianą sposobu, w jaki system wytwarza określoną trudność. Pierwsze może polegać na ograniczeniu skutków zjawiska, drugie zaś wymaga przekształcenia mechanizmów powodujących jego powtarzanie się. Program pomocy może łagodzić konsekwencje kryzysu, nie ograniczając jego przyczyn strukturalnych. Instytucja może sprawnie rozpatrywać skargi, nie zmniejszając częstotliwości naruszeń, które do nich prowadzą. Przedsiębiorstwo może poprawiać system napraw gwarancyjnych, nie usuwając wad konstrukcyjnych produktu.

Choć każde z tych działań ma wartość społeczną, ich znaczenie jest odmienne z perspektywy długofalowego rozwiązywania problemu. Właściwa teoria zmiany wymaga uznania, że nie każda interwencja musi prowadzić do reformy systemowej, ale każda powinna mieć jasno określony

zakres odpowiedzialności i nie przedstawiać łagodzenia skutków jako pełnego usunięcia mechanizmu przyczynowego. Najważniejszy wniosek z tych rozważań dotyczy konieczności zachowania epistemicznej pokory wobec rezultatów własnego działania. Instytucja może wierzyć, że jej przedsięwzięcie przyczyniło się do zmiany, lecz powinna odróżniać wiarygodność tego twierdzenia od jego społecznej atrakcyjności. Organizacja może odnosić sukcesy, nie będąc w stanie precyzyjnie oszacować własnego udziału w każdym z nich. Może wykazać wysoką skuteczność interwencji, nie rozstrzygając, czy podobne efekty wystąpią przy zmianie skali lub środowiska. Rzetelność naukowa wymaga przedstawiania zakresu dowodów, niepewności szacunków oraz warunków uogólniania rezultatów. Nie jest to oznaka słabości, lecz zdolność krytycznej oceny, która pozwala podejmować trafniejsze decyzje i ogranicza ryzyko marnowania zasobów na działania bez odpowiedniego uzasadnienia.

Teoria zmiany może być zatem uznana za szczególną formę odpowiedzialności za własną sprawczość. Człowiek lub organizacja nie ogranicza się do wskazania wartościowego celu czy wykazania intensywności pracy, lecz zobowiązuje się wyjaśnić, jak konkretne działania mają prowadzić do realnej zmiany, jakie warunki są niezbędne dla ich skuteczności i jakie obserwacje pozwolą ocenić rezultat. Odpowiedzialność ta nie wymaga nieomylności, lecz uczciwego rozpoznawania ograniczeń wiedzy i gotowości do korekty postępowania, gdy rzeczywistość podważa przewidywania. Nie oznacza to jednak bezwarunkowego podporządkowania celów kryteriom ilościowym, gdyż część wartości wymaga złożonych procedur oceny i wiąże się z prawami, których znaczenia nie da się sprowadzić do prostego rachunku korzyści.

Teoria zmiany jako fundament systemowej sprawczości

Właściwa teoria zmiany jest próbą pogodzenia skuteczności z odpowiedzialnością, naukowego rygoru z praktyczną potrzebą działania oraz rozpoznania ludzkich potrzeb z realnymi możliwościami ich urzeczywistniania. Przeprowadzona analiza prowadzi do wniosku, że kluczowym problemem współczesnej nauki o zmianie nie jest już samo wskazanie instrumentów poprawiających konkretny wskaźnik, lecz rozpoznanie pełnej architektury przyczynowej i instytucjonalnej, w której interwencja może doprowadzić do trwałego przekształcenia rzeczywistości. Wymaga to łączenia metod kontrfaktycznych, teorii mechanizmów, analizy wkładu, ewaluacji procesów i refleksji aksjologicznej, przy pełnej świadomości ograniczeń każdego z tych narzędzi.

Działanie skuteczne w jednym środowisku może tracić wartość w innym, a rezultat korzystny krótkookresowo nie musi oznaczać trwałej poprawy. Teoria zmiany nie jest zatem zamkniętą

receptą, lecz instrumentem ciągłego uczenia się o relacji między ludzkimi intencjami, strukturą systemów a rzeczywistymi konsekwencjami działań. Dopiero na tym poziomie możliwe staje się przejście do problematyki zmiany systemowej, rozumianej nie jako retoryczna deklaracja potrzeby głębokiej reformy, lecz jako celowe przekształcanie reguł, bodźców, relacji władzy i mechanizmów reprodukowania określonych zachowań. W tym kontekście kluczowe staje się zderzenie instytucjonalizmu Douglassa Northa, badań Elinor Ostrom nad samoorganizacją, teorii biurokracji pierwszego kontaktu Michaela Lipsky'ego, koncepcji sprzężeń zwrotnych Donelli Meadows oraz teorii zależności od ścieżki.

Wiedza o rzeczywistej przyczynowości pozwala nie tylko doskonalić pojedyncze interwencje, ale i zrozumieć, dlaczego instytucje utrzymują reguły generujące problemy, nawet gdy ich uczestnicy wiedzą o niekorzystnych skutkach dotychczasowych działań. Przejście od teorii zmiany do teorii zmiany systemowej wymaga odróżnienia lokalnego sukcesu od przekształcenia struktury możliwości dostępnych uczestnikom. Organizacja może stworzyć rozwiązanie skuteczne dla niewielkiej grupy, nie mając wpływu na reguły utrudniające jego szersze upowszechnienie. Jednostka może przezwyciężyć własną trudność, podczas gdy inni nadal będą napotykać te same bariery. Instytucja może poprawić jakość realizacji procedury, nie usuwając przyczyn, dla których ta procedura generuje zbędne obciążenia. W takich warunkach kluczowe jest rozpoznanie poziomu sprawczości oraz właściwego miejsca interwencji w strukturze systemu. Istotne staje się ustalenie, czy osiągnięta poprawa może zostać utrwalona przez zmianę norm, procedur, podziału zasobów lub sposobu reprezentowania interesów, czy pozostaje zależna od wyjątkowego wysiłku pojedynczych osób. Zagadnienie rozwiązywania problemów powraca więc do swojej podstawowej intuicji, lecz w wersji wzbogaconej o aparat nauk empirycznych, filozofii przyczynowości i teorii instytucjonalnej.

Nie wystarczy rozpoznać węzeł, znaleźć pozornie ukrytą szpilkę czy wykazać skuteczność pojedynczego sposobu jej usunięcia. Konieczne jest ustalenie, dlaczego system wytworzył określoną strukturę uwikłania, jakie mechanizmy podtrzymują jej trwałość, kto korzysta z obowiązujących reguł oraz czy proponowana zmiana nie doprowadzi do powstania nowych zależności, równie trudnych do przezwyciężenia. Rozwiązanie staje się wówczas zadaniem nie tylko poznawczym i technicznym, lecz również instytucjonalnym i moralnym, gdyż wymaga odpowiedzialnego przekształcania warunków, w których ludzie podejmują decyzje i współtworzą wspólną rzeczywistość.

> Czy potrafimy odróżnić moment, w którym faktycznie zmieniamy świat, od chwili, gdy jedynie sprawnie zarządzamy procesem jego powolnego niszczenia? Prawdziwe rozwiązanie problemu nie polega na znalezieniu jednej, magicznej szpilki do rozplątania

węzła, lecz na zrozumieniu, dlaczego system w ogóle go zawiązał. Być może największą formą sprawczości jest zatem nie pewność co do sukcesu, lecz pokora wobec mechanizmów, które czynią nas jedynie pasażerami własnych reform.

Inspiracje

[1] 

"The Knot Problems Can Be Solved" Seth Godin

Hachette UK | ISBN: 9781399645072

Seth W. Godin (ur. 10 lipca 1960) to amerykański autor, przedsiębiorca i mówca publiczny, powszechnie znany ze swojego przełomowego wkładu w rozwój nowoczesnego marketingu i przywództwa biznesowego. Po uzyskaniu licencjatu z informatyki i filozofii na Uniwersytecie Tufts oraz tytułu MBA w Stanford Graduate School of Business, Godin założył Yoyodyne, pionierską firmę marketingu bezpośredniego online, przejętą przez Yahoo! w 1998 roku, gdzie pełnił funkcję wiceprezesa ds. marketingu bezpośredniego. Później założył platformę Squidoo i uruchomił altMBA, intensywny internetowy warsztat dla liderów. Godin zrewolucjonizował strategię komercyjną, wprowadzając pionierski marketing za przyzwoleniem, który promował przewidywalne i oparte na konsensusie zaangażowanie konsumentów zamiast tradycyjnej reklamy przerywającej. W ponad dwudziestu bestsellerach ukształtował współczesne rozumienie innowacji, przywództwa plemiennego i pracy twórczej.

Seth W. Godin (born July 10, 1960) is an American author, entrepreneur, and public speaker widely recognized for his transformative contributions to modern marketing and business leadership. After earning a BS in computer science and philosophy from Tufts University and an MBA from Stanford Graduate School of Business, Godin founded Yoyodyne, a pioneering online direct-marketing firm acquired by Yahoo! in 1998, where he served as Vice President of Direct Marketing. He later founded the platform Squidoo and launched altMBA, an intensive online leadership workshop. Godin revolutionized commercial strategy by pioneering permission marketing, which advocated for anticipated and consensual consumer engagement over traditional interruption advertising. Across more than twenty bestselling books, he has shaped contemporary understanding of innovation, tribal leadership, and creative work.

altMBA

Literatura uzupełniająca

Książki

Książki autora i dzieła przywołane w tekście:



[1] "This Is Strategy"

Seth Godin

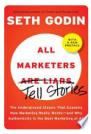
2024 | Simon and Schuster | ISBN: 9798893310306



[2] "Whatcha Gonna Do With That Duck?"

Seth Godin

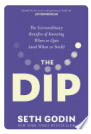
2013 | Penguin UK | ISBN: 9780670923007



[3] "All Marketers are Liars"

Seth Godin

2009 | Penguin | ISBN: 9781101184547



[4] "The Dip"

Seth Godin

2011 | Hachette UK | ISBN: 9780748129270



[5] "We are All Weird"

Seth Godin

2011 | Amazon Publishing | ISBN: 9781936719228

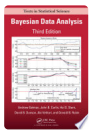


[6] "Purple Cow"

Seth Godin

2005 | Penguin UK | ISBN: 9780141924861

Książki autorów przywoływanych i literatura pokrewna:



[7] "Bayesian Data Analysis, Third Edition"

Donald Rubin

CRC Press | ISBN: 9781040062340

[8] "The Limits to Growth: The 30-Year Update"

Donella Meadows

[9] "The Miracle of Theism"

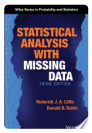
John Leslie Mackie



[10] "Ethics: Inventing Right and Wrong"

John Leslie Mackie

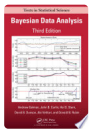
Penguin Group



[11] "Statistical analysis with missing data"

Donald Rubin

John Wiley & Sons | ISBN: 9781118596012



[12] "Bayesian Data Analysis"

Donald Rubin

CRC Press | ISBN: 9781439898208

[13] "Synthetic Control Methods for Comparative Case Studies"

Alberto Abadie, Jens Hainmueller

2007

[14] "Data on the paper: Assessing the validity of the Manifesto Common Space Scores"

Jens Hainmueller



[15] "Mostly Harmless Econometrics"

Joshua Angrist

Princeton University Press | ISBN: 9780691120355

[16] "Letter from John Mayne to W. Grierson"

John Mayne

1819

[17] "Letter from John Mayne to Henry Brougham"

John Mayne

1812

[18] "Letter from John Mayne to Sir Walter Scott"

John Mayne

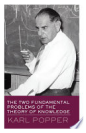
1831



[19] "Dissemination and Implementation Research in Health"

Enola Proctor

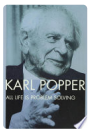
Oxford University Press | ISBN: 9780197660713



[20] "The Two Fundamental Problems of the Theory of Knowledge"

Karl Popper

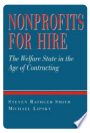
Routledge | ISBN: 9781135626761



[21] "All Life is Problem Solving"

Karl Popper

Routledge | ISBN: 9781135973056



[22] "Nonprofits for Hire: the Welfare State in the Age of Contracting"

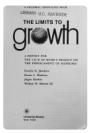
Michael Lipsky

Harvard University Press | ISBN: 9780674043817

[23] "Toward a Theory of Street-level Bureaucracy"

Michael Lipsky

1969



[24] "The Limits to Growth: A report for the Club of Rome's Project on the Predicament of Mankind"

Donella Meadows

New York : Universe Books

Artykuły naukowe

Prace autorów przywoływanych w tekście:

[1] "Evaluation : methods for studying programs and policies"

Carol Hirschon Weiss

1997

Źródło

[2] "Theory-based evaluation: Past, present, and future"

Carol Hirschon Weiss

1997 | New Directions for Evaluation | DOI: 10.1002/ev.1086

Źródło

[3] "EVALUATION RESEARCH: METHODS OF ASSESSING PROGRAM EFFECTIVENESS. By Carol H. Weiss. Englewood Cliffs: Prentice-Hall, 1972. 160 pp. $3.95"

Patricia P. Rieker, Carol Hirschon Weiss

1975 | Social Forces | DOI: 10.1093/sf/54.1.284

Źródło

[4] **"How Can Theory-Based Evaluation Make Greater Headway?"**
Carol Hirschon Weiss
1997 | Evaluation Review | DOI: 10.1177/0193841x9702100405
Źródło

[5] **"Theory-driven evaluations"**
Huey-Tsyh Chen
1990
Źródło

[6] **"Foundations of Program Evaluation: Theories of Practice."**
Huey-Tsyh Chen, William R. Shadish, Thomas D. Cook, Laura C. Leviton
1992 | Contemporary Sociology A Journal of Reviews | DOI: 10.2307/2074800
Źródło

[7] **"Evaluating With Sense: The Theory-Driven Approach"**
Huey-Tsyh Chen, Peter H. Rossi
1983 | Evaluation Review | DOI: 10.1177/0193841x8300700301
Źródło

[8] **"Practical program evaluation : assessing and improving planning, implementation, and effectiveness"**
Huey-Tsyh Chen
2005
Źródło

[9] **"Qualitative Data Analysis"**
Delwyn Goodrick, Patricia Rogers
2015 | Wiley eBooks | DOI: 10.1002/9781119171386.ch22
Źródło

[10] **"Using Programme Theory to Evaluate Complicated and Complex Aspects of Interventions"**
Patricia Rogers
2008 | Evaluation | DOI: 10.1177/1356389007084674
Źródło

[11] **"Purposeful Program Theory: Effective Use of Theories of Change and Logic Models"**

Sue Funnell, Patricia Rogers

2011 | RMIT Research Repository (RMIT University Library)

Źródło

[12] **"Program theory evaluation: Practice, promise, and problems"**

Patricia Rogers, Anthony J. Petrosino, Tracy A. Huebner, Timothy A. Hacsi

2000 | New Directions for Evaluation | DOI: 10.1002/ev.1177

Źródło

[13] **"A new framework for developing and evaluating complex interventions: update of Medical Research Council guidance"**

Kathryn Skivington, Lynsay Matthews, Sharon Anne Simpson, Peter Craig, Janis Baird

2021 | BMJ | DOI: 10.1136/bmj.n2061

Źródło

[14] **"Framework for the development and evaluation of complex interventions: gap analysis, workshop and consultation-informed update"**

Kathryn Skivington, Lynsay Matthews, Sharon Anne Simpson, Peter Craig, Janis Baird

2021 | Health Technology Assessment | DOI: 10.3310/hta25570

Źródło

[15] **"Effectiveness of workplace interventions in the prevention of upper extremity musculoskeletal disorders and symptoms: an update of the evidence"**

Dwayne Van Eerd, Claire Munhall, Emma Irvin, David M. Rempel, Shelley Brewer, Kathryn Skivington

2015 | Occupational and Environmental Medicine | DOI: 10.1136/oemed-2015-102992

Źródło

[16] **"COVID-19 pandemic and its impact on social relationships and health"**

Emily Long, Susan Patterson, Karen Maxwell, Carolyn Blake, Raquel Bosó Pérez, Kathryn Skivington

2021 | Journal of Epidemiology & Community Health | DOI: 10.1136/jech-2021-216690

Źródło

[17] **"IX. On the problem of the most efficient tests of statistical hypotheses"**

Jerzy Neyman, E. S. Pearson

1933 | Philosophical Transactions of the Royal Society of London Series A Containing Papers of a Mathematical or Physical Character | DOI: 10.1098/rsta.1933.0009

Źródło

[18] **"On the Application of Probability Theory to Agricultural Experiments. Essay on Principles. Section 9"**

Jerzy Neyman, D. M. Dabrowska, Terence Paul Speed

1990 | Statistical Science | DOI: 10.1214/ss/1177012031

Źródło

[19] **"ON THE PROBLEM OF THE MOST EFFICIENT TESTS OF STATISTICAL HYPOTHESES"**

Jerzy Neyman, E. S. Pearson

1967 | University of California Press eBooks | DOI: 10.1525/9780520339897-008

Źródło

[20] **"On the Two Different Aspects of the Representative Method: The Method of Stratified Sampling and the Method of Purposive Selection"**

Jerzy Neyman

1934 | Journal Of The Royal Statistical Society | DOI: 10.2307/2342192

Źródło

[21] **"Maximum Likelihood from Incomplete Data Via the EM Algorithm"**

A. P. Dempster, Nan M. Laird, Donald B. Rubin

1977 | Journal of the Royal Statistical Society Series B (Statistical Methodology) | DOI: 10.1111/j.2517-6161.1977.tb01600.x

Źródło

[22] **"The Central Role of the Propensity Score in Observational Studies for Causal Effects"**

Paul R. Rosenbaum, Donald B. Rubin

1983 | Biometrika | DOI: 10.1093/biomet/70.1.41

Źródło

[23] **"Bayesian Data Analysis"**

Andrew Gelman, John B. Carlin, Hal S. Stern, David B. Dunson, Aki Vehtari, Donald B. Rubin

1995 | Chapman & Hall/CRC eBooks | DOI: 10.1201/9780429258480

Źródło

[24] "Statistical Analysis With Missing Data"

Maureen Lahiff, Roderick J. A. Little, Donald B. Rubin

1989 | Journal of the American Statistical Association | DOI: 10.2307/2289883

Źródło

[25] "Demand analysis and tax reform in Pakistan"

Angus Deaton

1992

Źródło

[26] "Through the Darkness to a Brighter Future"

Angus Deaton

2014 | The MIT Press eBooks | DOI: 10.7551/mitpress/9404.003.0004

Źródło

[27] "The limitations of randomized controlled trials in predicting effectiveness"

Nancy Cartwright, Eileen Munro

2010 | Journal of Evaluation in Clinical Practice | DOI: 10.1111/j.1365-2753.2010.01382.x

Źródło

[28] "Mechanisms, laws and explanation"

N. Cartwright, J. Pemberton, S. Wieten

2020 | European Journal for Philosophy of Science | DOI: 10.1007/s13194-020-00284-y

Źródło

[29] "Difference-in-Differences Designs: A Practitioner's Guide"

Andrew Baker, Brantly Callaway, Scott Cunningham, Andrew Goodman-Bacon, Pedro H. C. Sant'Anna

2026 | Journal of Economic Literature | DOI: 10.1257/jel.20251650

Źródło

[30] "Quantile Treatment Effects in Difference in Differences Models with Panel Data"

Brantly Callaway, Tong Li

2017 | Quantitative Economics | DOI: 10.3982/qe935

Źródło

[31] "Quantile treatment effects in difference in differences models under dependence restrictions and with only two time periods"

Brantly Callaway, Tong Li, Tatsushi Oka

2018 | Journal of Econometrics | DOI: 10.1016/j.jeconom.2018.06.008

Źródło

[32] "Difference-in-Differences when Parallel Trends Holds Conditional on Covariates"

Carolina Caetano, Brantly Callaway

2024 | arXiv (Cornell University) | DOI: 10.48550/arxiv.2406.15288

Źródło

[33] "Efficient Estimation for Staggered Rollout Designs"

Jonathan Roth, Pedro H. C. Sant'Anna

2023 | Journal of Political Economy Microeconomics | DOI: 10.1086/726581

Źródło

[34] "Elections that inspire: Effects of Black mayors on educational attainment"

J. Ikawa, C. Martins, Pedro H. C. Sant'Anna, Rogerio Santarrosa

2026 | Journal of Public Economics | DOI: 10.1016/j.jpubeco.2026.105639

Źródło

[35] "Regression Discontinuity Designs in Economics"

David S Lee, Thomas Lemieux

2010 | Journal of Economic Literature | DOI: 10.1257/jel.48.2.281

Źródło

[36] "Regression discontinuity designs: A guide to practice"

Guido W. Imbens, Thomas Lemieux

2007 | Journal of Econometrics | DOI: 10.1016/j.jeconom.2007.05.001

Źródło

[37] "Unconditional Quantile Regressions"

SERGIO PINHEIRO FIRPO, Nicole M. Fortin, Thomas Lemieux

2009 | Econometrica | DOI: 10.3982/ecta6822

Źródło

[38] "Decomposition Methods in Economics"

Nicole M. Fortin, Thomas Lemieux, SERGIO PINHEIRO FIRPO

2011 | Handbook of labour economics | DOI: 10.1016/s0169-7218(11)00407-2

Źródło

[39] "Synthetic Control Methods for Comparative Case Studies: Estimating the Effect of California's Tobacco Control Program"

Alberto Abadie, Alexis Diamond, Jens Hainmueller

2010 | Journal of the American Statistical Association | DOI: 10.1198/jasa.2009.ap08746

Źródło

[40] "The Economic Costs of Conflict: A Case Study of the Basque Country"

Alberto Abadie, Javier Gardeazábal

2003 | American Economic Review | DOI: 10.1257/000282803321455188

Źródło

[41] "Large Sample Properties of Matching Estimators for Average Treatment Effects"

Alberto Abadie, Guido W. Imbens

2005 | Econometrica | DOI: 10.1111/j.1468-0262.2006.00655.x

Źródło

[42] "Comparative Politics and the Synthetic Control Method"

Alberto Abadie, Alexis Diamond, Jens Hainmueller

2014 | American Journal of Political Science | DOI: 10.1111/ajps.12116

Źródło

[43] "Genetic Matching for Estimating Causal Effects: A General Multivariate Matching Method for Achieving Balance in Observational Studies"

Alexis Diamond, Jasjeet Singh Sekhon

2012 | The Review of Economics and Statistics | DOI: 10.1162/rest_a_00318

Źródło

[44] "Synth : An R Package for Synthetic Control Methods in Comparative Case Studies"

Alberto Abadie, Alexis Diamond, Jens Hainmueller

2011 | Journal of Statistical Software | DOI: 10.18637/jss.v042.i13

Źródło

[45] "The Long-Term Impact of Employment Bans on the Economic Integration of Refugees"

Moritz Marbach, Jens Hainmueller, Dominik Hangartner

2017 | Science Advances | DOI: 10.1126/sciadv.aap9519

Źródło

[46] "Contribution analysis: Coming of age?"

John Mayne

2012 | Evaluation | DOI: 10.1177/1356389012451663

Źródło

[47] "Does ad hoc language training improve the economic integration of refugees? Evidence from Germany's response to the Syrian refugee crisis"

Moritz Marbach, Ehsan Vallizadeh, Niklas Harder, Dominik Hangartner, Jens Hainmueller

2025 | Journal of the Royal Statistical Society Series A (Statistics in Society) | DOI: 10.1093/jrsssa/qnae106

Źródło

[48] "Identification of Causal Effects Using Instrumental Variables"

Joshua D. Angrist, Guido W. Imbens, Donald B. Rubin

1996 | Journal of the American Statistical Association | DOI: 10.1080/01621459.1996.10476902

Źródło

[49] "Identification and Estimation of Local Average Treatment Effects"

Guido W. Imbens, Joshua D. Angrist

1994 | Econometrica | DOI: 10.2307/2951620

Źródło

[50] "Does Compulsory School Attendance Affect Schooling and Earnings?"

Joshua D. Angrist, Alan B. Krueger

1991 | The Quarterly Journal of Economics | DOI: 10.2307/2937954

Źródło

[51] "Instrumental Variables and the Search for Identification: From Supply and Demand to Natural Experiments"

Joshua D. Angrist, Alan B Krueger

2001 | The Journal of Economic Perspectives | DOI: 10.1257/jep.15.4.69

Źródło

[52] "Realist review - a new method of systematic review designed for complex policy interventions"

Ray Pawson, Trisha Greenhalgh, Gill Harvey, Kieran Walshe

2005 | Journal of Health Services Research & Policy | DOI: 10.1258/1355819054308530

Źródło

[53] "Evidence-based Policy"

Ray Pawson

2006 | SAGE Publications Ltd eBooks | DOI: 10.4135/9781849209120

Źródło

[54] "Evidence-based Policy: The Promise of `Realist Synthesis'"

Ray Pawson

2002 | Evaluation | DOI: 10.1177/135638902401462448

Źródło

[55] "Realist methods in medical education research: what are they and what can they contribute?"

Geoff Wong, Trisha Greenhalgh, Gill Westhorp, Ray Pawson

2011 | Medical Education | DOI: 10.1111/j.1365-2923.2011.04045.x

Źródło

[56] "Introducing EMMIE: an evidence rating scale to encourage mixed-method crime prevention synthesis reviews"

Shane D. Johnson, Nick Tilley, KATE J. BOWERS

2015 | Journal of Experimental Criminology | DOI: 10.1007/s11292-015-9238-7

Źródło

[57] "Understanding Car Parks, Crime and CCTV: Evaluation Lessons from Safer Cities"

Nick Tilley

1993 | OpenGrey (Institut de l'Information Scientifique et Technique)

Źródło

[58] "Caring Communities, Paradigm Polemics, Design Debates"

Ray Pawson, Nick Tilley

1998 | Evaluation | DOI: 10.1177/13563899822208400

Źródło

[59] "The diffusion of police innovation: A case study of problem-oriented policing in England and Wales"

Karen Bullock, A. Sidebottom, G. Laycock, N. Tilley

2022 | International Journal of Police Science & Management | DOI: 10.1177/14613557221106084

Źródło

[60] "Development of methodological guidance, publication standards and training materials for realist and meta-narrative reviews: the RAMESES (Realist And Meta-narrative Evidence Syntheses – Evolving Standards) project"

Geoff Wong, Trish Greenhalgh, Gill Westhorp, Ray Pawson

2014 | Health Services and Delivery Research | DOI: 10.3310/hsdr02300

Źródło

[61] "Protocol - realist and meta-narrative evidence synthesis: Evolving Standards (RAMESES)"

Trisha Greenhalgh, Geoff Wong, Gill Westhorp, Ray Pawson

2011 | BMC Medical Research Methodology | DOI: 10.1186/1471-2288-11-115

Źródło

[62] "Agroecology: Ecological Processes in Sustainable Agriculture. By Stephen R. Gliessman. 1998. CRC Press, Boca Raton, FL. xxi + 357 p. $54.95, ISBN 1-57504-043-3, hardcover."

John C Mayne

2000 | Renewable Agriculture and Food Systems | DOI: 10.1017/s0889189300008559

Źródło

[63] "Addressing Attribution through Contribution Analysis: Using Performance Measures Sensibly"

John Mayne

2001 | Canadian Journal of Program Evaluation | DOI: 10.3138/cjpe.016.001

Źródło

[64] "Contribution Analysis and the Long-term Perspective"

John Mayne

2020 | Long Term Perspectives in Evaluation | DOI: 10.4324/9781003058250-4

Źródło

[65] "Outcomes for Implementation Research: Conceptual Distinctions, Measurement Challenges, and Research Agenda"

Enola K. Proctor, Hiie Silmere, Ramesh Raghavan, Peter S. Hovmand, Gregory A. Aarons

2010 | Administration and Policy in Mental Health and Mental Health Services Research | DOI: 10.1007/s10488-010-0319-7

Źródło

[66] "A refined compilation of implementation strategies: results from the Expert Recommendations for Implementing Change (ERIC) project"

Byron J. Powell, Thomas J. Waltz, Matthew Chinman, Laura J. Damschroder, Jeffrey L. Smith, Enola K. Proctor

2015 | Implementation Science | DOI: 10.1186/s13012-015-0209-1

Źródło

[67] "Implementation strategies: recommendations for specifying and reporting"

Enola K. Proctor, Byron J. Powell, J. Curtis McMillen

2013 | Implementation Science | DOI: 10.1186/1748-5908-8-139

Źródło

[68] "Implementation Research in Mental Health Services: an Emerging Science with Conceptual, Methodological, and Training challenges"

Enola K. Proctor, John A. Landsverk, Gregory A. Aarons, David Chambers, Charles A. Glisson

2008 | Administration and Policy in Mental Health and Mental Health Services Research | DOI: 10.1007/s10488-008-0197-4

Źródło

[69] "Process evaluation of complex interventions: Medical Research Council guidance"

G. F. Moore, S. Audrey, M. Barker, L. Bond, C. Bonell

2015 | BMJ | DOI: 10.1136/bmj.h1258

Źródło

[70] "Process evaluation of complex interventions"

Elizabeth McGill, Graham Moore

2026 | Complex Interventions in Health | DOI: 10.4324/9781003495642-26

Źródło

[71] "Contact networks in the emergency department: Effects of time, environment, patient characteristics, and staff role"

Vicki Hertzberg, Jason L. Baumgardner, C. Christina Mehta, Lisa Elon, George Cotsonis, Douglas W. Lowery-North

2016 | Social Networks | DOI: 10.1016/j.socnet.2016.08.005

Źródło

[72] "People Believe that Morality Particularly Does Not Change"

Jinseok Chun, Michael S. North

2024 | DOI: 10.2139/ssrn.4766721

Źródło

[73] "The violence trap: a political-economic approach to the problems of development"

Gary W. Cox, D. North, Barry R. Weingast

2019 | Journal of Public Finance and Public Choice | DOI: 10.2139/ssrn.2370622

Źródło

[74] "A General Framework for Analyzing Sustainability of Social-Ecological Systems"

Elinor Ostrom B

2009 | Science | DOI: 10.1126/science.1172133

Źródło

[75] "Complexity of Coupled Human and Natural Systems"

Jianguo Liu, Thomas M. Dietz, Stephen R. Carpenter, Marina Alberti, Carl Folke, Elinor Ostrom B

2007 | Science | DOI: 10.1126/science.1144004

Źródło

[76] "Revisiting the Commons: Local Lessons, Global Challenges"

Elinor Ostrom B, Joanna Burger, Christopher B. Field, Richard B. Norgaard, David Policansky

1999 | Science | DOI: 10.1126/science.284.5412.278

Źródło

[77] "Property-Rights Regimes and Natural Resources: A Conceptual Analysis"

Edella C. Schlager, Elinor Ostrom B

1992 | Land Economics | DOI: 10.2307/3146375

Źródło

[78] **"Street-Level Bureaucracy and the Analysis of Urban Reform"**

Michael Lipsky

1971 | Urban Affairs Quarterly | DOI: 10.1177/107808747100600401

Źródło

[79] **"“Perspectives, Problems, and Models” from The Limits to Growth (1972)"**

D. Meadows, D. Meadows, J. Randers, W. Behrens

2014 | DOI: 10.4324/9781315770369-16

Źródło

Treść tworzy Fundacja Dobre Państwo.
Redaguje i publikuje APA ONE, autorski system redakcyjny Fundacji oparty na AI.
APA ONE Publishing System
Fundacja Dobre Państwo © 2026
Article ID: 74c1c9c3-fbbf-4975-8da1-c6ae06f6e627