Kultura uczenia się z porażek i ekonomia alokacji zasobów w zarządzaniu wpływem społecznym na podstawie koncepcji Matta Stevensona-Dodda

AUTOR

Fundacja Dobre Państwo

STRESZCZENIE / ABSTRACT

Tekst analizuje związek między kulturą organizacyjną a jakością danych w zarządzaniu wpływem społecznym. Autor wskazuje, że techniczna infrastruktura pomiaru (np. Theory of Change) jest nieskuteczna bez bezpieczeństwa psychologicznego, które pozwala na otwartość wobec porażek. Bez tego mechanizmu organizacje wpadają w pułapkę „testu Daily Mail", selekcjonując jedynie pozytywne informacje w celu ochrony reputacji. Przywołując badania Amy Edmondson, artykuł podkreśla, że bezpieczeństwo psychologiczne redukuje koszt ujawniania niewygodnych danych, co jest kluczowe dla przejścia od organizacji skupionej na działaniu („Do") do organizacji generującej realną zmianę („Difference"). Wskazuje również na konieczność rozróżnienia błędów podstawowych od inteligentnej eksploracji w procesie uczenia się.

The text analyzes the relationship between organizational culture and data quality in social impact management. The author indicates that technical measurement infrastructure (e.g., Theory of Change) is ineffective without psychological safety, which allows for openness toward failure. Without this mechanism, organizations fall into the 'Daily Mail test' trap, selecting only positive information to protect their reputation. Citing Amy Edmondson's research, the article emphasizes that psychological safety reduces the cost of disclosing uncomfortable data, which is crucial for transitioning from an organization focused on action ('Do') to one generating real change ('Difference'). It also points to the necessity of distinguishing basic errors from intelligent exploration in the learning process.

Artykuł stawia tezę, że prawdziwe zarządzanie wpływem społecznym (impact management) wymaga fundamentalnego przejścia od technicznego raportowania sukcesów do budowy kultury wysokiej rozdzielczości poznawczej. Autor argumentuje, że zaawansowane narzędzia analityczne są bezużyteczne w organizacjach „epistemicznie zamkniętych", gdzie lęk przed utratą reputacji wymusza selekcję danych i ukrywanie niepowodzeń. Kluczowym rozwiązaniem jest stworzenie infrastruktury bezpieczeństwa psychologicznego, która pozwala traktować tzw. inteligentne porażki jako cenne aktywa informacyjne, a nie dowody niekompetencji. Tekst wykracza poza prostą ewaluację, wprowadzając perspektywę ekonomii krańcowej i teorii portfelowej: sugeruje, że maksymalizacja wpływu społecznego to w rzeczywistości problem optymalnej alokacji rzadkich zasobów (czasu, kapitału, uwagi). W tym ujęciu dojrzałość organizacji przejawia się w zdolności do odróżnienia błędów od innowacji oraz odwadze w wygaszaniu programów o niskim wpływie marginalnym, nawet jeśli są one popularne lub historycznie zakorzenione. Ostatecznie wpływ społeczny zostaje zdefiniowany nie jako suma sukcesów, lecz jako świadomy proces zarządzania ryzykiem i wartością w warunkach ograniczeń etycznych i systemowych.

SŁOWA KLUCZOWE

impact management | bezpieczeństwo psychologiczne | inteligentna porażka | alokacja zasobów | Theory of Change | Impact Risk Appetite | negative case analysis | ekonomia alokacji zasobów | organizacja Difference | strategia małych strat | public value | resilience-adjusted efficiency | portfolio theory w sektorze społecznym | accountability for reasonableness

Bezpieczeństwo psychologiczne jako infrastruktura jakości danych

Porażka jako informacja: bezpieczeństwo psychologiczne, inteligentne ryzyko i odpowiedzialność organizacyjna

Aby organizacja mogła rzeczywiście uczyć się z danych, musi posiadać zdolność przyjmowania informacji sprzecznych z własnym interesem reputacyjnym, wcześniejszymi przekonaniami czy narracją sukcesu. To właśnie w tym punkcie techniczna infrastruktura impact management styka się z kulturą organizacyjną. Można bowiem zbudować znakomitą Theory of Change, zintegrować CRM, stworzyć aktualizowane w czasie niemal rzeczywistym dashboardy i zastosować walidowane mierniki, a mimo to pozostać organizacją epistemicznie zamkniętą, jeśli niekorzystny wynik wywołuje przede wszystkim poszukiwanie winnego, obronę reputacji albo reinterpretację danych. Porażka ma wówczas charakter podwójny: najpierw nie udaje się interwencja, a następnie organizacji nie udaje się dowiedzieć, dlaczego tak się stało.

Stevenson-Dodd traktuje ten problem jako jeden z konstytutywnych elementów przejścia od organizacji „Do" do organizacji „Difference". Organizacja raportująca wyłącznie sukcesy pozbawia się bowiem części danych niezbędnych do uczenia się i racjonalnej alokacji zasobów. Autor łączy tę deformację z reputacyjnym "Daily Mail Test": antycypacją tego, jak publiczna informacja o nieskutecznym projekcie, nieosiągniętym rezultacie albo utraconych środkach mogłaby zostać przedstawiona przez media, donatorów czy organy nadzorcze. W tej logice niepowodzenie programu staje się nie tyle materiałem empirycznym, co zagrożeniem dla legitymizacji. Powstaje systemowy bodziec do selekcji informacji pozytywnych, łagodzenia języka oraz przenoszenia ciężaru z pytania o skuteczność na pytanie o obronę organizacji.

Najmocniejszego teoretycznego uzupełnienia tej intuicji dostarcza dorobek Amy Edmondson. W klasycznym badaniu opublikowanym w **Administrative Science Quarterly** w 1999 r. zdefiniowała ona **team psychological safety** jako podzielane przez członków zespołu przekonanie, że środowisko jest bezpieczne dla podejmowania interpersonalnego ryzyka. Analiza 51 zespołów wykazała, że bezpieczeństwo psychologiczne wiąże się z zachowaniami służącymi uczeniu się — takimi jak poszukiwanie informacji zwrotnej, omawianie błędów, eksperymentowanie i refleksja nad praktyką; to one pośredniczyły w relacji między poczuciem bezpieczeństwa a wynikami zespołu. Interpretacja tej koncepcji wymaga jednak precyzji: psychological safety nie oznacza środowiska, w którym wszystko jest akceptowane, a nikt nie ponosi odpowiedzialności. Nie jest również synonimem miłej atmosfery.

Jego funkcja jest epistemiczna: redukuje interpersonalny koszt ujawnienia informacji niewygodnych, niepopularnych albo obciążających dla mówiącego. Pracownik powinien móc przyznać, że nie rozumie zadania, popełnił błąd, dostrzega zagrożenie, nie zgadza się z przełożonym lub uważa dominującą teorię za błędną — bez obawy, że sam akt ujawnienia problemu zostanie automatycznie uznany za dowód niekompetencji lub nielojalności. W organizacji mierzącej wpływ

społeczny znaczenie tej kategorii jest jeszcze większe niż w klasycznym środowisku produkcyjnym, ponieważ dane o outcomes nie są neutralnymi parametrami technicznymi.

Mogą one kwestionować wieloletnie programy, sens konkretnych stanowisk pracy, decyzje zarządu, strategie fundatora czy reputację organizacji. Analityk wykrywający brak efektu znajduje się więc w sytuacji interpersonalnie ryzykownej. Jeżeli struktura organizacyjna premiuje potwierdzanie sukcesu, najbardziej wartościowa informacja może stać się jednocześnie tą najbardziej niebezpieczną dla osoby, która ją przekazuje. Z tego względu bezpieczeństwo psychologiczne należy traktować jako element **infrastruktury jakości danych**. W systemie o niskim poziomie bezpieczeństwa błędy nie znikają; stają się jedynie mniej widoczne. Ludzie uczą się zmieniać język, opóźniać eskalację, selekcjonować korzystne obserwacje lub przedstawiać zły wynik jako chwilowe odchylenie. Dashboard może wtedy zachowywać wszystkie cechy technicznej poprawności, a mimo to stanowić wyłącznie wyrafinowaną reprezentację kultury obronnej.

Prowadzi to do istotnego paradoksu: organizacja może być bardziej przejrzysta technicznie, a jednocześnie mniej przejrzysta poznawczo. Im więcej informacji gromadzonych jest automatycznie, tym większa staje się rola tego, **jak są one interpretowane i które z nich uzyskują status problemu wymagającego działania**. Transparentność danych bez bezpieczeństwa interpretacji może prowadzić do nowego rodzaju rytualizmu: wszystko jest widoczne, ale niczego nie wolno właściwie nazwać.

Nie każda porażka jest lekcją: rozróżnienie błędów od inteligentnej eksploracji

Współczesna teoria uczenia się z niepowodzeń wyraźnie odcina się od populistycznego hasła "celebrate failure". Amy Edmondson rozróżnia trzy klasy porażek: podstawowe, złożone i inteligentne. Basic failure zachodzi w obszarze dobrze rozpoznanym – gdy proces jest znany, lecz pojawia się błąd, którego można było uniknąć. Complex failure wynika z konfiguracji wielu czynników, z których żaden samodzielnie nie doprowadziłby do negatywnego rezultatu. Intelligent failure powstaje natomiast podczas przemyślanej eksploracji nowego obszaru, w którym nie ma gotowych odpowiedzi; działanie ma tu wyraźny cel i hipotezę, a zakres eksperymentu jest ograniczony tak, by koszt niepowodzenia był proporcjonalny do wartości potencjalnej wiedzy.

Rozróżnienie to jest kluczowe dla impact management, gdyż chroni przed dwiema przeciwstawnymi patologiami. Pierwszą jest failure aversion – przekonanie, że każde

niepowodzenie świadczy o złym zarządzaniu. Drugą zaś failure romanticism – wiara, że każda porażka jest wartościowym doświadczeniem i zasługuje na pochwałę.

Oba podejścia są metodologicznie wadliwe. Błąd wynikający z rażącego lekceważenia procedur bezpieczeństwa nie staje się "eksperymentem" tylko dlatego, że organizacja może wyciągnąć z niego wnioski. Powtarzanie dobrze znanego błędu nie jest innowacją, a brak staranności nie zyskuje wartości poznawczej dzięki zmianie nazwy na "learning opportunity". Sama możliwość wyciągnięcia lekcji nie jest zatem wystarczającym kryterium „dobrej" porażki. Z każdej katastrofy można czegoś się nauczyć, lecz nie każda była racjonalnie uzasadnionym sposobem pozyskiwania wiedzy. Właśnie tutaj przebiega granica między ryzykiem poznawczym a zaniedbaniem.

Inteligentny eksperyment powinien dotyczyć rzeczywist〈 nierozpoznanego obszaru, odpowiadać na istotną niewiadomą i opierać się na dostępnej wiedzy. Musi posiadać jawne założenia oraz ograniczać ekspozycję ludzi i zasobów do poziomu niezbędnego do uzyskania informacji. Edmondson ujmuje tę logikę podobnie: inteligentna porażka powinna dotyczyć nowego terytorium, być ukierunkowana na ważny cel, wynikać z przemyślanej hipotezy oraz mieć jak najmniejszy zakres i koszt. Dla sektora społecznego ostatnie kryterium ma szczególne znaczenie, ponieważ „kosztem eksperymentu" nie są tylko pieniądze, lecz zdrowie, bezpieczeństwo, zaufanie, czas i prawa osób w słabszej pozycji. Etyka eksperymentowania społecznego wymaga więc ostrzejszej dyscypliny niż testowanie nowego interfejsu.

Już w 1992 roku Sim Sitkin sformułował koncepcję learning through failure: the strategy of small losses. Jego argument był jak na ówczesne standardy zarządzania radykalny: porażka może być koniecznym warunkiem adaptacji, podczas gdy powtarzający się sukces – choć zwiększa stabilność – może osłabiać poszukiwanie alternatyw i wrażliwość na informacje sprzeczne z utrwalonym modelem działania. Koncepcję tę warto połączyć z teorią zależności od ścieżki. Sukces wzmacnia nie tylko przekonanie o skuteczności działań, ale buduje wokół nich strukturę interesów. Program, który przyniósł prestiż, przychody i rozpoznawalność, jest znacznie trudniejszy do zakwestionowania niż eksperyment marginalny. Sukces tworzy więc epistemiczną rentę zasiedzenia: dotychczasowe rozwiązanie uzyskuje domniemanie słuszności, nawet jeśli zmienił się kontekst. Porażka może przerwać ten mechanizm, ale tylko wtedy, gdy pozostaje dostatecznie mała i informacyjna.

Sitkinowska strategia małych strat nie polega na mnożeniu niepowodzeń, lecz na projektowaniu organizacji tak, by wykrywać błędne hipotezy, zanim koszt ich kontynuacji stanie się zbyt wysoki. Jest to bliskie logice współczesnego eksperymentowania: lepiej uzyskać szybki i ograniczony negatywny wynik, niż inwestować wieloletni budżet w teorię, której podstawowych założeń nigdy nie przetestowano. W sektorze społecznym zasada ta ma szczególną doniosłość. Często spotykamy

programy utrzymywane latami dlatego, że zostały zinstytucjonalizowane, a nie dlatego, że każda kolejna analiza potwierdza ich skuteczność. Mały pilotaż pozwalający wcześnie odrzucić słabą teorię interwencji może być zatem formą odpowiedzialności finansowej i moralnej, a nie jej zaprzeczeniem.

Autor książki szczególną wagę przypisuje przypadkowi Street League. Opisuje on raport z 2016 roku, w którym organizacja otwarcie przyznała, że nie była w stanie pomóc 109 młodym ludziom, oraz wskazała rezultaty odrzucone podczas wewnętrznej weryfikacji ze względu na brak rygorów dowodowych. Niezależne publiczne omówienie raportu potwierdza kluczowy element tego przykładu: Street League rzeczywiście wyeksponowała liczbę 109 uczestników, którym nie zdołała zapewnić zakładanego outcome'u, obok 1281 osób, które osiągnęły wynik związany z zatrudnieniem lub szkoleniem.

Analiza przypadków negatywnych jako źródło wiedzy

W publikacji opisano trzy reguły raportowania organizacji: zakaz overclaimingu, podawanie liczb bezwzględnych obok procentów oraz wspieranie outcomes audytowalnymi dowodami. Pozornie mamy tu do czynienia jedynie z przykładem przejrzystości, lecz w istocie sprawa jest znacznie głębsza. W typowej narracji fundraisingowej przypadki niepowodzenia są traktowane jako statystycznym odpadem; Street League potraktowała je natomiast jako część pełnej dystrybucji rezultatów.

Odbiorca raportu nie otrzymał wyłącznie odpowiedzi na pytanie „komu pomogliśmy?", lecz również „komu nie pomogliśmy?". To przesunięcie jest kluczowe dla naukowego rozumienia wpływu, ponieważ wynik negatywny może być bardziej diagnostyczny niż kolejne potwierdzenie sukcesu. Jeżeli tysiąc osób osiąga rezultat, a sto nie, podstawowe pytanie zarządcze nie brzmi jedynie o to, czy wskaźnik sukcesu jest „wysoki". Należy ustalić, czy osoby bez rezultatu posiadają wspólną charakterystykę. Być może mechanizm nie działa wobec określonej grupy, wymagania logistyczne są zbyt wysokie lub kryteria rekrutacji powodują wykluczenie.

Możliwe, że program jest skuteczny dla 90% populacji i strukturalnie nieskuteczny dla pozostałych 10%. Właśnie ta mniejszość może ujawnić granicę Theory of Change. Autor książki ilustruje ten problem przeciwstawieniem historii sukcesu przypadkowi młodego człowieka, którego udział w programie został ograniczony przez kontekst związany z bezpieczeństwem i lokalizacją. Metodologicznie jest to przykład negative case analysis: zamiast usuwać odstępstwo od dominującego wzorca, wykorzystujemy je do zbadania warunków, w których teoria przestaje

działać. W naukach społecznych przypadek negatywny ma szczególną wartość, gdyż kwestionuje zbyt szerokie generalizacje.

Dziewięćdziesiąt przypadków zgodnych z modelem może zwiększać pewność, lecz jeden przypadek systematycznie niepasujący może ujawnić pominięty warunek konieczny. Organizacja „Difference" powinna więc posiadać nie tylko bazę sukcesów, lecz także rejestr kontrprzypadków.

Podobną funkcję pełni przywołany przez autora model CLIC Sargent (obecnie Young Lives vs Cancer). Opisuje on rozdzielenie raportowania na „The Good", „The Bad" oraz „The Ugly": rezultaty udane, błędy lub niedoskonałości wewnętrzne organizacji oraz zewnętrzne ograniczenia systemowe, nad którymi organizacja nie miała kontroli. Oficjalny komunikat organizacji potwierdza, że 29 października 2019 r. CLIC Sargent opublikowała raport **Impact and Accountability Report - The Good, The Bad and The Ugly**, kontynuując model otwartego raportowania o własnych niedoskonałościach. Analityczna wartość tej formuły polega nie tyle na efektownych nazwach kategorii, ile na alokacji przyczynowej odpowiedzialności. Kategorie „Bad" i „Ugly" nie powinny być ze sobą mylone.

Jeżeli organizacja miała kontrolę nad przyczyną problemu i popełniła błąd, właściwą odpowiedzią jest korekta procesu. Jeśli zaś przyczyna była zewnętrzna – recesja, zmiana regulacyjna, pandemia, niedobór mieszkań czy zerwanie łańcucha usług publicznych – odpowiedź wymaga adaptacji do kontekstu, współpracy systemowej albo realistycznego ograniczenia roszczeń attribution. Rozróżnienie to ma chronić przed dwoma rodzajami moral hazard: przerzucaniem każdego niepowodzenia na „system" oraz przypisywaniem organizacji odpowiedzialności za rezultaty pozostające poza jej realną sferą kontroli. Rzetelna analiza failure powinna zatem rozdzielać kontrolowalność, przewidywalność i odpowiedzialność. Nie wszystko, czego nie kontrolowano, było nieprzewidywalne, a nie wszystko, co przewidywalne, można było skutecznie wyeliminować. Przykładowo: recesja jest szokiem zewnętrznym, ale organizacja mogła mieć obowiązek przygotowania scenariuszy odporności. Niepowodzenie wynikające z samego szoku należy do „Ugly", lecz brak planu odporności może należeć do „Bad". Taka analiza zapobiega prostemu rozgrzeszaniu poprzez wskazywanie czynników zewnętrznych.

Najistotniejszym nieporozumieniem w dyskusji o failure culture jest przekonanie, że przeciwieństwem kultury winy jest kultura niskich wymagań. W rzeczywistości dojrzała organizacja ucząca się powinna łączyć wysokie bezpieczeństwo psychologiczne z wysokimi standardami odpowiedzialności. Bez bezpieczeństwa ludzie ukrywają problemy; bez odpowiedzialności – powtarzają błędy. Można tu wyróżnić cztery konfiguracje.

Niskie bezpieczeństwo i wysoka presja prowadzą do lęku i ukrywania informacji. Wysokie bezpieczeństwo przy niskiej odpowiedzialności może skutkować komfortem bez doskonalenia. Niskie wartości obu wymiarów sprzyjają apatii. Dopiero wysoki poziom obu tworzy warunki dla ambitnego uczenia się: można zgłosić błąd bez społecznego poniżenia, ale zgłoszenie to musi uruchomić analizę i korektę. Taka kultura wymaga odejścia od pytania „kto zawinił?" jako pierwszego odruchu organizacyjnego. Nie oznacza to zniesienia odpowiedzialności indywidualnej, lecz przyjęcie, że pierwszym przedmiotem analizy powinno być to, co się wydarzyło, w jakim mechanizmie i dlaczego system pozwolił na taki obrót spraw.

Edmondson przeciwstawia inteligentne eksperymentowanie błędom powstającym w znanym środowisku, podkreślając, że te drugie należy ograniczać poprzez szkolenia, dobre procesy i czujność. Granica między błędem a eksperymentem powinna być zatem ustalana przed zdarzeniem, a nie po nim. Jeżeli projekt był oznaczony jako pilotaż, posiadał hipotezę, zakres, kryteria zatrzymania, zabezpieczenia, plan zbierania danych i określony apetyt na ryzyko, to brak oczekiwanego rezultatu może być inteligentną porażką. Jeśli te elementy nie istniały, późniejsze nazywanie niepowodzenia „eksperymentem" jest zwykłą racjonalizacją.

Impact Risk Appetite jako narzędzie walki z paradoksem raportowania

Autor książki wprowadza interesującą kategorię **Impact Risk Appetite**, czyli formalne określenie poziomu ryzyka nieskuteczności, jaki organizacja jest gotowa zaakceptować w zamian za szansę na osiągnięcie ponadprzeciętnej zmiany społecznej. Koncepcja ta zasługuje na rozwinięcie, gdyż przenosi ryzyko społeczne z obszaru niejawnego tabu do ładu organizacyjnego. Każda organizacja posiada pewien apetyt na ryzyko, nawet jeśli nigdy go nie zdefiniowała. Instytucja finansująca wyłącznie sprawdzone programy również dokonuje wyboru: minimalizuje prawdopodobieństwo jednostkowej porażki, ale zwiększa ryzyko, że przez lata będzie osiągać jedynie przyrostowe rezultaty w obliczu problemu wymagającego radykalnej innowacji. Z kolei instytucja nastawiona na eksperymenty akceptuje więcej niepowodzeń projektowych, by zwiększyć szansę na odkrycie interwencji o wysokim efekcie.

Nie istnieje zatem stan całkowitego braku ryzyka; mamy do czynienia jedynie z transformacją jednego rodzaju ryzyka w inny. Rezygnacja z eksperymentu chroni reputację w krótkim terminie, lecz może zwiększać ryzyko strategicznej nieadekwatności. Kontynuowanie starego programu zapewnia stabilność operacyjną, ale grozi wieloletnim finansowaniem niskiego marginalnego

impactu. Właśnie dlatego rada nadzorcza powinna rozumieć nie tylko risk of action, lecz także **risk of inaction**. Impact Risk Appetite powinien zatem odpowiadać charakterowi portfela.

Program dotyczący bezpieczeństwa dzieci nie może mieć takiej samej tolerancji błędu jak test nowego sposobu komunikacji z darczyńcami. Eksperymenty w procesach administracyjnych są dopuszczalne przy wyższym ryzyku niż te, które mogłyby spowodować nieodwracalną szkodę zdrowotną. Ryzyko należy oceniać przez pryzmat prawdopodobieństwa szkody, jej ciężaru, odwracalności, liczby osób narażonych oraz wartości potencjalnej wiedzy. Takie podejście zbliża zarządzanie wpływem do **portfolio theory**. Nie trzeba wymagać identycznego profilu ryzyka od wszystkich programów; organizacja może utrzymywać rdzeń działań o wysokiej pewności, przeznaczając ograniczoną część zasobów na eksperymenty o większej niepewności. Portfel jako całość pozostaje racjonalny nawet przy niepowodzeniach poszczególnych projektów.

Autor książki słusznie przesuwa część odpowiedzialności za kulturę sukcesu z organizacji wykonawczych na fundatorów. Mechanizm ten można pogłębić poprzez teorię principal-agent. Grantodawca powierza środki organizacji, lecz nie obserwuje bezpośrednio wszystkich jej działań i posiada uboższą wiedzę o warunkach terenowych, co daje organizacji przewagę informacyjną.

Fundator próbuje ją ograniczyć za pomocą kontraktów, wskaźników, raportów i audytów. Badania nad nonprofit governance wskazują, że sektor pozarządowy cechuje kombinacja asymetrii informacji po stronie zarządzających oraz nierównowagi władzy finansowej po stronie fundatorów. Kontrakty oparte na wynikach są odpowiedzią na ten problem, lecz wpływają one również na to, co organizacje mierzą i komunikują. Jeśli finansowanie zależy od wykazania sukcesu, agent otrzymuje bodziec do prezentowania danych sprzyjających utrzymaniu środków. Principal, dążąc do uzyskania wiarygodnej wiedzy, może zwiększać liczbę wskaźników i rygor raportowania, na co agent odpowiada profesjonalizacją narracji.

Powstaje wyścig kontroli i prezentacji, w którym obie strony dysponują coraz większą liczbą dokumentów, lecz niekoniecznie większą ilością prawdy. To klasyczny paradoks accountability. System zbudowany, by zmniejszyć asymetrię informacji, może paradoksalnie zwiększyć asymetrię pomiędzy rzeczywistością a raportem, jeśli cena każdego niepowodzenia jest zbyt wysoka. Niezależne badania governance organizacji nonprofit zwracają uwagę, że redukowanie accountability do wąskiej logiki principal-agent może premiować oczekiwania najpotężniejszych interesariuszy i wzmacniać compliance kosztem szerszej misji społecznej.

Projektowanie bodźców zamieniających porażkę w aktywo poznawcze

Dojrzały fundator powinien zatem nie tylko żądać informacji, lecz projektować bodźce do prawdziwej informacji. Jeśli jawne ujawnienie ograniczeń programu zmniejsza szansę na kolejne finansowanie, racjonalny grantobiorca uczy się je ukrywać. Jeżeli natomiast wiarygodna analiza niepowodzenia, terminowa korekta i przejrzyste udokumentowanie wiedzy są traktowane jako element jakości, motywacja zmienia się radykalnie. Właśnie dlatego autor proponuje rozróżnienie pomiędzy ocenianiem samego rezultatu a ocenianiem także jakości wiedzy uzyskanej z eksperymentu. W przypadku projektu eksploracyjnego rezultat zerowy może mieć dużą wartość, o ile pozwala wiarygodnie wykluczyć kosztowną hipotezę. Nie oznacza to jednak, że brak rezultatu zawsze jest cenny – wiedza musi być rzeczywiście nowa, istotna i wiarygodnie wygenerowana. Można to ująć poprzez pojęcie expected value of experimentation: organizacja angażuje zasoby nie tylko w oczekiwany bezpośredni efekt, ale również w informację, która usprawni przyszłe decyzje. Jeśli eksperyment kosztuje sto tysięcy złotych i nie pomaga bezpośrednio beneficjentom, lecz pozwala uniknąć wdrożenia nieskutecznego programu za dziesięć milionów, jego wartość społeczna pozostaje wysoka.

Warunkiem jest jednak to, że wynik eksperymentu faktycznie wpływa na kolejny krok. Porażka, z której nic nie wynika, pozostaje wyłącznie porażką. Porażka staje się aktywem poznawczym dopiero poprzez ekstrakcję, zapisanie, dystrybucję i wykorzystanie wiedzy. Organizacja powinna zatem dokumentować nie tylko to, „co nie wyszło", ale konkretnie: które założenie zostało obalone, jakie dowody to uzasadniają, co zmieniono w Theory of Change i czego odtąd nie należy powtarzać. Pojawia się tu kategoria organizacyjnej amortyzacji błędu.

Wydatek na nieskuteczny eksperyment nie musi być całkowicie stracony, jeśli wygenerowana wiedza zmniejsza prawdopodobieństwo przyszłych strat. Jeśli natomiast informacja pozostaje w zespole, który się rozwiązuje, a po kilku latach inna jednostka powtarza tę samą próbę, organizacja płaci za tę samą lekcję wielokrotnie. Autor wykorzystuje DARPA jako metaforę instytucji zdolnej finansować wysokie ryzyko i sugeruje model wydzielonego „failure fund". Tę analogię należy jednak rozdzielić na element historyczny i normatywny.

Autor podaje, że DARPA powstała w 1957 r., podczas gdy oficjalna historia agencji wskazuje, że Sputnik wystrzelono w październiku 1957 r., a ARPA utworzono formalnie 7 lutego 1958 r. jako część amerykańskiej odpowiedzi na strategiczne zaskoczenie technologiczne. Nie ma również podstaw, aby proponowany przez autora „20% Failure Fund" przedstawiać jako historyczny mechanizm DARPA. Należy traktować go jako propozycję inspirowaną logiką portfelowego

finansowania ryzykownej innowacji, a nie opis faktycznego modelu budżetowego agencji. Oficjalne materiały DARPA potwierdzają orientację na przełomy transformacyjne i unikanie wyłącznie przyrostowego postępu, lecz nie ustanawiają uniwersalnej zasady dwudziestoprocentowej transzy na niepowodzenia. Sama analogia pozostaje jednak heurystycznie wartościowa. Historia technologii rozwijanych pod auspicjami ARPA/DARPA pokazuje, że projekty przełomowe często przechodzą przez nieudane iteracje; przykładowo program Centaur doświadczył kilku niepowodzeń, zanim osiągnął udany lot orbitalny i późniejszą trwałą użyteczność. Nie oznacza to, że sektor społeczny powinien bezkrytycznie kopiować model wojskowej agencji badawczej.

Różne reżimy oceny dla rutyny i innowacji

Wynika z tego jednak ogólniejsza zasada: innowacja eksploracyjna wymaga innego reżimu oceny niż rutynowe świadczenie znanej usługi. Program o ustalonym mechanizmie należy oceniać przede wszystkim przez pryzmat skuteczności, jakości i efektywności. Eksperyment typu frontier powinien być natomiast oceniany także przez rygor postawionej hipotezy, jakość testu, bezpieczeństwo, tempo uczenia się oraz informacyjność negatywnego wyniku. Stosowanie identycznych kryteriów do obu tych typów działalności jest błędem kategorii.

Popularna formuła **fail fast** jest szczególnie podatna na banalizację; w sektorze społecznym sama szybkość porażki nie stanowi wartości. Rozsądniejsza wersja brzmiałaby: fail small, safe and informative. Porażka powinna być możliwie mała, ograniczona, odwracalna i zaprojektowana tak, by dostarczyć odpowiedź na istotne pytanie. O ile w eksperymencie dotyczącym sposobu wysyłania przypomnień można łatwo stosować szybkie iteracje, o tyle przy badaniu intensywności terapii, świadczeń dla dzieci czy interwencji wobec osób w kryzysie bezdomności margines ryzyka musi być znacznie mniejszy.

Wartość informacji nie znosi zasady non-maleficence. Człowiek nie może stawać się minimalnie opłacalnym prototypem. Z tego względu adaptacja metodologii Lean Startup do polityki społecznej wymaga etycznego filtra. MVP może być rozsądnym narzędziem uczenia w projektowaniu usługi, ale musi istnieć minimum ethical product – poziom, poniżej którego eksperyment nie może zejść, niezależnie od potencjalnych korzyści poznawczych. Standard ten obejmuje bezpieczeństwo, godność, zgodę, podstawowe normy profesjonalne oraz możliwość przerwania procesu po pojawieniu się sygnałów o wyrządzonej szkodzie.

Takie rozróżnienie pozwala zbudować precyzyjniejszą teorię failure culture. Niepowodzenie może być informacyjnie wartościowe i moralnie dopuszczalne, jeśli eksperyment był racjonalny, proporcjonalny i bezpieczny. Może być informacyjnie wartościowe, lecz moralnie niedopuszczalne,

gdy wiedzę uzyskano kosztem nieproporcjonalnego ryzyka. Może być moralnie neutralne i poznawczo jałowe, jeśli organizacja nie wyciąga z niego wniosków. Wreszcie może być zarówno poznawczo, jak i moralnie naganne, gdy polega na powtarzaniu znanego już błędu.

Taka macierz jest istotniejsza niż proste przeciwstawienie sukcesu porażce. Sukces również bywa epistemicznie słaby – program może osiągnąć rezultat przypadkiem, dzięki zewnętrznej koniunkturze lub pomimo błędnej Theory of Change. Jeżeli organizacja interpretuje taki wynik jako potwierdzenie własnego modelu, sukces może uczyć gorzej niż inteligentna porażka.

Jest to jeden z najbardziej subtelnych wniosków całego podejścia. Nie należy pytać wyłącznie o to, czy program odniósł sukces, lecz co konkretnie wynik mówi o modelu przyczynowym. Porażka potrafi obalić fałszywe założenie, podczas gdy sukces często nie pozwala rozróżnić między kilkoma konkurencyjnymi mechanizmami. Z punktu widzenia nauki najcenniejszy nie zawsze jest wynik pozytywny; najcenniejszy jest wynik informacyjny. Stevenson-Dodd słusznie łączy w tym kontekście tabu porażki z kwestią reputacji.

Transparentność jako budowanie reputacji epistemicznej

Organizacja społeczna opiera się w wyjątkowo wysokim stopniu na zaufaniu. Nie może ona zaoferować donatorowi klasycznego prawa własności ani pełnej kontroli nad produktem; donator finansuje obietnicę, że środki zostaną wykorzystane zgodnie z misją. Reputacja staje się zatem zasobem ekonomicznym, co wyjaśnia racjonalność zachowawczej komunikacji. Ujawnienie porażki może bowiem prowadzić do utraty finansowania, problemów medialnych lub reakcji organu nadzoru. Nie należy więc moralizatorsko zakładać, że każda organizacja ukrywająca niekorzystne informacje robi to z braku odwagi – często jest to reakcja na realną architekturę bodźców.

Jednocześnie długookresowa reputacja oparta na permanentnym prezentowaniu sukcesów jest krucha. Jeśli wszystkie programy organizacji „działają”, wskaźniki rosną, beneficjenci są zadowoleni, a żaden projekt nigdy nie wymaga zamknięcia, wiarygodność raportowania powinna budzić poznawczy sceptycyzm. W złożonym świecie bezbłędność jest sama w sobie informacją podejrzaną. Dlatego transparentność może pełnić funkcję kosztownego sygnału wiarygodności. Organizacja przyznająca, że część programu nie przyniosła rezultatów, rezygnuje z krótkoterminowej korzyści komunikacyjnej, a właśnie ten koszt sprawia, że jej twierdzenia o sukcesach stają się bardziej przekonujące. Przypadki Street League i CLIC Sargent są wartościowe właśnie z tej perspektywy: jawność niepowodzenia nie jest przeciwieństwem budowania reputacji;

może być mechanizmem budowania reputacji epistemicznej – przekonania, że organizacja raportuje również to, czego wolałaby nie raportować.

Radykalna przejrzystość ma jednak swoje granice. Organizacja dysponuje danymi o beneficjentach, pracownikach, sporach prawnych czy zdarzeniach potencjalnie stygmatyzujących. Publikowanie „wszystkiego" nie jest przejawem dojrzałości, lecz mogłoby stanowić naruszenie praw osób trzecich. Transparentność należy zatem podporządkować materialności, proporcjonalności, ochronie danych, bezpieczeństwu i obowiązkom powierniczym. Dojrzała zasada brzmi raczej: publikuj wszystko, co materialne dla uczciwej oceny wpływu, o ile publikacja nie narusza praw lub uzasadnionych interesów innych osób. Informację osobową można agregować, problem operacyjny opisać bez identyfikowania pracownika, a spór HR-owy nie musi stać się publicznym spektaklem, by organizacja uczciwie przyznała, że istniały problemy zarządcze. Transparentność nie jest więc maksymalizacją ilości informacji, lecz maksymalizacją uczciwości obrazu.

W złożonych interwencjach społecznych szczególnie ważne jest unikanie personalizacji systemowych niepowodzeń. Jeżeli program zawodzi wskutek współdziałania dziesięciu drobnych czynników, wskazanie jednego pracownika może być organizacyjnie wygodne, ale poznawczo fałszywe. Complex failure wymaga analizy interakcji, a nie poszukiwania kozła ofiarnego. Jednocześnie hasło „system zawiódł" nie może znosić odpowiedzialności indywidualnej tam, gdzie istniała realna możliwość i obowiązek działania. Teoria systemowa nie jest teorią niewinności wszystkich uczestników. Jej funkcją jest prawidłowe rozmieszczenie odpowiedzialności pomiędzy projekt procesu, nadzór, komunikację, zasoby, warunki zewnętrzne i decyzje jednostek.

Można tu mówić o just culture, rozwiniętej szczególnie w lotnictwie i ochronie zdrowia: kulturze rozróżniającej pomyłkę ludzką, ryzykowne zachowanie i świadome lekceważenie standardów. Jej logika jest analogiczna do potrzeb impact managementu. Reakcja powinna odpowiadać typowi zdarzenia, a nie samemu faktowi, że rezultat był negatywny.

Dojrzałość organizacji to zdolność odróżniania błędów od innowacji

Powyższe rozważania prowadzą do daleko idących konsekwencji dla ekonomii finansowania społecznego. Jeśli fundator płaci wyłącznie za z góry określony rezultat, organizacja będzie optymalizować ten wynik oraz sposób jego sprawozdawania. Jeżeli jednak fundator finansuje również zdolność organizacyjną do wykrywania błędów, ewaluowania hipotez i adaptacji, powstaje zupełnie inny typ relacji. Nie chodzi tu o finansowanie nieefektywności. Przeciwnie – organizacja,

która wcześnie wykryje nieskuteczny program i go zamknie, może okazać się bardziej efektywna niż ta, która przez pięć lat dostarcza formalnie poprawne wskaźniki (outputs). Kluczowy jest horyzont oceny.

W krótkiej perspektywie pierwsza z nich „straciła" część budżetu na eksperyment. W długiej – uniknęła znacznie wyższych kosztów kontynuowania błędu. Dlatego portfolio fundatora powinno uwzględniać różne klasy kapitału: finansowanie stabilnego świadczenia sprawdzonej usługi, skalowanie interwencji opartych na silnych dowodach oraz finansowanie eksperymentów obarczonych wysoką niepewnością. Każda z tych klas powinna podlegać innemu reżimowi dowodowemu, gdyż wymaganie od eksperymentu takiej samej przewidywalności jak od dojrzałego programu eliminuje go jeszcze przed rozpoczęciem.

W kontekście całego artykułu wyłania się zatem głębsza definicja organizacji „Difference". Nie jest nią instytucja osiągająca wyłącznie sukcesy, lecz taka, która posiada epistemiczne mechanizmy odróżniania sukcesu od błędu, niepewności czy zewnętrznej bariery i potrafi dostosować swoje działanie do tego rozpoznania. Organizacja, która nigdy nie raportuje porażek, może być wyjątkowo skuteczna, lecz bardziej prawdopodobne jest, że posługuje się definicją sukcesu na tyle elastyczną, by każde działanie można było ex post przedstawić jako wartościowe. W takim systemie falsyfikacja misji staje się niemożliwa: każda liczba znajduje dobre wyjaśnienie, każdy projekt „buduje relacje", każde niepowodzenie „dostarcza doświadczenia", a każdy wydatek zostaje powiązany z jakimś aspektem szeroko rozumianej zmiany społecznej.

Jest to przeciwieństwo postawy naukowej, nawet jeśli język raportu jest nasycony danymi. Najbardziej dojrzała kultura wpływu potrzebuje zatem warunków przegranej hipotezy. Przed rozpoczęciem interwencji należy jasno określić, jaki wynik skłoni organizację do uznania, że program nie działa zgodnie z założeniami. Bez takiego kryterium Theory of Change staje się niefalsyfikowalna. Porażka ma więc paradoksalnie znaczenie konstytutywne dla dowodu sukcesu. Jeśli organizacja nie dopuszcza wyniku, który uznałaby za niepowodzenie, jej sukcesy tracą wartość informacyjną – przypomina to eksperyment naukowy skonstruowany tak, by gwarantował potwierdzenie hipotezy. Rezultat przestaje cokolwiek rozróżniać. Przejście od safety do accountability oznacza tutaj dojrzałość polegającą na rozróżnianiu, a nie na usprawiedliwianiu.

Ostatecznie kultura uczenia się z porażek nie powinna być kulturą afirmacji niepowodzeń, lecz kulturą wysokiej rozdzielczości poznawczej. Taka organizacja potrafi odróżnić błąd możliwy do uniknięcia od intelligent failure, złożony splot przyczyn od zaniedbania, wpływ kontekstu od wadliwego projektu programu, brak rezultatu od braku wiedzy, a eksperyment od improwizacji.

Wymaga to zarówno bezpieczeństwa psychologicznego i odpowiedzialności, jak i odwagi w raportowaniu oraz rygoru dowodowego. Stevenson-Dodd słusznie zauważa, że organizacja wyłącznie celebrująca sukcesy traci zdolność uczenia się. Należy jednak dodać, że organizacja celebrująca wszystkie porażki traci zdolność odróżniania innowacji od niekompetencji.

Dojrzałość leży pomiędzy tymi skrajnościami. W tej perspektywie „The Good, the Bad and the Ugly" nabiera znaczenia wykraczającego poza formułę raportową; jest próbą przywrócenia sprawozdawczości pełnego rozkładu rzeczywistości: tego, co zadziałało, co organizacja zrobiła źle oraz co ograniczało ją z zewnątrz. Street League dodaje do tego rygor audytowalności, Edmondson – bezpieczeństwo i klasyfikację porażek, Sitkin – poznawczą wartość małych strat, a teoria principal-agent – analizę bodźców do ukrywania niewygodnych informacji. Razem tworzą one znacznie silniejszą teorię uczenia organizacyjnego niż proste hasło „nie bójmy się błędów". Najważniejszym kapitałem takiej organizacji nie jest obraz nieomylności, lecz zdolność do wykrywania własnej omylności, zanim jej koszt stanie się zbyt wysoki.

Maksymalizacja wpływu jako problem alokacji zasobów i kosztu alternatywnego

Kolejnym krokiem jest przejście od wyciągania wniosków z pojedynczych programów do maksymalizacji wpływu jako problemu alokacji zasobów. Pozwala to na głębszą analizę Stevenson-Doddowskiego prawa 80/20 i "paradoksu 95/5", marginalnej efektywności społecznej, kosztu alternatywnego (opportunity cost) oraz zarządzania portfelem interwencji. Prowadzi to do postawienia najtrudniejszego pytania przed zarządem: nie „czy ten program robi coś dobrego?", lecz "czy przy tej samej godzinie pracy i tej samej złotówce istnieje działanie, które wytworzyłoby większą, trwalszą lub bardziej sprawiedliwie rozłożoną zmianę?".

Po zbudowaniu teorii zmiany, systemu pomiaru, mechanizmów przyczynowego wnioskowania, pętli uczenia oraz kultury pozwalającej ujawniać niepowodzenia pojawia się pytanie jeszcze trudniejsze niż ustalenie, czy pojedynczy program działa. Organizacja musi bowiem zdecydować, które spośród skutecznych działań powinny otrzymywać zasoby, w jakiej skali, kosztem których innych interwencji i według jakiej koncepcji wartości społecznej. Właśnie na tym poziomie zaczyna się właściwe znaczenie Stevenson-Doddowskiego **Maximise**. Maksymalizacja wpływu nie jest bowiem prostym zwiększaniem skali pomocy.

Jest to problem alokacji skończonego kapitału finansowego, czasu, uwagi kierownictwa, kompetencji personelu, infrastruktury, reputacji i tolerancji ryzyka pomiędzy konkurencyjne

zastosowania. Autor książki opisuje dwa elementy maksymalizacji: wiarygodne komunikowanie rezultatów oraz bieżące przesuwanie zasobów ku działaniom przynoszącym największą wartość społeczną, z możliwością wygaszania programów nieskutecznych. Podejście to jest intuicyjnie przekonujące, lecz wymaga pogłębienia ekonomicznego. Samo stwierdzenie, że „program A działa lepiej niż program B", pozostaje niewystarczające, dopóki nie określimy, co dokładnie oznacza słowo „lepiej": większą liczbę rezultatów, głębokość zmiany, niższy koszt jednostkowy, trwałość, dodatkowość, pomoc osobom w najgorszej sytuacji, możliwość skalowania, mniejsze ryzyko czy potencjał wygenerowania wiedzy dla przyszłych działań.

Ekonomia rozpoczyna się od rzadkości. Gdyby zasoby były nieograniczone, organizacja mogłaby finansować wszystkie interwencje o dodatniej wartości. W realnym świecie każda alokacja posiada jednak koszt alternatywny. Środki przeznaczone na program A nie mogą zostać jednocześnie wykorzystane na program B; godzina pracy analityka poświęcona rozwojowi dashboardu nie może zostać wykorzystana na pracę nad nową interwencją; miejsce w programie przyznane jednej osobie może oznaczać brak miejsca dla innej. Z tego powodu koszt społeczny działania nie jest tożsamy z księgowym kosztem jego wykonania. Do kosztu należy bowiem doliczyć wartość najlepszej niewykorzystanej alternatywy. To pojęcie opportunity cost stanowi jeden z fundamentów ekonomii dobrobytu i ekonomiki zdrowia. W metodzie program budgeting and marginal analysis (PBMA) decyzje o zasobach są konstruowane wokół dwóch kategorii: kosztu alternatywnego oraz efektu krańcowego. Jeżeli dodatkowa jednostka zasobu przynosi większą korzyść w programie A niż w B, realokacja z B do A zwiększy całkowitą wartość przy zachowaniu tego samego budżetu.

Maksymalizacja wpływu to problem marginalny, a nie średni

Ta zasada fundamentalnie zmienia pytanie zarządcze. Nie wystarczy już ustalić, czy projekt przynosi pozytywny rezultat. Należy pytać: czy marginalna złotówka wydana na ten projekt wytworzy więcej wartości niż marginalna złotówka wykorzystana w następnym najlepszym zastosowaniu? Program może być skuteczny i jednocześnie nie zasługiwać na zwiększenie finansowania, jeśli osiągnął punkt malejących przychodów krańcowych. Inny program może mieć mniejszą dotychczasową skalę wpływu, ale znacznie większy potencjał dodatkowego rezultatu przy niewielkim zwiększeniu zasobów. Maksymalizacja wpływu jest zatem problemem marginalnym, a nie wyłącznie średnim. Średnia skuteczność historyczna mówi jedynie o tym, co organizacja osiągała przy dotychczasowej skali.

Decyzja budżetowa dotyczy natomiast kolejnej jednostki zasobu. To rozróżnienie jest często pomijane w sektorze publicznym: budżety kopiuje się z roku na rok, ponieważ dotychczasowe programy mają historię, personel i beneficjentów. Tymczasem właściwe pytanie ekonomiczne powinno dotyczyć korzyści utraconych przez zachowanie historycznej alokacji. Nauki ekonomiczne pozwalają rozróżnić dwa poziomy sprawności, które w dyskusjach o wpływie bywają mylone. Efektywność techniczna oznacza osiąganie danego rezultatu przy najmniejszym możliwym zużyciu zasobów lub osiąganie maksymalnego rezultatu przy określonych nakładach. Organizacja staje się technicznie sprawniejsza, gdy ten sam outcome osiąga taniej, szybciej lub przy mniejszym obciążeniu personelu.

Efektywność alokacyjna dotyczy czegoś innego: tego, czy zasoby znajdują się w tych zastosowaniach, które wytwarzają największą wartość względem alternatyw. Można więc być perfekcyjnie efektywnym technicznie w realizacji programu, który nie powinien otrzymywać tak dużej części budżetu. PBMA zostało zaprojektowane właśnie jako metoda porównująca możliwości zwiększenia inwestycji oraz zmniejszenia lub zakończenia części istniejących działań. Rozróżnienie to można przełożyć bezpośrednio na model "Do-Difference".

Organizacja typu "Do" może z roku na rok zwiększać wydajność techniczną: obsługiwać więcej uczestników tym samym zespołem, skracać procedury czy automatyzować dokumentację. Organizacja "Difference" musi dodatkowo pytać, czy ten rodzaj działania w ogóle pozostaje najlepszym wykorzystaniem zasobu wobec innych sposobów realizacji purpose. Jest to moment intelektualnie niewygodny, ponieważ oznacza możliwość likwidacji programu, który jest dobrze zarządzany, popularny i rzeczywiście pomaga ludziom. Ekonomia krańcowa nie pyta bowiem tylko o wartość absolutną programu, lecz o względną wartość alternatyw. W świecie niedoboru "dobry" nie zawsze oznacza "najlepszy do dalszego finansowania".

Szczególnie ważna staje się tutaj zasada malejących korzyści krańcowych. Pierwsze środki mogą trafiać do osób, którym stosunkowo łatwo zapewnić istotny rezultat. Kolejne rozszerzanie programu może wymagać dotarcia do bardziej rozproszonych populacji, zatrudnienia droższych specjalistów, otwierania nowych placówek albo intensywniejszego wsparcia osób o bardziej złożonych potrzebach. Koszt uzyskania kolejnego outcome może więc wzrastać. Nie oznacza to, że takie rozszerzenie jest niewłaściwe; przeciwnie, osoby wymagające większego nakładu mogą posiadać silniejsze roszczenie moralne do pomocy. Oznacza to jedynie, że krańcowy koszt rezultatu zmienia się wraz ze skalą. Używanie jednego historycznego współczynnika cost-per-outcome do dowolnego poziomu ekspansji jest więc błędem. Analogicznie: pierwsza złotówka przeznaczona na komunikację może silnie zwiększyć dotarcie, pięćdziesiąta kampania reklamowa już znacznie mniej. Pierwszy psycholog w zespole może radykalnie zwiększyć zdolność obsługi, kolejny jedynie

ograniczyć kolejkę. Pierwsza integracja informatyczna może usunąć główny silos danych, kolejne rozwiązanie BI może tylko marginalnie poprawić komfort analityczny.

Zarządzanie portfelem interwencji jako balans między eksploatacją a eksploracją

Dlatego sprawne zarządzanie wpływem powinno opierać się na konstruowaniu **krzywych marginalnego wpływu**, nawet jeśli początkowo mają one charakter przybliżony. Nie jest konieczne, by znać je z matematyczną precyzją. Już samo pytanie o to, jak zmienia się dodatkowa wartość wraz ze zwiększaniem budżetu, niesie znacznie więcej informacji niż automatyczne stosowanie zasady: „skoro działa, finansujmy więcej".

95/5 jako heurystyka uwagi, nie odkryte prawo zarządzania

Stevenson-Dodd wskazuje na tzw. „paradoks 95/5": obserwację, że rady i zarządy poświęcają zdecydowaną większość czasu finansom oraz operacjom, a zaledwie niewielką część realnemu wpływowi. W analizie naukowej proporcję 95/5 należy traktować jako **heurystykę autora**, a nie zweryfikowaną, uniwersalną regularność empiryczną. Stevenson-Dodd nie przedstawia bowiem próby, metodologii ani rozkładu obserwacji, które pozwalałyby interpretować tę wartość jako statystykę populacyjną.

Sama intuicja pozostaje jednak poznawczo wartościowa i można ją osadzić w teorii uwagi organizacyjnej. Herbert Simon zauważał, że w warunkach nadmiaru informacji rzadkim zasobem staje się uwaga. Rada nadzorcza nie tylko rozdziela pieniądze; rozdziela także **czas poznawczy**. To, co znajduje się stale na początku agendy, otrzymuje regularną interpretację, wskaźniki, rytuały kontrolne i przypisanych odpowiedzialnych. To, co zostaje zepchnięte na koniec spotkania, staje się strategicznie peryferyjne, nawet jeśli oficjalna misja głosi coś przeciwnego. Z tej perspektywy agenda zarządu jest ukrytym budżetem wartości. Jeśli każdej złotówce przypisuje się miesięczny raport, a wpływowi społecznemu dwudziestominutowy punkt raz na kwartał, organizacja ujawnia rzeczywistą hierarchię poznawczą.

Przesunięcie proponowane przez Stevensona-Dodda nie powinno jednak oznaczać literalnego „odwrócenia" proporcji i ograniczenia nadzoru finansowego. W organizacji społecznej stabilność finansowa jest warunkiem możliwości realizacji misji. Dojrzałe governance powinno raczej **zintegrować finanse z wpływem**, tak aby dyskusja budżetowa nie sprowadzała się do pytania „ile wydaliśmy?", lecz obejmowała również kwestie: „jakiej zmiany oczekujemy po tej alokacji i czego nie sfinansujemy, wybierając właśnie ją?". Autor przedstawia także regułę 80/20, według

której około 80% zasobów powinno wspierać działania sprawdzone i relatywnie bezpieczne, a około 20% – eksperymenty o większej niepewności i potencjale innowacyjnym. Podobnie jak w przypadku 95/5, proporcji tej nie należy przedstawiać jako rezultatu prawa Pareta ani empirycznie optymalnego podziału dla każdej organizacji.

Jest to **heurystyka portfelowa**, która próbuje pogodzić eksploatowanie wiedzy już posiadanej z eksplorowaniem nowych możliwości. To właśnie tutaj propozycja Stevensona-Dodda spotyka się z jedną z najważniejszych teorii organizacyjnego uczenia – rozróżnieniem Jamesa Marcha między **exploration i exploitation**. Exploration obejmuje poszukiwanie, eksperymentowanie, odkrywanie i podejmowanie ryzyka; exploitation – doskonalenie, standaryzację, zwiększanie efektywności i wykorzystywanie tego, co już działa. March pokazał, że oba procesy konkurują o zasoby i mają odmienne profile czasowe: exploitation szybciej przynosi wymierne korzyści, przez co organizacje mogą systematycznie niedoinwestowywać exploration, choć to właśnie ono warunkuje długookresową adaptację.

To rozróżnienie dostarcza mocnej podstawy dla samej idei 80/20, lecz nie dla konkretnych liczb. Optymalna proporcja zależy od środowiska. Organizacja prowadząca stabilną usługę krytyczną może racjonalnie przeznaczać na eksplorację mniej niż 20%. Laboratorium innowacji społecznych może przeznaczać znacznie więcej.

W okresie gwałtownej zmiany technologicznej lub regulacyjnej zbyt niska eksploracja zwiększa ryzyko strategicznej dezaktualizacji. W środowisku stabilnym nadmiar eksperymentowania może marnować zasoby. Właściwa interpretacja zasady 80/20 brzmi więc: **portfel organizacji powinien celowo rozdzielać zasoby pomiędzy eksploatację sprawdzonych mechanizmów i eksplorację nowych**, zamiast oczekiwać, że innowacja wydarzy się spontanicznie w systemie premiującym wyłącznie przewidywalne rezultaty.

Perspektywa finansowa dostarcza kolejnego użytecznego języka. Inwestor nie ocenia rozsądku portfela wyłącznie poprzez pytanie, czy każda pozycja osobno jest atrakcyjna. Interesuje go relacja ryzyka, oczekiwanej stopy zwrotu, korelacji między składnikami, płynności i horyzontu. Analogicznie organizacja społeczna może posiadać **portfel interwencji** różniących się skalą, ryzykiem, dojrzałością dowodową, kosztem jednostkowym, możliwością skalowania i zależnością od warunków zewnętrznych. Jedne programy mogą cechować się wysoką pewnością przy niskim potencjale przełomu. Inne mogą mieć niepewne outcomes, ale tworzyć opcję strategiczną. Jeszcze inne pełnią funkcję infrastrukturalną: same nie generują bezpośredniego impactu, lecz zwiększają skuteczność innych programów. Pomiar każdego z nich tym samym wskaźnikiem byłby błędem.

Zarządzanie portfelowe sugeruje również znaczenie **dywersyfikacji epistemicznej**. Organizacja całkowicie zależna od jednego mechanizmu może stać się bardzo skuteczna w określonym środowisku, lecz niezwykle krucha po jego zmianie. Portfel kilku różnych podejść może mieć niższą średnią efektywność krótkoterminową, ale większą odporność wobec zmian otoczenia. Nie wszystkie „duplikacje" są zatem marnotrawstwem; czasem redundancja pełni funkcję resilience. System oparty na jednej niezwykle efektywnej metodzie może być podatny na awarię, podczas gdy kilka częściowo nakładających się zdolności zwiększa odporność. Teoria wpływu musi więc uważać, aby mechaniczne dążenie do maksymalizacji efficiency nie wyeliminowało buforów niezbędnych do przetrwania szoku.

Realne opcje i walka z pułapką kosztów utopionych

Teoria finansów wprowadza kategorię real options. W tym ujęciu wartość inwestycji nie zawsze przejawia się w bieżącym zwrocie, lecz w stworzeniu możliwości przyszłego działania. Mały pilotaż może nie generować dziś znaczącego wpływu, ale buduje wiedzę, partnerstwa, infrastrukturę lub kompetencje umożliwiające późniejsze skalowanie. Zamknięcie takiego projektu wyłącznie z powodu niskiego obecnego SROI może oznaczać utratę cennej opcji. W warunkach wysokiej niepewności sama możliwość odroczenia, rozszerzenia, zmniejszenia albo porzucenia programu posiada realną wartość. Oznacza to, że architektura finansowania powinna preferować rozwiązania etapowe wszędzie tam, gdzie wiedza jest ograniczona. Zamiast angażować pełny budżet na cztery lata, organizacja może nabywać kolejne „opcje": pilotaż, etap rozszerzony, skalowanie regionalne i dopiero na końcu wdrożenie systemowe.

Takie podejście jest spójne z wcześniej omawianą inteligentną porażką. Niewielka inwestycja informacyjna pozwala uniknąć ogromnego, nieodwracalnego zaangażowania w błędną teorię. Maksymalizacja wpływu obejmuje zatem również zarządzanie odwracalnością decyzji. W idealnym modelu ekonomicznym decyzja o przyszłym finansowaniu zależy od przyszłych kosztów i korzyści, a nie od tego, ile wydano w przeszłości. Organizacje ludzkie regularnie naruszają tę zasadę wskutek sunk cost effect. Im więcej czasu, pieniędzy i reputacji zainwestowano w program, tym trudniej go zamknąć, mimo że wcześniejszych kosztów nie da się już odzyskać. Efekt ten łączy ekonomię behawioralną z teorią instytucji: program przestaje być jedynie pozycją budżetową, a staje się częścią tożsamości. Posiada dyrektora, personel, interesariuszy, historię sukcesów, nazwę i własną komunikację.

Jego likwidacja może zostać odczytana jako przyznanie się do błędu w poprzednich decyzjach. W tej sytuacji organizacja racjonalizuje kontynuację argumentem: „zainwestowaliśmy już tak dużo".

Impact management wymaga odwrócenia tego odruchu. Koszt historyczny może pomóc zrozumieć genezę programu, ale nie powinien samodzielnie uzasadniać dalszej alokacji. Właściwe pytanie brzmi: co stanie się z kolejną złotówką, jeśli wydamy ją tutaj zamiast gdzie indziej?

To jeden z momentów, w których kultura porażki staje się warunkiem efektywności ekonomicznej. Organizacja niepotrafiąca przyznać, że interwencja osiągnęła granicę użyteczności, jest skazana na eskalację zaangażowania. Podobny problem występuje w administracji publicznej i dużych organizacjach, gdzie budżet ma często charakter przyrostowy. Historyczne wydatki stają się punktem odniesienia, a dyskusja ogranicza się do ich zwiększenia lub zmniejszenia o kilka procent. Taki system jest politycznie i administracyjnie stabilny, lecz konserwuje proporcje, których nikt dziś nie wybrałby świadomie od początku.

PBMA jest interesujące właśnie dlatego, że nie wymaga pełnego budżetowania zerowego, lecz nakazuje badać margines inwestycji i dezinwestycji. Przegląd praktyk priority setting wskazuje PBMA, HTA oraz wielokryterialne ramy wartości jako trzy główne typy formalnych metod alokacji zasobów; wspólnym problemem pozostaje jednak ograniczona instytucjonalizacja i słaby follow-up decyzji. To ważna korekta wobec czysto technokratycznej wizji: nawet najbardziej racjonalny model nie reorganizuje budżetu samodzielnie. Decyzje alokacyjne są osadzone w prawie, polityce, interesach organizacyjnych, zobowiązaniach pracowniczych, oczekiwaniach społeczności oraz wcześniejszych kontraktach. Maksymalizacja wpływu jest zatem nie tylko problemem matematycznym, lecz również problemem instytucjonalnego przejścia. Programu nie zawsze można „wyłączyć" natychmiast tylko dlatego, że wykres wskazuje niższy marginalny outcome. Istnieją prawa nabyte, zobowiązania wobec ludzi, koszty zamknięcia, ryzyko utraty kompetencji i efekty sieciowe. Racjonalna dezinwestycja musi uwzględniać także koszt transformacji. Najtrudniejszy problem zaczyna się jednak wcześniej: zanim przystąpimy do optymalizacji zasobów, musimy zdefiniować funkcję celu.

Wagi aksjologiczne jako fundament sprawiedliwej alokacji

Ekonomia dobrobytu dowodzi, że maksymalizacja nie jest neutralna aksjologicznie. Możemy dążyć do maksymalizacji sumy korzyści, minimalizowania najgorszego położenia, redukcji nierówności, zapewnienia minimów prawnych czy preferowania największej potrzeby – albo stosować kombinację tych kryteriów. Jeśli dwie interwencje kosztują tyle samo, a pierwsza nieznacznie poprawia życie stu osób, podczas gdy druga radykalnie zmienia sytuację dziesięciu osób w skrajnie trudnym położeniu, to która z nich ma większy impact?

Sama arytmetyka nie daje odpowiedzi, dopóki nie ustalimy wag. W klasycznej logice utylitarnej naturalnym kryterium byłaby suma wytworzonego dobrostanu. Podejście priorytarystyczne zakłada natomiast, że korzyść osiągnięta przez osobę w najgorszym położeniu ma większą wagę niż identyczna korzyść osoby dobrze sytuowanej. Tradycja rawlsowska zwraca uwagę na położenie najmniej uprzywilejowanych i nie pozwala sprowadzić sprawiedliwości do prostej sumy użyteczności, z kolei koncepcja capabilities Sena przesuwa punkt ciężkości z samych zasobów i użyteczności ku realnym możliwościom funkcjonowania. Literatura dotycząca alokacji zasobów zdrowotnych pokazuje, że decydenci w praktyce uwzględniają szeroki zbiór kryteriów: skuteczność, cost-effectiveness, bezpieczeństwo, siłę dowodów, potrzeby, fairness, equity, zdolność organizacyjną oraz interesy interesariuszy. W jednym z przeglądów equity i fairness należały do najczęściej pojawiających się kryteriów. Oznacza to, że realna praktyka alokacyjna rzadko przypomina czysty algorytm maksymalizujący jeden wskaźnik. Często upraszcza się tę dyskusję do przeciwstawienia efektywności i sprawiedliwości, co jest częściowo błędne.

Wiele reform może poprawiać oba te wymiary jednocześnie: usunięcie zbędnej procedury potrafi obniżyć koszty i jednocześnie zwiększyć dostępność dla grup gorzej obsługiwanych. Dobra profilaktyka może zarówno zmniejszać nierówności, jak i ograniczać wydatki na późniejsze leczenie. Literatura wskazuje zatem na możliwość traktowania equity i efficiency jako elementów komplementarnych, pod warunkiem jawnego zdefiniowania obu celów.

Istnieją jednak sytuacje realnego trade-off. Interwencja skierowana do grupy o największej potrzebie może być droższa i generować mniej mierzalnych outcomes w przeliczeniu na jednostkę budżetu. Maksymalizacja samego cost-effectiveness mogłaby wówczas zwiększyć całkowity efekt, ale jednocześnie pogłębić nierówności. System musi zatem zdecydować, ile efektywności jest gotów poświęcić dla sprawiedliwszej dystrybucji. Nowsza literatura ekonomiki zdrowia rozwija w tym zakresie **distributional cost-effectiveness analysis**, equity weighting i inne metody pozwalające włączyć dystrybucję do rachunku. Przegląd z 2024 r. wskazuje jednak na brak konsensusu co do najlepszego sposobu formalizacji equity, ograniczenia danych oraz konieczność uwzględniania fairness i opportunity cost. Ma to kluczowe znaczenie dla Stevenson-Doddowskiego języka "maximise". Maksymalizacja bez jawnych wag może ukrywać własną aksjologię pod pozorem technicznej neutralności.

Naturalnym narzędziem w sytuacji wielokryterialnej jest **Multi-Criteria Decision Analysis**. Zamiast sprowadzać wszystkie wartości do jednej jednostki pieniężnej, MCDA pozwala zdefiniować kilka kryteriów, przypisać im znaczenie, ocenić warianty i uczynić kompromisy bardziej przejrzystymi. Systematyczny przegląd z 2023 r. wskazuje na rosnące zastosowanie MCDA w ochronie zdrowia, szczególnie przy ustalaniu priorytetów, choć zauważa jednocześnie istotne

różnice metodologiczne oraz niepełne raportowanie sposobu doboru skal i wag. Dla organizacji społecznej kryteriami mogą być nie tylko liczba outcomes i koszt, ale także trwałość, additionality, głębokość zmiany, potrzeba grupy, equity, siła dowodu, możliwość skalowania, ryzyko szkody, zgodność z purpose i potencjał uczenia. Największą zaletą MCDA nie jest matematyka, lecz **ujawnienie wartości**. Jeżeli zarząd przyznaje equity wagę 30%, a cost-effectiveness 25%, nie oznacza to odkrycia obiektywnej prawdy, lecz sprawia, że jego priorytety stają się widoczne i możliwe do zakwestionowania. Ukryte sądy normatywne zostają przekształcone w jawne parametry decyzji.

Największym ryzykiem MCDA jest z kolei pseudoprecyzja. Waga "17%" może sprawiać wrażenie naukowej dokładności, choć wynika jedynie z deliberacji grupy. Suma punktów może przykrywać fakt, że niektóre kryteria są trudno kompensowalne. Organizacja może uznać na przykład, że naruszenia praw człowieka nie da się zrekompensować wyższą efektywnością ekonomiczną.

Wpływ społeczny jako optymalizacja z ograniczeniami

W takim przypadku niezbędne jest wprowadzenie ograniczenia leksykograficznego lub twardego progu, a nie zwykłej wagi w sumie punktów. Aksjologia wpływu wymaga bowiem oddzielenia dóbr podlegających optymalizacji od norm, które wyznaczają granice tego procesu. Jeśli prawo do niedyskryminacji, bezpieczeństwa czy ochrony godności traktujemy wyłącznie jako jedno z wielu kryteriów o określonej wadze, teoretycznie można je „przeważyć" znaczną korzyścią w innych obszarach. W wielu systemach normatywnych byłoby to nieakceptowalne. Z perspektywy teorii optymalizacji właściwsze jest traktowanie części wartości jako constraints. Maksymalizujemy impact pod warunkiem zachowania określonych minimów prawnych, etycznych i jakościowych. Nie wolno na przykład zwiększać liczby osiąganych outcomes poprzez arbitralne wykluczenie konkretnej grupy, jeśli naruszałoby to zasadę równego traktowania, ani poprawiać skuteczności terapii kosztem bezpieczeństwa pacjenta poza dopuszczalny próg.

Taka konstrukcja jest bardziej adekwatna niż uproszczony język „największego dobra", ponieważ społeczeństwa konstytucyjne nie są jednowymiarowymi maszynami użyteczności. Godność, prawa, autonomia i równość stanowią często warunki brzegowe racjonalnej alokacji, a nie jedynie dodatkowe punkty w rankingu. Kolejnym błędem jest utożsamianie maksymalizacji impactu ze skalowaniem. Niektóre interwencje osiągają wysoką jakość właśnie dlatego, że są relacyjne, lokalne i intensywne; ich mechanizm może ulec degradacji przy gwałtownym zwiększeniu liczby uczestników. Skalowanie zmienia bowiem ratio personelu, charakter relacji, kulturę organizacyjną

oraz selekcję beneficjentów. To, co sprawdzało się w grupie stu osób, nie musi działać identycznie przy dziesięciu tysiącach.

W teorii innowacji mówi się o problemie scale-up fidelity versus adaptation. Z jednej strony należy zachować elementy mechanizmu odpowiedzialne za skuteczność, z drugiej zaś program musi dostosować się do nowego kontekstu. Zbyt wierne kopiowanie może ignorować lokalne różnice, natomiast zbyt daleka adaptacja grozi usunięciem komponentu generującego rezultat. Dlatego decyzja o skali powinna rozstrzygać co najmniej trzy kwestie: czy efekt marginalny się utrzymuje, czy istnieje capacity do zachowania jakości oraz czy kontekst nowej populacji uruchamia ten sam mechanizm. W przeciwnym razie scale-up może zwiększyć outputs, jednocześnie zmniejszając impact per capita.

Wpływ społeczny posiada również co najmniej trzy osie: breadth, depth i duration. Można pomóc bardzo wielu osobom w niewielkim stopniu, mniejszej liczbie w sposób głęboki albo osiągnąć zmianę umiarkowaną, lecz trwałą. Te wymiary nie mają naturalnego wspólnego mianownika. Program krótkiej interwencji informacyjnej może dotrzeć do miliona osób, podczas gdy intensywny program terapeutyczny obejmie kilkaset.

Porównywanie ich wyłącznie przez pryzmat liczby beneficjentów byłoby absurdalne, jednak analiza oparta jedynie na głębokości zmiany mogłaby ignorować skalę. Niektóre modele impact investing próbują agregować te elementy, lecz każda taka próba wymaga założeń dotyczących wymienialności tych wartości. Właśnie dlatego zarząd powinien odróżnić maksymalizację od monomaniakalnej optymalizacji jednego miernika. Prawdziwa maksymalizacja w systemie społecznym jest zazwyczaj problemem wielowymiarowym.

Inwestycja w zdolność produkcyjną wpływu zamiast finansowania projektowego

Przykład organizacji Gordon Moody służy jako studium przypadku wdrożenia PGSI, CORE-10 oraz dashboardów czasu rzeczywistego, co miało przełożyć się na wzrost przychodów z około 2 mln GBP do 5,93 mln GBP. Dane publiczne potwierdzają zarówno zastosowanie tych narzędzi, jak i znaczący skok finansowy: Charity Commission wykazuje dochód na poziomie 2,18 mln GBP w roku zakończonym w marcu 2021, 5,93 mln GBP w 2024 oraz 7,87 mln GBP w 2025. Gordon Moody publikuje również dane wskazujące na znaczne spadki wyników PGSI i CORE-10 w trakcie procesu leczenia.

Wymaga to jednak precyzyjniejszej interpretacji. Po pierwsze, wzrost z 2,00 do 5,93 mln GBP nie wynosi „niemal 300%” – końcowa wartość stanowi około 296,5% wartości początkowej, co oznacza, że sam wzrost wynosi około 196,5%. Po drugie, współwystępowanie wdrożenia analityki i wzrostu przychodów nie dowodzi, że system ten był jedyną przyczyną sukcesu. Trust Impact przedstawia własne case study, w którym kierownictwo Gordon Moody przypisuje pracy nad purpose, Theory of Change i wizualizacją danych istotną rolę w usprawnieniu decyzji oraz komunikacji z fundatorami; wskazano tam również wzrost dochodu z 1,4 mln GBP do prognozowanych 7 mln GBP. Jest to wartościowy materiał ilustracyjny, obarczony jednak naturalnym konfliktem interesów, gdyż pochodzi od podmiotu świadczącego usługę. Rzetelna interpretacja brzmi zatem: transformacja danych współwystępowała z rozwojem organizacji i według świadectwa kierownictwa wspierała decyzje oraz fundraising, lecz bez projektu kontrfaktycznego nie można przypisać jej całego wzrostu przychodów.

Taka korekta jest w istocie idealnym zastosowaniem doktryny "never over-claim". Autor książki wiąże maksymalizację wpływu również z krytyką restrykcyjnego finansowania projektowego, postulując większy udział core funding, który pozwala utrzymać infrastrukturę i zdolność organizacyjną. Teza ta znajduje silne uzasadnienie w teorii kosztów transakcyjnych i organizational capacity. Finansowanie projektowe skłania do traktowania organizacji jako zbioru osobnych kontraktów, gdzie każdy grant ma własny budżet, formularz, wskaźniki, ramy czasowe i personel. W rzeczywistości organizacja opiera się jednak na wspólnej księgowości, IT, danych, zarządzaniu, kompetencjach, reputacji i relacjach. Jeśli fundatorzy finansują wyłącznie „bezpośrednie działania”, infrastruktura staje się pozornie kosztem administracyjnym, choć w istocie jest produkcyjnym warunkiem impact. Powstaje wówczas znany problem nonprofit starvation cycle: organizacje zaniżają koszty pośrednie, fundatorzy przyjmują te wartości za normę, a chroniczne niedoinwestowanie infrastruktury pogarsza zdolność organizacji do skutecznego działania i pomiaru. Z perspektywy maksymalizacji wpływu finansowanie rdzenne nie jest więc formą dobroczynności wobec organizacji, lecz może być inwestycją w zdolność produkcyjną wpływu. Nie oznacza to oczywiście, że każdy overhead jest wartościowy.

Maksymalizacja wpływu wymaga horyzontu intertemporalnego i świadomego zarządzania kosztem informacji

Wymaga to postawienia pytania marginalnego: czy kolejna złotówka przeznaczona na analitykę, kompetencje zarządcze, systemy danych lub rezerwę płynności zwiększy przyszły wpływ bardziej

niż ta sama kwota wydana bezpośrednio na usługę. Ekonomia klasyczna często analizuje efektywność w konkretnym punkcie czasowym, jednak organizacje społeczne potrzebują szerszego pojęcia – **dynamic efficiency**.

Inwestycja może bowiem przejściowo zmniejszać bieżący impact, by zwiększyć zdolność operacyjną w przyszłości. Szkolenie personelu, budowa interoperacyjnej infrastruktury danych, tworzenie rezerw finansowych czy prowadzenie eksperymentów nie generują natychmiastowych rezultatów, ale podnoszą przyszłą produktywność. Jeśli system finansowania premiuje wyłącznie bezpośrednie efekty w cyklu dwunastomiesięcznym, racjonalna organizacja będzie niedoinwestowywać tych aktywów. Otrzymujemy wówczas sytuację analogiczną do przedsiębiorstwa, które maksymalizuje wynik kwartalny kosztem badań i konserwacji.

Maksymalizacja wpływu musi zatem posiadać **horyzont intertemporalny**. Organizacja powinna oceniać nie tylko dzisiejszy impact, ale wartość ścieżki wpływu w czasie. Na tym etapie pojawiają się problemy dyskontowania przyszłości, niepewności oraz sprawiedliwości międzypokoleniowej. W programach klimatycznych, edukacyjnych czy profilaktycznych koszt ponosi się dziś, podczas gdy znacząca część korzyści materializuje się po latach. Wysoka stopa dyskontowa systematycznie dewaluuje przyszłe dobro; nie jest to zatem jedynie parametr ekonomiczny, lecz decyzja etyczna dotycząca wagi dobra przyszłych pokoleń.

Stevenson-Doddowska idea maksymalizacji powinna zostać skonfrontowana także z koncepcją **bounded rationality** Herberta Simona. Realne organizacje nie dysponują pełną informacją o wszystkich alternatywach, kosztach, przyszłych stanach świata czy funkcjach dobrobytu, przez co nie są w stanie matematycznie wyznaczyć globalnego maksimum. Często stosują zatem **satisficing** – poszukują rozwiązania wystarczająco dobrego względem przyjętego progu aspiracji. Z perspektywy zarządzania wpływem społecznym nie musi to być wada. Przy bardzo wysokich kosztach pozyskania danych, poszukiwanie teoretycznie najlepszego wariantu może pochłonąć więcej zasobów, niż przyniesie dodatkowej wartości.

Dlatego „maximise” należy interpretować jako **systematyczne zwiększanie wartości poprzez świadome porównywanie alternatyw**, a nie obietnicę odnalezienia obiektywnie optymalnego portfela. Organizacja powinna dążyć do poprawy jakości decyzji, ale jednocześnie znać koszt samego procesu optymalizacji. W warunkach niepewności staje przed wyborem: natychmiast sfinansować program albo najpierw zainwestować w badanie redukujące ryzyko. Teoria decyzji opisuje to jako **value of information**. Jeśli dwie interwencje wydają się podobnie skuteczne, lecz różnią się stopniem niepewności, niewielki pilotaż może mieć wysoką wartość, gdyż jego wynik pozwoli skierować większą część budżetu do właściwego działania. Informacja staje się wtedy samodzielnym dobrem produkcyjnym.

Nie oznacza to jednak, że każda ewaluacja jest opłacalna. Jeśli decyzja pozostanie niewrażliwa na prawdopodobny wynik badania, koszt jego przeprowadzenia przewyższy jego wartość. W tym punkcie sceptycyzm Stevenson-Dodda wobec rozbudowanych ewaluacji spotyka się z formalną teorią decyzji: **nie warto kupować informacji, która nie zmieni wyboru**.

W sektorze publicznym problem maksymalizacji jest jeszcze trudniejszy, gdyż administracja nie posiada jednego, jednoznacznego odpowiednika zysku.

Maksymalizacja wpływu musi być ograniczona ramami legalności i etyki

Działania państwa muszą realizować prawo, gwarantując równość, bezpieczeństwo, ciągłość, dostępność, transparentność oraz odpowiedzialność demokratyczną. Program o najwyższym mierzalnym impact może okazać się niedopuszczalny, jeśli narusza zasadę legalności lub prowadzi do arbitralnego różnicowania obywateli. Dlatego w analizie polityk publicznych użyteczna jest tradycja **public value**, która traktuje skuteczność jako jeden z elementów legitymowanego działania państwa, a nie jako wartość całkowicie autonomiczną. Administracja musi porównywać efekty, lecz jednocześnie respektować konstytucyjne ograniczenia i proceduralną sprawiedliwość.

Krytyka formalizmu nie może prowadzić do postulatu, by organ publiczny po prostu przenosił środki tam, gdzie algorytm wskazuje największy outcome. Administracja operuje w przestrzeni prawa, odpowiedzialności parlamentarnej, kontroli finansowej i równego dostępu. Zadaniem jest zatem **połączenie legalności z informacją o efekcie**, a nie zastąpienie jednej przez drugą. Im bardziej rozbudowujemy racjonalną metodę alokacji, tym ostrzej powraca problem polityczny: **kto ma prawo zdecydować, co maksymalizujemy?** Analityk potrafi obliczyć relacje koszt-efekt, ale nie może samodzielnie przesądzić, że jeden rodzaj ludzkiego dobra ma dwukrotnie większą wartość od innego. Dlatego mechanizmy priority setting wymagają procedury legitymizacyjnej. Literatura dotycząca alokacji w ochronie zdrowia podkreśla bowiem nie tylko cost-effectiveness, lecz także jawność kryteriów, udział interesariuszy i fairness procesu. W praktyce PBMA buduje się wielodyscyplinarne panele definiujące kryteria oraz rozważające inwestycje i dezinwestycje; sama analiza ekonomiczna nie zastępuje deliberacji.

W tym kontekście warto przywołać koncepcję **Accountability for Reasonableness** Normana Danielsa i Jamesa Sabina. W pluralistycznym społeczeństwie nie zawsze istnieje zgoda co do jednej teorii sprawiedliwej alokacji, zatem legitymizację można budować poprzez warunki racjonalnego procesu: publiczność przesłanek, ich relewantność dla rozsądnych uczestników, możliwość

odwołania oraz mechanizmy egzekwujące reguły. Maksymalizacja impact nie jest więc wyłącznie rachunkiem, lecz **instytucjonalizacją uzasadniania decyzji o tym, kto otrzyma zasób, a kto go nie otrzyma**.

Jedną z najbardziej kuszących idei jest przeliczenie wszystkich programów na wspólny współczynnik społecznego rezultatu przypadającego na złotówkę. Metodologicznie byłoby to niezwykle wygodne, jednak problem polega na tym, że różne outcomes mogą być niewspółmierne. Rok życia, uniknięcie bezdomności, poprawa samodzielności, wzrost czytelnictwa i ochrona prawna ofiary przemocy nie należą naturalnie do jednej skali. Monetyzacja, QALY, wellbeing valuation i SROI są próbami rozwiązania tego problemu, lecz każda z tych metod wnosi określone założenia aksjologiczne.

Nie oznacza to jednak konieczności rezygnacji z tych narzędzi. Wskazuje raczej na to, że wspólny współczynnik powinien być traktowany jako **instrument pomocniczy**, a nie ontologia wartości społecznej. W wielu przypadkach lepsze od fałszywa unifikacja może być podejście wielokryterialne. Autor książki posługuje się także pojęciem "impact velocity", sugerując możliwość dynamicznego przekierowywania kapitału do przedsięwzięć generujących wpływ szybciej. Jest to intuicyjnie atrakcyjne, lecz wymaga ostrożności, gdyż szybkość nie jest tożsama z wielkością ani trwałością.

Program może generować szybkie, lecz nietrwałe outcomes, podczas gdy inny, o długim okresie dojrzewania, przyniesie efekt systemowy. Edukacja wczesnodziecięca, profilaktyka zdrowotna, badania naukowe czy reforma instytucjonalna mają często odroczoną krzywą rezultatu. Jeśli fundator zacznie stale premiować "impact velocity", może doprowadzić do selekcji przeciwko inwestycjom długoterminowym. Szybkość powinna być zatem jednym z parametrów portfela, a nie nadrzędnym kryterium. Nauka o systemach złożonych uczy również, że system maksymalnie odchudzony może stać się kruchy. Redundancja, rezerwa, różnorodność i zapas zdolności wyglądają nieefektywnie w normalnych warunkach, ale zyskują ogromną wartość po wystąpieniu szoku. Fundacja utrzymująca rezerwę płynności może w danym roku pomóc mniejszej liczbie osób niż podmiot wydający sto procent przychodu, jednak po kryzysie finansowym pierwsza może kontynuować misję, podczas gdy druga przestać istnieć.

Maksymalizacja wpływu jako systemowa ekonomia polityczna

Szpital dysponujący wolnymi mocami przerobowymi w normalnym okresie może wydawać się mniej efektywny niż placówka stale pracująca na pełnych obrotach; jednak w czasie epidemii sytuacja ta ulega odwróceniu. Maksymalizacja wpływu musi zatem uwzględniać resilience-adjusted efficiency. Pytanie nie brzmi jedynie: „ile wpływu generujemy w warunkach stabilnych?", lecz: „jaką zdolność do realizacji misji zachowujemy w obliczu prawdopodobnych szoków?". Jeszcze bardziej wymagający poziom analizy pojawia się, gdy organizacja wykracza poza własny budżet. Może się bowiem okazać, że największy wpływ osiągnie nie poprzez rozbudowę własnej działalności, lecz poprzez wzmocnienie innego podmiotu, zmianę regulacji, udostępnienie danych, szkolenie partnerów lub zamknięcie własnego programu i przekazanie zasobów komuś skuteczniejszemu.

Jest to silny test purpose. Organizacja skupiona na własnym wzroście utożsamia maksymalizację wpływu z powiększaniem własnego footprintu. Z kolei organizacja skoncentrowana na problemie społecznym potrafi zaakceptować, że jej optymalny wkład ma charakter pośredni. W ekonomii można to opisać jako przejście od optymalizacji lokalnej do systemowej. Decyzja racjonalna lokalnie nie musi maksymalizować wyniku całego ekosystemu; każda organizacja może poprawiać własne KPI, podczas gdy system jako całość generuje duplikację działań, luki terytorialne i nadmierne koszty transakcyjne. Stąd idea collective impact i sieciowego governance: rozwiązywanie złożonych problemów wymaga wspólnej infrastruktury danych, podziału ról, uzgodnionej Theory of Change i wzajemnego uznania wkładu. W takim układzie pytanie o to, „czyja była zasługa?", staje się mniej istotne niż kwestia tego, czy system jako całość działa lepiej.

Budżet nie jest jedynym zasobem podlegającym maksymalizacji. Równie ważne są czas kierownictwa, zdolność personelu, kapitał relacyjny i reputacja. Program może być tani finansowo, lecz pochłaniać ogromną część uwagi CEO. Partnerstwo może przynosić niewielkie bezpośrednie rezultaty, ale budować sieć niezbędną do późniejszej zmiany systemowej. Kampania z kolei może zużywać kapitał reputacyjny, który nie jest nieskończony. Dlatego pełna teoria alokacji wpływu powinna uwzględniać wiele budżetów równolegle.

Można dysponować funduszami, lecz brakować kompetencji wykonawczych. Można mieć ludzi, ale nie posiadać mandatu politycznego. Można prowadzić świetny program, lecz nie mieć zdolności do jego skalowania. Najbardziej ograniczający zasób – bottleneck – może zmieniać się w czasie. Dołożenie finansowania do systemu ograniczonego liczbą specjalistów nie zwiększy proporcjonalnie wpływu; może jedynie podnieść koszty lub obniżyć jakość. Właśnie dlatego zwiększanie budżetu nie jest uniwersalnym odpowiednikiem maksymalizacji dobra. Można zatem

sformułować bardziej rygorystyczną wersję etapu Maximise: organizacja nie powinna pytać tylko o to, który program ma najwyższy wynik, lecz badać, jak zmieni się oczekiwana wartość społeczna, dystrybucja korzyści, ryzyko, wiedza i odporność po przesunięciu krańcowej jednostki zasobu z jednego zastosowania do innego. To konstrukcja znacznie bogatsza niż prosty ranking.

W praktyce wymaga to połączenia Theory of Change z informacją o kosztach, heterogeniczności rezultatów, additionality, marginalnych returns, equity, uncertainty, capacity, czasie i konsekwencjach dezinwestycji. Metody takie jak MCDA, PBMA, ekonomika zdrowia, teoria portfela, behawioralna teoria organizacji czy etyka dystrybucyjna nie są tu konkurencyjnymi szkołami, lecz warstwami jednego problemu. W tym świetle koncepcja Maximise autorstwa Stevensona i Dodda zachowuje dużą wartość, choć wymaga naukowego rozszerzenia. Ich najważniejsza intuicja jest trafna: pomiar wpływu powinien prowadzić do decyzji o alokacji, a nie kończyć się na raporcie. Organizacja, która wie, że dany program generuje słabe rezultaty, a mimo to automatycznie go finansuje, nie zarządza wpływem – jedynie go opisuje. Zasada 80/20 może tu służyć jako prosta heurystyka zabezpieczająca exploration, 95/5 jako metafora zaniedbania strategicznej uwagi, a Live Data jako mechanizm skracający czas między informacją a realokacją. Autorzy słusznie kierują więc analizę ku zasobom, a nie wyłącznie pomiarowi. Jednak nauki ekonomiczne i społeczne pokazują, że „maksymalizować" nie oznacza prostego przesuwania kapitału do najwyższego KPI. Należy uwzględnić marginalność, opportunity cost, diminishing returns, equity, rights constraints, uncertainty, learning value, resilience, irreversibility, distributional consequences oraz zależności systemowe.

Najbardziej dojrzałe pytanie zarządu nie brzmi więc: „który program jest skuteczny?", lecz: „jaki portfel działań, przy naszych ograniczeniach i wartościach, maksymalizuje uzasadnioną społecznie wartość w czasie, nie przerzucając kosztów na grupy, których nie widzi nasz podstawowy KPI?".

To właśnie tutaj zarządzanie wpływem przechodzi z poziomu ewaluacji do poziomu ekonomii politycznej organizacji. Każda realokacja tworzy beneficjentów i przegranych, każda dezinwestycja narusza istniejący układ interesów, każda waga w systemie MCDA materializuje konkretną teorię dobra, a każdy budżet jest jednocześnie dokumentem ekonomicznym i normatywnym. Warto w tym miejscu przejść do analizy różnic sektorowych. Ta sama architektura Understand-Measure-Maximise działa inaczej w przedsiębiorstwie, administracji publicznej, NGO, przedsiębiorstwie społecznym, fundacji grantodawczej, funduszu impact investing czy organizacji sportowej. Odmienna jest struktura bodźców, odpowiedzialności, horyzontów czasowych i dostępu do danych, a co za tym idzie – inne są patologie pomiaru i możliwości rzeczywistego maksymalizowania wpływu.

> Czy organizacja, która nigdy nie raportuje porażżek, jest rzeczywiście skuteczna, czy jedynie posługuje się definicją sukcesu tak elastyczną, że stała się ona niefalsyfikowalna? Prawdziwa dojrzałość nie polega na osiąganiu nieskazitelnych wyników, lecz na posiadaniu odwagi, by zapytać: kogo nasz system pomiaru nie widzi i jaką cenę płacimy za iluzję nieomylności. W świecie złożonych problemów społecznych najcenniejszym kapitałem nie jest obraz perfekcji, lecz zdolność do wykrywania własnej omylności, zanim jej koszt stanie się zbyt wysoki.

Inspiracje

[1] 

"The Difference How to Understand Measure and Maximise Social Impact" Matt Stevenson-Dodd

Practical Inspiration Publishing | ISBN: 9781788609562

Matt Stevenson-Dodd jest brytyjskim dyrektorem organizacji non-profit, konsultantem ds. wpływu społecznego i autorem specjalizującym się w przywództwie charytatywnym i ocenie wpływu społecznego. Jest założycielem i dyrektorem generalnym Trust Impact, firmy konsultingowej wspierającej organizacje charytatywne, przedsiębiorstwa społeczne oraz podmioty publiczne i komercyjne w określaniu i monitorowaniu ich efektów społecznych. W latach 2010-2019 Stevenson-Dodd pełnił funkcję dyrektora generalnego Street League, brytyjskiej organizacji charytatywnej wykorzystującej sport do walki z bezrobociem wśród młodzieży. Był pionierem w raportowaniu danych o wpływie sportu na żywo i publicznie opowiadał się za transparentnością i wyciąganiem wniosków z porażek organizacji. Jako były przewodniczący Football Beyond Borders i członek zarządu St John Ambulance, został przyjęty w poczet Czcigodnego Zakonu Świętego Jana w 2020 roku. Ukończył z wyróżnieniem studia MBA w Nottingham University Business School.

Matt Stevenson-Dodd *is a British non-profit executive, social impact consultant, and author specializing in charity leadership and social impact evaluation. He is the founder and chief executive officer of Trust Impact, a consultancy that supports charities, social enterprises, and public and commercial bodies in clarifying and tracking their social outcomes. From 2010 to 2019, Stevenson-Dodd served as chief executive of Street League, a UK charity using sport to tackle youth unemployment, where he pioneered live impact data reporting and publicly advocated for transparency and learning from organizational failure. A former chair of Football Beyond Borders and trustee of St John Ambulance, he was admitted as a Member of the Venerable Order of Saint John in 2020. He holds an MBA with Distinction from Nottingham University Business School.*

Trust Impact

Literatura uzupełniająca

Książki

Książki autorów przywoływanych i literatura pokrewna:



[1] "The Fearless Organization: Creating Psychological Safety in the Workplace for Learning, Innovation, and Growth"

Amy Edmondson

John Wiley & Sons | ISBN: 9781119477228

[2] "Organizations"

James March

[3] "behavioral theory of the firm"

James March



[4] "Strauss and Mayer's Emergency Department Management"

Robert W. Strauss, Thom A. Mayer

2013 | McGraw Hill Professional | ISBN: 9780071762397

Artykuły naukowe

Prace autorów przywoływanych w tekście:

[1] "Not So Different After All: A Cross-Discipline View Of Trust"

Denise M. Rousseau, Sim B. Sitkin, Ronald S. Burt, Colin F. Camerer

1998 | Academy of Management Review | DOI: 10.5465/amr.1998.926617

Źródło

[2] "Psychological Safety and Learning Behavior in Work Teams"

Amy C. Edmondson

1999 | Administrative Science Quarterly | DOI: 10.2307/2666999

Źródło

[3] "Methodological fit in management field research"

Amy C. Edmondson, Stacy E. McManus

2007 | Academy of Management Review | DOI: 10.5465/amr.2007.26586086

Źródło

[4] "Psychological Safety: The History, Renaissance, and Future of an Interpersonal Construct"

Amy C. Edmondson, Zhike Lei

2014 | Annual Review of Organizational Psychology and Organizational Behavior | DOI: 10.1146/annurev-orgpsych-031413-091305

Źródło

[5] "Making it safe: the effects of leader inclusiveness and professional status on psychological safety and improvement efforts in health care teams"

Ingrid M. Nembhard, Amy C. Edmondson

2006 | Journal of Organizational Behavior | DOI: 10.1002/job.413

Źródło

[6] "Reconceptualizing the Determinants of Risk Behavior"

Sim B. Sitkin, Amy L. Pablo

1992 | Academy of Management Review | DOI: 10.2307/258646

Źródło

[7] "Explaining the Limited Effectiveness of Legalistic “Remedies” for Trust/Distrust"

Sim B. Sitkin, Nancy L. Roth

1993 | Organization Science | DOI: 10.1287/orsc.4.3.367

Źródło

[8] "DETERMINANTS OF RISKY DECISION-MAKING BEHAVIOR: A TEST OF THE MEDIATING ROLE OF RISK PERCEPTIONS AND PROPENSITY."

Sim B. Sitkin, Laurie R. Weingart

1995 | Academy of Management Journal | DOI: 10.2307/256844

Źródło

[9] "Exploration and Exploitation in Organizational Learning"

James G. March

1991 | Organization Science | DOI: 10.1287/orsc.2.1.71

Źródło

[10] "A Garbage Can Model of Organizational Choice"

Michael D. Cohen, James G. March, Johan Pelck Olsen

1972 | Administrative Science Quarterly | DOI: 10.2307/2392088

Źródło

[11] "Organizational Learning"

Barbara Levitt, James G. March

1988 | Annual Review of Sociology | DOI: 10.1146/annurev.so.14.080188.001535

Źródło

[12] "The New Institutionalism: Organizational Factors in Political Life"

James G. March, Johan Pelck Olsen

1983 | American Political Science Review | DOI: 10.2307/1961840

Źródło

[13] "Wide Reflective Equilibrium and Theory Acceptance in Ethics"

Norman Daniels

1979 | The Journal of Philosophy | DOI: 10.2307/2025881

Źródło

[14] "From Chance to Choice"

Allen E. Buchanan, Dan W. Brock, Norman Daniels, Daniel I. Wikler

2000 | Cambridge University Press eBooks | DOI: 10.1017/cbo9780511806940

Źródło

[15] "Limits to Health Care: Fair Procedures, Democratic Deliberation, and the Legitimacy Problem for Insurers"

Norman Daniels, James E. Sabin

1997 | Philosophy &amp Public Affairs | DOI: 10.1111/j.1088-4963.1997.tb00082.x

Źródło

[16] "Bystander Ethics and Good Samaritanism: A Paradox for Learning Health Organizations."

J. Sabin, N. Cocoros, Crystal Garcia, J. Goldsack, K. Haynes

2019 | The Hastings center report | DOI: 10.1002/hast.1031

Źródło

[17] "How Can Clinical Ethics Committees Take on Organizational Ethics? Some Practical Suggestions"

J. Sabin

2016 | Journal of Clinical Ethics | DOI: 10.1086/jce2016272111

Źródło

[18] **"Using Moral Distress for Organizational Improvement"**

J. Sabin

2017 | Journal of Clinical Ethics | DOI: 10.1086/jce2017281033

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: f9f3ee6f-7bd1-437d-9bfe-913717bbb51f