

*Marta Marcjanik*¹

*Jakub Krawiec*²

*Wojciech Białaszek*³

Ukryte koszty skuteczności interwencji nudge

Abstrakt

Choć interwencje nudge zyskały dużą popularność w polityce publicznej dzięki swojej skuteczności, niskim kosztom wdrożenia i wizerunkowi „miękkiego” oddziaływania, coraz więcej badań ujawnia ich mniej oczywiste i często pomijane koszty. Artykuł ma charakter przeglądowo-koncepcyjny i stanowi próbę zwrócenia uwagi projektantów polityk publicznych na fakt, że skuteczność nie zawsze oznacza poprawę dobrostanu. Autorzy przedstawiają empirycznie osadzony katalog kosztów interwencji behawioralnych. Te zostały uporządkowane w cztery obszary oddziaływania: 1) emocjonalne i poznawcze, 2) behawioralne, 3) społeczne i relacyjne oraz 4) wtórne. Proponowana typologia może stanowić punkt wyjścia do lepszego rozpoznawania potencjalnych kosztów i projektowania bardziej świadomych działań w polityce publicznej.

Słowa kluczowe: interwencje behawioralne, nudge, polityka publiczna, ekonomia behawioralna

Kody klasyfikacji JEL: D03, D60, D91, I38

¹ Uniwersytet SWPS, e-mail: mmarcjanik@swps.edu.pl, <https://orcid.org/0009-0007-8879-8545>

² Uniwersytet SWPS, e-mail: jkrawiec@swps.edu.pl, <https://orcid.org/0000-0002-8422-8090>

³ Uniwersytet SWPS, e-mail: wbialaszek@swps.edu.pl, <https://orcid.org/0000-0002-4672-4376>

Hidden Costs of the Effectiveness of Nudge Interventions

Abstract

Although behavioural interventions have gained significant popularity in public policy due to their effectiveness, low implementation costs and reputation as a “soft” form of influence, a growing body of research reveals their less obvious – and often overlooked – costs. This article adopts a conceptual and narrative review approach, aiming to draw policymakers’ and decision architects’ attention to the fact that effectiveness does not always equate to improved well-being. The authors present an empirically grounded catalogue of the costs of nudges, organised into four domains of impact: emotional and cognitive costs, behavioural costs, social and relational costs, and secondary costs. The proposed typology provides a starting point for better identifying potential unintended costs and for designing more mindful public policy interventions.

Keywords: behavioural interventions, nudge, public policy, behavioural economics

JEL Classification Codes: D03, D60, D91, I38

Wprowadzenie

Interwencje behawioralne zdobyły w ostatnich dekadach silną pozycję w projektowaniu polityk publicznych. Postrzegane jako skuteczne, tanie i subtelne narzędzia wpływu – w odróżnieniu od twardych regulacji prawnych – zyskały reputację miękkich oddziaływań wspierających obywateli w podejmowaniu lepszych decyzji (Sunstein et al., 2018). Jednym z najbardziej znanych przykładów takiego podejścia jest zastosowanie opcji domyślnej w partycypacji w programach oszczędnościowych: automatyczny zapis sprawia, że pozostanie w programie staje się stanem wyjściowym, co wykorzystuje ludzką skłonność do inercji i utrzymywania *status quo*. Badania pokazują, że takie ustawienie znacząco zwiększa partycypację w planach emerytalnych, co potwierdzają analizy programów w USA (Thaler i Benartzi, 2004), jak i w Polsce (Jedynak, 2019).

Publikacja książki *Nudge* Richarda Thalera i Cassa Sunsteina (2008) zapoczątkowała systematyczne stosowanie interwencji behawioralnych w politykach publicznych, ułatwiając skalowanie tych praktyk. Od tego czasu wzrosła świadomość ich potencjału i zakres zastosowań – od zdrowia, przez edukację, po ochronę środowiska.

Obecnie najpopularniejszą interwencją behawioralną, czyli interwencją zmieniającą zachowanie bez stosowania kar i nagród (szczególnie ekonomicznych), wyko-

rzyszywaną w politykach publicznych, jest nudging (Sunstein et al., 2018). Zgodnie z definicją zaproponowaną przez Thalera i Sunsteina (2008) nudging to „każdy aspekt architektury sytuacji wyboru, który przewidywalnie zmienia zachowanie, bez zabrania jakichkolwiek opcji, ani istotnej zmiany korzyści ekonomicznych”.

W polskiej literaturze słowo „nudge” przyjęło określenie „impuls” (Thaler i Sunstein, 2008). Zgodnie z definicją słownikową angielsko-polską rzeczownik „nudge” może być rozumiany jako „kuksaniec” lub „szturchnięcie”, natomiast jego wersja czasownikowa oznacza „trącić” lub „szturchnąć” (*Cambridge Dictionary*, b.d.). Żadne z powyżej przedstawionych tłumaczeń nie oddaje w pełni pierwotnego znaczenia tego terminu. W związku z tym zdecydowano pozostawić oryginalną angielską formę.

Dotychczasowe zastosowania nudge’y można zaobserwować m.in. w programach polityki społecznej USA i Wielkiej Brytanii, gdzie odegrały kluczową rolę dzięki wdrożeniom Behavioural Insights Team (Hallsworth et al., 2018). Znaczący impuls nadał im także *Executive Order* prezydenta Obamy (Obama, 2015), promujący wykorzystanie nauk behawioralnych w administracji publicznej, z kolei na poziomie międzynarodowym – inicjatywy takie jak UN Behavioral Science Group oraz WHO Behavioral and Cultural Insights.

Paradoksalnie, właśnie skuteczność nudge’y i ich niska inwazyjność mogą sprzyjać słabej refleksji nad potencjalnymi kosztami wprowadzanych oddziaływań, które łatwo mogą wymykać się na etapie oceny zastosowanego podejścia. Określamy wspomniane koszty jako ukryte, ale nie w sensie dosłownym. Pomimo że są one mierzalne, dotyczą dobrostanu psychicznego, relacji społecznych czy zaufania do instytucji, to bardzo często, w ocenie zastosowanej interwencji, dominuje skupienie się na jej skuteczności, której wskaźnikiem jest siła zmiany danego zachowania. Koszty mają charakter ukryty, ponieważ często ich pomiar jest pomijany. Natomiast, na przykład na poziomie jednostki, koszty mają charakter jawny i bezpośredni.

Benartzi i współautorzy (2017) uznają interwencje nudge za efektywne, gdy prowadzą one do oczekiwanych zmian w zachowaniu. Autorzy podkreślają jednak, że istotna jest także efektywność kosztowa, czyli opłacalność interwencji. Z takiej perspektywy interwencja mająca prowadzić do zwiększenia oszczędności emerytalnych będzie uznana za efektywną, jeśli osiągnie zakładany wzrost oszczędności, nawet w sytuacji, gdy zmiana zachowania jest niewielka, ale jest równocześnie opłacalna z perspektywy polityki.

W niniejszym artykule interwencję nudge uznajemy za efektywną, gdy realizuje cele określone przez architekta decyzji. Jeśli celem jest natychmiastowa zmiana zachowania, jej osiągnięcie stanowi podstawę pozytywnej oceny, nawet jeżeli z perspektywy dłuższej ramy czasowej, ocena ogólnej efektywności budzi wątpliwości. Może tak być, gdy ewaluacja oddziaływania ogranicza się jedynie do

wybranych wskaźników, pomijając (celowo lub nie) inne konsekwencje wprowadzonej interwencji.

Celem artykułu jest pokazanie szerokiej gamy kosztów związanych z interwencjami behawioralnymi. Tekst odwołuje się do hipotezy, że skuteczność interwencji nie zawsze przekłada się na poprawę dobrostanu, a ignorowanie mniej oczywistych kosztów może prowadzić do zagrożeń dla jakości życia oraz relacji między obywatelami a instytucjami publicznymi. Oryginalność artykułu polega na zaproponowaniu katalogu kosztów ubocznych, który może stanowić podstawę do dalszych badań.

Nudge w politykach publicznych

Popularność nudge'y

Do głównych zalet nudge'y należy ich skuteczność, niski koszt oraz subtelny charakter. Della Vigna i Linos (2020) wykazali, że średni efekt interwencji w publikowanych badaniach wynosił 33,5% (wzrost reakcji w porównaniu do grupy kontrolnej), a w działaniach *nudge units* – 8,1%.

Mertens i współautorzy (2022) wskazują, że interwencje oparte na architekturze wyboru są co najmniej równie skuteczne jak tradycyjne, bardziej kosztowne metody. Z ich metaanalizy wynika, że skutkują one małym do średniego efektu zmiany zachowania, porównywalnego z kampaniami edukacyjnymi czy zachętami finansowymi. Jednocześnie ponowna analiza tych wyników sugeruje, że obserwowane efekty mogą być częściowo związane z efektem publikacyjnym lub realizacją badań w warunkach laboratoryjnych (Maier et al., 2022).

Analizując koszty, Benartzi z zespołem (2017) wykazali, że, pod względem stosunku efektu do kosztu, nudge często wypadają taniej niż tradycyjne narzędzia, np. ulgi podatkowe.

Interwencje tego typu są również stosunkowo łatwe i szybkie w implementacji, gdyż niekoniecznie muszą wymagać uruchamiania procedur prawnych czy ustawodawczych, tak jak to może mieć miejsce w przypadku klasycznych narzędzi polityk publicznych (Krawiec et al., 2021).

Kolejną zaletą jest brak negatywnego wpływu na poczucie autonomii jednostki (Van Roekel et al., 2023; Wachner et al., 2020, 2021b), również w sytuacji świadomości oddziaływania interwencji, czyli jej transparentności.

Ważną zaletą dla polityk publicznych może być też możliwość „anulowania” efektów interwencji. Jak wskazują Sunstein i zespół (2018), po osiągnięciu rezultatu, w momencie usunięcia interwencji nudge ze środowiska decyzyjnego, zachowanie

jednostki powinno powrócić do stanu sprzed interwencji. Jest to jeden z argumentów autorów: gdy stosowanie nudgingu może być dyskusyjne etycznie, w momencie kiedy cele społeczne będą osiągnięte, można interwencje „wygasić”, bez ich dalszego, długotrwałego wpływu.

Podejście krytyczne do nudge’y: teoria, ewaluacja i wyniki badań empirycznych

Dwie dekady wprowadzania i ewaluacji różnego rodzaju nudge’y dostarczają rozległej wiedzy nie tylko na temat tego, jakie pożądane rezultaty mogą przynosić, ale też z jakimi ryzykami się wiążą. Powstało wiele paradygmatów i metodologii, które wspierają ich projektowanie zgodnie z ustrukturyzowanymi zasadami (m.in. Dolan et al., 2012; Lades i Delaney, 2022; Mills i Whittle, 2024; Tummers, 2019), a także rozbudowanych debat teoretycznych (m.in. De Ridder et al., 2024; Kuyer i Gordijn, 2023; Schmidt i Engelen, 2020) koncentrujących się na potencjalnie negatywnych aspektach tych interwencji.

Krytyka nudge’y formułowana jest najczęściej w kategoriach etycznych i normatywnych, takich jak naruszenie autonomii, brak transparentności, manipulacja czy paternalizm. Refleksje te mają zazwyczaj wysoki poziom ogólności i koncentrują się na problemach systemowych, rzadziej analizując bezpośrednie koszty funkcjonowania interwencji. W tym kontekście Hertwig i Grüne-Yanoff (2017) wskazują na *boosting* jako alternatywę ukierunkowaną na wzmacnianie kompetencji decyzyjnych, a nie kierowanie wyborem w określoną stronę.

W badaniach nad interwencjami typu nudge dominującym przedmiotem oceny jest wielkość zmiany zachowania, która może być zoperacjonalizowana poprzez osiągnięcie założonego kryterium (zewnętrzne lub analityczne). Najczęściej stosuje się podejście eksperymentalne, które może uwzględniać analizę ogólnej efektywności kosztowej lub nie (DellaVigna i Linos, 2020; Benartzi et al., 2017; Mertens et al., 2022), ale też analizy danych zastanych (Beshears et al., 2015; Richter et al., 2018).

W efekcie istniejące narzędzia ewaluacyjne nie uwzględniają nieprzewidzianych kosztów, które są tematem niniejszego artykułu – mimo że w literaturze coraz częściej podkreśla się potrzebę uwzględniania również mniej oczywistych, ale realnych konsekwencji tych interwencji (Hummel i Maedche, 2019; Mavrot et al., 2025).

Zaproponowane w tym artykule podejście, skupiając się na zestawieniu kosztów interwencji nudge, wnosi kilka uzupełniających perspektyw. Proponowany katalog uszczegóławia analizę kosztów, pozwalając identyfikować konkretne mechanizmy emocjonalne, poznawcze, behawioralne i społeczne, zamiast ograniczać się do ogólnych zarzutów. Wyróżnienie czterech poziomów kosztów umożliwia ocenę interwencji

zarówno w perspektywie krótkoterminowej, jak i długofalowej, z uwzględnieniem dobrostanu i relacji społecznych. Typologia ta ma także potencjał operacjonalizacji – może wspierać projektowanie badań i wskaźników ewaluacyjnych oraz sprzyjać bardziej odpowiedzialnemu planowaniu interwencji.

Rodzaje kosztów nudge’y

Katalog kosztów

Tabela 1 prezentuje wszystkie zidentyfikowane koszty wraz z przypisanym typem, przykładami oraz źródłem, z którego pochodzą.

Tabela 1. Koszty nudge’y

Typ kosztu	Rodzaj kosztu	Przykłady	Źródła
1) Koszty emocjonalne i poznawcze	negatywne emocje (wstyd, poczucie winy, frustracja, irytacja – tzw. <i>emotional tax</i>)	poczucie winy przy etykietach kalorycznych	Allcott i Kessler (2019); Barev et al. (2021); Jensen et al. (2020); Rusmevichientong et al. (2021); Thunström (2019); Vitek i Syed Shah (2019)
	lęk/poczucie zagrożenia prywatności	ostrzeżenia o ochronie danych wywołujące strach	Barev et al. (2021); Dorison et al. (2022)
	obciążenie zasobów poznawczych/spadek jakości wykonania zadania	gorsze wyniki w czytaniu przy takim samym czasie poświęconym na czytanie	Di Matola (2020); Kalil et al. (2023); Nafziger (2020)
	zniechęcenie do refleksyjnego działania/obniżenie sprawczości	upraszczające domyślne opcje mniej świadomych decyzji	Chesterley (2015)
	błędna interpretacja informacji	nadmierne oszczędzanie w wyniku utrzymania zbyt wysokiej domyślnej składki emerytalnej	Tor (2020)
	osłabienie samooceny	uświadomienie sobie niskiego statusu ekonomicznego	Beshears et al. (2015)
2) Koszty behawioralne	działania kontrproduktywne (działania pogarszające dobrostan)	nadmierne ubezpieczenie, pogorszenie jakości diety	Harrison et al. (2024); P.C. Medina (2021); Richter et al. (2018); Thunström (2018); Tor (2020)
	efekty kompensacyjne/licencja moralna (poczucie przyzwolenia na mniej etyczne działania po wcześniejszym dobrym uczynku)	„nagroda” po pożądanym zachowaniu, mniej działań pro-eko po jednym działaniu	Clot et al. (2022); Marchiori et al. (2017)
	unikanie/opór (reaktancja)	wypisy z list mailingowych, ignorowanie komunikatów	Campigotto et al. (2025); Damgaard i Gravert (2018); Shah et al. (2021)

Typ kosztu	Rodzaj kosztu	Przykłady	Źródła
	nietrwałość/zanik efektów	efekty krótkookresowe w wyborach żywieniowych/ekologicznych	Gravert i Kurz (2021)
	przeniesienie	zwiększenie spożycia zarówno zdrowych, jak i niezdrowych produktów	Wilson et al. (2015)
3) Koszty społeczne i relacyjne	spadek prospołeczności/wsparcia dla dobra wspólnego	mniej wsparcie dla podatków klimatycznych, spadek altruizmu przy obserwowalności	Bolton et al. (2020)
	efekty dystrybucyjne/pogłębianie nierówności	większe obciążenie osób oszczędnych	Thunström (2018)
	spadek zaufania do ekspertów/nadawcy	obniżone zaufanie do lekarzy, negatywna ocena doradcy finansowego	Avitzour et al. (2018); Meunier et al. (2024)
	percepcja manipulacji/braku przejrzystości	odbiór nudge'a jako manipulacji, utrata poczucia autonomii	Kuyer i Gordijn (2023); Michaelsen et al. (2024); Schmidt i Engelen (2020); Wachner et al. (2021a)
4) Koszty wtórne	erozja norm etycznych	normalizacja i racjonalizacja korupcji	Dimant i Schulte (2016)
	spadek legitymizacji polityk publicznych	niższe poparcie dla podatków węglowych	Hagmann et al. (2018)

Źródło: opracowanie własne.

Typy kosztów nudge'y – omówienie szczegółowe

1) Koszty emocjonalne i poznawcze

W serii badań dotyczących informacji żywieniowych zauważono, że etykiety z danymi o wartości odżywczej (Rusmevichientong et al., 2021) oraz oznaczenia kaloryczności dań w menu (Thunström, 2019) mogą wywoływać u konsumentów poczucie winy. Efekty emocjonalne zauważył również zespół Jensena (2020), Vitek i Syed Shah (2019), Barev et al. (2021), Dorison et al. (2022) oraz Allcott i Kessler (2019).

Kalil i współautorzy (2023) wykazali, że przypomnienia wysyłane do rodziców w celu zwiększenia czasu spędzonego na czytaniu przez ich dzieci, choć skuteczne pod względem ilości czasu, prowadziły do gorszych wyników w czytaniu, co sugeruje spadek jakości tej aktywności. Chesterley (2015) zaobserwował, że połączenie upraszczających interwencji (takich jak domyślna zgoda na usługę lub zintegrowanie kilku decyzji w jedną) może prowadzić do zniechęcenia do aktywnego i świadomego podejmowania decyzji, co z kolei może podważać rozwój kompetencji obywatelskich i samodzielności decyzyjnej. Tego typu koszty zauważali również Nafziger (2020), Di Matola (2020), Tor (2020) oraz Beshears et al. (2015).

Negatywne reakcje emocjonalne, takie jak poczucie winy, frustracja czy irytacja, osłabiają dobrostan odbiorców, a więc działają wbrew jednemu z kluczowych celów polityk publicznych, jakim jest poprawa jakości życia. Oznacza to, że interwencja mająca na celu naprawę jednego problemu społecznego może jednocześnie niepostrzeżenie pogarszać inne aspekty funkcjonowania jednostki. Co więcej, koszty poznawcze, wynikające z działania interwencji na poziomie automatycznym, bez pełnego zaangażowania i zrozumienia, mogą obniżyć jakość realizowanych zadań lub podejmowanych decyzji. Choć interwencje te często prowadzą do pożądanego zachowań, to brak refleksyjnej motywacji sprawia, że działania te bywają powierzchowne, mniej przemyślane i mniej efektywne z punktu widzenia odbiorcy. W konsekwencji, nawet skuteczna z perspektywy behawioralnej interwencja może nie przynosić oczekiwanych korzyści w praktyce. Częściowym rozwiązaniem tej kwestii jest naturalna ewolucja rozumienia interwencji *nudge*. Przykładowo koncepcja *nudge plus* włącza elementy refleksji i samodzielnego rozważenia decyzji przez odbiorcę, zwiększając tym samym głębię i skuteczność interwencji (Banerjee i John, 2024).

2) Koszty behawioralne

W badaniu Richter i współautorów (2018) informacja w miejscu sprzedaży o tym, ilu konsumentów wybiera owoce morza ze zrównoważonych hodowli, zwiększyła ogólną sprzedaż kategorii owoców morza ogólnie, ale nie przełożyła się na wybór produktów zrównoważonych, co rodzi pytania o precyzję sygnałów społecznych.

Thunström (2018) wykazała, że przypomnienia o kosztach alternatywnych, choć zaprojektowane z myślą o ograniczeniu wydatków, sprawiły, że to osoby już wcześniej oszczędne jeszcze bardziej zmniejszyły swoje zakupy, co może być niekorzystne z perspektywy jakości życia. Na koszty behawioralne zwracali uwagę również Marchiori et al. (2017), Damgaard i Gravert (2018), Wilson et al. (2015), Clot et al. (2022), Shah et al. (2021), Harrison et al. (2024), Medina (2021), Campigotto et al. (2025) oraz Gravert i Kurz (2021).

Koszty behawioralne pokazują, że choć określone zachowanie zostaje osiągnięte, może ono jednocześnie wywołać efekt uboczny w innym obszarze, np. poprzez kompensacyjne lub licencyjne działania, które znoszą lub osłabiają ogólny bilans korzyści. Dodatkowo, ponieważ interwencje są często wdrażane na szeroką skalę, bez precyzyjnego dopasowania do potrzeb i cech odbiorców, mogą generować wyraźne koszty dla części populacji (potrzebę dopasowania *nudge* do cech demograficznych zauważa Szkudlarek (2022)). Nawet jeśli ich ogólny efekt jest pozytywny, to z perspektywy konkretnych grup lub jednostek mogą one być źródłem kosztów, skłaniając

do refleksji zarówno z perspektywy etyki, jak i z perspektywy praktycznej odnośnie do równowagi stosowanych rozwiązań.

3) Koszty społeczne i relacyjne

Bolton et al. (2020) wykazali, że sama świadomość bycia obserwowanym podczas zadania może prowadzić do osłabienia działań prospołecznych. W eksperymencie, w którym uczestnicy wiedzieli, że są obserwowani przez anonimowego obserwatora podczas podejmowania decyzji o przekazaniu pieniędzy, w przeciwieństwie do warunków prywatnych, przekazywali mniej na rzecz innych. Meunier et al. (2024) pokazali, że zastosowanie domyślnych ustawień promujących ryzykowne inwestycje może skutkować negatywną oceną doradcy finansowego, podważając zaufanie do profesjonalnych rekomendacji. W warunkach promujących ryzykowne inwestycje doradca majątkowy proponował domyślną alokację wynoszącą 100% inwestycji w aktywa ryzykowne (i 0% na bezpiecznym rachunku). Skutkowało to niższą oceną porady doradcy w porównaniu do warunków, gdy doradzana była alokacja 0% w ryzykowne aktywa. Z kolei Avitzour et al. (2018) stwierdzili, że stosowanie nudge'y przez lekarzy w relacjach z pacjentami może osłabić zaufanie do lekarzy. Koszty społeczne zaobserwowali również: Thunström (2018), Kuyer i Gordijn (2023), Michaelsen et al. (2024), Schmidt i Engelen (2020), Wachner, Adriaanse i De Ridder (2021).

Bodźce oparte na porównaniach społecznych lub presji obserwacyjnej, choć często stosowane w celu promowania pożądanych zachowań, mogą osłabiać normy prospołeczne, zwiększać zachowania antyspołeczne lub obniżać poparcie dla polityk publicznych. W efekcie interwencji te, zamiast wzmacniać współpracę i zaufanie w społeczeństwie, mogą stopniowo podważać fundamenty spójności społecznej, ograniczając tym samym długofalową skuteczność działań publicznych (Allcott i Kessler, 2019).

4) Koszty wtórne

Dimant i Schulte (2016) pokazali, że porównania społeczne mogą prowadzić do normalizacji i racjonalizacji korupcji, gdy uczestnicy postrzegają nieetyczne zachowania jako powszechnie akceptowalne w ich otoczeniu. O kosztach wtórnych pisał też Hagmann ze współpracownikami (2018).

Nudge mogą więc również generować efekty wtórne wobec trzech wymienionych typów kosztów. Mogą one obejmować stopniową erozję norm etycznych, normalizację zachowań niepożądanych, a także osłabienie legitymizacji polityk publicznych. W efekcie nawet skuteczna interwencja może prowadzić do spadku poparcia dla działań wymagających współpracy społecznej, takich jak regulacje środowiskowe

czy podatki węglowe. Koszty wtórne pokazują więc, że pewne interwencje – mimo osiągnięcia zamierzonego efektu behawioralnego – mogą w dłuższej perspektywie podważać zaufanie instytucjonalne i wspólne wartości społeczne, ograniczając stabilność i trwałość zmian.

Dyskusja

Przegląd literatury i praktyk ewaluacyjnych wskazuje na istotną lukę: brakuje systematycznego podejścia uwzględniającego potencjalne koszty interwencji behawioralnych *nudge*. Nawet selektywna analiza przypadków pokazuje, że koszty te są realne i zróżnicowane – dotyczą emocji, procesów poznawczych, zachowań i zjawisk społecznych – co podkreśla potrzebę szerszej, bardziej krytycznej refleksji wykraczającej poza samą skuteczność, obejmującej koszty psychologiczne, jakość zmiany i wpływ na relacje między obywatelami a instytucjami.

Należy podkreślić, że przedstawione zestawienie nie zakłada hierarchizacji kosztów. Obecny stan badań nie dostarcza wystarczających podstaw do przeprowadzenia analizy porównawczej poszczególnych kosztów ani do określenia, w jakim stopniu przekładają się one na skuteczność interwencji. Brakuje systematycznych przeglądów i metaanaliz pozwalających ocenić, które koszty mają charakter incydentalny i kontekstowy, a które powodują trwałe skutki na poziomie indywidualnym lub społecznym. Tym samym proponowana typologia stanowi punkt wyjścia do dalszych badań nad relacją między skutecznością behawioralną a kosztami.

Oceniając interwencje behawioralne, kluczowe wydaje się nie tylko to, czy działają, ale dla kogo, w jakim kontekście i jakim kosztem. Wiele kosztów, które traktowane są jako efekt uboczny, wynika z tego, że interwencje są projektowane i wdrażane w oderwaniu od szerszego kontekstu życia osoby, której dotyczą – jej zasobów, ograniczeń, wcześniejszych doświadczeń czy innych interakcji z instytucjami publicznymi. Co więcej, koncentrowanie się w badaniach terenowych na obserwowalnych i behawioralnych zmiennych zależnych może sprowadzać analizy skuteczności interwencji tylko i wyłącznie do aspektów obserwowalnych zachowania, ignorując konsekwencje dla zdrowia psychicznego jednostki. Skuteczne działanie w jednej sferze może wywołać koszty w innej (zob. np. P.C. Medina, 2021); jednorazowa zmiana zachowania może nie przełożyć się na trwałe nawyki (zob. np. Gravert i Kurz, 2021), a wywołana reakcja może pozostać powierzchowna i nieoparta na rzeczywistej motywacji (zob. np. Kalil et al., 2023). Dopiero dobre zrozumienie kontekstu – społecznego, emocjonalnego i poznawczego – w którym funkcjonuje odbiorca interwencji, pozwala na sensowne wkomponowanie bodźców behawioralnych w szersze strategie

wsparcia, rozwoju kompetencji i budowania sprawczości. Innymi słowy, skuteczność interwencji wzrasta, gdy nie jest ona działaniem jedynym ani samotnym, ale gdy staje się częścią bardziej złożonego i ugruntowanego procesu wspierania obywatela w dokonywaniu realnych i trwałych zmian.

Podsumowanie

Skupienie się na realizacji politycznych celów może prowadzić do pominięcia kosztów psychologicznych i jakościowych, które z perspektywy obywateli są równie istotne. W ocenie skuteczności interwencji nie wystarczy więc deklaratywna uczciwość intencji – konieczne jest uwzględnienie rzeczywistego doświadczenia odbiorców oraz większa wrażliwość na zróżnicowane reakcje i potencjalne koszty, jakie może wywoływać nawet najlepiej zaprojektowana polityka.

Należy podkreślić także, że interwencje behawioralne typu nudge przynoszą również szereg pozytywnych efektów. W wymiarze emocjonalnym i poznawczym mogą one zmniejszać obciążenie poznawcze dzięki upraszczaniu wyborów oraz wzmacniać poczucie sprawczości poprzez jasne wskazówki, a także wywoływać pozytywne emocje związane z satysfakcją z podjętej decyzji (Benartzi et al., 2017; Thaler i Sunstein, 2008; Van Roekel et al., 2023).

Na poziomie behawioralnym badania konsekwentnie pokazują, że nudge skutecznie wspierają pożądane zmiany w obszarze zdrowia, finansów czy ekologii, prowadząc do wzmacniania korzystnych nawyków (DellaVigna i Linos, 2020; Dolan et al., 2012; Mertens et al., 2022).

W domenie społecznej i relacyjnej interwencje oparte na informacjach społecznych czy porównaniach mogą wzmacniać prospołeczne normy i sprzyjać współpracy, a także ułatwiać dostęp do rozwiązań, które zmniejszają nierówności (Allcott i Kessler, 2019; Bicchieri i Dimant, 2022; Lades i Delaney, 2022).

Wreszcie, dobrze zaprojektowane bodźce mogą wzmacniać zaufanie do instytucji oraz zwiększać legitymizację polityk publicznych, dzięki widocznym, pozytywnym efektom dla obywateli (Hagmann et al., 2018; Michaelsen et al., 2024). Pokazuje to, że choć analiza kosztów jest konieczna, równie istotne jest uwzględnianie potencjalnych korzyści, które mogą wspierać dobrostan jednostek i trwałość polityk publicznych.

Przedstawiona analiza nie prowadzi do wniosku, że interwencje typu nudge należy ograniczać lub porzucać. Ich potwierdzona zdolność do zmiany zachowania oraz relatywnie niski koszt wdrożenia czynią je istotnym narzędziem polityki publicznej. Konieczne jest jednak rozszerzenie standardów projektowania i ewaluacji stosowanych

interwencji. Odpowiedzialne stosowanie nudge wymaga uwzględniania nie tylko zmiany zachowania, lecz także potencjalnych kosztów, w perspektywie krótko- i długookresowej. Oznacza to monitorowanie efektów ubocznych, analizę zróżnicowanych reakcji odbiorców oraz uwzględnienie wskaźników dobrostanu w praktykach ewaluacyjnych.

Aby typologia miała charakter aplikacyjny, potrzebne są dalsze badania nad wagą poszczególnych kosztów i ich wpływem na jednostki oraz instytucje. Wymaga to systematycznych analiz skutków ubocznych. Dopiero na tej podstawie możliwe będzie wypracowanie pełniejszych standardów ewaluacji. W pierwszej kolejności zasadna jest identyfikacja i operacjonalizacja kosztów opisanych w literaturze oraz włączanie ich pomiaru do badań nad interwencjami.

Oświadczenie o wkładzie poszczególnych autorów

Wszyscy wymienieni autorzy wnieśli znaczący, bezpośredni i intelektualny wkład w artykuł i zatwierdzili go do publikacji.

Oświadczenie o konflikcie interesów

Autorzy deklarują, że badanie zostało przeprowadzone przy braku jakichkolwiek powiązań komercyjnych lub finansowych, które można by zinterpretować jako potencjalny konflikt interesów.

Oświadczenie o etyce badawczej

Autorzy poświadczają, że opublikowane w tekście badania zostały przeprowadzone zgodnie z etyką badawczą afiliowanej uczelni.

Bibliografia

- Allcott, H., Kessler, J.B. (2019). The Welfare Effects of Nudges: A Case Study of Energy Use Social Comparisons, *American Economic Journal: Applied Economics*, 11(1): 236–276, <https://doi.org/10.1257/app.20170328>.
- Avitzour, D., Barnea, R., Avitzour, E., Cohen, H., Nissan-Rozen, I. (2018). Nudging in the Clinic: The Ethical Implications of Differences in Doctors' and Patients' Point of View, *Journal of Medical Ethics*, 45: 181–189.
- Banerjee, S., John, P. (2024). Nudge Plus: Incorporating Reflection into Behavioral Public Policy, *Behavioural Public Policy*, 8(1): 69–84, <https://doi.org/10.1017/bpp.2021.6>.

- Barev, T., Schwede, M., Janson, A. (2021). *The Dark Side of Privacy Nudging – An Experimental Study in the Context of a Digital Work Environment*, Hawaii International Conference on System Sciences, <https://doi.org/10.24251/HICSS.2021.500>.
- Benartzi, S., Beshears, J., Milkman, K.L., Sunstein, C.R., Thaler, R.H., Shankar, M., Tucker-Ray, W., Congdon, W.J., Galing, S. (2017). Should Governments Invest More in Nudging?, *Psychological Science*, 28(8): 1041–1055, <https://doi.org/10.1177/0956797617702501>.
- Beshears, J., Choi, J.J., Laibson, D., Madrian, B.C., Milkman, K.L. (2015). The Effect of Providing Peer Information on Retirement Savings Decisions, *The Journal of Finance*, 70: 1161–1201.
- Bicchieri, C., Dimant, E. (2022). Nudging with Care: The Risks and Benefits of Social Information, *Public Choice*, 191(3): 443–464, <https://doi.org/10.1007/s11127-019-00684-6>.
- Bolton, G.E., Dimant, E., Schmidt, U. (2020). *When a Nudge Backfires: Combining (Im) Plausible Deniability with Social and Economic Incentives to Promote Behavioral Change*, Working Paper Nr 8070. Munich: Center for Economic Studies and ifo Institute (CESifo).
- Cambridge Dictionary. (b.d.). <https://dictionary.cambridge.org/pl/dictionary/english-polish/nudge> (dostęp: 14.01.2026).
- Campigotto, N., Gioia, C., Ploner, M. (2025). Salience and Information Avoidance in Voluntary Carbon Offsetting Decisions: Evidence from Online Experiments, *Ecological Economics*, 233: 108577, <https://doi.org/10.1016/j.ecolecon.2025.108577>.
- Chesterley, N. (2015). Nudges and Welfare: Optimal Behavioural Policy with Decision Costs, *SSRN Electronic Journal*, <https://doi.org/10.2139/ssrn.2567190>.
- Clot, S., Della Giusta, M., Jewell, S. (2022). Once Good, Always Good? Testing Nudge's Spillovers on Pro Environmental Behavior, *Environment and Behavior*, 54(3): 655–669, <https://doi.org/10.1177/00139165211060524>.
- Damgaard, M.T., Gravert, C. (2018). The Hidden Costs of Nudging: Experimental Evidence from Reminders in Fundraising, *Journal of Public Economics*, 157: 15–26, <https://doi.org/10.1016/j.jpubeco.2017.11.005>.
- De Ridder, D., Feitsma, J., Van Den Hoven, M., Kroese, F., Schillemans, T., Verweij, M., Venema, T., Vugts, A., De Vet, E. (2024). Simple Nudges That Are Not So Easy, *Behavioural Public Policy*, 8(1): 154–172, <https://doi.org/10.1017/bpp.2020.36>.
- DellaVigna, S., Linos, E. (2020). RCTs to Scale: Comprehensive Evidence From Two Nudge Units, *Econometrica*, 90(1): 81–116, <https://doi.org/10.3982/ECTA18709>.
- Di Matola, S.M. (2020). *A Nudge Too Much: Nudging Healthy Food Choices in Restaurants through its Menu and Eating Environment*. Twente: University of Twente. <https://essay.utwente.nl/81303/> (dostęp: 14.01.2026).
- Dimant, E., Schulte, T. (2016). The Nature of Corruption: An Interdisciplinary Perspective, *German Law Journal*, 17(1): 53–72, <https://doi.org/10.1017/S2071832200019684>.
- Dolan, P., Hallsworth, M., Halpern, D., King, D., Metcalfe, R., Vlaev, I. (2012). Influencing Behaviour: The Mindspace Way, *Journal of Economic Psychology*, 33(1): 264–277, <https://doi.org/10.1016/j.joep.2011.10.009>.
- Dorison, C.A., Lerner, J.S., Heller, B.H., Rothman, A.J., Kawachi, I.I., Wang, K., Rees, V.W., Gill, B.P., Gibbs, N., Ebersole, C.R., Vally, Z., Tajchman, Z., Zsido, A.N., Zrimsek, M.,

- Chen, Z., Ziano, I., Gialitaki, Z., Ceary, C.D., Lin, Y., ..., Coles, N.A. (2022). In COVID-19 Health Messaging, Loss Framing Increases Anxiety with Little-to-No Concomitant Benefits: Experimental Evidence from 84 Countries, *Affective Science*, 3(3): 577–602, <https://doi.org/10.1007/s42761-022-00128-3>.
- Ewert, B. (2019). Moving Beyond the Obsession with Nudging Individual Behaviour: Towards a Broader Understanding of Behavioural Public Policy, *Public Policy and Administration*, 35(3): 337–360, <https://doi.org/10.1177/0952076719889090>.
- Gravert, C., Kurz, V. (2021). Nudging à La Carte: A Field Experiment on Climate-Friendly Food Choice, *Behavioural Public Policy*, 5(3): 378–395, <https://doi.org/10.1017/bpp.2019.11>.
- Hagmann, D., Ho, E.H., Loewenstein, G. (2018). Nudging out Support for a Carbon Tax, *Nature Climate Change*, 9: 484–489.
- Harrison, G., Morsink, K., Schneider, M. (2024). *Do No Harm? The Welfare Consequences of Behavioural Interventions*, <https://cear.gsu.edu/files/2024/05/CEAR-WP-2020-12-Do-No-Harm-The-Welfare-Consequences-of-Behavioural-Interventions-MAY-2024.pdf> (dostęp: 14.01.2026).
- Hertwig, R., Grüne-Yanoff, T. (2017). Nudging and Boosting: Steering or Empowering Good Decisions, *Perspectives on Psychological Science*, 12(6): 973–986, <https://doi.org/10.1177/1745691617702496>.
- Hummel, D., Maedche, A. (2019). How Effective Is Nudging? A Quantitative Review on the Effect Sizes and Limits of Empirical Nudging Studies, *Journal of Behavioral and Experimental Economics*, 80: 47–58, <https://doi.org/10.1016/j.socec.2019.03.005>.
- Jedynak, T. (2019). Pracownicze plany kapitałowe jako instrument mobilizacji dodatkowych oszczędności emerytalnych w Polsce, *Krakow Review of Economics and Management/ Zeszyty Naukowe Uniwersytetu Ekonomicznego w Krakowie*, 2(980): 113–130, <https://doi.org/10.15678/ZNUEK.2019.0980.0207>.
- Jensen, C.L., Andersen, L.M., Hansen, L.G., Henningsen, G. (2020). Is Social Nudging Too Emotionally Taxing? A Field Experiment of Public Utilities and Electricity Consumers in Denmark, *Energy Research & Social Science*, 67: 101515, <https://doi.org/10.1016/j.erss.2020.101515>.
- Kalil, A., Liu, H., Mayer, S., Rury, D., Shah, R. (2023). Nudging or Nagging? Conflicting Effects of Behavioral Tools, *SSRN Scholarly Paper*, 4318177. Chicago: University of Chicago, Becker Friedman Institute for Economics.
- Krawiec, J.M., Piaskowska, O.M., Piesiewicz, P.F., Białaszek, W. (2021). Tools for Public Health Policy: Nudges and Boosts as Active Support of the Law in Special Situations Such as the COVID-19 Pandemic, *Globalization and Health*, 17(1): 132, <https://doi.org/10.1186/s12992-021-00782-5>.
- Kuyer, P., Gordijn, B. (2023). Nudge in Perspective: A Systematic Literature Review on the Ethical Issues with Nudging, *Rationality and Society*, 35(2): 191–230, <https://doi.org/10.1177/10434631231155005>.
- Lades, L.K., Delaney, L. (2022). Nudge FORGOOD, *Behavioural Public Policy*, 6(1): 75–94, <https://doi.org/10.1017/bpp.2019.53>.
- Loewenstein, G., Chater, N. (2017). Putting Nudges in Perspective, *Behavioural Public Policy*, 1(1): 26–53, <https://doi.org/10.1017/bpp.2016.7>.

- Maier, M., Bartoš, F., Stanley, T.D., Shanks, D.R., Harris, A.J.L., Wagenmakers, E.-J. (2022). No Evidence for Nudging after Adjusting for Publication Bias, *Proceedings of the National Academy of Sciences*, 119(31): e2200300119, <https://doi.org/10.1073/pnas.2200300119>.
- Marchiori, D.R., Adriaanse, M.A., De Ridder, D.T.D. (2017). Unresolved Questions in Nudging Research: Putting the Psychology Back in Nudging, *Social and Personality Psychology Compass*, 11(1): e12297, <https://doi.org/10.1111/spc3.12297>.
- Mavrot, C., Potluka, O., Balzer, L., Eicher, V., Haunberger, S., Heuer, C., Viallon, F.-X. (2025). What Evaluation Criteria Are Used in Policy Evaluation Research: A Cross-Field Literature Review, *Evaluation and Program Planning*, 108: 102512, <https://doi.org/10.1016/j.evalprogplan.2024.102512>.
- Medina, J. de Q. (2021). What is Wrong with Nudges? Addressing Normative Objections to the Aims and the Means of Nudges, *Gestión y Análisis de Políticas Públicas*, 25: 23–37.
- Medina, P.C. (2021). Side Effects of Nudging: Evidence from a Randomized Intervention in the Credit Card Market, *The Review of Financial Studies*, 34(5): 2580–2607, <https://doi.org/10.1093/rfs/hhaa108>.
- Mertens, S., Herberz, M., Hahnel, U.J.J., Brosch, T. (2022). The Effectiveness of Nudging: A Meta-Analysis of Choice Architecture Interventions Across Behavioral Domains, *Proceedings of the National Academy of Sciences*, 119(1): e2107346118, <https://doi.org/10.1073/pnas.2107346118>.
- Meunier, L., Bashirzadeh, Y., Ohadi, S. (2024). Framing the Default Option Right, *Journal of Behavioral Decision Making*, 37(3): e2395, <https://doi.org/10.1002/bdm.2395>.
- Michaelsen, P., Johansson, L.-O., Hedesström, M. (2024). Experiencing Default Nudges: Autonomy, Manipulation, and Choice-Satisfaction as Judged by People Themselves, *Behavioural Public Policy*, 8(1): 85–106, <https://doi.org/10.1017/bpp.2021.5>.
- Mills, S., Whittle, R. (2024). Seeing the Nudge from the Trees: The 4S Framework for Evaluating Nudges, *Public Administration*, 102(2): 580–600, <https://doi.org/10.1111/padm.12941>.
- Nafziger, J. (2020). Spillover Effects of Nudges, *Economics Letters*, 190: 109086, <https://doi.org/10.1016/j.econlet.2020.109086>.
- Obama, B. (2015). *Executive Order – Using Behavioral Science Insights to Better Serve the American People*, <https://obamawhitehouse.archives.gov/the-press-office/2015/09/15/executive-order-using-behavioral-science-insights-better-serve-American> (dostęp: 14.01.2026).
- Richter, I., Thøgersen, J., Klöckner, C.A. (2018). A Social Norms Intervention Going Wrong: Boomerang Effects from Descriptive Norms Information, *Sustainability*, 10(8): 8, <https://doi.org/10.3390/su10082848>.
- Rusmevichientong, P., Ebrahim, R., Nila, J., Cheng, I., Weiss, J. (2021). The New FDA Nutrition Facts Labels and Consumer Purchase Intentions: An Effect of Emotional Tax, *Journal of Consumer Behaviour*, 20(6): 1422–1435, <https://doi.org/10.1002/cb.1944>.
- Schmidt, A.T., Engelen, B. (2020). The Ethics of Nudging: An Overview, *Philosophy Compass*, 15(4): e12658, <https://doi.org/10.1111/phc3.12658>.
- Shah, J., Shah, R., Liang, P. (2021). Too Much Nudging: Can it Cause a Decrease in the Desired Response?, *Journal of Student Research*, 10(4), <https://doi.org/10.47611/jsrhs.v10i4.2477>.

- Sunstein, C.R., Reisch, L.A., Rauber, J. (2018). A Worldwide Consensus on Nudging? Not Quite, but Almost, *Regulation & Governance*, 12(1): 3–22, <https://doi.org/10.1111/rego.12161>.
- Szkudlarek, P. (2022). *Regulacja z perspektywy ekonomii instytucjonalnej i ekonomii behawioralnej na przykładzie rynku usług telekomunikacyjnych w Polsce*. Szczecin: Wydawnictwo Naukowe Uniwersytetu Szczecińskiego.
- Thaler, R.H., Benartzi, S. (2004). Save More Tomorrow™: Using Behavioral Economics to Increase Employee Saving, *Journal of Political Economy*, 112 (S1): S164–S187, <https://doi.org/10.1086/380085>.
- Thaler, R.H., Sunstein, C.R. (2008). *Nudge: Improving Decisions about Health, Wealth, and Happiness*. New Haven: Yale University Press.
- Thunström, L. (2019). Welfare Effects of Nudges: The Emotional Tax of Calorie Menu Labeling, *Judgment and Decision Making*, 14(1): 11–25, <https://doi.org/10.1017/S1930297500002874>.
- Thunström, L. (2018). Incidence of an Emotional Tax: The Case of Calorie Menu Labeling, *AEA Randomized Controlled Trials*, <https://doi.org/10.2139/ssrn.3272036>.
- Tor, A. (2020). Nudges that Should Fail?, *Behavioural Public Policy*, 4(3): 316–342, <https://doi.org/10.1017/bpp.2019.5>.
- Tummers, L. (2019). Public Policy and Behavior Change, *Public Administration Review*, 79(6): 925–930, <https://doi.org/10.1111/puar.13109>.
- Heijden, J. van der, Kosters, M. (2015). From Mechanism to Virtue: Evaluating Nudge-Theory, *Evaluation*, 21, <https://doi.org/10.1177/1356389015590218>.
- Van Roekel, H., Giurge, L.M., Schott, C., Tummers, L. (2023). Nudges Can Be Both Autonomy-Preserving and Effective: Evidence from a Survey and Quasi-Field Experiment, *Behavioural Public Policy*, 1–24, <https://doi.org/10.1017/bpp.2023.18>.
- Vitek, V., Syed Shah, T. (2019). Implementing a Nudge to Prevent Email Phishing, <https://urn.kb.se/resolve?urn=urn:nbn:se:kth:diva-259403> (dostęp: 14.01.2026).
- Wachner, J., Adriaanse, M.A., De Ridder, D.T.D. (2020). And How Would That Make You Feel? How People Expect Nudges to Influence Their Sense of Autonomy, *Frontiers in Psychology*, 11: 607894, <https://doi.org/10.3389/fpsyg.2020.607894>.
- Wachner, J., Adriaanse, M.A., De Ridder, D.T.D. (2021a). The Effect of Nudges on Autonomy in Hypothetical and Real Life Settings, *Plos One*, 16(8): e0256124, <https://doi.org/10.1371/journal.pone.0256124>.
- Wachner, J., Adriaanse, M., De Ridder, D. (2021b). The Influence of Nudge Transparency on the Experience of Autonomy, *Comprehensive Results in Social Psychology*, 5 (1–3): 49–63, <https://doi.org/10.1080/23743603.2020.1808782>.
- Wilson, A.L., Bogomolova, S., Buckley, J.D. (2015). Lack of Efficacy of a Salience Nudge for Substituting Selection of Lower-Calorie for Higher-Calorie Milk in the Work Place, *Nutrients*, 7(6): 4336–4344, <https://doi.org/10.3390/nu7064336>.