

UNIWERSYTET SWPS

WYDZIAŁ PSYCHOLOGII W WARSZAWIE

**Identyfikacja czynników decydujących o skuteczności interwencji
behawioralnych w decyzjach ryzykownych**

Jakub M. Krawiec

Praca doktorska napisana pod kierunkiem

dra hab. Wojciecha Białaszka, prof. Uniwersytetu SWPS

Warszawa, 2024

Dziękuję,

Rodzicom, za inspirację do pracy naukowej,

Mamie, za pomoc w redakcji niniejszej pracy doktorskiej,

Julii, za codzienne, ogromne wsparcie i dobre słowo,

Wojtkowi, za poświęcony czas, serdeczność i dobrą radę.

SPIS TREŚCI

WPROWADZENIE	6
1. CZĘŚĆ TEORETYCZNA	9
1.1 Wstęp teoretyczny	9
1.1.1 Ryzyko i niepewność w ujęciu teorii ograniczonej racjonalności	9
1.1.2 Dwie szkoły badania heurystyk i błędów poznawczych.....	10
1.1.3 Dwie kategorie interwencji behawioralnych: Nudging i Boosting.....	14
1.1.4 Program Save More Tomorrow – więcej niż pojedyncza interwencja behawioralna ...	23
1.1.5 Rola różnic indywidualnych i świadomości oddziaływań jako moderatorów efektywności interwencji behawioralnych.....	29
1.1.6 Behawioralne Laboratoria Polityk Publicznych	33
1.2 Podsumowanie części teoretycznej	37
1.2.1 Wyzwania i ograniczenia badań nad interwencjami behawioralnymi.....	37
1.2.2 Koncepcja badań własnych.....	40
2. CZĘŚĆ EMPIRYCZNA	44
2.1 BADANIE 1 <i>Metaanaliza badań porównujących efektywność interwencji nudge i boost w zmianie zachowania</i>	44
2.1.1 Metoda	45
2.1.2 Wyniki.....	46
2.1.3 Podsumowanie	50
2.1.4 Dyskusja.....	51
2.2 Badania 2A i 2B – zarys ogólny	54
2.2.1 Metoda ogólna badań 2A i 2B	54
2.3 Badanie 2A <i>Transparentność a efektywność interwencji nudge i boost w zmianie zachowania w decyzjach ryzykownych</i>	59
2.3.1 Osoby badane.....	59
2.3.2 Procedura	60
2.3.4 Wyniki.....	61
2.3.4 Podsumowanie	75
2.4 Badanie 2B <i>Trwałość efektów interwencji behawioralnych nudge i boost w decyzjach ryzykownych</i>	78
2.4.1 Osoby badane.....	78
2.4.2 Procedura	79

2.4.3 Wyniki.....	80
2.4.4 Podsumowanie	89
2.5 Dyskusja Badań 2A i 2B	91
2.6. Badania 3A i 3B – zarys ogólny.....	96
2.6.1 Metoda ogólna badań 3A i 3B	96
2.7 Badanie 3A <i>Efektywność interwencji nudge i boost w zmianie zachowań oszczędnościowych</i>	101
2.7.1 Osoby badane.....	101
2.7.2 Narzędzia/Procedura	101
2.7.3 Wyniki.....	106
2.7.4 Podsumowanie	121
2.8 Badanie 3B <i>Efekt synergii interwencji nudge i boost w zmianie zachowań oszczędnościowych</i>	123
2.8.1 Osoby badane.....	123
2.8.2 Procedura	124
2.8.3 Wyniki.....	125
2.8.4 Podsumowanie	136
2.9 Dyskusja Badań 3A i 3B	138
2.10 Dyskusja ogólna	141
3. PODSUMOWANIE	157
4. BIBLIOGRAFIA.....	161
5. STRESZCZENIE W JĘZYKU POLSKIM.....	177
6. STRESZCZENIE W JĘZYKU ANGIELSKIM	178
7. ZAŁĄCZNIKI.....	179
7.1 Procedura FFR.....	179
7.2 Procedura PDQ	181
7.3 Procedura paradygmatu wyborów dyskretnych (ang. <i>Discrete Choice Experiment</i> ; DCE)	183

WPROWADZENIE

Prezentowana praca dotyczy zagadnienia skuteczności interwencji behawioralnych pod kątem zmiany zachowania w decyzjach ryzykownych. Projekt badawczy koncentruje się na dwóch kategoriach interwencji behawioralnych: nudgingu i boostingu. Są one coraz częściej stosowane jako narzędzia polityki społecznej w Polsce, jak i za granicą. Celem pracy jest identyfikacja czynników sytuacyjnych i cech indywidualnych, decydujących o skuteczności interwencji nudge i boost.

Tym, co różni niniejszy projekt badawczy od dotychczasowych badań, jest ocena kilku rodzajów interwencji nudge i boost w domenie decyzji finansowych oraz ich komplementarne ujęcie, dążące do zbadania potencjalnych efektów synergii. W pracy podjęto badania nad interwencjami boost soczewkowy i boost wg Reguły Franklina, które do tej pory nie były testowane w domenie decyzji finansowych. Dodatkowo przeprowadzona została ewaluacja mniej popularnych interwencji typu nudge, uwzględniając cechy psychologiczne oraz czynniki sytuacyjne, takie jak transparentność interwencji i utrzymywanie się efektów interwencji w czasie. Nowatorskim aspektem pracy jest wykorzystanie paradygmatu wyboru dyskretnego do oceny dwóch interwencji nudge i boost w zmianie zachowań oszczędnościowych. Wybrane do tego projektu rodzaje interwencji behawioralnych nie były do tej pory badane z wykorzystaniem podejścia decyzji wieloatrybutowych, szczególnie w domenie sytuacji finansowych.

Rozprawa doktorska składa się z dwóch części: teoretycznej i empirycznej. Pierwsze rozdziały prezentują podstawy teoretyczne interwencji behawioralnych nudge i boost oraz systematyczny przegląd narzędzi do nich zaliczanych. Następnie zostały przedstawione przykłady implementacji interwencji nudge i boost w politykach publicznych i biznesie. Część teoretyczna zawiera również uzupełnienie dotychczas mało eksplorowanych zagadnień w

dziedzinie interwencji behawioralnych oraz opisuje wyzwania stojące przed proponentami nudge'y i boostów. Część teoretyczna kończy się umotywowaniem i przedstawieniem koncepcji badań własnych.

Celem prezentowanych badań empirycznych jest empiryczna weryfikacja zagadnień związanych z identyfikacją czynników skuteczności interwencji behawioralnych w zmianie zachowania. Część empiryczna jest podzielona na trzy bloki tematyczne: metaanalizę dotychczasowych badań skupiających się na porównaniu interwencji typu nudge i boost (Badanie 1A) oraz dwa bloki tematyczne zawierające po dwa badania empiryczne. W ramach Badań 2A i 2B testowano interwencje behawioralne w warunkach ryzyka oraz uwzględniono czynniki sytuacyjne transparentności i trwałości efektów interwencji w czasie. Z kolei poprzez Badania 3A i 3B testowano efektywność interwencji behawioralnych w zmianie zachowania w warunkach niepewności, na przykładzie zachowań oszczędnościowych. Ważnym elementem badania 3B był również potencjalny efekt synergii pomiędzy interwencjami nudge i boost.

Badania z zakresu interwencji behawioralnych nudge i boost można umiejscowić przede wszystkim w obszarze zainteresowań ekonomii behawioralnej, psychologii decyzji i psychologii ekonomicznej. Dwie kategorie interwencji wywodzą się z dwóch różnych podejść badawczych do ludzkiego zachowania. Podejście nudge ma swoje początki w szkole „Heurystyk i Błędów poznawczych” (Kahneman i Tversky, 1974). Program badawczy Kahnemana i Tversky'ego postrzega ludzi jako często stosujących uproszczone reguły decyzyjne, które, z punktu widzenia normatywnych teorii ekonomicznych, mogą prowadzić do nieoptymalnych i nieracjonalnych decyzji. Z kolei boosting, wywodzi się ze szkoły „Prostych Heurystyk” (Gigerenzer i Todd, 1999). To podejście postrzega heurystyki w bardziej pozytywnym świetle, jako efektywne i adaptacyjne narzędzia, które pozwalają podejmować dobre decyzje w warunkach niepewności i

ograniczonej informacji. Choć obydwie szkoły wyrastają z koncepcji „ograniczonej racjonalności” Herberta Simona, prezentują odmiennie spojrzenie na znaczenie heurystyk i błędów poznawczych dla oceny podejmowanych decyzji. Należy zaznaczyć, że rosnąca popularność nudgingu i boostingu, jest skorelowana z sukcesami polityk publicznych je wykorzystujących. Prace takich zespołów jak Behavioural Insights Team, spowodowały szerokie zastosowanie interwencji behawioralnych w rozwiązaniach polityk społecznych na całym świecie, również w Polsce. Na rynku polskim, najbardziej zauważalny przykład interwencji typu nudge został zaimplementowany w ramach programu oszczędnościowego Polskich Planów Kapitałowych. Tym samym, niniejsza praca dostarcza informacji ważnych nie tylko z punktu widzenia dziedziny psychologii i ekonomii behawioralnej, ale również dyscyplin administracji publicznej czy polityki publicznej.

1. CZĘŚĆ TEORETYCZNA

1.1 Wstęp teoretyczny

1.1.1 Ryzyko i niepewność w ujęciu teorii ograniczonej racjonalności

W 1954 roku, opisując procesy ekonomiczne i decyzyjne, Leonard Savage wprowadził rozróżnienie na „małe” i „duże” światy. Termin „małe światy” odnosi się do scenariuszy, w których decydenci mają pełną wiedzę na temat dostępnych opcji, ich konsekwencji oraz powiązanych prawdopodobieństw lub „ryzyk” (Savage, 1954). Z kolei „duże światy” reprezentują sytuacje, gdzie istotna część informacji jest nieznana lub musi być ekstrapolowana z ograniczonych danych. W takich warunkach, decyzje są podejmowane w stanie „niepewności”, co Savage częściowo opierał na wcześniejszych rozważaniach Franka Knighta, który jako pierwszy dokonał rozróżnienia między ryzykiem a niepewnością (Knight, 1921). Według Savage’a i Knighta, większość naszych codziennych decyzji jest podejmowana w stanie niepewności, podczas gdy sytuacje, w których ryzyko jest w pełni znane, są stosunkowo rzadkie i ograniczają się głównie do hazardu i rzadkich decyzji finansowych (Kozyreva i in., 2019). Takie podejście do ryzyka, stojące w opozycji do klasycznego podejścia w dziedzinie ekonomii, zainspirowało Herberta Simona do prac nad teorią „ograniczonej racjonalności” (ang. *bounded rationality*; Simon, 1972) zakładającą, że proces decyzyjny ludzi jest ograniczony przez dostępną wiedzę, dostępne informacje, ograniczenia czasowe oraz zdolność umysłu do przetwarzania tych informacji, takich jak choćby możliwość sprawnego obliczania ryzyka, czy też prawdopodobieństwa. Simon argumentował przeciwko klasycznemu modelowi racjonalności, który zakłada nieograniczony dostęp do informacji i zdolności obliczeniowych umożliwiających maksymalizację użyteczności, sugerując, że w rzeczywistości ludzkie zachowanie jest ograniczone przez wiele czynników (Simon, 1972; zobacz także: Hertwig i in., 2019).

Wskazywał, że głównym problemem dominujących, normatywnych modeli wyboru, takich jak teoria oczekiwanej użyteczności Bernoulliego (Simon, 1972), jest ich nieadekwatność do opisywania zachowań w realnych środowiskach. Teorie te, określane przez niego mianem „modeli olimpijskich” (ang. *Olympian models*), opierały się na idealistycznym założeniu, że decydenci są wszechwiedzący i dysponują nieograniczonymi zasobami poznawczymi (Simon, 1990). Simon argumentował, że skuteczniejsze narzędzia racjonalności ludzkiej muszą znacząco różnić się od tych oferowanych przez istniejące podejścia normatywne. W sytuacji, gdy koszty obliczeniowe znalezienia optymalnego rozwiązania są zbyt wysokie lub gdy optymalne rozwiązanie nie jest dostępne, decydent może zamiast tego poszukiwać rozwiązania przybliżonego i „wystarczająco dobrego”, związanego z regułą „poszukiwania wystarczającej opcji” (ang. *satisficing*), a nie optymalnej (Simon, 1997). Poglądy Simona na temat rozwiązań zadowolających i ograniczonej racjonalności zainicjowały badania psychologiczne nad racjonalnością ograniczoną, heurystykami i błędami.

1.1.2 Dwie szkoły badania heurystyk i błędów poznawczych

Dwoma programami badawczymi, które bezpośrednio wywodzą się z prac Simona i szczególnie wpłynęły na rozwój psychologii decyzji są program „Heurystyk i błędów poznawczych” (ang. *Heuristics i biases*) spopularyzowany przez Daniela Kahnemana i Amosa Tversky’ego (Kahneman, 2011; Tversky i Kahneman, 1974) oraz program „Prostych Heurystyk” (ang. *Simple Heuristics*) rozpoczęty przez Gerda Gigerenzera i zespół ABC (Gigerenzer i Todd, 1999b), rozwijanego w dalszym ciągu przez jego ucznia Ralpha Hertwiga (Hertwig i in., 2019). Choć szkoły badawcze Kahnemana i Tversky’ego oraz Gigerenzera, mają wspólne podwaliny, autorzy postrzegają ludzkie zachowanie w nieco inny sposób (Hertwig i Grüne-Yanoff, 2017). Jak wspomina Kahneman, „jego badania miały na celu utworzenie mapy ograniczonej

racjonalności poprzez eksplorowanie systematycznych błędów poznawczych, które oddzielają przekonania ludzi i ich wybory od optymalnych przekonań i wyborów zakładanych przez modele ekonomicznej racjonalności” (Kahneman, 2003). Inspirując się pracami Simona, Kahnemana i Tversky oraz ich następcy od blisko pięćdziesięciu lat tworzą katalog systematycznych odstępstw od norm klasycznie i ekonomicznie rozumianej racjonalności, badając ludzki osąd i wybory w odniesieniu do reguł prawdopodobieństwa, statystyki i aksjomów racjonalnego wyboru. W tym nurcie człowiek jest przedstawiany jako nieefektywny podmiot decyzyjny, który ze względu na ograniczenia poznawcze nie jest do końca w stanie podejmować dobrych dla siebie decyzji w ujęciu racjonalnych modeli wyboru (Thaler i Sunstein, 2008; Hertwig i Grüne-Yanoff, 2017). Dwie najpopularniejsze heurystyki analizowane w badaniach Kahnemana i Tversky’ego, które wskazują na odstępstwa od ekonomicznie racjonalnego wyboru to heurystyka dostępności i heurystyka reprezentatywności. Pierwsza z nich polega na ocenie prawdopodobieństwa zdarzeń na podstawie łatwości, z jaką ich przykłady przychodzą decydentom do głowy. Jeśli coś łatwo przychodzi nam na myśl, możemy uznać, że jest to bardziej prawdopodobne. Na przykład, po obejrzeniu wiadomości o wypadkach lotniczych, ludzie mogą przecenić ryzyko katastrof lotniczych, ponieważ tragiczne zdarzenia są łatwo dostępne w ich pamięci (Kahneman, 2011). Z kolei heurystyka reprezentatywności jest tendencją do oceny prawdopodobieństwa zdarzeń na podstawie tego, jak bardzo są one reprezentatywne lub typowe dla kategorii, do której je przypisujemy, ignorując przy tym istotne prawdopodobieństwa bazowe. Na przykład, ludzie mogą uważać, że ktoś, kto jest nieśmiały i lubi czytać, jest bardziej prawdopodobnie bibliotekarzem niż sprzedawcą, mimo że liczba sprzedawców w populacji jest znacznie większa niż liczba bibliotekarzy (Kahneman, 2011; Tversky i Kahneman, 1974). Dla niniejszych rozważań ważne jest by zwrócić uwagę na efekt wspomnianych heurystyk. Opisywane przez Kahnemana i Tversky’ego heurystyki i błędy powodują, że wnioski, które decydenci wyciągają

nie są w pełni prawdziwe czy racjonalne, ponieważ ignorują część informacji lub są stronnicze w ich ocenie. W rezultacie, heurystyki są oceniane raczej jako coś niepożądanego czy suboptymalnego w dążeniach do decyzji optymalnych i obiektywnie racjonalnych (Grüne-Yanoff i Hertwig, 2016; Kahneman, 2011).

Istotna z punktu widzenia interwencji behawioralnych¹ i związana z programem badawczym „Heurystyk i błędów poznawczych” jest również spopularyzowana przez Kahnemana koncepcja dotycząca Systemu 1 i Systemu 2, które reprezentują dwa różne tryby ludzkiego myślenia (Kahneman, 2011). System 1 charakteryzuje się szybkim, intuicyjnym działaniem, opartym na automatycznych reakcjach i heurystykach, działając bez świadomego wysiłku i często nieświadomie. Jest używany w sytuacjach, które wymagają szybkich decyzji lub rozwiązują prostsze problemy. Z kolei System 2 to wolniejszy, bardziej racjonalny sposób myślenia, angażujący świadomy wysiłek, logiczne rozumowanie i analizę danych. Jest stosowany w sytuacjach, które wymagają głębszej analizy, rozwiązania skomplikowanych problemów lub podejmowania decyzji wymagających uwagi i koncentracji. Według założeń Kahnemana, koncepcja ta pomaga zrozumieć, jak różne tryby myślenia wpływają na proces podejmowania decyzji oraz jak świadome korzystanie z obu systemów może poprawić jakość naszych wyborów (Kahneman, 2003). Choć Kahneman nie był pierwszym autorem tej koncepcji, często zwanej „dualnym” systemem przetwarzania informacji, a wcześniej opisywała to choćby Chaiken (1980) czy Epstein (1994), to zdecydowanie Kahneman spopularyzował to podejście dotychczas najszerszej (Weber i Johnson, 2009; Lejarraga i Hertwig, 2021).

¹ W niniejszej pracy „podejście behawioralne” odnosi się szeroko do nauk o zachowaniu (ang. *behavioral science*; APA Dictionary of Psychology, 2024), których przedmiotem badania jest ludzkie zachowanie. Zgodnie z tym podejściem, interwencje behawioralne są definiowane jako szeroko pojęte narzędzia zmieniające ludzkie zachowanie (Hertwig i Grüne-Yanoff, 2017; Sunstein i Reisch, 2023) i niekoniecznie muszą odnosić się do tradycji psychologii behawioralnej i analizy zachowania (Tagliabue i in., 2019).

Mimo, że program Kahnemana i Tversky'ego wydaje się być dominującym podejściem do badania ograniczonej racjonalności, to warto podkreślić, że nie jest jedynym (Lejarraga i Hertwig, 2021). Program „Prostych Heurystyk” Gerda Gigerenzera postrzega ograniczoną racjonalność jako model deskryptywny, a nie normatywny ludzkich zachowań i dodatkowo kładzie nacisk na koncepcję „racjonalności ekologicznej” czy też „racjonalności środowiskowej” (ang. *ecological rationality*), postulującej, że rozważania na temat racjonalności danych wyborów muszą być podejmowane z uwzględnieniem środowiska/otoczenia decyzyjnego, w którym dany decydent funkcjonuje, jego przekonań i priorytetów. Dodatkowo, Gigerenzer wskazuje, że kluczowe jest zrozumienie, jakie reguły wnioskowania są najbardziej adekwatne w danej sytuacji decyzyjnej, sugerując, aby metody, teorie i badania ukierunkowane na ich tworzenie skupiały się na konkretnych dziedzinach, np. finansów, zdrowia czy ekologii (Kozyreva i in., 2019). Te metody to właśnie „Proste Heurystyki”, czyli proste reguły postępowania, które można wykorzystać, by ustrukturyzować informacje w danej sytuacji decyzyjnej i podjąć najbardziej trafny i szybki wybór zgodny ze swoimi preferencjami krótko lub długoterminowymi. Ze względu na inną naturę heurystyk prostych niż tych wyróżnianych przez Kahnemana i Tversky'ego, warto podać dwa przykłady najczęściej opisywanych prostych heurystyk w celu ich dokładniejszego omówienia: regułę „Wybierz najlepszą” (ang. *Take-the-Best*) oraz heurystykę rozpoznania (ang. *recognition heuristic*; Gigerenzer i Todd, 1999).

Heurystyka „Wybierz najlepszą” odnosi się do procesu, w którym podczas podejmowania decyzji, zamiast rozważać i porównywać wszystkie dostępne informacje dotyczące alternatyw, uwzględniane jest tylko najważniejsze kryterium kontekstu (Gigerenzer i Goldstein, 1999). Na przykład, decydując, który z dwóch programów oszczędnościowych wybrać, można skoncentrować się wyłącznie na takim czynniku, jak ekspercka ocena tego programu, zamiast na

wielu innych potencjalnie ważnych, ale mniej istotnych danych. Jeśli jeden program wyróżnia się pod tym względem, wybieramy go, nie zagłębiając się dalej w porównania (Gigerenzer i Goldstein, 1999). Druga heurystyka, heurystyka rozpoznania, opiera się na założeniu, że gdy z dwóch opcji rozpoznajemy tylko jedną, automatycznie przyznajemy jej wyższą wartość w procesie oceny. Przykładowo, kiedy ludzie stają przed wyborem między dwoma miastami, jako potencjalnym miejscem do życia i znamy nazwę tylko jednego z nich, heurystyka rozpoznania skłania nas do wyboru tego rozpoznanego miasta, zakładając jego wyższość. Ta heurystyka jest wyjątkowo przydatna w sytuacjach, gdzie nasza wiedza na temat dostępnych opcji jest ograniczona (Goldstein i Gigerenzer, 1999). Proste heurystyki, mimo swoich pozornych ograniczeń, okazały się bardzo skuteczne w wielu rzeczywistych sytuacjach decyzyjnych. Gigerenzer argumentuje, że w niektórych przypadkach, proste heurystyki mogą być tak samo dobre, a czasami nawet lepsze niż skomplikowane metody analizy decyzyjnej, zwłaszcza gdy bierze się pod uwagę niepewność i ograniczoną ilość informacji, z którą musi się zmierzyć decydent (Gigerenzer i in., 2022).

1.1.3 Dwie kategorie interwencji behawioralnych: Nudging i Boosting

W oparciu o dwie przedstawione szkoły badawcze, w ciągu ostatnich kilkunastu lat wyrosły dwa podejścia o charakterze aplikacyjno-wdrożeniowym, starające się w praktyce wykorzystać wiedzę o heurystykach i błędach poznawczych do zmiany zachowania na poziomie polityk publicznych, zachowań konsumenckich i obywatelskich. Pierwszy z nich to nudging, wywodzący się ze szkoły Kahnemana i Tversky'ego, za sprawą książki opublikowanej przez noblistę Richarda Thaler'a oraz Cassa Sunsteina (2008), a drugi to boosting opisany przez kontynuatorów myśli Gerda Gigerenzera, Ralpha Hertwiga i Tilla Grüne-Yanoffa (2016). Mimo wielu różnic pomiędzy dwoma podejściami, które zostaną opisane poniżej, warto zauważyć, że to

co je łączy, to nie tylko wspólne fundamenty w postaci teorii „ograniczonej racjonalności” Simona, ale również ostateczny cel, do którego mogą być wykorzystane przez architektów decyzji. Zarówno nudging, jak i boosting dąży do „zmiany zachowania” poprzez oddziaływanie na architekturę sytuacji decyzyjnej (Hertwig i Grüne-Yanoff, 2017). W związku z tym, że Simon poprzez wprowadzenie koncepcji ekologicznej racjonalności, włączył do rozumienia architektury sytuacji decyzyjnej również ograniczenia, wiedzę i zdolności poznawcze decydenta, nudging i boosting odnoszą się nie tylko do zmiany cech architektury decyzji fizycznej czy zewnętrznej, ale również do potencjalnej zmiany poznawczej czy wewnętrznej decydenta, np. poprzez wyposażenie go w określone „filtry” poznawcze, czy po prostu umiejętności rozumienia danej sytuacji. Ważne jest, by zauważyć, że obydwie kategorie interwencji nie dążą do zmiany przekonań czy postaw decydentów i w konsekwencji kształtując ich wybór, a w bardziej behawioralny sposób. Dlatego coraz częściej w literaturze te dwie kategorie interwencji nazywa się „interwencjami behawioralnymi” (Hertwig i Grüne-Yanoff, 2017; Nisa i in., 2019). Dodatkowo, co szczególnie istotne w porównaniu z klasycznymi rozwiązaniami polityk publicznych, ani nudging ani boosting nie stosuje finansowych kar ani nagród. Interwencje nie opierają się również na zabranianiu jednostce żadnych opcji decyzyjnych (Grüne-Yanoff i Hertwig, 2016; Sunstein i Reisch, 2023).

Kilka lat starszym, a przez to powszechniejszym podejściem do zmiany zachowań na poziomie jednostki i społeczeństwa jest nudging². Definicja nudge przyjmuje różne formy w literaturze, zwykle różniąc się liczbą kategorii interwencji kwalifikujących się jako nudging,

² 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 działanie poprzez „popychanie”, „kłucie” lub „łokciowanie” (Cambridge Dictionary, bd). Żadne z powyżej przedstawionych tłumaczeń nie oddaje w pełni pierwotnego znaczenia tego terminu. W związku z tym autor tego rozdziału zdecydował się pozostawić oryginalną angielską formę.

funkcją interwencji behawioralnych oraz wpływem nudge na architekturę podejmowania decyzji oraz samych decydentów. Jednakże główne koncepcje związane z podejściem nudge opierają się na definicji zaproponowanej przez Thaler i Sunsteina (2008), którzy definiują nudging jako „każdy aspekt architektury sytuacji wyboru, który przewidywalnie zmienia zachowanie, bez zabrania jakiegokolwiek opcji, ani istotnej zmiany korzyści ekonomicznych”. Wprowadzenie nudgingu spowodowało intensywną dyskusję na temat etyki, nadziei i potencjalnych zastosowań tej koncepcji w polityce społecznej. Ze względu na zazwyczaj niskie koszty wdrożenia, podejście nudge, które z definicji nie uwzględnia zachęt finansowych ani kar, intensywnie się rozwija, zwłaszcza w obszarach decyzji finansowych, zdrowotnych i środowiskowych (Benartzi i in., 2017; Franklin i in., 2019).

Mimo, że nudging został wypromowany jako skuteczny sposób zmiany zachowań poprzez ingerencję w architekturę sytuacji podejmowania decyzji, nie jest to jedyny sposób, który może być wykorzystany do efektywnej zmiany zachowania. Drugim, bardziej nowoczesnym, ale równie obiecującym podejściem jest metoda boost, zaproponowana przez Grüne-Yanoffa i Hertwiga (2016). Podczas, gdy główny wpływ nudge'y dotyczy architektury sytuacji podejmowania decyzji, co skutkuje wpływem na zachowanie decyzyjne, głównym celem boostów jest zwiększenie kompetencji decydenta w podejmowaniu najlepszych możliwych decyzji, zgodnie z preferencjami decydenta oraz odpowiednimi czynnikami w danej sytuacji. Pomimo, że nudging i boosting działają w podobnych obszarach architektury wyboru, zwolennicy podejścia boost wskazują na kilka różnic między nimi. Przede wszystkim twierdzą, że nudging w przeciwieństwie do boostingu może wpływać na autonomię decydenta, może mieć możliwie odwracalne efekty po wycofaniu interwencji oraz opiera się na błędach poznawczych jako podstawie swojego wpływu (Hertwig i Grüne-Yanoff, 2017). Ponadto zwolennicy boostów

twierdzą, że nie zakłócają one autonomii decydenta, nie wywołują żadnych długoterminowych efektów, ani nie wpływają na kompetencje decydenta czy poprawę sposobu podejmowania decyzji. Mimo, że zwolennicy podejścia boost twierdzą, że nie zakłócają one autonomii decydentów, autor podziela zdanie, że nie jest to całkowicie prawdziwe (Krawiec i in., 2021). Jeśli interwencje typu boost są wybierane na subiektywnym, indywidualnym poziomie, zachowana jest wolność wyboru, jednak, gdy do promocji boostów używane są bardziej argumenty polityczne lub ideologiczne, mogą one ograniczać autonomię wyboru, podobnie jak interwencje typu nudge. Choć interwencje typu boost wydają się bardziej neutralne, gdy boost jest proponowany przez podmiot będący wyżej w hierarchii, może zostać narzucony określony styl zachowania.

Przedstawione poniżej taksonomie nudge'y i boostów, zostały zaadaptowane z artykułu Krawiec i in. (2021) i opierają się na pracach Sunstein i Reisch (2023) oraz Hertwiga i Grüne-Yanoffa (2017). Podział uwzględnia również prace Hummela i Maedche (2019) oraz Kozyreva i in. (2020). W swojej metaanalizie, Hummel i Maedche (2019) rozwinęli taksonomię nudge'y, wprowadzając kategoryzację nudge'y ze względu na ich zasadę działania. Innymi słowy, taksonomia przedstawiona w Tabeli 1 zakłada, że architekci decyzji w różnych kontekstach mogą chcieć skłaniać ludzi do konkretnych opcji prezentując im wybraną, ograniczoną liczbę opcji do wyboru lub poprzez przedstawienie opcji w określony sposób. Kategoria nudge'y „oddziaływujących na strukturę sytuacji decyzyjnej” odnosi się do tego, co prezentować, podczas gdy kategoria nudge'y „opisująca możliwe opcje decyzyjne” reprezentuje sposób, w jaki opcje do wyboru są prezentowane. Warto zauważyć, że narzędzie implementacji intencji jest uwzględniane oryginalnie zarówno przez jeden, jak i drugi zespół badaczy. Jednakże, zdaniem autora, gdyby kierować się spójnością definicji interwencji nudge i boost, technika implementacji

intencji (Gollwitzer, 1999) bardziej pasuje do kategorii boostów, którą można przekazać decydentom jako rodzaj kompetencji do wykorzystania zgodnie ze swoimi preferencjami, a nie automatyczny nudge, który jest włączany do sytuacji decyzyjnej jako jej element przez architekta decyzji. Technika implementacji intencji to strategia opracowana pierwotnie przez Gollwitzera (1999), która łączy sytuacje w których chcemy przykładowo uniknąć określonego zachowania z przygotowanymi reakcjami; na przykład: "Zawsze, gdy wystąpi sytuacja x, rozpocznę działanie y". Skuteczne wykorzystanie tego narzędzia wymaga indywidualnego świadomego planowania, co jest umiejętnością bardziej prawdopodobną do zastosowania przez osobę świadomą swoich preferencji, a więc elementu kluczowego dla boostów. Z drugiej jednak strony, budowanie taksonomii jako dwóch oddzielnych kategorii interwencji, komplikuje wprowadzenie przez Sunsteina (2016) dodatkowej kategorii „nudge’y edukacyjnych”, które mają bliźniaczą charakterystykę do interwencji typu boost Hertwiga i Grüne-Yanoffa (2017). Sunstein stara się w tej publikacji wskazać, że boosty są właśnie pewnym rodzajem nudge’y edukacyjnych, co nie spotkało się ze zrozumieniem i aprobatą autorów interwencji typu boost (Hertwig i Ryall, 2020).

Tabela 1. Systematyka interwencji typu nudge

Rodzaj interwencji	Charakterystyka/przykłady działania
Nudge oddziałujące na strukturę sytuacji decyzyjnej	
1. Nudge opcji domyślnej (ang. <i>default</i>)	Automatyczny zapis do programów takich jak programy oszczędnościowe, programy dawców organów czy edukacyjnych
2. Uproszczenie (ang. <i>simplification</i>)	Zmniejszanie złożoności cech sytuacji decyzyjnych
3. Nawiązanie do normy społecznej (ang. <i>social norm/social reference</i>)	Podkreślanie, co robi większość osób w tej sytuacji, tj. nawiązywanie do deskryptywnej normy społecznej np. „Większość osób planuje zagłosować” lub „9/10 klientów hotelu używa ręczników kilkakrotnie
4. Zmiana w dostępności/ułatwienie (ang. <i>increases in ease and convenience</i>)	Wysuwanie tańszych lub zdrowszych opcji na pierwszy plan decyzyjny
5. Ujawnianie (ang. <i>disclosure</i>)	Prezentowanie ekonomicznych lub środowiskowych kosztów związanych z opcjami do wyboru
Nudge opisujące możliwe opcje decyzyjne	

6. Ostrzeżenia/grafiki (ang. <i>warnings/graphics</i>)	Prezentowanie graficzne ryzyka i nawiązywanie do emocji, np. przy okazji skutków palenia na paczkach papierosów
7. Przypomnienia (ang. <i>reminders</i>)	Przypomnienia e-mailowe lub smsowe, np. dotyczące zaległych rachunków lub umówionych wizyty lekarskich
8. Pre-zobowiązania (ang. <i>precommitment</i>)	Zobowiązanie się do konkretnych decyzji/zachowań z wyprzedzeniem, np. w celu ograniczenia spożywania alkoholu lub oszczędzania
9. Informowanie decydentów o naturze i konsekwencjach ich przeszłych decyzji (ang. <i>informing of past choices/feedback</i>)	Dostarczanie obywatelowi jego historii rachunków, wydatkowania energii, historii leczenia, np. w oparciu o bazy danych korporacji i podmiotów publicznych
10. Implementacja intencji	Tworzenie wcześniej przygotowanych scenariuszy działania na wypadek wystąpienia określonej sytuacji „Jeśli x, to y”

Adnotacja: Tabela zaadaptowana z Krawiec i in. (2021) i uzupełniona w oparciu o Sunstein i Reisch (2023).

Taksonomia interwencji typu boost (Tabela 2) skupia się z kolei na trzech kategoriach: boostach wzmacniających umiejętności odczytywania i rozumienia ryzyka, boostach pomagających zarządzać niepewnością oraz boostach motywacyjnych. Te interwencje są podzielone na osobne kategorie, aby podkreślić ich istotne cele: pierwsza kategoria ma na celu rozwijanie kompetencji pomagających jednostkom lepiej zrozumieć informacje statystyczne, boosty pomagające zarządzać niepewnością wprowadzają plany działania, które krok po kroku umożliwiają podejmowanie decyzji zgodnych z preferencjami jednostki, a boosty motywacyjne mają na celu szeroko pojęte rozwijanie umiejętności jednostki w indywidualnym wzmocnieniu swojej motywacji, kontroli poznawczej i samoograniczenia.

Tabela 2. Systematyka interwencji typu boost

Rodzaj interwencji	Charakterystyka/przykłady działania
Boosty wzmacniające umiejętności odczytywania i rozumienia ryzyka	
1. Reprezentacje graficzne ryzyka	Prezentowanie ryzyka w sposób ułatwiający jego zrozumienie, np. przekształcając ryzyko z formatu absolutnego „5% osób może doświadczyć efektów ubocznych” na „50 z 1000 osób może doświadczyć efektów ubocznych”

2. Reprezentacje oparte na doświadczeniu	Zmiana prezentacji ryzyka deskrypcyjnego na ryzyko prezentowane w sposób doświadczeniowy – decydent dowiaduje się o ryzyku wystąpienia danego obiektu samodzielnie losując z puli obiektów
3. Reprezentacje unikające wpływu zniekształceń ramowania (ang. <i>framing</i>)	Przekształcanie informacji na niezależne od ramy zysków lub strat
4. Krótkie szkolenie w transformacji nieprzejrzystych reprezentacji	Trening rozumienia ramowania i sposobów jego przekształcania
5. Szkolenie ogólnych umiejętności matematycznych	Przekazywanie takich umiejętności jak obliczanie wartości oczekiwanej (ang. <i>Expected value</i>) lub reguły Franklina (ang. <i>Weighted Additive rule</i>)

Boosty wspierające zarządzanie niepewnością

6. Proste metody wnioskowania aktuarialnego	Proste metody matematyczne stosowane przy obliczaniu ryzyka wypadków losowych
7. Proste zasady inteligencji zbiorowej	Wykorzystywanie reguł „mądrości tłumu” (ang. <i>wisdom of the crowd</i>). Np. wykorzystywanie średniej odpowiedzi lub prognozy uzyskanej poprzez agregację różnych opinii lub szacunków członków grupy
8. Reguły proceduralne: szybkie i oszczędne drzewa decyzyjne (ang. <i>Fast and Frugal trees</i>), heurystyki proste, rutynowe procedury	Drzewa decyzyjne składające się z kilku warstw, z których każda reprezentuje kolejny krok decyzyjny. Heurystyki proste i procedury. Przykłady: akronimy pomagające zapamiętać kolejne kroki jak „Trzy C” (ang. <i>Three Cs</i> ; WHO, 2020) czy reguła dywersyfikująca 1/N
9. Strategie inokulacji	Odnoszą się do mentalnego „szczepienia”, podawania jednostkom mentalnej szczepionki poprzez szkolenie ich w procesie rozpoznawania określonych informacji

Boosty motywacyjne

10. Pisanie ekspresywne	Forma pisania koncentrująca się na wyrażaniu i eksplorowaniu własnych myśli, uczuć i doświadczeń
11. Ćwiczenia związane z mindsetem wzrostu (ang. <i>growth mindset</i>)	Ćwiczenia oparte na koncepcji mindsecie wzrostu Carol Dweck
12. Treningi uwagowe	Treningi rozwijające zdolności uwagowe w zadaniach poznawczych
13. Treningi wzmacniające poczucie związku z Ja-przyszłościowym	Rozwijanie kompetencji osób w zakresie wzmacniania ich poczucia związku/łączności/powiązania z Ja-przyszłościowym (ang. <i>Future-self</i>)
14. Ćwiczenia wiązania nagród i wysiłku (ang. <i>Reward-/Temptation-bundling</i>)	Sposób wypracowania nawyków łączenia awersyjnych działań z przyjemniejszymi zachowaniami

15. Implementacja intencji i tworzenie automatycznych wzorców zachowania	Tworzenie wcześniej przygotowanych scenariuszy działania na wypadek wystąpienia określonej sytuacji „Jeśli x, to y”
16. Strategie pre-zobowiązania	Zobowiązanie się do konkretnych decyzji/zachowań z wyprzedzeniem, np. w celu ograniczenia spożywania alkoholu lub oszczędzania

Adnotacja: Tabela zaadaptowana z Krawiec i in. (2021) i uzupełniona w oparciu o Hertwig i Grüne-Yanoff (2017) oraz Kozyreva i in. (2020)

Dotychczasowe zastosowania nudge’y i boostów można zanotować w programach polityki społecznej USA i Wielkiej Brytanii, rozslawionych przez działania rządowej agencji badawczo-wdrożeniowej Behavioural Insights Team (Hallsworth i in., 2018). Szczególnie dużym echem odbiły się sukcesy programów oszczędnościowych i zdrowotnych wykorzystujących nudge opcji domyślnej (ang. *default nudge*). Interwencja ma na celu nakłonienie ludzi do lepszego wyboru z perspektywy architekta sytuacji decyzyjnej poprzez wyznaczenie jednej z alternatyw jako automatyczny wybór decydenta. Nudge opcji domyślnej wpływa na jednostki poprzez wstępny wybór jednej opcji, co sprawia, że decydenci chętniej przy niej pozostają, mimo możliwości rezygnacji w każdej chwili (Jachimowicz i in., 2019). Thaler i Benartzi (2004) wskazują, że dzieje się tak m.in. ze względu na efekt inercji i błąd poznawczy typu „status quo”, który zniechęca osoby do podjęcia działania i w konsekwencji innej decyzji. Dotychczasowe badania przeglądowe, wskazały, że to właśnie nudge opcji domyślnej/zapis automatyczny jest najbardziej efektywną interwencją typu nudge (Jachimowicz i in., 2019; Sunstein, 2014) oraz najczęściej badaną (Hummel i Maedche, 2019). Drugim najczęściej badanym rodzajem interwencji typu nudge są ostrzeżenia/grafiki, które obejmują różne sygnały mające na celu skierowanie uwagi jednostek na kluczowe informacje (Hummel i Maedche, 2019). Jest to bardzo szeroka kategoria interwencji, głównie obejmująca bodźce wykorzystujące kolory, emocje i wszelkie aspekty wyróżniające daną cechę z otoczenia. Jak wskazują Sunstein i Reisch (2023), właśnie ten rodzaj nudge’y bazuje na takich błędach poznawczych jak opisywana wcześniej

heurystyka dostępności czy błąd poznawczy istotności (ang. *salience bias*) opierający się na fizycznej i poznawczej dostępności i w konsekwencji więcej „wadze” w procesie podejmowania decyzji.

Według Hummel i Maedche (2019) obok nudge’ a opcji domyślnej i grafik, trzecią najczęściej badaną interwencją typu nudge jest rodzaj nudge’ a odwołującego się do norm społecznych, a konkretnie do deskryptywnej normy społecznej (Mol i in., 2020). Główną ideą tej interwencji jest opisanie opcji zawartych w sytuacji decyzyjnej w odniesieniu do kontekstu społecznego, przykładowo podkreślając zachowanie wykonywane przez większość osób w tej sytuacji. Jeden z najbardziej skutecznych nudge’ y dotyczących norm społecznych został opisany przez Hallswortha i in. (2017), gdzie głównym celem było zwiększenie płatności za zaległe podatki. Stało się to możliwe dzięki zastosowaniu listów przypominających, które zawierały komunikaty wykorzystujące normy społeczne. Najbardziej skuteczny komunikat zawierał deskryptywną normę społeczną, odnoszącą się do zasady, której przestrzegała większość osób w danym regionie Wielkiej Brytanii - w tym kontekście terminowej płatności podatków. Dzięki wprowadzonej zmianie badacze z Behavioural Insights Team, byli w stanie zwiększyć spłatę podatków o dodatkowe 9,3 mln funtów w ciągu pierwszych 23 dni eksperymentu.

W związku z tym, że interwencje typu boost, są o 8 lat młodsze i jak wskazują autorzy (Hertwig i Grüne-Yanoff, 2017) mogą być nieco bardziej czasochłonne w implementacji w politykach publicznych od interwencji typu nudge, przykładów ich zastosowania w politykach publicznych jest zdecydowanie mniej. Niemniej jednak, jako jeden z najbardziej znaczących boostów można wskazać strategię inokulacji, wykorzystywane do wzmocnienia zdolności jednostek do rozpoznawania dezinformacji, takich jak fałszywe wiadomości w internecie. Interwencja stworzona przez Roozenbeeka i Van der Lindena (2019) odnosi się do teorii

inokulacji opracowanej przez McGuire'a i Papageorgisa (1961). Autorzy uważają, że podobnie jak szczepionka biologiczna, osłabiona wersja dezinformacji może wzmocnić jednostki i uczynić je odpornymi na fałszywe wiadomości. W swoim projekcie badacze poszli nawet o krok dalej - stworzyli grę przeglądarkową, pozwalającą uczestnikom na doświadczanie produkcji fałszywych informacji od środka. Aby zdobywać punkty, muszą używać sześciu technik powszechnie stosowanych w dezinformacji, m.in. polaryzacji, wywoływania emocji, szerzenia teorii spiskowych czy trollingu w internecie. Jak dotąd wyniki pokazują, że po takim szkoleniu uczestnicy są znacznie lepsi w rozpoznawaniu i oporze wobec fałszywych wiadomości napotkanych po rozgrywce (Basol i in., 2020). Innym przykładem interwencji typu boost, którą można zakwalifikować do kategorii „Boost zarządzający niepewnością” jest reguła DDM, którą w trakcie pierwszej fali pandemii COVID-19 wprowadziło do masowej komunikacji polskie Ministerstwo Zdrowia. Rozwinięcie skrótu (Dystans, Dezynfekcja, Maseczki) miał być prostą instrukcją postępowania – heurystyką – pozwalającą obywatelom na zmniejszenie ryzyka zarażenia się wirusem COVID-19 w przestrzeni publicznej (Ministerstwo Zdrowia, 2020). Podobne akronimy pod postacią „Avoid three C’s” sugerujące kolejne etapy decyzyjne rozpowszechniała również Światowa Organizacja Zdrowia (WHO, 2020; World Health Organization, 2022).

1.1.4 Program Save More Tomorrow – więcej niż pojedyncza interwencja behawioralna

Jedną z kluczowych domen leżących w obrębie zainteresowań wielu ekonomistów behawioralnych jest oszczędzanie (Benartzi i in., 2017). Według Richarda Thaler (2018), teorie ekonomiczne można podzielić na trzy kategorie: normatywne, deskryptywne i preskryptywne. W kontekście oszczędzania, normatywną teorią jest Teoria Cyklu Życia, która zakłada, że gospodarstwa domowe chcą wyrównać konsumpcję na przestrzeni całego cyklu życia,

funkcjonowania. Zgodnie z tą teorią oczekuje się, że gospodarstwa domowe i indywidualni decydenci będą rozwiązywać odpowiedni problem optymalizacyjny w każdym okresie, zanim zdecydują, ile konsumować i ile oszczędzać. Jak wskazuje Thaler (2018), teoria cyklu życia, zakłada, że zwykli obywatele podejmują decyzje, tak jak zrobiliby to ekonomiści. Innymi słowy, opisuje zachowanie decydentów jako racjonalne ekonomicznie i dążące do optymalizacji przy nieograniczonych zasobach obliczeniowych. W rzeczywistości, zachowanie gospodarstw domowych może różnić się od tego optymalnego planu przynajmniej z dwóch powodów. Po pierwsze, problem rozkładania bieżących wydatków jest trudny, nawet dla ekonomisty, więc gospodarstwa domowe mogą nie być w stanie obliczyć poprawnej kwoty oszczędności, jaką co miesiąc czy pół roku powinni odkładać, żeby osiągnąć zakładany kapitał po 20 czy 30 latach. Po drugie, nawet jeśli znana byłaby poprawna kwota okresowego oszczędzania, gospodarstwa domowe mogą nie mieć wystarczającej samokontroli, aby zmniejszyć bieżącą konsumpcję na rzecz przyszłej konsumpcji (Shefrin i Thaler, 1988).

Jako alternatywę dla teorii normatywnej, Shefrin i Thaler (1988), proponują Behawioralną Teorię Cyklu Życia, która jako teoria deskryptywna stara się opisać, jak ludzie faktycznie dokonują wyborów, podkreślając systematyczne odstępstwa od teorii normatywnej. Teoria opiera się na dwóch kluczowych aspektach: samokontroli oraz mentalnym księgowaniu (ang. *mental accounting*). W ramach samokontroli, autorzy ujmują konflikt pomiędzy „sprawcą” (ang. *Doer*) a planistą (ang. *Planner*), które należy rozumieć analogicznie do ujęcia dualistycznego przetwarzania informacji Kahnemana (2011) czy Epsteina (1994), którzy rozróżniali działania człowieka na kierowane Systemem 1 i Systemem 2. W kontekście Behawioralnej Teorii Cyklu Życia, to moduł sprawcy dąży do szybkiego konsumowania zasobów i zniechęca do odraczania gratyfikacji, a rolą modułu planisty jest

kontrola bieżących impulsów i próby odkładania dóbr, w tym pieniędzy, na później. Drugim elementem omawianej teorii, są mechanizmy mentalnego księgowania, które przypominają księgowanie w przedsiębiorstwach. Podobnie, jak w budżecie przedsiębiorstwa, Shefrin i Thaler wyróżniają trzy umysłowe rachunki: rachunek bieżących dochodów, rachunek majątku i rachunek przyszłych dochodów. Autorzy zakładają, że jeśli któreś środki finansowe zostaną zakwalifikowane przez decydenta jako rachunek majątku lub jako rachunek przyszłych dochodów, to rosną szanse, że zostaną zaoszczędzone (Shefrin i Thaler, 1988; Zaleśkiewicz, 2011). Jak wskazują badacze z obszaru psychologii ekonomicznej, teoria Shefrina i Thalera lepiej od teorii cyklu życia, tłumaczy dlaczego ludzie m.in. nie oszczędzają wystarczająco funduszy w wieku produkcyjnym i dlaczego na emeryturze nie wydają swoich środków w całości (Zaleśkiewicz, 2011).

Zgodnie z nomenklaturą Thalera (2018), trzecim rodzajem teorii w naukach społecznych są teorie preskryptywne, które dążą do tego by oferować doradztwo w zakresie tego, jak decydenci mogą poprawić swoje decyzje i zbliżyć się do ideału opisywanego przez teorie normatywne. Przykładem teorii preskryptywnej w domenie oszczędzania jest projekt Save More Tomorrow (Thaler i Benartzi, 2004), którego celem była i jest pomoc tym pracownikom, którzy chcieliby oszczędzać więcej, ale brakuje im silnej woli, aby działać zgodnie z tym pragnieniem. Choć program Save More Tomorrow (SMT) został opisany przez Thalera i Benartziego w 2004 roku, do tej pory żaden program próbujący łączyć różne rodzaje interwencji behawioralnych z innymi instrumentami polityk publicznych nie odniósł tak efektywnych rezultatów (Alt i in., 2024). Rdzeniem programu SMT, jest zawarty w Tabeli 1, jeden z najbardziej znanych i najefektywniejszych nudge'y – nudge opcji domyślnej (Jachimowicz i in., 2019). Program ma na celu zwiększenie liczby osób oszczędzających na emeryturę w ramach nieobowiązkowego planu

emerytalnego, w trakcie którego co miesiąc pracownik i pracodawca odkładają określoną wpłatę na indywidualne konto partycypującego (Thaler i Benartzi, 2004). Bazując na wynikach badań Madrian i Shea (2001), Thaler i Benartzi opracowali program składający się z trzech elementów mających na celu ograniczenie behawioralnych i psychologicznych czynników zmniejszających partycypację w programach emerytalnych:

1. Pracownicy są automatycznie zapisywani do planu oszczędnościowego w momencie rozpoczęcia pracy w danej firmie, lub przystąpienia podmiotu do programu (nudge opcji domyślnej). W zamyśle Thalera i Benartziego, rezultatem opcji domyślnej jest występowanie inercji i błędu poznawczego status quo, które zmniejszają szanse na podjęcie wysiłku przez pracowników i wypisanie się z programu, co zgodnie z założeniami programu, mogą w każdej chwili uczynić (Thaler i Benartzi, 2004).
2. Wysokość comiesięcznej raty wpłacanej przez pracownika, jest definiowana możliwie z jak największym wyprzedzeniem przed rozpoczęciem faktycznego rozpoczęcia oszczędzania, co jak wskazują autorzy, ma na celu zmniejszenie dyskontowania hiperbolicznego. Zapis oraz faktyczne rozpoczęcie oszczędzania powinno dzielić jak najwięcej czasu – najlepiej kilka miesięcy. W nawiązaniu do Tabeli 1, temu etapowi najbliższą jest do interwencji typu „pre-zobowiązania” (Sunstein, 2014; Thaler i Benartzi, 2004).
3. Wysokość comiesięcznej raty jest zwiększana z każdą kolejną podwyżką, aż do osiągnięcia maksymalnego poziomu raty. Krok ma na celu zmniejszenie awersji do ryzyka, potencjalnie pojawiającego się w momencie zauważenia odkładanych pieniędzy, a w tym momencie „ukrytego” w rezultacie podwyżki. Do tego etapu,

uczestnicy się „pre-zobowiązują” podczas przystępowania do programu (Thaler i Benartzi, 2004).

Wyniki pierwszego wdrożenia programu SMT przyniosły bardzo obiecujące rezultaty: po 40 miesiącach od wprowadzenia programu, 80% pracowników w nim pozostało, a średnia miesięczna wysokość raty wzrosła z 3,5% do 13,6%. W ciągu 40 miesięcy trwania programu miały miejsce cztery coroczne podwyżki, którym towarzyszyło automatyczne zwiększenie miesięcznej raty odkładanej przez pracownika. W związku z osiągniętymi rezultatami, program SMT z kluczowym nudgem opcją domyślną jest uznawany za jedną z najbardziej efektywnych interwencji behawioralnych (Jachimowicz i in., 2019). Co dodatkowo ważne, dla przyszłych rozważań na temat synergii, program opracowany przez Benartziego i Thalera łączy przynajmniej dwa rodzaje nudge'y: opcję domyślną oraz pre-zobowiązanie. W porównaniu do poprzednich badań Madrian i Shea (2001), które jako pierwsze raportowały o automatycznym zapisie do programów, program SMT nie tylko osiągnął podobną partycypację w wysokości 80% osób uprawnionych do udziału, ale również wskazał na ciągły udział w programie pomimo kolejnych wzrostów w wysokości raty oszczędzania z 3,5% do 13,6%. Z perspektywy normatywnej teorii cyklu życia, jest to najbardziej pożądanym wynikiem, gdyż osoby badane w ten sposób zwiększały przyrost swoich środków na przestrzeni lat, realizując w ten sposób dwa cele: nie tylko stałe oszczędzanie, ale również systematyczne podnoszenie tempa oszczędzania (Thaler i Benartzi, 2004).

Jako alternatywę dla oryginalnej wersji programu SMT opartej na podejściu nudge, Hertwig i Grüne-Yanoff (2017) proponują zastosowanie dwóch rodzajów boostów: regułę proceduralną ułatwiającą dywersyfikację środków 1/N oraz interwencję wzmacniającą poczucie łączności z Ja-przyszłościowym (ang. *future-self connectedness*). Reguła 1/N opiera się na

równomiernym dzieleniu posiadanych funduszy i inwestowaniu ich w N różnych branż, funduszy czy domen. W przypadku oszczędzania oznaczałoby to odłożenie równowartości 1 z N kategorii naszych wydatków. Przykładowo, jeśli korzystając z mentalnego księgowania przyporządkujemy wydatkom etykiety takie jak produkty spożywcze, wyjścia kulturalne czy odzież, średnią, podobną kwotę wydawaną na te kategorie moglibyśmy przeznaczać na konto oszczędnościowe (Drexler i in., 2014; Hertwig i Grüne-Yanoff, 2017). Druga interwencja miałaby za zadanie wzmocnić psychologiczną łączność z Ja-przyszłościowym. W kontekście oszczędności, oznaczałoby to, że „im bardziej ktoś jest świadomy tego, że będzie przyszłym beneficjentem dzisiejszych oszczędności, tym bardziej ta osoba będzie przygotowana do oszczędzania na emeryturę. Zgodnie z tą samą logiką, osoba, która jest oderwana od swojego przyszłego ja — przez brak wiary lub wyobraźni — jest mniej skłonna do oszczędzania” (Hertwig i Grüne-Yanoff, 2017). Badacze powołują się w tej propozycji na badania Zaval i in. (2015), gdzie osoby badane pisały krótki esej na temat tego, jak chciałyby być zapamiętane przez przyszłe pokolenia. Ta metoda okazała się pomocna w skłanianiu ludzi do rozważania długoterminowej perspektywy i promowania proekologicznych intencji oraz zachowań. Hertwig i Grüne-Yanoff (2017) wskazują, że jeśli ludzie są wyposażeni w kompetencje psychologiczne, umożliwiające im mentalne przekraczanie długich horyzontów czasowych, sami, a nie w wyniku oddziaływania architektów wyboru, mogą zdecydować o wykorzystaniu tej kompetencji, kiedykolwiek dostrzegają asymetrię między korzyściami krótkoterminowymi a kosztami długoterminowymi. Autorzy nie oczekują uzyskania większych efektów tej „boostowej” interwencji w porównaniu do oryginalnego programu SMT. Jednakże, w sytuacji gdy nie byłoby możliwe zaimplementowanie SMT, interwencje boost zakładające mniejszą ingerencję strukturalną niż interwencje nudge, mogłyby okazać się bardziej elastycznym i zasadniejszym wyborem (Hertwig i Grüne-Yanoff, 2017).

1.1.5 Rola różnic indywidualnych i świadomości oddziaływań jako moderatorów efektywności interwencji behawioralnych

Choć początki nudgingu datowane są na 2008 rok, to zainteresowania badawcze wokół moderatorów efektywności interwencji behawioralnych uwarunkowanych stałymi cechami decydentów pojawiły się stosunkowo niedawno. Marchiori i in. (2017) jako pierwsi zwrócili uwagę, że badania nad behawioralnymi narzędziami polityk publicznych ograniczają się do świadomości oddziaływania i wskaźników socjo-demograficznych, natomiast rzadko analizuje się zmienne psychologiczne (Marchiori i in., 2017). Należy wyróżnić przynajmniej dwa obszary czynników, które do tej pory przyciągały uwagę badaczy jako potencjalne moderatory interwencji behawioralnych: świadomość interwencji (De Ridder i in., 2022) oraz czynniki osobowościowe (Ingendahl i in., 2021). W związku z tym, że boosting jest znacznie młodszym podejściem, przedstawiane przykłady w większości dotyczą badań na kategorią interwencji typu nudge.

Bovens (2009) jako jeden z pierwszych autorów opisujących interwencje behawioralne typu nudge wskazywał, że w celu zagwarantowania odpowiedniej skuteczności, nudge „muszą działać w cieniu”, lub przynajmniej być niezauważane przez tych, na których oddziałują (Steffel i in., 2016). To założenie wiązało się również z przekonaniem wielu autorów, że w obliczu paternalistycznych rozwiązań, jakimi są nudge, odbiorcy interwencji mogą być im niechętni, co w konsekwencji może prowadzić do efektu reaktancji (De Ridder i in., 2022). De Ridder (2023), powołuje się na szereg badań wskazujących, że jest dokładnie odwrotnie – świadomość interwencji nie prowadziła do spadku efektywności interwencji. Jednakże, jako o ograniczeniach dotychczasowych badań należy wspomnieć o tym, że większość z nich wykorzystywała hipotetyczne scenariusze, koncentrowała się w znacznej mierze na

interwencjach nudge opcji domyślnej i nie brała pod uwagę interwencji typu boost (de Ridder, 2023). Szczególnie w porównaniu z kategorią boostów, które z założenia powinny być zawsze transparentne, aby decydent mógł je świadomie użyć, różnice w efektywności nudge'y mogą zaistnieć.

W przypadku różnic indywidualnych, do tej pory tylko kilka badań poruszało zmienne osobowościowe jak model Wielkiej Piątki (Kamilçelebi i Songül, 2020; Warberg i in., 2019), potrzebę poznania czy potrzebę unikalności (Ingendahl i in., 2021). Jednym z wyjątków były badania Kamilçelebi i Songül (2020), którzy zidentyfikowali sumienność jako istotny predyktor zwiększający efektywność nudge'a opcji domyślnej. Z kolei badania Ingendahla i inni (2021) nad rolą potrzeby poznania, wskazały na niewielki lub brak związku tego czynnika z efektami nudge'a opcji domyślnej i nudge'a normatywnego. Brak efektów cech osobowościowych wchodzących w skład modelu Wielkiej Piątki zanotowali z kolei Warberg i in. (2019), którzy badali efekt dopasowywania interwencji pod kątem profilu osobowościowego do decydenta. Innym potencjalnym konstruktym uzupełniającym badania z wykorzystaniem modelu Wielkiej Piątki, mogłaby być Regulacyjna Teoria Temperamentu (Cyniak-Cieciura i in., 2018). Choć jeszcze było badań uwzględniających cechy temperamentalne w ocenie działania interwencji behawioralnych, dotychczasowe projekty badawcze poruszające Regulacyjną Teorię Temperamentu w decyzjach finansowych, wskazywały już na istotną rolę perseweratywności (Gąsiorowska, 2011). Niewielka liczba badań nad różnicami indywidualnymi w obszarze interwencji behawioralnych jest dowodem na obecną lukę badawczą. Dotychczasowe badania porównujące ze sobą nudge i boosty, również koncentrowały się jedynie na porównaniu efektywności interwencji między sobą, bez poszukiwania potencjalnych moderatorów wśród różnic indywidualnych (por. Bradt, 2019; Franklin i in., 2019; Ruggeri i in., 2021).

Wśród innych różnic indywidualnych, które mogą potencjalnie moderować efektywność interwencji behawioralnych są również zdolności numeryczne, zdolności poznawcze czy potrzeba poznania (Hertwig i Grüne-Yanoff, 2017). Autorzy podejścia boost, zauważają, że wykorzystywanie prostych heurystyk w bardziej skomplikowanych środowiskach decyzyjnych, może niekiedy wymagać pewnego poziomu poznawczych zdolności w danej domenie, by skutecznie się posługiwać heurystykami (Hertwig i Grüne-Yanoff, 2017). Wydaje się to tym bardziej zasadne w świetle badań Caballero i Plonera (2022), gdzie w przypadku osób o niższym dochodzie, boosty negatywnie wpływały na oszczędzanie energii, co nie miało miejsca w przypadku osób o wyższym dochodzie. Autorzy tłumaczyli to związkiem właśnie między niższymi zdolnościami poznawczymi a niskim poziomem zamożności (Caballero i Ploner, 2022). Kolejnymi aspektami, które mogłyby się okazać się pomocne przy okazji oceny interwencji behawioralnych w zachowaniach finansowych, są skłonność do ryzyka oraz perspektywa czasowa. Jak wskazywały jedne z dotychczasowych badań, w warunkach decyzji ryzykownych sformułowanych jako zyski lub straty, jedni decydenci mogą preferować podążanie za interwencjami boost, a inni za interwencjami typu nudge (Folke i in., 2021). Dotychczas nie zbadanym aspektem ryzyka w interwencjach behawioralnych, jest domena, której odbywa się decyzja. W dotychczasowych badaniach nad skłonnością do ryzyka, podkreślano kontekstowość i domenę podejmowanych decyzji (Blais i Weber, 2006), wydaje się więc zasadne, zbadać czy domenowa skłonność do ryzyka będzie moderować efektywność interwencji. Z kolei gdy decyzja ma odroczone konsekwencje jak np. w programie oszczędnościowym Save More Tomorrow, to wówczas perspektywy czasowe mogą być związane z określonym kierunkiem zachowań finansowych (Jacobs-Lawson i Hershey, 2005).

W kontekście czynników wpływających na interwencje, świeżym pojęciem jest pojęcie „podatności na nudging” (ang. *nudgeability*), które docelowo ma określać podatność jednostki na wpływ nudgingu, pomagając lepiej zrozumieć, jak ta kategoria interwencji wpływa na wybór (De Ridder i in., 2022). Wśród pytań badawczych oscylujących wokół definicji „nudgeability” pojawiły się takie kwestie, jak to 1) czy świadomość oddziaływania interwencji ma wpływ na jej efektywność, 2) jaki jest wpływ preferencji i celów personalnych na gotowość do zachowywania się zgodnie z kierunkiem wskazywanym przez nudge, 3) jaki jest wpływ stanu umysłu osoby w momencie dokonywania wyboru, tj. czy jest bardziej skłonna podejmować decyzję korzystając z Systemu 1, czy angażować się w powolne rozważania Systemu 2. Należy w tym miejscu nadmienić, że większość literatury przedmiotu utożsamia skuteczne działanie nudgingu z funkcjonowaniem w danym momencie w „ramach” Systemu 1 (np. De Ridder i in., 2022; Hertwig i Grüne-Yanoff, 2017; Ingendahl i in., 2021; Marchiori i in., 2017). W definicji „nudgeability” brak jest odniesień do różnic indywidualnych, takich jak osobowość czy właściwości poznawcze związane ze zdolnościami numerycznymi (Cokely i in., 2016). Choć powoli pojawiają się prace starające się badać określone różnice indywidualne w kontekście wybranych interwencji należących do grupy nudge (jak rola osobowości w opcjach domyślnych; Kamilçelebi i Songül, 2020), to w oparciu o pojedyncze badanie jednego typu interwencji próbują generalizować również na inne interwencje z kategorii nudge. Dotychczas przeprowadzono nieliczne badania porównujące więcej niż jedną interwencję pod kątem roli różnic indywidualnych (Ingendahl i in., 2021). Natomiast w badaniach porównujących nudge i boosty badacze poszukują potencjalnych moderatorów głównie wśród zmiennych socjo-demograficznych, na razie bez większych rezultatów (Ruggeri in., 2021; Ruggeri i Cafarelli, 2023).

1.1.6 Behawioralne Laboratoria Polityk Publicznych

Laboratoria polityki publicznej, nazywane również jako „Behawioralne Laboratoria Polityk Publicznych” (ang. *Behavioural Insights/Behavioural Policy Labs*; Olejniczak i in., 2020), są innowacyjnymi placówkami badawczymi, które integrują dane behawioralne z randomizowanych eksperymentów laboratoryjnych, jak i polowych, z procesem tworzenia, wdrażania oraz oceny polityk publicznych. Przykłady takich laboratoriów, jak MindLab w Danii, Behavioural Insights Team z Wielkiej Brytanii, czy GovLab w Stanach Zjednoczonych, pokazują różnorodność podejść i szeroki zasięg geograficzny, gdzie są stosowane (Hallsworth i in., 2018). Powstały głównie w ciągu ostatnich dwóch dekad i zyskały uznanie za skuteczne łączenie teorii behawioralnych z praktyką polityczną, czego przykładem może być wprowadzenie automatycznego zapisywania do planów emerytalnych w Wielkiej Brytanii (Dolan i in., 2012). Znaczenie korzystania z wglądów behawioralnych (ang. *behavioural insights*) w politykach publicznych wynika z potrzeby bardziej efektywnego kształtowania zachowań ludzkich w sposób, który promuje dobro społeczne i indywidualne. Wglądy te, oparte na najnowszych badaniach z psychologii, ekonomii behawioralnej i nauk społecznych, pozwalają na głębsze zrozumienie mechanizmów decyzyjnych, które kierują ludźmi. Dzięki temu, możliwe jest projektowanie polityk, które są nie tylko bardziej skuteczne, ale również mniej kosztowne i lepiej dostosowane do realnych potrzeb i oczekiwań społeczeństwa (Dolan i in., 2012; Olejniczak i in., 2019).

Wśród narzędzi opracowanych przez najbardziej znaną agencję badawczą, Behavioural Insights Team, najbardziej znanym jest paradygmat MINDSPACE, który jest rekomendowany jako jedno z narzędzi wykorzystywanych przez laboratoria polityk publicznych do analizy i projektowania interwencji behawioralnych (Hallsworth i in., 2018). Akronim MINDSPACE

składa się z dziewięciu elementów, które służą do projektowania komunikatów i interwencji w taki sposób, aby były one bardziej przystępne, atrakcyjne, społecznie akceptowalne i skuteczne w zachęcaniu do pożądanых zachowań z perspektywy architekta decyzji. MINDSPACE pomaga twórcom polityk zrozumieć, jak wykorzystać naturalne skłonności ludzkie do poprawy skuteczności polityk publicznych (Dolan i in., 2012). Proponowany paradygmat jest zbiorem zasad, które nawiązują do taksonomii nudge'y opublikowanej przez Sunsteina (2014), łącząc takie interwencje, jak opcje domyślne czy normy społeczne. Oprócz nich, w uproszczony sposób prezentują efekty znane w psychologii społecznej, takie jak efekt torowania, bez głębszego ich eksplorowania i tłumaczenia (por. Tabela 3). Sposób prezentacji modelu jest nastawiony na popularyzację szczególnie w gronach specjalistów ds. polityk publicznych, dlatego autorzy ograniczają naukowy i szczegółowy język.

Tabela 3. Paradygmat MINDSPACE

Rodzaj interwencji/bodźca	Charakterystyka/zachowanie
1. Nadawca (ang. <i>Messenger</i>)	Na zachowanie osób podejmujących decyzje może silnie oddziaływać to, kto jest nadawcą informacji
2. Zachęty finansowe (ang. <i>Incentives</i>)	Zachowanie decydentów może być zmienione w obliczu strat i zysków pieniężnych
3. Normy społeczne (ang. <i>Norms</i>)	Decydenci mogą sugerować się podczas podejmowania decyzji funkcjonującymi normami społecznymi
4. Opcje domyślne (ang. <i>Defaults</i>)	Opcje domyślne mogą wyznaczać kierunek podejmowania decyzji bez ich kontestowania
5. Waga/Ekspozycja informacji (ang. <i>Salience</i>)	Uwaga decydentów koncentruje się na informacjach eksponowanych w centralnym miejscu, zwiększając tym samym wagę informacji dla decyzji
6. Efekt torowania (ang. <i>Priming</i>)	Zachowanie osób podejmujących decyzje może być uwarunkowane bodźcami eksponowanymi jako pierwsze w sytuacji decyzyjnej
7. Afekt (ang. <i>Affect</i>)	Decydenci mogą być podatni na emocje współwystępujące z innymi elementami sytuacji decyzyjnej
8. Zaangażowanie (ang. <i>Commitments</i>)	Osoby podejmujące decyzje często postępują zgodnie ze swoimi publicznymi deklaracjami

Adnotacja: Tabela zaadaptowana z (Dolan i in., 2012)

Drugim narzędziem promowanym przez Behavioural Insights Team, które zyskało duży zasięg jest podejście EAST (Hallsworth i in., 2016), opierające się na zasadach: Łatwości (ang. *Make it Easy*), Atrakcyjności (ang. *Make it Attractiveness*), Społeczności (ang. *Make it Social*) oraz Terminowości (ang. *Make it Timely*). Proponowany paradygmat jest jeszcze prostszym schematem niż MINDSPACE, mającym na celu ułatwienie projektowania i wdrażania interwencji behawioralnych. Zasada Łatwości dotyczy zmniejszenia wysiłku potrzebnego do wykonania pożądanego zachowania, Atrakcyjność odnosi się do przyciągania uwagi do pożądanego zachowania, Społeczność wzmocnienia zachowań przez normy społeczne i wzajemne zobowiązania, a Terminowość wykorzystania odpowiednich momentów na implementację wprowadzanej zmiany. EAST wskazuje na praktyczne sposoby, dzięki którym można zwiększyć efektywność polityk poprzez uwzględnienie behawioralnych czynników wpływających na wybory decydentów (Hallsworth i in., 2016).

Nowe technologie, takie jak okulografia, międzynarodowe panele badawcze czy eksperymenty z wykorzystaniem decyzji wieloatrybutowych, oferują laboratoriom polityk publicznych możliwość przeprowadzania zaawansowanych badań nad zachowaniami ludzkimi w kontrolowanych warunkach (Kroese i in., 2016; Ruggeri, Ojinaga-Alfageme, i in., 2021; van den Broek-Altenburg i Atherly, 2020). Do niedawna, czymś nowym dla polityk publicznych były również randomizowane eksperymenty z grupami kontrolnymi, natomiast praktyka RCT (ang. *randomised-controlled trials*), została już na stałe wdrożona do repertuaru laboratoriów zajmujących się zmianą społeczną (Hallsworth i in., 2018). Nowe technologie pozwalają na precyzyjne mierzenie wpływu różnych czynników na wybory dokonywane przez ludzi, co może

przyczynić się do lepszego zrozumienia, jak skutecznie promować pożądane zachowania w ramach polityk publicznych. Wykorzystanie takich narzędzi umożliwia nie tylko głębsze badania nad mechanizmami decyzyjnymi (Ruggeri, Ojinaga-Alfageme, i in., 2021), ale również testowanie różnych interwencji w warunkach zbliżonych do rzeczywistych, co znacząco zwiększa potencjał tworzenia skutecznych i innowacyjnych polityk publicznych (Kroese i in., 2016). W kontekście stosowania wglądów behawioralnych do projektowania polityk publicznych, laboratoria polityki stają przed istotnym wyzwaniem: znalezieniem równowagi między ograniczeniem kosztów pre-testowania interwencji a dokładnym ocenianiem ich efektywności w kontrolowanych warunkach laboratoryjnych. Jest to kompromis między potrzebą oszczędności a koniecznością zapewnienia, że interwencje behawioralne będą skuteczne po wdrożeniu w rzeczywistych warunkach społecznych (Hallsworth i in., 2018). Wykorzystanie nowych technologii w procesie testowania interwencji w laboratorium oferuje obiecującą ścieżkę do osiągnięcia tego kompromisu. Rośnie również zainteresowanie symulowaniem bardziej realistycznych sytuacji decyzyjnych, w których twórcy polityk publicznych chcieliby odkryć preferencje obywateli w bardziej behawioralny a nie deklaracyjny sposób, stosując metody takie jak paradygmat wyborów dyskretnych (Knudsen i Havens, 2021; van den Broek-Altenburg i Atherly, 2020). Kwantyfikując jakościową hierarchię atrybutów w wyborach wieloatrybutowych, architekci decyzji są w stanie znacznie dokładniej przewidzieć potencjalne zachowanie obywateli i odbiór polityk publicznych (van den Broek-Altenburg i Atherly, 2020; Ruggeri i in., 2021). Rozwój nowych technologii ułatwia badanie preferencji za pomocą paradygmatu wyboru dyskretnego (Traets i in., 2020) i w konsekwencji wdrażanie tej metodologii w politykach publicznych, co jeszcze niedawno było znacznie bardziej kosztowne i czasochłonne (van den Broek-Altenburg i Atherly, 2020).

1.2 Podsumowanie części teoretycznej

Dokładnie 70 lat temu, Leonard Savage wprowadził rozróżnienie na „małe światy” i „duże światy”, żeby jeszcze lepiej opisać procesy podejmowania decyzji. Idee Savage’a są wciąż kontynuowane w pracach współczesnych badaczy zajmującymi się m.in. psychologią podejmowania decyzji (Gigerenzer i in., 2022) i wspólnie z teoriami Knighta (1921) i Simona (1972) stworzyły podstawy dla rozwoju deskryptywnych programów badawczych „Heurystyk i błędów poznawczy” (Kahneman, 2002) i „Prostych heurystyk” (Gigerenzer i Todd, 1999b). Z kolei te dwa programy badawcze, umożliwiły powstanie preskryptywnych podejść interwencji behawioralnych nudgingu i boostingu. W kolejnych sekcjach zostaną omówione wyzwania i ograniczenia tych dwóch najpopularniejszych kategorii interwencji behawioralnych, a następnie przedstawiona koncepcja badań własnych, odpowiadająca na bieżące, kluczowe pytania badawcze w tym obszarze.

1.2.1 Wyzwania i ograniczenia badań nad interwencjami behawioralnymi

Jednym z wyzwań, z którymi zmagają się laboratoria polityki publicznej i instytucje wprowadzające interwencje behawioralne, było przekonanie osób pełniących wysokie stanowiska w sektorze publicznym i prywatnym do ich stosowania obok klasycznych rozwiązań politycznych opartych na ekonomii i prawie (Krawiec i in., 2021; Marchiori i in., 2017). Początkowy sceptycyzm zmalął po przyznaniu Richardowi Thalerowi Nagrody Nobla w 2016 roku za jego wkład w ekonomię behawioralną. Do zwiększenia przekonania o efektywności interwencji behawioralnych typu nudge przyczyniły się także obiecujące wyniki projektów realizowanych przez takie organizacje jak Behavioural Insights Team (Benartzi i in., 2017; Sunstein i Reisch, 2023). To zainteresowanie zaowocowało szerszym przyjęciem podejścia behawioralnego, w tym w Polsce, gdzie wprowadzenie Pracowniczych Planów Kapitałowych,

wykorzystujących nudge opcji domyślnej, zwiększyło uczestnictwo w programach oszczędnościowych (Jedynak, 2018). Promowanie tych rozwiązań wymaga jednak ciągłego podkreślania znaczenia podejścia opartego na dowodach (ang. *evidence-base*), co podkreślają Ruggeri i in. (2021), wskazując, że jedynie empiryczne podejście do interwencji behawioralnych zapewni solidną podstawę dla przyszłych inicjatyw.

Kolejnym wyzwaniem stojącym przed proponentami podejścia interwencji behawioralnych jest kwestia etyki i przejrzystości nudgingu. Badacze tacy jak Bruns i in. (2018) oraz Rebonato (2014) podkreślają znaczenie jasnego komunikowania obecności, źródeł, celów i mechanizmów działania interwencji behawioralnych. Transparentność może zwiększyć akceptację i zaufanie do tych interwencji, co jest kluczowe dla ich skuteczności (Bruns i in., 2018). W kontekście badań wskazujących na niezależność efektywności interwencji behawioralnych od świadomości oddziaływania (de Ridder, 2023), należy podkreślić, że wyniki tych badań były oparte przede wszystkim tylko na jednym rodzaju nudgingu, opcjach domyślnych, podczas gdy nie sprawdzały czy inne interwencje typu nudge jak np. nudge normy społecznej, czy interwencje typu boost spotkają się z akceptacją obywateli (de Ridder, 2023). Jest to kolejna nisza badawcza wymagająca badań.

Trzecim wyzwaniem i jednocześnie ograniczeniem dotychczasowych wyników badań nad interwencjami behawioralnymi, jest pomijanie roli różnic indywidualnych w ocenie efektywności interwencji. Marchiori i in. (2017) zwracają uwagę na potrzebę uwzględniania szerokiego spektrum czynników psychologicznych, aby zwiększyć trafność i skuteczność tych interwencji. Zrozumienie, jak osoby charakteryzujące się różnymi wynikami na skalach osobowości, skłonności do ryzyka czy postrzegania czasu reagują na różne typy interwencji behawioralnych,

może doprowadzić do bardziej spersonalizowanych i skutecznych rozwiązań w politykach publicznych.

Analiza i porównanie efektów interwencji behawioralnych z uwzględnieniem efektu publikacyjnego (ang. *publication bias*) stanowią kolejne wyzwanie dla oceny przydatności interwencji behawioralnych (Maier i in., 2022; Mertens i in., 2022). Istnieje potrzeba przeprowadzania systematycznych badań, które pozwolą na porównanie skuteczności różnych interwencji niezależnie od przyjętych metod analitycznych i domen podejmowania decyzji. Takie porównania mogą pomóc w identyfikacji najbardziej efektywnych strategii oraz w zrozumieniu mechanizmów leżących u podstaw zachowań ludzkich. Jak wskazuje Maier i in. (2022), po uwzględnieniu tendencji publikacyjnej (ang. *publication bias*), średni efekt interwencji może być mały, na poziomie d Cohena równy 0,11.

Ostatecznie, kompromis pomiędzy kosztami testowania a dokładną oceną efektywności interwencji behawioralnych w warunkach laboratoryjnych stanowi istotny dylemat dla projektantów polityk (Hallsworth i in., 2018). Przeznaczenie wystarczających funduszy na etapie testowania interwencji może zapobiec dalszym kosztom podczas wdrażania interwencji o małej sile efektu (Mertens i in., 2022). Zastosowanie nowych technologii i metod badawczych takich jak paradygmat decyzji wieloatrybutowych czy okulografia (Kroese i in., 2016; van den Broek-Altenburg i Atherly, 2020) może pomóc w rozwiązaniu tego problemu, umożliwiając bardziej precyzyjne i kosztowo bardziej efektywne testowanie interwencji przed ich wdrożeniem. Takie podejście nie tylko zwiększa prawdopodobieństwo skuteczności interwencji, ale także przyczynia się do bardziej efektywnego wykorzystania zasobów publicznych, co jest kluczowe w kontekście ciągłego dążenia do poprawy dobrostanu społecznego.

1.2.2 Koncepcja badań własnych

Popularyzacja wyników interwencji behawioralnych, szczególnie z kategorii nudgingu, doprowadziła do wzrostu zainteresowania podmiotów z sektora publicznego i biznesu prowadząc do zakładania kolejnych (po Behavioural Insights Team) zespołów zajmujących się badaniami behawioralnymi w nurcie „nudge” na całym świecie, a ich liczba przekroczyła już kilka lat temu 200 podmiotów (Hallsworth i in., 2018). Było to podyktowane również wzrostem funduszy przeznaczanych przez sektor publiczny na badania i wdrożenia interwencji reklamowanych jako „tanie, proste i efektywne” w swojej aplikacji (Benartzi i in., 2017). Jednakże, w wyniku znaczącego wzrostu wysiłków w tym obszarze (Sunstein i Reisch, 2023) i funduszy (Benartzi i in., 2017), zaobserwowano szereg problemów szczególnie związanych z interwencjami behawioralnymi typu nudge. Do czterech najczęściej poruszanych aspektów należą: 1) aspekty etyczne oraz transparentność interwencji behawioralnych typu nudge (Bruns i in., 2018; Rebonato, 2014), 2) brak uwzględniania czynników psychologicznych takich jak różnice indywidualne w badaniach nad interwencjami behawioralnymi (Marchiori i in., 2017), 3) stosunkowo mała liczba systematycznych badań porównujących siły efektów interwencji, które w zależności od przyjętych metod analitycznych charakteryzujących się różną wrażliwością prowadzą do różnych wyników (Mertens i in., 2022; Maier i in., 2022), 4) kompromis pomiędzy oszczędnością testowania interwencji a trafną oceną efektywności interwencji behawioralnych w warunkach laboratoryjnych przed wdrożeniem interwencji. Niniejsza praca doktorska porusza wszystkie cztery przedstawione aspekty i stara się na nie odpowiedzieć formułując przedstawione poniżej pytania badawcze i hipotezy:

1. **Pyt. 1:** Jakie różnice istnieją pomiędzy oddziaływaniami typu boost i nudge?
 - a. **H1.1:** Istnieją różnice w sile wpływu oddziaływania interwencji typu boost i nudge na decyzje w ryzyku i niepewności

2. **Pyt. 2:** Jakie czynniki moderują różnice w oddziaływaniach typu boost i nudge?
 - a. **H2.1:** Istnieją różnice w sile wpływu oddziaływania interwencji typu boost i nudge w domenie zysków
 - b. **H2.2:** Istnieją różnice w sile wpływu oddziaływania interwencji typu boost i nudge w domenie strat
 - c. **H2.3:** Różnica w oddziaływaniu boostów i nudge'y jest moderowana przez przejrzystość (transparentność) interwencji dla osób badanych
 - d. **H2.4:** Różnica w oddziaływaniu boostów i nudge'y jest moderowana przez trwałość interwencji w czasie

3. **Pyt. 3:** Czy łączenie interwencji typu nudge i boost, będzie charakteryzowało się wyższą skutecznością niż efekty osiągnane za pomocą osobnych interwencji?
 - a. **H3:** Istnieją różnice w sile wpływu oddziaływania pojedynczych interwencji behawioralnych i interwencji łączonych

Dodatkowo, ważnym elementem projektu będzie także analiza różnic indywidualnych jako potencjalnych moderatorów działania interwencji nudge i boost. W związku z brakiem jednoznacznej teorii dotyczącej tego jak różnice indywidualne zmieniają efektywność nudge'y i boostów, aspekt różnic indywidualnych jest traktowany w tym projekcie eksploracyjnie, bez formułowania jednoznacznych hipotez.

W celu odpowiedzi na powyższe pytania i weryfikacji postawionych hipotez przeprowadzono cztery badania empiryczne i metaanalizę. Celem badania 1, meataanalizy, było

znalezienie odpowiedzi na pytanie badawcze 1, poprzez porównanie oszacowanych sił efektów dla dwóch kategorii nudge i boost. Mając na uwadze wyniki metaanalizy dotyczącej interwencji typu boost (Mertens i in., 2022) i krytykę obejmującą potencjalny efekt publikacyjny, po którego uwzględnieniu wyniki interwencji behawioralnych zanikają (Maier i in., 2022), w badaniu pierwszym ważnym elementem było zastosowanie analizy potencjalnego efektu publikacyjnego.

Badanie 2 zostało zaplanowane tak aby sytuacją decyzyjną odpowiadać „małemu światu” wyróżnianego w teorii Savage’a (1954). Zastosowana procedura loterii, umożliwia obliczenie ryzyka stojącego u podstaw każdego z wyborów, w postaci różnej wartości oczekiwanej. Procedura loterii została zaadaptowana z badań Franklina i in. (2019), którzy jako jedni z pierwszych porównywali interwencje typu nudge i boost w wyborach finansowych obarczonych ryzykiem. Celem dwuetapowego Badania 2 była odpowiedź na pierwsze i drugie pytanie badawcze. Szczególnymi elementami planu badawczego uwzględnionymi w Badaniu 2 były transparentność interwencji, ich trwałość w czasie i różnice indywidualne.

Schemat badania 3 został zaplanowany tak aby sytuacja decyzyjna była bardziej zbliżona do „dużego świata”, w którym ryzyko wystąpienia efektów naszych decyzji jest nieznane lub niejawne (Gigerenzer i in., 2022). Jako przykładowa sytuacja decyzyjna z „dużego świata” została tutaj wybrana sytuacja wyboru programu oszczędnościowego, opisana jako wybór wieloatrybutowy. Za pomocą procedury dyskretnego wyboru (rodzaj procedury badającej decyzje wieloatrybutowe, ang. *Discrete Choice Experiment*) zostały przebadane interwencje nudge i boost, które miały na celu zmianę wagi atrybutów decyzji o oszczędzaniu. Wśród atrybutów decyzyjnych znalazły się aspekty programu Save More Tomorrow, dotychczas nie przebadane za pomocą procedury dyskretnego wyboru, a podkreślane jako kluczowe dla sukcesów programu m.in.: waga odroczonego startu oszczędzania, zaplanowana podwyżka czy

wysokość własnej raty miesięcznej. W badaniu 3 został również uwzględniony aspekt ważny dla rynku programów oszczędnościowych w Polsce, czyli potencjalne zaangażowanie podmiotów publicznych w plany oszczędnościowe. Wybór sytuacji oszczędnościowej nawiązującej do sytuacji wykorzystywanej przy okazji programu Save More Tomorrow, miał na celu porównanie interwencji nudge i boost w warunkach laboratoryjnych, możliwie zbliżonej pod kątem wagi atrybutów do realnej sytuacji decyzyjnej. Badanie 3 starało się odpowiedzieć na pytanie, które interwencje będą bardziej efektywne w danej sytuacji decyzyjnej (pyt. 1) oraz czy łączenie interwencji nudge i boost może prowadzić do efektów synergii (pyt. 3). W charakterze eksploracyjnym, uwzględniało również szereg różnic indywidualnych. Badanie 3 stosujące podejście wieloatrybutowe w badaniach nad interwencjami behawioralnymi jest jedną z pierwszych takich prób szczególnie w kontekście zachowań finansowych. Nieliczne, dotychczasowe próby implementacji interwencji behawioralnych w badaniach nad decyzjami wieloatrybutowymi nie dotyczyły decyzji ryzykownych w domenie finansów oraz ograniczały się jedynie do interwencji typu nudge (Gottselig i in., 2023; Hazlewood i in., 2020; Ouvrard i in., 2020). Dodatkowo, zastosowanie metodologii badającej decyzje wieloatrybutowe, może końcowo doprowadzić do zmniejszenia kosztów testowania interwencji poprzez osiągnięcie dokładniejszych danych na temat preferencji osób badanych i ich zachowania w obliczu realnych atrybutów opisujących daną sytuację decyzyjną. W kontekście prac Thaler i Benartziego (2004) nad programem Save More Tomorrow, proponowane w niniejszej pracy interwencje nudge i boost są preskryptywnymi metodami wspomagającymi wybór, zbliżającymi do wyborów nakreślalnych przez teorie normatywne lub wyborów racjonalnych ekologicznie w danej sytuacji decyzyjnej.

2. CZĘŚĆ EMPIRYCZNA

Część empiryczna jest podzielona na trzy bloki tematyczne: metaanalizę badań skupiających się na porównaniu interwencji typu nudge i boost (Badanie 1A), oraz dwa bloki tematyczne zawierające po dwa badania empiryczne. Badanie 2A i 2B testowało interwencje behawioralne w warunkach ryzyka oraz uwzględniało czynniki sytuacyjne transparentności i upływu czasu dla trwałości efektów interwencji. Z kolei Badanie 3A i 3B badało efektywność interwencji behawioralnych w zmianie zachowania w niepewności, na przykładzie zachowań oszczędnościowych. Ważnym elementem badania 3B był również potencjalny efekt synergii pomiędzy interwencjami nudge i boost. Część empiryczna kończy się dyskusją ogólną, podsumowującą uzyskane wyniki w świetle dostępnej literatury i dotychczasowych badań.

2.1 BADANIE 1 *Metaanaliza badań porównujących efektywność interwencji nudge i boost w zmianie zachowania*

Celem badania była ocena skuteczności zmiany zachowania dwóch rodzajów interwencji behawioralnych zaliczanych do kategorii boostów i nudge'y. Badanie miało charakter metaanalizy i zostało przeprowadzone w październiku 2021 roku. Oszacowanie efektywności badanych interwencji zostało przeprowadzone w oparciu o miarę wielkość efektu *d*-Cohena. Definicje, charakterystyka interwencji oraz taksonomie, zostały zaadaptowane z badań Hertwiga i Grüne-Yanoff (2017) oraz Sunsteina (2014). Zgodnie z wiedzą autora, jest to pierwsza metaanaliza porównująca ze sobą te dwa rodzaje interwencji behawioralnych. Przegląd badań obejmował artykuły, które zostały opublikowane do końca października 2021 roku. Na podstawie wyników metaanalizy zostały wykonane analizy mocy, w celu oszacowania wielkości próby osób badanych w Badaniach 2A i 2B.

2.1.1 Metoda

Przegląd literatury i selekcja materiału do badania

Artykuły były wyszukiwane poprzez bazy Web of Science, EBSCO oraz Google Scholar. Ze względu na to, że koncepcje nudgingu i boostingu powstały kolejno w 2008 i 2016 roku (Hertwig i Grüne-Yanoff, 2017; Thaler i Sunstein, 2008), kwerenda artykułów objęła swoim zasięgiem początek 2008 roku i kończyła się w październiku 2021 roku. Używane słowa kluczowe obejmowały następujące frazy wraz z ich odmianami: boosting, boost, nudging, nudge, behavioral interventions. W wyniku wyszukiwania znaleziono 1352 artykuły, z których następnie wyselekcjonowano 671 artykułów odnoszących się w pierwszym kroku do dyscypliny psychologii i kolejno w drugim kroku do interwencji behawioralnych typu nudge i boost. Następnie, ze względu na charakterystykę literatury powstającej wokół interwencji behawioralnych, która obejmuje wiele tekstów teoretycznych, wyodrębnione zostały artykuły empiryczne. W kolejnym kroku, wzięto także pod uwagę artykuły cytujące wyodrębnione artykuły empiryczne. Końcowa baza zawierała 138 artykułów odnoszących się do nudge'y i 7 artykułów odnoszących się do boostów.

Pierwszym pytaniem badawczym postawionym w niniejszym projekcie było „Jakie różnice istnieją pomiędzy oddziaływaniami typu boost i nudge?”. W związku z tym, w metaanalizie skoncentrowano się na badaniach porównujących bezpośrednio te dwie interwencje oraz badaniach, które testowały jedną z dwóch kategorii w podobnej domenie. W ten sposób wyodrębniono 7 artykułów odnoszących się do nudge'y i 7 artykułów odnoszących się do boostów. Z wybranych artykułów wyselekcjonowano 15 warunków badawczych w których badano wpływ nudge'y na zachowanie i 17 warunków badawczych w których badano wpływ boostów. W sumie zidentyfikowano 32 pojedynczych warunków badawczych. Następnie, dane zostały pozyskane z artykułów, dokonując niezbędnych przeliczeń na wskaźnik *d*-Cohena.

Analiza danych

Siły efektów z poszczególnych badań zostały sprowadzone do jednej statystyki w postaci współczynnika *d*-Cohena (Rycina 1). Następnie, średni efekt populacyjny został obliczony dla: wszystkich zebranych interwencji, oddzielnie dla interwencji typu nudge, oddzielnie dla interwencji typu boost. Ostatnim etapem analiz, został obliczony potencjalny efekt tendencyjności publikacyjnej. Obliczenia zostały wykonane za pomocą paczki *metafor* w programie RStudio (Viechtbauer i Viechtbauer, 2015). Średni efekt populacyjny został wyestymowany metodą ograniczonej maksymalnej wiarygodności z użyciem modelu efektów losowych (ang. *restricted maximum likelihood model*; REML; Viechtbauer, 2010). Model efektów losowych został wybrany ze względu na specyfikę badanej domeny. Jak wskazuje Hedges i Vevea (1998), w naukach społecznych stosowanie modelu efektów losowych jest najbardziej adekwatnym rozwiązaniem, gdyż taki zabieg pozwala na wnioskowanie o rozkładzie danych ogólnej populacji, na co model efektów losowych bierze dodatkową poprawkę. Dodatkowo, warto zaznaczyć, że reguła kciuka, którą sugerują Jackson i Turner (2017), oznacza, że w większości przypadków, wymagana minimalna liczba obserwacji odnoszących się do sił efektu, w celu przeprowadzenia metaanalizy dla modelu efektów losowych, wynosi pięć.

2.1.2 Wyniki

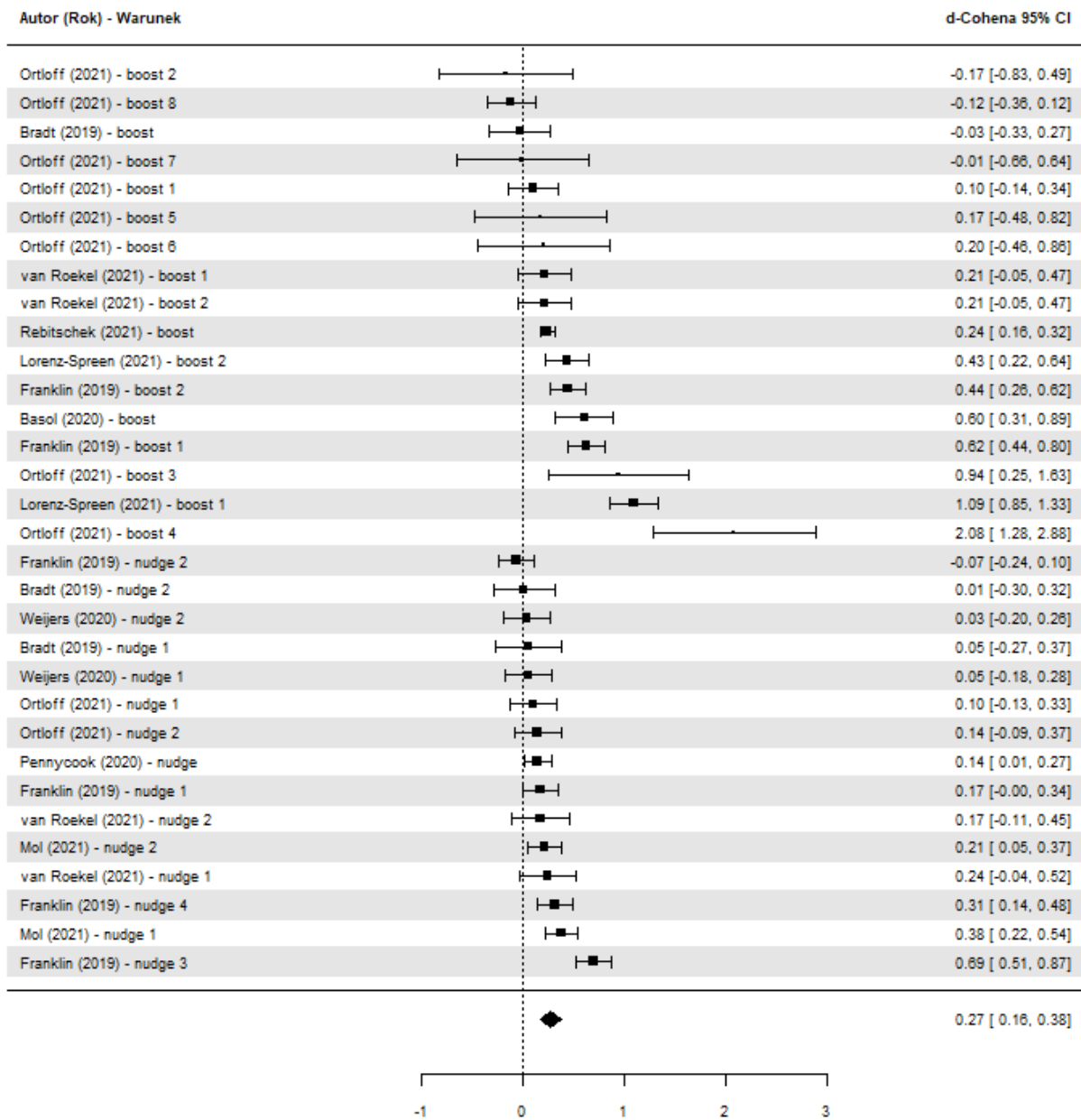
Średni efekt populacyjny dla wszystkich interwencji

W analizie uwzględniono łącznie $k = 32$ warunków badawczych, pochodzących z . Zaobserwowane wartości *d*-Cohena wahały się od -0,17 do 2,07, przy czym większość szacunków była dodatnia (84%). Szacowana średnia wartość wyniku na podstawie modelu z efektami losowymi wynosiła $\hat{\mu} = 0,27$ (95% CI: 0,16 do 0,37). Zatem średni wynik różnił się znacząco od zera ($z = 4,94$, $p < 0,001$). Wykres typu „forest” przedstawiający zaobserwowane

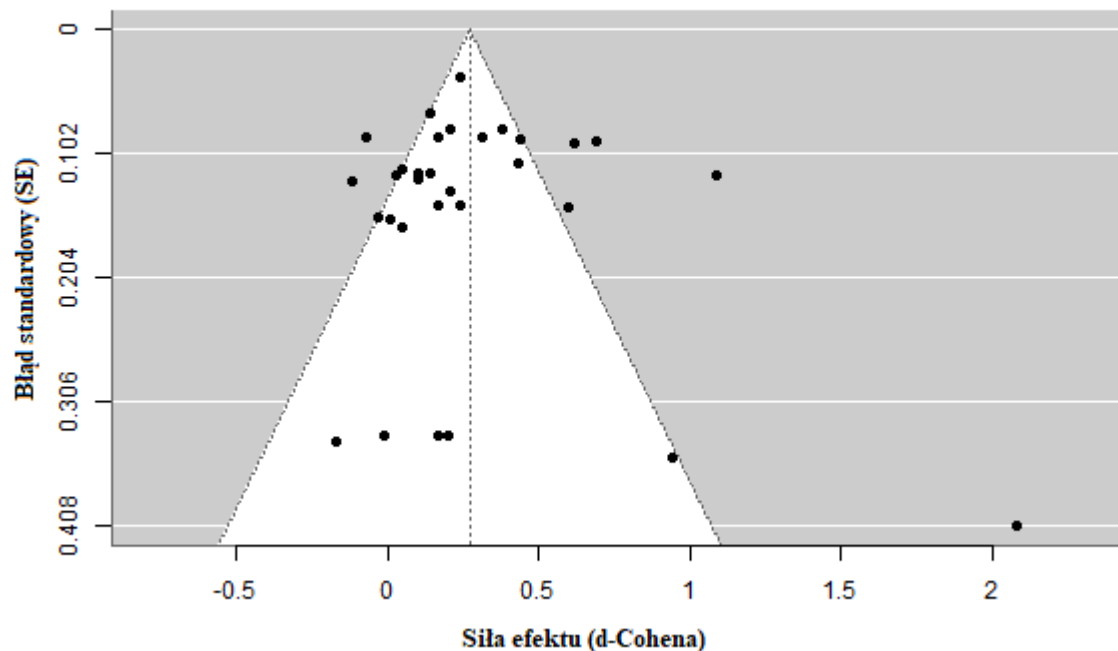
wyniki oraz szacunek na podstawie modelu z efektami losowymi został przedstawiony na Rycinie 1.

Według testu Q, wyniki okazały się heterogeniczne ($Q(31) = 173,66$, $p < 0,001$, $\tau^2 = 0,07$, $I^2 = 86,34\%$), co wskazuje na różny efekt zebranych interwencji i jest argumentem za oddzielną analizą dla nudge'y i boostów lub analizą z moderatorem. Test ten służy do sprawdzenia, czy zaobserwowane różnice w wynikach między badaniami są większe niż te, które można by oczekiwać w wyniku samego przypadku (Higgins i Thompson, 2002). Przedział predykcyjny na poziomie 95% dla prawdziwych wyników mieścił się w zakresie od -0,26 do 0,81. Stąd, chociaż średni wynik jest szacowany jako dodatni, w niektórych badaniach prawdziwy wynik może być faktycznie ujemny.

Analiza reszt ujawniła, że dwa trzy siły efektów, Lorenz-Spreen i in. (2021), Ortloff (2021) boost nr. 1 oraz Ortloff (2021) boost nr. 4, miały wartości większe niż $\pm 3,1628$ i mogą być potencjalnymi wartościami odstającymi w kontekście tego modelu. Zgodnie z odległościami Cooka, jedno badanie (Lorenz-Spreen (2021) - boost 1) mogłoby być uznane za nadmiernie wpływowe. Końcowym etapem analizy była kwestia wykrycia potencjalnego efektu publikacyjnego. Wykres lejkowy dla wszystkich obserwacji przedstawiono na Rycinie 2. Ani test korelacji rangowej, ani test regresji nie wykazały żadnej asymetrii wykresu lejkowego (odpowiednio $p = 0,509$ i $p = 0,244$; Viechtbauer i Cheung, 2010).



Rycina 1. Wykres typu „forest”. Przedstawia wyniki interwencji behawioralnych nudge i boost z ich szacowaną wielkością efektu d Cohena i odpowiadającymi im 95% przedziałami ufności (dane zostały posortowane względem siły efektu osobno dla interwencji typu boost i typu nudge). Każdy wiersz oznacza inny warunek badawczy i jego wynik względem grupy kontrolnej (przykładowo, oznaczenie w pierwszym wierszu *boost 2* oznacza warunek drugi z badania Ortloff (2021)). Ujemne siły efektu wskazują na odwrotny kierunek efektu interwencji. Średni efekt ogólny wyniósł 0,27 [0,16; 0,38].



Rycina 2. Wykres lejkowy dla wszystkich badań uwzględnionych w metaanalizie

Obliczenie średnich efektów interwencji nudge i boost.

W celu zbadanie potencjalnego efektu moderacyjnego oraz porównania ze sobą dwóch rodzajów interwencji, zbudowano model uwzględniający czynnik losowy publikacji oraz informacje o rodzaju interwencji jako moderatora. Do dopasowania modelu wykorzystano podobnie jak wcześniej metodę REML. Model nie wykazał istotnie lepszego dopasowania niż model bazowy ($p = 0,188$) jak zostało to przedstawione w tabeli 4. Jednocześnie trudno go jednoznacznie przedstawić jako gorsze rozwiązanie – choć AIC i BIC jest nieco wyższe to została zanotowana poprawa pod kątem wskaźnika logLikelihood. W związku z niejednoznacznymi wynikami, zdecydowano o ocenie pozostałych aspektów modelu alternatywnego. Test Cochra z poprawką dla jednego moderatora wskazał na brak podstaw do

odrzućcia hipotezy o heterogeniczności co wskazywało na brak dodatkowej wariacji między badaniami poza tą objętą przez moderatory $Q_E(30) = 24,39$, $p = 0,754$, $\tau^2 < 0,001$). Test moderatorów dla dwóch zmiennych (interwencja nudge i interwencja boost) był istotny statystycznie ($F(2, 30) = 10,94$, $p < 0,001$), co sugeruje znaczący wpływ tych zmiennych na wyniki. Współczynnik dla interwencji boost wyniósł 0,36 (SE = 0,09, $t(30) = 4,09$, $p < 0,001$, 95% CI [0,18, 0,54]), a dla interwencji nudge 0,19 (SE = 0,08, ($t(30) = 2,28$, $p = 0,030$, 95% CI [0,02, 0,36]) wskazując, że obydwie grupy interwencji wykazywały się średnim (boosty) lub małym (nudge) efektem według nomenklatury Cohena (1992). Zastosowany kontrast wskazał na brak istotnych różnic pomiędzy dwoma interwencjami -0,17 ($t(30) = 0,12$, $p = 0,165$, 95% CI [-0,42, 0,08])

Tabela 4. Porównanie metaanalitycznego modelu bazowego i modelu z moderatorem

Model	AIC	BIC	DF	logLik	LRT	Istotność (p)
Model bazowy	25,36	28,22	2	-10,68		-
Model Alternatywny (zmienna moderująca)	25,63	29,83	3	-9,81	1,72	0,188

Adnotacja: LRT – Likelihood Ratio Test

2.1.3 Podsumowanie

Przeprowadzona metaanaliza objęła swoim zakresem 32 warunki badawcze, które testowały efekty interwencji behawioralnych nudge i boost. Dane statystyczne zostały wyodrębnione z artykułów naukowych opublikowanych do końca października 2021 roku.

Otrzymane wyniki, wskazują na statystycznie istotną, ale niską skuteczność nudge'y i boostów, ujętych jako jedną kategorię interwencji behawioralnych ($d = 0,27$). Wedle nomenklatury przyjętej przez Cohena (1992), wartości wskaźnika d od 0,2 do 0,3 oznaczają małą siłę efektu.

W drugiej części analiz, przeprowadzono analizę moderacji, która ujawniła istotny efekt kategorii interwencji ujętej jako moderatora, co oznacza, że przynależność do danej kategorii interwencji, tłumaczyła zróżnicowanie zebranych danych. Model uwzględniający czynnik moderacyjny, został oceniony jako lepiej dopasowany do danych od modelu bazowego. Analiza kontrastów wykazała, że nie było istotnych różnic pomiędzy nudge'ami i boostami, choć interwencje typu boost ($d = 0,36$) odnotowały prawie dwa razy silniejszy efekt niż interwencje typu nudge ($d = 0,19$). Najwyższa siła efektu w kategorii boost, została odnotowana w badaniach Ortloff i in. (2021; $d = 2,08$), a w kategorii nudge w badaniach Franklina i in. (2019; $d = 0,69$). W przypadku badań Ortloff i in (2021) interwencja boost bazowała na przekazaniu heurystyki wspierającej monitorowanie środowiska internetowego w celu ochrony danych osobowych, a z kolei najbardziej skuteczna interwencja nudge stosowana przez Franklina i in. (2019) wykorzystywała deskryptywną normę społeczną (nudge normatywny). Warto zaznaczyć, że drugie badanie z największą siłą efektu w kategorii nudge, również wykorzystywało interwencję typu nudge normatywny. Wykonane analizy nie wykazały potencjalnej tendencji publikacyjnej.

2.1.4 Dyskusja

Przedstawiona metaanaliza wykazała, że średnia siła efektu dla interwencji behawioralnych ujętych jako jedna kategoria interwencji, osiągnęła małą wielkość efektu. Jak wskazują Thaler i Sunstein (2008) oraz Hertwig i Grune-Yanoff (2017), nawet niewielkie istotne efekty mogą mieć znaczące znaczenie w przypadku stosowania interwencji w politykach

publicznych. Według Thaler i Sunsteina (2008), nawet niewielki efekt może oznaczać znaczącą zmianę zachowania wśród kilkuset tysięcy osób, jeśli interwencja zostanie zaimplementowana np. na populacji całego kraju. Dodatkowo, na uwagę zasługuje różne siły efektów dla kategorii interwencji: w przypadku interwencji typu nudge średnia siła efektu wynosiła $d = 0,19$, a dla interwencji boost $d = 0,36$. Tym samym, wynik analiz wskazał na blisko dwukrotnie większą siłę efektów boostów. Jest to bardzo obiecujący efekt, który może uzasadniać potrzebę ponoszenia wyższych kosztów wdrożeniowych interwencji typu boost. Zazwyczaj, potencjalny koszt interwencji boost jest związany z wyższym wysiłkiem poznawczym, pojawiającym się podczas zdobywania nowych kompetencji. Jednakże, gdy średnia siła efektu dla boostów wynosi $d = 0,36$, jest to już wynik bliski poziomowi określanego przez Cohena jako „średni”. Wskaźnik d Cohena równy $0,36$, oznacza, że 64% grupy „interwencyjnej” będzie powyżej średniej grupy kontrolnej (U_3 Cohena), 86% przypadków z grupy interwencyjnej i kontrolnej będzie się pokrywać oraz istnieje 60% szans, że osoba wybrana losowo z grupy interwencyjnej będzie miała wyższy wynik niż osoba wybrana losowo z grupy kontrolnej (Magnusson, 2024). Równocześnie, należy mieć na uwadze, że zebrane interwencje pochodzą z bardzo różnorodnych obszarów. Analizowane dane dotyczyły podejmowania decyzji w obszarze finansów (hipotetyczne wybory finansowe; wybory dot. ubezpieczenia przed powodzią), obszarze zdrowia (decyzje o szczepieniach; decyzje o podążaniu za wytycznymi higieny w związku z pandemią COVID-19), obszarze nowych technologii i internetu (zachowanie zgodne ze sponsorowanymi reklamami; decyzje dot. ochrony prywatności w internecie; rozpoznawanie fake-newsów). Niestety, z uwagi na nowatorską dyscyplinę i niewielką liczbę badań w każdym z obszarów, nie było możliwe by zbadać średnią siłę efektów wewnątrz tych obszarów.

Wyniki prezentowanej metaanalizy wskazały, że interwencje typu boost odnotowywały dwukrotnie wyższą siłę efektu niż interwencje typu nudge. Jest to tym bardziej obiecujące w porównaniu do dotychczasowych metaanaliz uwzględniających jedynie interwencje nudge (Hummel i Maedche, 2019; Mertens i in., 2022) lub interwencje behawioralne jako ogólną, homogeniczną kategorię (Nisa i in., 2019). Choć oryginalne badania Mertens i in. (2022) wskazywały na średnią siłę efektu ($d = 0,43$), to po nałożeniu na wyniki poprawki wynikającej z tendencji publikacyjnej, efekt się zmniejszył (Szaszi i in., 2022), wahając się od braku efektu ($d = 0,04$, przy założeniu bardzo konserwatywnego podejścia; Maier i in., 2022) przez bardzo małą siłę efektu ($d = 0,08$; Maier i in., 2022; Szaszi i in., 2022) do średniej siły efektu ($d = 0,31$, przy założeniu najmniej konserwatywnego podejścia; Szaszi i in., 2022). Z kolei badania Nisy i in. (2019), poświęcone metaanalizie nudge'y i innych interwencji behawioralnych w domenie środowiskowej, również wskazały na istotną, ale stosunkowo małą siłę efektu ($d = 0,09$) podobną do badań Szaszi i in. (2022). Warto zaznaczyć, że w przypadku Mertens i in. (2022) zgromadzone dane obejmowały wyniki z ponad 440 warunków badawczych, umożliwiając przeprowadzenie analiz moderacyjnych uwzględniających subdomeny interwencyjne. Badanie Mertens i in. (2022) wykazało, że siła oddziaływania zależy od tego w jakiej domenie interwencja występuje i jaki rodzaj nudge'a był wykorzystywany. Największe siły efektów notowano w domenie środowiska ($d = 0,43$) i zdrowia ($d = 0,65$), a wśród rodzajów interwencji w przypadku nudge'y opcji domyślnej ($d = 0,62$; Mertens i in., 2022). Z uwagi na znacznie mniejszą próbę warunków badawczych ($k = 32$), w przedstawionym autorskim Badaniu 1, analiza takich moderatorów nie była możliwa. Warto natomiast zauważyć, że zarówno wyniki Badania 1 jak i przywołanych badań Mertens i in. (2022) oraz Nisy i in. (2019) **wskazują na heterogeniczność oddziaływania interwencji behawioralnych**. W zależności od wyboru interwencji i jej odpowiedniego dopasowania do środowiska decyzyjnego, architekt decyzji może

uzyskać wyniki wahające się od braku efektu do nawet średniej siły efektu stosowanego oddziaływania.

W kontekście niniejszej rozprawy, wyniki metaanalizy odgrywają ważną funkcję informatywną dla kolejnych etapów projektu: wyboru interwencji behawioralnych do dalszych badań oraz pozwalają na oszacowanie prób badawczych w ramach analizy mocy. Wyniki metaanalizy wskazują na rodzaj interwencji, które powinny być w pierwszej kolejności testowane w kolejnych badaniach. Rezultaty odnoszą się przede wszystkim do rodzajów nudge'y, które na tle pozostałych interwencji ze swojej kategorii okazały się najbardziej obiecujące. Była to kategoria nudge'y normatywnych (Franklin i in. 2019, Mola i in., 2021), opierająca się na wykorzystaniu deskryptywnej normy społecznej. Drugim rodzajem interwencji nudge zauważalnym w powyższym zestawieniu był nudge dyskluzywny, opierający się na odpowiednim upraszczaniu informacji i podkreślaniu kilku najważniejszych aspektów środowiska informacyjnego. Z kolei w kategorii boostów, wiele interwencji notujących wyższe siły efektów, opierały swoje działanie na przekazywaniu reguły proceduralnych takich jak szybkie i oszczędne drzewa decyzyjne (ang. *Fast and Frugal trees*) i prostych heurystyk umożliwiających analizę sytuacji decyzyjnej. **Należy również nadmienić, że wyniki metaanalizy w postaci uśrednionego efektu *d*-Cohena, posłużyły do oszacowania wielkości próby w ramach analizy mocy dla Badań 2A i 2B.**

2.2 Badania 2A i 2B – zarys ogólny

2.2.1 Metoda ogólna badań 2A i 2B

Celem Badań 2A i 2B, było odpowiedzenie na Pyt. Badawcze 2: **Jakie czynniki moderują różnice w oddziaływaniach typu boost i nudge?** Cel ten był realizowany poprzez replikację efektów otrzymanych przez Franklina i in. (2019) dotyczących efektu sformułowania

(Hipotezy H2.1 i H2.2), oraz zweryfikowanie, czy efekty te utrzymają się mimo (1) świadomości oddziaływania interwencji behawioralnej (Hipoteza H2.3), (2) trwałości interwencji w czasie (Hipoteza H2.4) oraz 3) czy efekty interwencji będą moderowane przez różnice indywidualne osób badanych (komponent eksploracyjny). Wśród najistotniejszych rezultatów badań Franklina i współpracowników, które starano się zreplikować w niniejszym projekcie, należy wymienić wyższą skuteczność boostów w domenie zysków oraz o wyższej skuteczności nudge'y dyskluzyjnych w domenie strat. Dodatkowo, osoby, które dokonywały wyborów o wyższej wartości oczekiwanej (ang. *expected value*; EV) podczas fazy przed-interwencyjnej, dokonywały również wyborów o wyższej EV w fazie po-interwencyjnej. Co więcej, efekt ten utrzymywał się zarówno w domenie strat jak i zysków (Franklin i in., 2019). **Dla przejrzystości przedstawianych wyników, w niniejszej pracy procedura zaczerpnięta z badań Franklina i in. (2019) jest określana skrótem „FFR” od nazwisk autorów: Franklina, Folke i Ruggeri’ego.**

W celu wyizolowania efektów badanych czynników, aspekt świadomości oddziaływania interwencji został sprawdzony w ramach badania 2A, a wpływu czasu był przedmiotem badania 2B. Oprócz sposobów pomiaru odpowiedzi zaadaptowanych z Franklina i in., (2019), zostały również włączone kwestionariusze pomiaru cech osobowościowych. W celu potwierdzenia związku cech osobowościowych z interwencjami behawioralnymi wskazywanymi w dotychczasowych badaniach (Ingendahl i in., 2021; Kamilçelebi i Songül, 2020; Warberg i in., 2019), wprowadzone następujące kwestionariusze: FCZ-KT (R) (Cyniak-Cieciura i in. 2017), Kwestionariusz Potrzeby Poznania (Matusz i in., 2011) oraz TIPI (ang. *Ten-Item Personality Inventory*; Sorokowska i in., 2014). W celu weryfikacji czy zdolności poznawcze mogą moderować efektywność interwencji (jak w badaniach Caballero i Plonera, 2022), wprowadzono

Berliński Test Numeryczny (ang. *Berlin Numeracy Test*; Cokely i in., 2016). Dodatkowo, w badaniu został włączony alternatywny pomiar zmiennej zależnej, stworzony w oparciu o procedurę Probability Discounting Questionnaire (PDQ; Madden i in., 2009). W tym alternatywnym pomiarze zmiennej zależnej, pozycje uwzględniające kwoty oraz skalę ryzyka zostały zaczerpnięte bezpośrednio z PDQ. Celem tego kroku, było zweryfikowanie efektów osiągniętych przez Franklina w procedurze opartej na loteriach (dalej zwaną procedurą FFR), które systematycznie będą różnić się wielkością kwot i ryzykiem. W oryginalnych badaniach, wybór zastosowanych wielkości kwot i wartości ryzyka wydaje się arbitralny, a Franklin i in. (2019) nie przedstawiają konkretnej motywacji takich decyzji. Dodatkowo, liczba pozycji, a tym samym wyborów w procedurze FFR była niewielka – cztery pozycje w procedurze dla strat i cztery pozycje w procedurze dla zysków. W świetle ostatnich badań Lejarraga i Hertwig (2021), liczba decyzji, które osoby badane muszą podejmować może mieć znaczenie na badany efekt, dlatego w niniejszym badaniu zaplanowano w związku z tym drugi pomiar w oparciu o procedurę PDQ liczącą w sumie 60 pozycji decyzyjnych. Choć procedura PDQ oryginalnie została stworzona w celu badania tempa dyskontowania, w tym badaniu była traktowana jako systematyczny zbiór loterii pieniężnych o różnej wartości oczekiwanej.

Do badań 2A i 2B wybrano trzy interwencje behawioralne, które zostały następnie zaimplementowane w loteriach finansowych: procedurze FFR (Franklin i in., 2019) oraz procedurze PDQ (Madden i in., 2009). Wśród wybranych interwencji znalazła się jedna interwencja boost i dwie interwencje nudge: nudge dyskluzyjny oraz nudge normatywny. Końcowy kierunek działania każdej z interwencji był jednakowy – jeśli osoba badana podążała za opcją, która została desygnowana przez interwencję, miała możliwość wyboru opcji o wyższej wartości oczekiwanej (ang. *expected value*; EV), zwanej tutaj opcją optymalną (za Franklin i in.,

2019). O ile interwencje nie różniły się kierunkiem w którym zachęcały osoby badane do podążania, różniły się schematem działania. W przypadku interwencji typu boost, osoby badane na samym początku przechodziły przez krótką fazę treningową – dowiadywały się o tym czym jest wartość oczekiwana oraz jaką heurystykę mogą zastosować do jej obliczenia. Następnie, dopiero po przejściu trzech loterii treningowych i poprawnym wyborze opcji o wyższej EV były przepuszczane do właściwego pomiaru eksperymentalnego. W pomiarze eksperymentalnym w grupie boost, pozycje loterii nie zawierały żadnych dodatkowych informacji (Rycina 3, panel C) – osoby badane miały za zadanie korzystając z heurystyki obliczania EV dokonać wyboru. Dozwolone było korzystanie z kalkulatora.

Schemat działania interwencji nudge'a dyskluzywnego i nudge'a normatywnego był nieco odmienny od interwencji boost: w przypadku obydwu nudge'y, osoby badane widziały pod każdą loterią dodatkową informację, która miała na celu je lepiej nakierować na opcję o wyższym EV. W przypadku nudge'a dyskluzywnego, była to dokładna informacja o tym jaka jest EV dla każdej z opcji. Warto nadmienić, że tych osób nie informowano czym jest EV. Przykładowa pozycja eksperymentalna z grupy nudge'a dyskluzywnego jest na Rycinie 3 (panel A). Z kolei interwencja nudge'a normatywnego była oparta na prezentacji dodatkowej informacji o tym, co rzekomo wybierała większość osób badanych. W ten sposób, interwencja wykorzystywała deskryptywną normę społeczną, która opisywała zachowanie większości osób badanych (Sunstein, 2014). Informacja o wyborze „Większości” towarzyszyła zawsze opcji o wyższej wartości oczekiwanej. Wygląd pozycji eksperymentalnych w grupie nudge'a normatywnego został przedstawiona w panelu B Ryciny 3.

A) Masz 230 złotych. Które z poniższych rozwiązań wybierasz?

100% szansy na utratę 30 zł Przewidywana wartość oczekiwana tej opcji to 200 zł <input type="radio"/>	75% szansy na zachowanie wszystkiego lub 25% szansy na utratę wszystkiego Przewidywana wartość oczekiwana tej opcji to 172,5 zł <input type="radio"/>
---------------------------------------------------------------------------------------------------------------------	---------------------------------------------------------------------------------------------------------------------------------------------------------------------

B) Co wybierasz?

100% szansy na wygranie 200 zł <input type="radio"/>	75% szansy na wygranie 303,3 zł lub 25% szansy na wygranie 0 zł Większość uczestników badania wybierało tę opcję <input type="radio"/>
-------------------------------------------------------------	------------------------------------------------------------------------------------------------------------------------------------------------------

C) Masz 690 zł. Który z poniższych wariantów wolisz?

100% szansy na utratę 490 zł <input type="radio"/>	25% szansy na zachowanie wszystkiego lub 75% szansy na utratę wszystkiego <input type="radio"/>
-----------------------------------------------------------	--------------------------------------------------------------------------------------------------------

Rycina 3. Przykładowe pozycje z procedury FFR w poszczególnych grupach

eksperymentalnych. Panel A: Nudge dyskluzyjny; loteria w domenie strat; Panel B: Nudge normatywny; loteria w domenie zysków; Panel C: Boost i warunek kontrolny; loteria w domenie strat.

Wybór interwencji do badania odbył się dwutorowo: po pierwsze, wspomniane interwencje były również zastosowane w badaniach Franklin i in. (2019), Folke i in. (2021) oraz Ruggeri i in. (2021). Po drugie, przegląd literatury i dokonana metaanaliza wskazała na szerokie zainteresowanie badaczy, a co za tym idzie również specjalistów ds. polityk publicznych nie tylko nudge'ami i boostami, oferującymi upraszczanie reguł decyzyjnych (jak boost proponujący heurystykę obliczania wartości oczekiwanej czy nudge dyskluzyjny upraszczający format

interwencji) ale również szeroko zakrojone stosowanie interwencji opartych o normy społeczne (Constantino i in., 2022; Mertens i in., 2022).

Podsumowując, celem Badań 2A i 2B, było odpowiedzenie na Pyt. Badawcze 2: **Jakie czynniki moderują różnice w oddziaływaniach typu boost i nudge?** Cel ten był realizowany przede wszystkim poprzez replikację efektów otrzymanych przez Franklina i in. (2019) oraz zbadanie potencjalnych czynników moderujących efektywność interwencji jak ich transparentność i trwałość efektów interwencji w czasie. Dodatkowym komponentem eksploracyjnym była rola różnic indywidualnych w efektywności interwencji behawioralnych. W badaniu 2A zostały przetestowane interwencje z uwzględnieniem przede wszystkim roli transparentności oraz różnic indywidualnych, a w badaniu 2B, interwencje przy udziale wpływu czasu i ponownie tych samych różnic indywidualnych.

2.3 Badanie 2A *Transparentność a efektywność interwencji nudge i boost w zmianie zachowania w decyzjach ryzykownych*

2.3.1 Osoby badane

W ramach przygotowań do Badania 2A, została przeprowadzona analiza mocy dla minimalnej wymaganej wielkości próby, odpowiedniej do wykrycia efektów wieloczynnikowej analizy wariancji z kowariantami (ANCOVA), dla planu 4x2 (4 czynnikowa zmienna grupy interwencyjnej x 2 czynnikowa zmienna transparentności). Poziom alfa ustalono na 0,05, moc na poziomie 0,8, stopnie swobody na 3, a siłę efektu $f = 0,135$. Siła efektu dla przyjętych parametrów została wyliczona ze wzoru $f = d \text{ Cohena} / 2$ (Cohen, 2013) dla średniej siły efektu interwencji behawioralnych otrzymanej w Badaniu 1 (Metaanaliza), która wyniosła $d \text{ Cohena} = 0,27$. Analizy mocy dokonano za pomocą programu GPower, a rekomendowana próba wyniosła 603 osoby.

Zgodnie z analizą mocy, przebadano 685 osoby, z czego 543 kobiet i 127 mężczyzn. Piętnaście osób wolało nie udzielać odpowiedzi na temat swojej płci. Średnia wieku wyniosła 27,73 lat (SD= 8,65). Badanie przestrzegało wytycznych etycznych, zatwierdzonych przez Komisję Etyki Wydziału Psychologii w Warszawie Uniwersytetu SWPS. Rekrutacja osób badanych odbywała się za pośrednictwem systemu SONA Uniwersytetu SWPS, a osoby badane otrzymywały za udział w badaniu uczelniane punkty za aktywność.

2.3.2 Procedura

Badanie 2A, można podzielić na dwie fazy: fazę pierwszą przed-interwencyjną oraz fazę drugą interwencyjną. W pierwszej fazie, osoby badane miały z zadanie udzielić odpowiedzi na trzy pozycje zaczerpnięte z procedury FFR (Franklin i in., 2019), oparte o loterie o różnej wartości oczekiwanej. Podobnie jak w oryginalnych badaniach, wynik w tym pomiarze w analizach został wykorzystany jako predyktor wyników zmiennej zależnej. W pomiarze przed-interwencyjnym, wprowadzono również alternatywną zmienną zależną – procedurę PDQ (Madden i in., 2009). Z procedury PDQ zostało zaczerpniętych w sumie 60 pozycji, po 10 pozycji na każdą grupę: małe straty, małe zyski, średnie straty, średnie zyski oraz duże straty i duże zyski.

Pomiar interwencyjny składał się z pełnej procedury FFR oraz ponownie procedury PDQ z uwzględnionymi interwencjami behawioralnymi. Procedura FFR składała się z ośmiu pozycji, z czego cztery były osadzone w domenę strat a cztery w domenę zysków. By uniknąć efektów torowania, kwestionariusze mierzące różnice indywidualne były prezentowane po pozycjach behawioralnych. Wśród kwestionariuszy mierzących różnice indywidualne znalazły się: FCZ-KT (R) (Cyniak-Cieciura i in. 2017; skale: Żwawość, Perseweratywność, Wrażliwość sensoryczna, Wytrzymałość, Reaktywność emocjonalna i Aktywność, oraz Rytmiczność), Kwestionariusz

Potrzeby Poznania (Matusz i in., 2011), TIPI (ang. *Ten-Item Personality Inventory*; Sorokowska i in., 2014; skale: Otwartość na doświadczenie, Sumienność, Ekstrawersja, Ugodowość, Neurotyczność) oraz Berliński Test Numeryczny (ang. *Berlin Numeracy Test*; Cokely i in., 2016) Na sam koniec osoby badane zostały poproszone o wypełnienie pytań metryczkowych. Badania przeprowadzono w całości online za pomocą platformy Qualtrics (Qualtrics, Provo, UT; qualtrics.com).

2.3.4 Wyniki

Replikacja wyników Franklina i in. (2019) za pomocą procedury FFR. W celu replikacji efektów oryginalnego badania, zaplanowano serię modeli liniowych wyjaśniających liczbę wyborów optymalnych (o wyższym EV, oddzielnie dla domeny zysków i strat:

- 1) Model bazowy, zakładający stałą wartość dla zmiennej zależnej
- 2) Model 1 – dodający do modelu bazowego jako zmienną niezależną wybór przed-interwencyjny
- 3) Model 2 – dodający do modelu 1 zmienną interwencji behawioralnej
- 4) Model 3 – dodający do modelu 2 interakcję między interwencją behawioralną a wyborem przed-interwencyjnym

W celu oceny, czy dodanie kolejnych zmiennych poprawia dopasowanie modelu do danych, przeprowadzono analizę wariancji (ANOVA) dla serii modeli liniowych. W domenie zysków, wyniki sugerowały, że dodanie zmiennej zawierającej wybory przed-interwencyjne do modelu bazowego znacząco poprawiało dopasowanie modelu do danych, ($F(1, 683) = 61,12, p < 0,001$). Dodatkowo, uwzględnienie efektu interwencji behawioralnej w modelu 2 również prowadziło do znaczącej poprawy dopasowania w porównaniu z modelem zawierającym tylko wybór przed-interwencyjny, ($F(3, 680) = 7,90, p < 0,001$). Niemniej jednak, dodanie interakcji

między wyborem przed-interwencyjnym a interwencją behawioralną w modelu 3 nie przyniosło znaczącej poprawy dopasowania w stosunku do modelu 2, ($F(3, 677) = 1,88, p = 0,132$).

Szczegółowe statystyki analizy ANOVA, przedstawiono w tabeli 5.

W domenie strat zaobserwowano podobne zależności, co w domenie zysków: dodanie zmiennej wyboru przed-interwencyjnego do modelu bazowego znacząco poprawiło dopasowanie modelu, ($F(1, 683) = 50,81, p < 0,001$). Uwzględnienie zmiennej interwencji behawioralnej w modelu 2 również przyniosło znaczącą poprawę dopasowania w porównaniu z modelem tylko z zmienną wyboru przed-interwencyjnego, ($F(3, 680) = 3,60, p = 0,013$). Jednakże, dodanie interakcji między wyborem przed-interwencyjnym a interwencją behawioralną w modelu 3 nie prowadziło do znaczącej poprawy dopasowania względem modelu 2, ($F(3, 677) = 1,10, p = 0,347$).

Tabela 5. Porównanie modeli liniowych w domenie zysków i strat (Procedura FFR).

Porównanie modeli	Suma kwadratów reszt	Różnica sumy kwadratów reszt	Statystyka F	Istotność (p)
Zyski				
Model bazowy	1298,4	-	-	-
Model 1 – M bazowy	1195,0	103,405	61,1234	< 0,001
M2 – M1	1154,8	40,116	7,9043	< 0,001
M3 – M2	1145,3	9,528	1,8773	0,132
Straty				
Model bazowy	1099,2	-	-	-

Model 1 – M bazowy	1023,9	75,274	50,8095	< 0,001
M2 – M1	1007,9	15,999	3,5997	0,013
M3 – M2	1003,0	4,902	1,1029	0,347

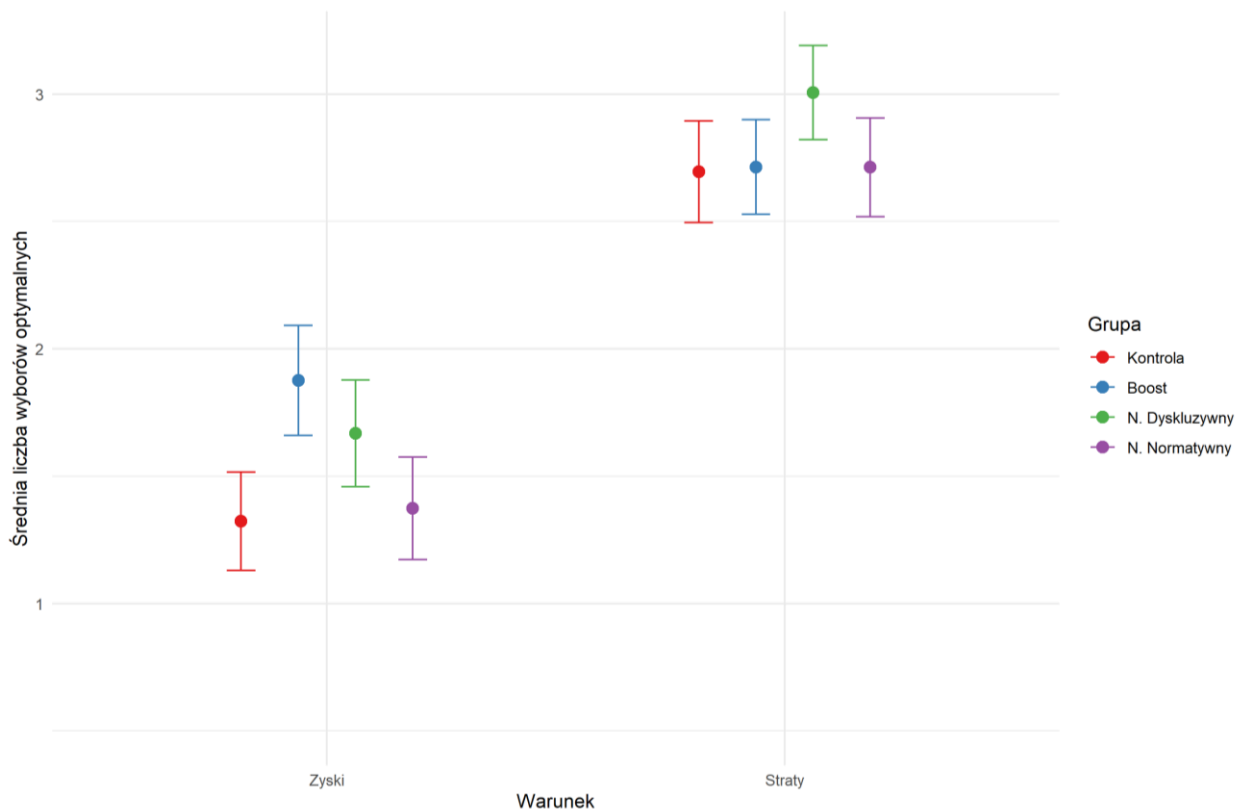
W związku z tym, że modele interakcyjne nie różniły się pod kątem dopasowania od modeli uwzględniających interwencje behawioralne, w dalszych analizach skupiono się na analizie modeli „2”. W domenie zysków model regresji przewidujący liczbę wyborów optymalnych w domenie zysków, na podstawie zmiennych wyboru przed-interwencyjnego i interwencji behawioralnej, wyjaśniał 11,05% zmienności w zmiennej zależnej. Test F dla całego modelu był istotny statystycznie ($F(4, 680) = 21,13, p < 0,001$), wskazując, że model przewiduje liczbę wyborów optymalnych w domenie zysków lepiej niż by to było oczekiwane przy założeniu, że żadna z zmiennych niezależnych istotnie nie przewiduje zmiennej zależnej. Zmienna wyboru przed-interwencyjnego również okazała się istotnie przewidywać wynik zmiennej zależnej ($b = 1,00, SE = 0,13, t(680) = 7,99, p < 0,001$). Analizując poziomy zmiennej interwencji behawioralnej w odniesieniu do grupy kontrolnej (której wartości są reprezentowane przez stałą: $b = 1,13, SE = 0,10, t(680) = 10,85, p < 0,001$), liczba wyborów optymalnych w grupie boost była średnio o 0,59 punktu wyższa w porównaniu z grupą kontrolną ($SE = 0,14, t(680) = 4,19, p < 0,001$). To sugeruje, że interwencja typu boost skutkowała istotnie wyższą liczbą wyborów optymalnych w domenie zysków w porównaniu z grupą kontrolną. Dla interwencji nudge dyskluzyny, liczba wyborów optymalnych była średnio o 0,30 punktu wyższa w porównaniu z grupą kontrolną ($SE = 0,14, t(680) = 2,08, p = 0,037$). Oznacza to, że osoby z tej grupy również miały istotnie wyższą liczbę wyborów optymalnych w porównaniu z grupą kontrolną, choć efekt ten był mniej wyraźny niż w grupie boost. Dla grupy nudge’a

normatywnego, różnica w liczbie wyborów optymalnych w porównaniu z grupą kontrolną nie była statystycznie istotna ($b = 0,02$, $SE = 0,14$, $t(680) = 0,15$, $p = 0,882$).

W domenie strat, model regresji opisywał liczbę wyborów optymalnych na podstawie zmiennych wyboru przed-interwencyjnego i interwencji behawioralnych wyjaśniał 8,3% zmienności zmiennej zależnej. Test F dla całego modelu był istotny statystycznie, ($F(4, 680) = 15,4$, $p < 0,001$), sugerując, że model przewiduje liczbę wyborów optymalnych w domenie strat lepiej niż by to było oczekiwane przy założeniu, że żaden z predyktorów nie jest istotny.

Analizując poziomy zmiennej interwencji behawioralnej w odniesieniu do grupy kontrolnej (której wartości są reprezentowane przez stałą: $b = 2,32$, $SE = 0,11$, $t(680) = 21,53$, $p < 0,001$).

Dla grupy boost, liczba wyborów optymalnych była średnio o 0,08 punktu wyższa w porównaniu z grupą kontrolną, ($SE = 0,13$, $t(680) = 0,63$, $p = 0,532$), co oznacza, że nie zaobserwowano istotnej statystycznie różnicy między grupą boost a grupą kontrolną. W grupie nudge'a dyskluzyjnego liczba wyborów optymalnych była średnio o 0,40 punktu wyższa w porównaniu z grupą kontrolną, ($SE = 0,13$, $t(680) = 3,01$, $p = 0,003$). Interwencja nudge'a dyskluzyjnego skutkowała istotnie wyższą liczbą wyborów optymalnych w porównaniu z grupą kontrolną. W przypadku osób w grupie z interwencją nudge'a normatywnego, liczba wyborów optymalnych była podobna do wyniku osób z grupy boost, ($b = 0,08$, $SE = 0,13$, $t(680) = 0,62$, $p = 0,535$). Tak samo jak w przypadku grupy boost, nie zaobserwowano istotnej statystycznie różnicy między grupą nudge'a normatywnego a grupą kontrolną. Wyniki zostały przedstawione na rycinie 4.



Rycina 4. Różnice między grupami w domenach zysków i strat w oparciu o procedurę FFR. Kropki na rycinie przedstawiają średnie dla każdej z grup eksperymentalnej: grupy kontrolnej, grupy boost, grupy nudge’a dyskluzywnego i grupy nudge’a normatywnego. Wąsy przedstawiają 95% przedziały ufności. Pomiar odnosi się do znaku wypłaty otrzymywanej w pozycjach decyzyjnych.

Efekt wpływu transparentności i związku różnic indywidualnych z wyborami optymalnymi w procedurze FFR. Kolejnym krokiem była analiza uzyskanych efektów włączająca czynnik transparentności interwencji. Głównym celem było sprawdzenie czy świadomość oddziaływania interwencji wśród uczestników będzie zmieniać liczbę wyborów optymalnych. Do modelu uwzględniającego wybór przed-interwencyjny oraz interwencję behawioralną, wprowadzono zmienną transparentności. W przypadku domeny zysków, nowy model nie okazał się lepiej dopasowany do danych niż model z mniejszą liczbą zmiennych:

różnica w sumie kwadratów reszt między modelami wyniosła zaledwie 0,07, co nie było istotne statystycznie, ($F(1, 516) = 0,04$, $p = 0,840$). Podobne wyniki zaobserwowano w domenie strat. Dodanie transparentności jako dodatkowego predyktora do modelu nie przyniosło istotnej poprawy w jego dopasowaniu. Różnica w sumie kwadratów reszt wyniosła 0,65, co nie było istotne statystycznie, ($F(1, 516) = 0,44$, $p = 0,509$).

W ramach dalszych kroków analizy o charakterze eksploracyjnym, zbadano potencjalny związek pozostałych różnic indywidualnych ze zmienną zależną. Na podstawie najlepiej dopasowanego i jednocześnie najprostszego modelu liniowego uwzględniającego czynnik wyborów przed-interwencyjnych i interwencji behawioralnych, zbudowano pięć oddzielnych modeli oddzielnie dla każdej domeny zysków i strat:

- a) Model A – uwzględniający zmienne metryczkowe: wiek, płeć, status związku, posiadanie dzieci, poziom wykształcenia, status pracy
- b) Model B – uwzględniający zmienne osobowościowe zgodnie z modelem Wielkiej Piątki: ekstrawersja, otwartość na doświadczenie, sumienność, neurotyczność i ugodowość.
- c) Model C – uwzględniający cechy temperamentalne z kwestionariusza FCZ-KT (R): Żwawość, Perseweratywność, Wrażliwość sensoryczna, Wytrzymałość, Reaktywność emocjonalna i Aktywność, oraz Rytmiczność.
- d) Model D – uwzględniający wyniki na skali potrzeby poznania
- e) Model E – uwzględniający zdolności numeryczne

Dla porównania modelu bazowego z modelem A, analiza ANOVA wskazała na istotność różnicy między dwoma modelami, ($F(6,674) = 2,85$, $p = 0,0096$). Oprócz zmiennej wyborów przed-interwencyjnych, grupy boost i nudge'a dyskluzycznego tak jak w modelu bazowym, istotną zmienną w modelu okazała się również zmienna wieku. Wiek uczestnika miał negatywny

wpływ na wybór optymalnych opcji, ($b = -0,033$, $t(674) = -3,397$, $p = 0,001$). Inne zmienne demograficzne nie osiągnęły istotnych wyników. Z kolei w domenie strat, model uwzględniający zmienne demograficzne, był znacznie gorzej dopasowany do danych od modelu bazowego, odnotowując większą sumę kwadratów reszt ($RSS = 1044,20$) od modelu bazowego ($RSS = 1005,40$).

W przypadku modelu B, uwzględniającego zmienne osobowościowe, porównanie względem modelu bazowego wykazało brak lepszego dopasowania zarówno dla domeny zysków jak i strat. W przypadku domeny zysków, model z cechami osobowościowymi nie był istotnie lepszy od modelu bazowego ($F(8, 666) = 1,24$, $p = 0,272$), a w przypadku domeny strat, odnotował większą sumę reszt kwadratów ($RSS = 1032,20$; dla modelu bazowego $RSS = 1005,40$). Podobne rezultaty odnotowano w przypadku dopasowania modeli uwzględniających cechy temperamentalne (Model C). W przypadku domeny zysków, porównanie modeli regresji wykazało, że dodanie cech temperamentu nie przyniosło istotnej poprawy w dopasowaniu modelu ($F(7, 657) = 1,01$, $p = 0,429$), a w domenie strat, model odnotował większą sumę reszt kwadratów ($RSS = 1026,88$) niż model bazowy ($RSS = 987,92$).

Następnie sprawdzono dopasowanie modelu D, uwzględniającego wyniki na skali potrzeby poznania. Zarówno w domenie zysków jak i strat, model D okazał się gorzej dopasowany od modelu bazowego: w domenie zysków ($F(4, 665) = 0,89$, $p = 0,471$); w domenie strat RSS dla modelu D wyniosło 1024,50, a dla modelu bazowego 999,00. Ostatnim krokiem była weryfikacja modelu E, uwzględniającego zdolności numeryczne. Porównanie modeli regresji dla domeny zysków wskazało, że dodanie zmiennej zdolności numerycznych nie poprawiło istotnie dopasowania modelu ($F(1, 679) = 0,09$, $p = 0,767$). Podobnie w domenie strat, suma reszt kwadratów była wyższa dla modelu alternatywnego ($RSS = 1042,40$) niż dla modelu bazowego

(RSS = 1007,90), co oznaczało brak lepszego dopasowania modelu. Podsumowując, poza zmienną wieku, różnice indywidualne nie odgrywały większej roli w badanej relacji.

Replikacja wyników Franklin i in. (2019) za pomocą procedury PDQ. W celu systematycznej replikacji efektów uzyskanych za pomocą procedury FFR dla loterii o innych poziomach prawdopodobieństwa i wielkości kwot, wykorzystano procedurę PDQ, składającą się w sumie z 60 pytań, podzielonych na 6, dziesięcio-pozycyjnych zestawów: małe straty, małe zyski, średnie straty, średnie zyski, duże straty i duże zyski. W porównaniu do procedury Franklina, zarówno w fazie przed-interwencyjnej jak i po-interwencyjnej, wykorzystano wszystkie 60 pozycji, z tą różnicą, że w fazie drugiej, pozycjom towarzyszyły interwencje analogiczne do wcześniej przedstawionych: boost, nudge dyskluzyny i nudge normatywny. Czwartą grupą była grupa kontrolna. Analizy dokonano metodą mieszanych modeli liniowych za pomocą pakietu „lmerTest” w języku R, która jest rozwinięciem paczki „lmer” i pozwala na określenie istotności dla poszczególnych zmiennych w modelu (Kuznetsova i in., 2015). W ramach analiz zaplanowano serię mieszanych modeli liniowych, uwzględniających efekty losowe dla osób badanych. Podobnie jak w przypadku analiz dla procedury FFR, analizy zostały wykonane oddzielnie dla domeny zysków i strat:

- 1) Model bazowy – model bazowy uwzględniał tylko stałą i losowe efekty dla osób badanych
- 2) Model 1 – dodawał do modelu zmienną zawierającą informacje o tym czy pomiar został dokonany w fazie przed-interwencyjnej czy po-interwencyjnej
- 3) Model 2 – dodawał do modelu zmienną zawierającą informacje o interwencji behawioralnej oraz interakcję między interwencją behawioralną a punktem pomiaru

W domenie zysków, w porównaniu z modelem bazowym, model 1 uwzględniający wpływ zmiennej punktu pomiaru nie przyniósł znaczącej poprawy dopasowania, a z kolei model 2, uwzględniający interakcję między interwencją behawioralną a pomiarem, przyniósł znaczącą poprawę dopasowania. W domenie strat, zarówno model 1 jak i 2, w porównaniu z modelem bazowym, przyniosły znaczącą poprawę dopasowania. Wskaźniki dopasowania opisanych modeli zostały przedstawione w Tabeli 6.

Tabela 6. Porównanie mieszanych modeli liniowych w domenie zysków i strat (Procedura PDQ).

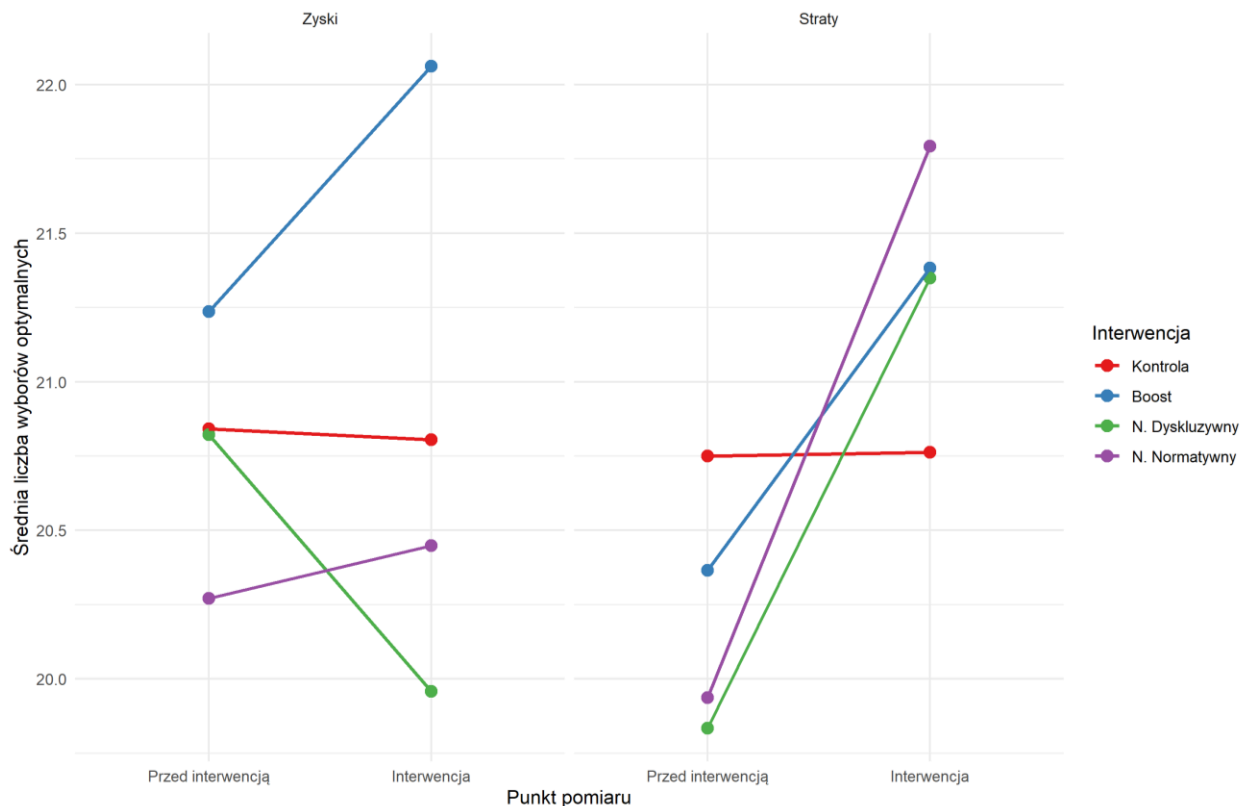
Porównanie modeli	AIC	BIC	LogWorth	Statystyka X ²	Istotność (p)
Zyski					
Model bazowy	7789,9	7805,6	-3892,0	-	-
Model 1 – M bazowy	7791,9	7812,7	-3891,9	0,0585	0,808
M2 – M1	7777,8	7830,0	-3878,9	26,0890	< 0,001
Straty					
Model bazowy	7755,2	7770,9	-3874,6	-	-
Model 1 – M bazowy	7708,0	7728,9	-3850,0	49,228	< 0,001
M2 – M1	7699,7	7751,9	-3839,8	20,344	0,002

W związku z tym, że najlepsze dopasowanie osiągnęły modele serii 2 zarówno w domenie zysków i strat, w kolejnych etapach analizy, przyjrano się istotności statystycznej poszczególnych predyktorów, w obrębie każdego z modeli. W domenie zysków, analizując efekty stałe, zaobserwowano istotny wpływ interwencji boost na liczbę wyborów optymalnych ($t(989,56) = 2,506$, $p = 0,012$, d -Cohena = 0,44), z oszacowaną wartością parametru

niestandaryzowanego $b = 1,26$. Oznacza to, że osoby z grupy boost miały średnio o 1,26 jednostki wyższe zyski w porównaniu do grupy kontrolnej. Grupa nudge'a dyskluzyjnego nie wykazała różnic względem grupy kontrolnej ($t(989,56) = -1,67, p = 0,096$). Podobnie jak grupa nudge'a normatywnego nie różniła się istotnie od grupy kontrolnej pod względem zysków ($t(989,56) = -0,71, p = 0,479$). Zaobserwowano także pewną tendencję w przypadku interakcji między punktem pomiarowym a interwencją boost ($t(989,56) = -1,96, p = 0,051$). Oznacza to, że różnice między grupą boost a grupą kontrolną zmieniały się w zależności od punktu pomiarowego. Podobny, choć mniej istotny trend zaobserwowano dla interakcji zmiennej punktu pomiarowego i nudge'a dyskluzyjnego ($t(681) = 1,85, p = 0,064$). Wyniki są zbliżone do tych odnotowanych w przypadku procedury zaczerpniętej z badania Franklin i in. (2019), gdzie również interwencja typu boost odnosiła najlepsze rezultaty w domenie zysków.

W domenie strat, oceniając efekty stałe, zarówno efekt interwencji boost i nudge'a dyskluzyjnego okazały się nieistotne (kolejno: $p = 0,199, p = 0,230$). Natomiast istotny efekt odnotowano w przypadku nudge'a normatywnego $b = 1,03, (t(1009) = 2,13, p = 0,034)$. Warto również zwrócić uwagę na interakcje między poszczególnymi interwencjami behawioralnymi a punktem pomiarowym – w przypadku każdej z interwencji, ta interakcja była istotna. W grupie boost, interakcja z punktem pomiarowym wskazała na $b = -1,00$, co było istotne statystycznie ($t(681) = -2,30, p = 0,022$). Dla grupy nudge'a dyskluzyjnego, interakcja ta wykazała się jeszcze silniejsza, z istotnym oszacowaniem pomiarowym wynoszącym $b = -1,50 (t(681) = -3,403, p < 0,001)$. Największa interakcja ze zmienną punktu pomiarowego została odnotowana w grupie nudge'a normatywnego, z oszacowaniem parametru równym $b = -1,84 (t(681) = -4,21, p < 0,001)$. Oznacza to, że największy wzrost liczby z większą wartością oczekiwaną pomiędzy pierwszym i drugim pomiarem nastąpił w grupie nudge'a normatywnego (Rycina 5). Wynik w

domenie strat, również częściowo potwierdzały wyniki uzyskane za pomocą procedury Franklina, gdzie najbardziej efektywną interwencją okazała się interwencja z grupy nudge, z tą różnicą, że w przypadku procedury Franklina był to nudge dyskluzyjny. Wizualizacja średniej liczby wyborów optymalnych w procedurze PDQ znajduje się na Rycinie 5.



Rycina 5. Średnia liczba wyborów optymalnych w procedurze PDQ w zależności od rodzaju interwencji, punktu pomiaru i w podziale na domenę zysków i strat. Kropki na rycinie przedstawiają średnie dla każdej z grup eksperymentalnych: grupy kontrolnej, grupy boost, grupy nudge’a dyskluzyjnego i grupy nudge’a normatywnego. Punkt pomiaru odnosi się do wyniku w procedurze PDQ przed wystąpieniem interwencji i w momencie działania interwencji.

Podobnie jak w przypadku procedury FFR, podczas analizy wyników procedury PDQ, w ramach eksploracji, zbadano potencjalny związek pozostałych różnic indywidualnych ze zmienną zależną. Na podstawie najlepiej dopasowanego i jednocześnie najprostszego modelu liniowego

uwzględniającego czynnik wyborów przed-interwencyjnych i interwencji behawioralnych, zbudowano sześć modeli oddzielnie dla każdej domeny zysków i strat:

- a) Model A – uwzględniający zmienną transparentność
- b) Model B - uwzględniający zmienne demograficzne: wiek, status związku, posiadanie dzieci, poziom wykształcenia, status pracy
- c) Model C – uwzględniający zmienne osobowościowe zgodnie z modelem Wielkiej Piątki: ekstrawersja, otwartość na doświadczenie, sumienność, neurotyczność i ugodowość.
- d) Model D – uwzględniający cechy temperamentalne z kwestionariusza FCZ-KT (R): Żwawość, Perseweratywność, Wrażliwość sensoryczna, Wytrzymałość, Reaktywność emocjonalna i Aktywność, oraz Rytmiczność.
- e) Model E – uwzględniający wyniki na skali potrzeby poznania
- f) Model F - uwzględniający zdolności numeryczne

Analizie poddano dopasowanie poszczególnych modeli względem modelu bazowego.

Porównanie modelu A uwzględniającego transparentność interwencji z modelem bazowym, wskazało na brak lepszego dopasowania zarówno w przypadku domeny zysków ($\chi^2(6) = 1,84$, $p = 0,934$, $AIC = 5927,40$, $BIC = 5996,70$) jak i w domenie strat ($\chi^2(6) = 5,55$, $p = 0,476$, $AIC = 5888,1$, $BIC = 5957,4$).

W przypadku modelu B, uwzględniającego zmienne demograficzne, porównanie względem modelu bazowego wskazało na lepsze dopasowanie modeli w domenie zysków ($\chi^2(5) = 16,97$, $p = 0,005$, $AIC = 7770,8$, $BIC = 7849,1$) i strat ($\chi^2(5) = 23,66$, $p < 0,001$, $AIC = 7686,0$, $BIC = 7764,4$). W domenie zysków istotnymi predyktorami okazały się zmienne statusu związku $b = -0,37$ ($t(676) = -2,049$, $p = 0,041$) oraz poziom wykształcenia $b = -0,55$ ($t(676) = -2,148$, $p = 0,032$). W domenie strat istotnymi predyktorami okazały się zmienne statusu związku

$b = -0,40$ ($t(676) = -2,34$, $p = 0,019$), edukacji $b = -0,53$ ($t(676) = -2,184$, $p = 0,029$) oraz status pracy $b = -0,20$ ($t(676) = -2,43$, $p = 0,016$). W celu dalszej eksploracji potencjalnych moderatorów interwencji, zbudowano modele uwzględniające interakcje pomiędzy istotnymi predyktorami demograficznymi a interwencjami. Żaden z istotnych predyktorów demograficznych nie wchodził w interakcję z interwencjami.

W modelu C, który uwzględniał zmienne osobowościowe, dopasowanie do danych okazało się lepsze zarówno w przypadku domeny zysków ($\chi^2(5) = 44,16$, $p < 0,001$, AIC = 7666,4, BIC = 7744,6) i strat ($\chi^2(5) = 29,69$, $p < 0,001$, AIC = 7611,7, BIC = 7689,9). W domenie zysków istotnymi predyktorami okazały się zmienne summienność $b = 0,24$ ($t(670) = 2,20$, $p = 0,028$), ugodowość $b = 0,40$ ($t(670) = 2,997$, $p = 0,003$) oraz otwartość na doświadczenia $b = 0,57$ ($t(670) = 3,903$, $p < 0,001$). W domenie strat istotnymi predyktorami okazały się zmienne ugodowość $b = 0,39$ ($t(670) = 3,01$, $p = 0,003$) oraz otwartość na doświadczenia $b = 0,36$ ($t(670) = 2,61$, $p = 0,009$). W celu dalszej eksploracji potencjalnych moderatorów interwencji, zbudowano modele uwzględniające interakcje pomiędzy istotnymi predyktorami osobowościowymi a interwencjami. Żaden z istotnych predyktorów osobowościowych nie wchodził w interakcję z interwencjami.

Model D, uwzględniający zmienne temperamentalne pochodzące z kwestionariusza FCZ-KT, również wykazał lepsze dopasowanie do danych niż model bazowy w domenie zysków ($\chi^2(7) = 46,32$, $p < 0,001$, AIC = 7594,60, BIC = 7683,10) i strat ($\chi^2(7) = 49,98$, $p < 0,001$, AIC = 7542,80, BIC = 7631,30). W domenie zysków istotnymi predyktorami zmiennej zależnej okazały się żwawość $b = 1,58$ ($t(663) = 3,76$, $p < 0,001$), perseweratywność $b = 1,27$ ($t(663) = 2,71$, $p = 0,007$) oraz wytrzymałość $b = -0,97$ ($t(663) = -2,08$, $p = 0,038$). W domenie strat istotnymi predyktorami okazały się zmienne żwawość $b = 1,40$ ($t(663) = 3,51$, $p < 0,001$),

perseweratywność $b = 1,27$ ($t(663) = 2,85$, $p = 0,005$), wytrzymałość $b = -1,41$ ($t(663) = -3,15$, $p = 0,002$) oraz aktywność $b = -0,90$ ($t(663) = -2,72$, $p = 0,007$). W celu dalszej eksploracji potencjalnych moderatorów interwencji, zbudowano modele uwzględniające interakcje pomiędzy istotnymi predyktorami temperamentalnymi a interwencjami. Model w domenie zysków okazał się lepiej dopasowany do danych od modelu bazowego ($\chi^2(56) = 91,12$, $p < 0,001$, AIC = 7647,80, BIC = 7991,40) i zanotował trzy istotne efekty interakcyjne. W grupie nudge'a dyskluzyjnego, wyższe wyniki zmiennej zależnej notowały osoby badane charakteryzujące się wyższą wytrzymałością $b = 3,23$ ($t(940) = 2,17$, $p = 0,030$), w grupie nudge'a normatywnego wyższe wyniki osiągały osoby charakteryzujące się wyższymi wynikami na skali rytmiczności $b = 2,08$ ($t(940) = 2,08$, $p = 0,038$), a w grupie boost wyższe wyniki notowały osoby o wyższej emocjonalności $b = 3,01$ ($t(940) = 2,09$, $p = 0,037$). W domenie strat model z interakcjami również był lepiej dopasowany od modelu bazowego ($\chi^2(56) = 106,18$, $p < 0,001$, AIC = 7584,60, BIC = 7928,20), a w jego obrębie zanotowano trzy interakcje z interwencjami behawioralnymi: w grupie nudge'a normatywnego dodatkową rolę odgrywała rytmiczność $b = 2,08$ ($t(940) = 2,08$, $p = 0,038$), a z grupą nudge'a dyskluzyjnego i boostu w interakcję wchodziła emocjonalność, kolejno notując współczynniki $b = 3,28$ ($t(940) = 2,47$, $p = 0,013$), $b = 3,01$ ($t(940) = 2,09$, $p = 0,037$).

W modelu E, jako dodatkową zmienną wzięto pod uwagę potrzebę poznania. Zarówno w zyskach ($\chi^2(1) = 51,64$, $p < 0,001$, AIC = 7578,00, BIC = 7635,30) jak i w stratach ($\chi^2(1) = 51,64$, $p < 0,001$, AIC = 7522,4, BIC = 7579,7) model okazał się lepiej dopasowany. Potrzeba poznania była istotnym predyktorem zmiennej zależnej w zyskach $b = 2,77$ ($t(669) = 7,31$, $p < 0,001$) i stratach $b = 2,65$ ($t(669) = 7,30$, $p < 0,001$). W celu dalszej eksploracji potrzeby poznania jako potencjalnego moderatora interwencji, zbudowano modele uwzględniające jej interakcje z

interwencjami. W żadnej z domen potrzeba poznania nie wchodziła w interakcję z interwencjami nudge i boost.

Model F, uwzględniający pomiar zdolności numerycznych, odznaczał się lepszym dopasowaniem niż model bazowy zarówno w zyskach ($\chi^2(1) = 13,37$, $p < 0,001$, AIC = 7766,4, BIC = 7823,8) jak i w stratach ($\chi^2(1) = 23,07$, $p < 0,001$, AIC = 7678,6, BIC = 7736,0).

W obydwu domenach, zdolności numeryczne były istotnym predyktorem zmiennej zależnej, notując kolejno w zyskach $b = 0,53$ ($t(680) = 3,66$, $p < 0,001$) i stratach $b = 0,66$ ($t(680) = 4,83$, $p < 0,001$). W żadnej z domen zdolności numeryczne nie wchodziły w interakcję z interwencjami nudge i boost.

2.3.4 Podsumowanie

Przeprowadzone badanie dotyczyło Hipotez H2.1, H2.2 i H2.3: 1) czy istnieją różnice w sile wpływu oddziaływania interwencji typu boost i nudge w domenie zysków, 2) czy istnieją różnice w sile wpływu oddziaływania interwencji typu boost i nudge w domenie strat i oraz 3) czy różnice w oddziaływaniu boostów i nudge'y są moderowane przez przejrzystość (transparentność) interwencji. W celu zweryfikowania Hipotez H2.1 i H2.2 wprowadzono do badania procedurę FFR (Franklin i in., 2019) oraz dodano procedurę PDQ. Procedura PDQ została wykorzystana do zreplicowania efektów procedury FFR. Zmienne zależne uzyskane w rezultacie procedur FFR i PDQ miały taką samą strukturę – była to suma wyborów o wyższej wartości oczekiwanej. Dodatkowo, w celu weryfikacji hipotezy H2.2 wprowadzono czynnik transparentności wyboru, połowa osób badanych była od początku świadoma działających interwencji, a druga połowa dowiedziała się o nich dopiero na koniec badania.

W ramach pierwszego etapu, testującego procedurę FFR, zostały potwierdzone zarówno efekty i kierunki działania interwencji raportowane w literaturze (Franklin i in., 2019). Podobnie jak w badaniach Franklina, interwencje typu boost przekładały się na większą liczbę wyborów optymalnych (charakteryzujących się wyższą wartością oczekiwaną; EV) w porównaniu do grupy kontrolnej, gdy wybór odbywał się w domenie zysków. Z kolei gdy wybór miał miejsce w domenie strat, to interwencja nudge charakteryzowała się większą liczbą wyborów o wyższym EV względem grupy kontrolnej. Pojawiły się także rozbieżności względem oryginalnego badania. Różnica, w porównaniu do wyników Franklina, wystąpiła w tym, że w przeprowadzonych badaniach, to interwencja nudge dyskluzyny prowadziła do większej liczby wyborów optymalnych. Natomiast w badaniu wcześniejszym, interwencją prowadzącą do większej liczby wyborów optymalnych w domenie strat był nudge normatywny. W przypadku procedury PDQ, interwencje typu boost również prowadziły do większej liczby wyborów o wyższej EV w porównaniu do grupy kontrolnej w domenie zysków. Różnica w porównaniu do procedury FFR pojawiła się w domenie strat. Procedura PDQ wskazała, że to nudge normatywny prowadził do większej liczby wyborów o wyższej EV w porównaniu do grupy kontrolnej w domenie strat, co było również replikacją efektu uzyskanego w poprzednich badaniach (Franklin i in., 2019). Tym samym, wyniki uzyskane w oparciu o procedury FFR i PDQ pozwalają na **przyjęcie Hipotez H2.1 H2.2:** w domenie zysków, boosty miały istotnie wyższą liczbę wyborów optymalnych niż grupa kontrolna, a w domenie strat to interwencje typu nudge miały istotnie wyższą liczbę wyborów optymalnych niż grupa kontrolna. Należy natomiast zaznaczyć, że przeprowadzone analizy nie wykazały bezpośrednich różnic między interwencjami nudge i boost ani w domenie zysków, ani w domenie strat. Zauważone różnice dotyczyły porównań między wskazanymi interwencjami a grupą kontrolną.

Ważnym elementem Badania 2A był czynnik transparentności. Zarówno w przypadku procedury FFR jak i PDQ, transparentność interwencji nie była istotnym predyktorem wyborów osób badanych. **Wyniki pozwalają na odrzucenie Hipotezy H2.3.**

Badanie 2A miało również charakter eksploracyjny w kontekście roli różnic indywidualnych i zmiennych demograficznych jako potencjalnych czynników moderujących efektywność interwencji behawioralnych. Procedura FFR nie wykazała żadnych istotnych zależności między różnicami indywidualnymi i zmiennymi demograficznymi, poza wiekiem osób badanych, który był negatywnie związany z liczbą optymalnych decyzji. Inaczej było w przypadku procedury PDQ. Wśród zmiennych demograficznych, liczbę optymalnych wyborów istotnie przewidywał status związku (podobnie jak we Franklin i in. 2019), poziom wykształcenia oraz status pracy. W przypadku zmiennych osobowościowych, istotnymi predyktorami były sumienność, ugodowość i otwartość na doświadczenia, a wśród zmiennych temperamentalnych były to żwawość, perseweratywność, wytrzymałość i aktywność. Dalsza analiza interakcji wykazała, że w domenie zysków osoby badane notujące wyższe wyniki na skali wytrzymałości charakteryzowały się wyższymi wynikami w przypadku oddziaływania nudge'a dyskluzyjnego, osoby charakteryzujące się wyższymi wynikami na skali rytmiczności osiągały większą liczbę optymalnych decyzji w grupie nudge'a normatywnego, a boosty lepiej oddziaływały gdy osoby badane miały wyższe wyniki na skali emocjonalności. Podobne wyniki zaobserwowano w przypadku domeny strat, z tą różnicą, że to emocjonalność a nie wytrzymałość wchodziła w interakcję z nudgem dyskluzyjnym. Warto również zwrócić uwagę, że zarówno w domenie zysków jak i strat zmienne potrzeba poznania i zdolności numeryczne zwiększały liczbę optymalnych wyborów, natomiast nie wchodziły w interakcję z interwencjami behawioralnymi.

2.4 Badanie 2B *Trwałość efektów interwencji behawioralnych nudge i boost w decyzjach ryzykownych*

Celem Badania 2B było zbadanie wpływu czasu na trwałość efektów interwencji behawioralnych przedstawionych w Badaniu 2A. Badanie było dwuetapowe, oddzielone od siebie tygodniem przerwy. Interwencja występowała w pierwszym tygodniu, natomiast w drugim tygodniu uczestnicy podchodzili ponownie do tych samych loterii, tym razem już bez towarzyszących interwencji.

2.4.1 Osoby badane

Przed badaniem przeprowadzono analizę mocy dla minimalnej wymaganej wielkości próby, odpowiedniej do wykrycia efektów wieloczynnikowej analizy wariancji w schemacie mieszanym z interakcjami na poziomie między- i wewnątrzgrupowym (ANOVA), dla planu 4x2 (4 czynnikowa zmienna grupy interwencyjnej x 2 czynnikowa zmienna odzwierciedlająca powtarzany pomiar). Poziom alfa ustalono na poziomie 0,05, moc na poziomie 0,8, korelację pomiędzy pomiarami na 0,5, a siłę efektu $f = 0,135$. Siła efektu na tym poziomie została wyliczona ze wzoru $f = d \text{ Cohena} / 2$ (Cohen, 2013) dla średniej siły efektu interwencji behawioralnych otrzymanej w Badaniu 1 (Metaanaliza), która wyniosła $d \text{ Cohena} = 0,27$. Analizy mocy dokonano za pomocą programu GPower, a rekomendowana próba wyniosła 156 osób.

W badaniu 2B wzięło udział 733 osoby, natomiast jedynie 300 osób, zostało zakwalifikowanych do końcowych analiz, ze względu na udział w obydwu etapach badania. Wszystkie osoby zostały zrekrutowane przez panel badawczy SONA Uniwersytetu SWPS, w związku z tym, wszyscy uczestnicy mieli status studenta w momencie udziału w procedurze.

Wśród uczestników było 239 kobiet, 58 mężczyzn oraz 3 osoby, które wolały nie podawać informacji o swojej płci. Średnia wieku wyniosła 26,04 lat ($SD=8,32$).

2.4.2 Procedura

Badanie 2B, bazowało na planie badawczym wcześniejszego Badania 2A, które oprócz dwóch pierwszych faz odbywających się w trakcie pierwszego etapu, posiadało również trzecią fazę po upływie tygodnia: fazę pierwszą przed-interwencyjną, fazę drugą interwencyjną oraz fazę po-interwencyjną. W pierwszej fazie, osoby badane miały za zadanie udzielić odpowiedzi na trzy pozycje zaczerpnięte z procedury FFR (Franklin i in., 2019), oparte o loterie o różnej wartości oczekiwanej. Podobnie jak w oryginalnych badaniach, wynik w tym pomiarze w analizach został wykorzystany jako predyktor wyników zmiennej zależnej. W pomiarze przed-interwencyjnym, wprowadzono również alternatywną zmienną zależną – procedurę PDQ (Madden i in., 2009). Z procedury PDQ zostało zaczerpniętych w sumie 60 pozycji, po 10 pozycji na każdą grupę: małe straty, małe zyski, średnie straty, średnie zyski oraz duże straty i duże zyski.

Pomiar interwencyjny składał się z procedury FFR (Franklin i in., 2019) oraz ponownie procedury PDQ z uwzględnionymi interwencjami behawioralnymi. Procedura FFR składała się z osiem pozycji, z czego cztery były osadzone w domenę strat a cztery w domenę zysków. By uniknąć efektów torowania, kwestionariusze mierzące różnice indywidualne były prezentowane po pozycjach behawioralnych. Wśród kwestionariuszy mierzących różnice indywidualne znalazły się: FCZ-KT (R) (Cyniak-Cieciura i in. 2017; skale: Żwawość, Perseweratywność, Wrażliwość sensoryczna, Wytrzymałość, Reaktywność emocjonalna i Aktywność, oraz Rytmiczność), Kwestionariusz Potrzeby Poznania (Matusz i in., 2011), TIPI (ang. *Ten-Item Personality Inventory*; Sorokowska i in., 2014; skale: Otwartość na doświadczenie, Sumiennność, Ekstrawersja, Ugodowość, Neurotyczność) oraz Berliński Test Numeryczny (ang. *Berlin*

Numeracy Test; Cokely i in., 2016) Na sam koniec fazy drugiej osoby badane zostały poproszone o wypełnienie pytań metryczkowych. Po upływie tygodnia, uczestnicy zostali zaproszeni do fazy trzeciej, w której zostali poproszeni o wypełnienie jeszcze raz procedur FFR oraz PDQ, tym razem bez bieżącego oddziaływania żadnej interwencji. Celem takiego schematu badawczego było sprawdzenie, czy efekty interwencji utrzymają się po ich usunięciu. Badania przeprowadzono w całości online za pomocą platformy Qualtrics (Qualtrics, Provo, UT; qualtrics.com).

2.4.3 Wyniki

Procedura FFR. W celu replikacji efektów badania 2A oraz sprawdzenia czy wyniki osób badanych utrzymają się po upływie czasu, analiz dokonano metodą mieszanych modeli liniowych za pomocą pakietu „lmerTest” w języku R, która jest rozwinięciem paczki „lmer” i pozwala na określenie istotności dla poszczególnych zmiennych w modelu (Kuznetsova i in., 2015). W ramach analiz zaplanowano serię mieszanych modeli liniowych, uwzględniających efekty losowe dla osób badanych. Podobnie jak w przypadku badania 2A, analizy zostały wykonane oddzielnie dla domeny zysków i strat:

- 1) Model bazowy – model bazowy uwzględniał tylko stałą i losowe efekty dla osób badanych
- 2) Model 1 – dodawał do modelu zmienną zawierającą informacje o interwencji behawioralnej
- 3) Model 2 – dodawał do modelu zmienną zawierającą informacje o punkcie pomiarowym (faza interwencyjna oraz po-interwencyjna) oraz interakcję między interwencją behawioralną a punktem pomiaru

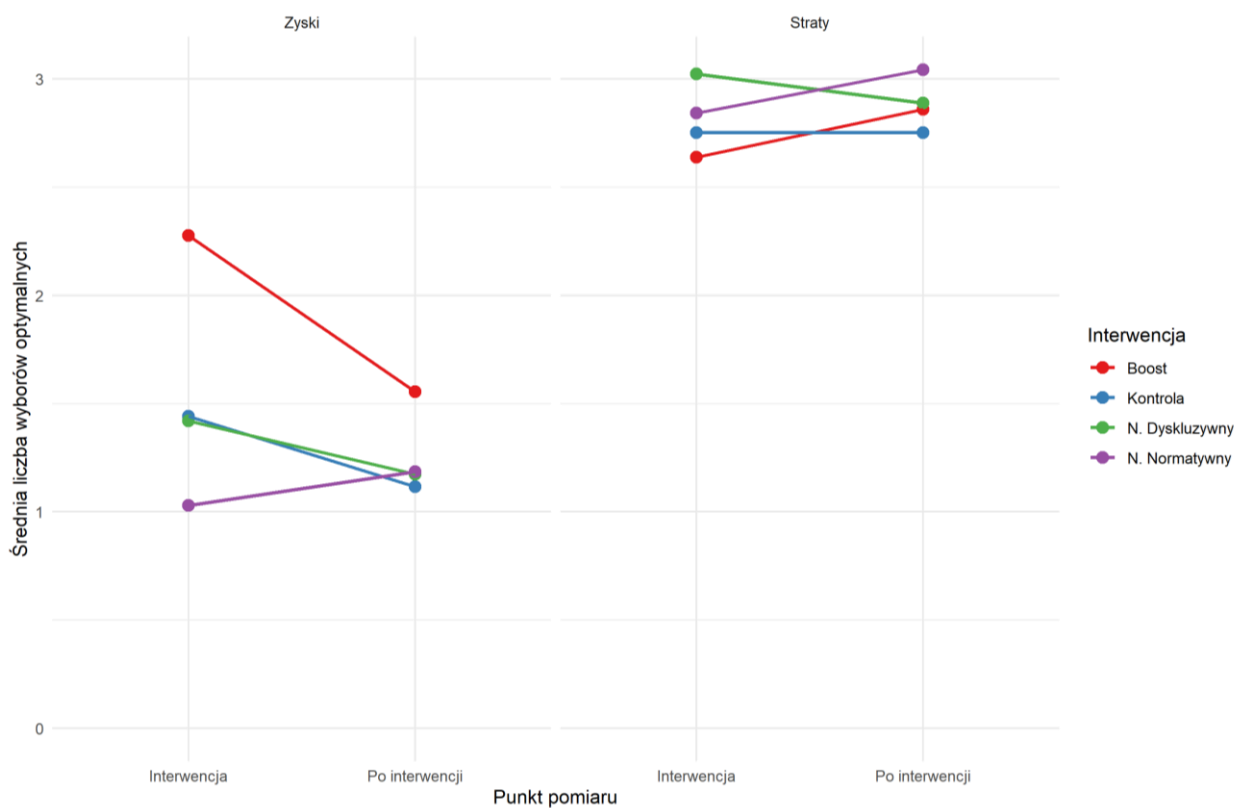
W domenie zysków, w porównaniu z modelem bazowym, zarówno model 1 uwzględniający wpływ interwencji behawioralnej przyniósł znaczącą poprawę dopasowania, jak i model 2, uwzględniający interakcję między interwencją behawioralną a pomiarem. Z kolei w domenie strat, w porównaniu z modelem bazowym, ani model 1 ani 2 nie przyniosły poprawy dopasowania. Wskaźniki dopasowania modeli zostały przedstawione w Tabeli 7.

Tabela 7. Porównanie mieszanych modeli liniowych w domenie zysków i strat dla danych uzyskanych z procedury FFR

Porównanie modeli	AIC	BIC	LogWorth	Statystyka X ²	Istotność (p)
Zyski					
Model bazowy	1952,99	1966,18	-973,50	-	-
Model 1 – M bazowy	1935,60	1961,98	-961,80	23,39	< 0,001
M2 – M1	1914,16	1958,13	-947,08	29,44	< 0,001
Straty					
Model bazowy	1876,53	1889,73	-935,27	-	-
Model 1 – M bazowy	1880,08	1906,46	-934,04	2,46	0,483
M2 – M1	1882,59	1926,56	-931,29	5,49	0,241

W związku z tym, że najlepsze dopasowanie osiągnął jedynie model 2 w domenie zysków, w kolejnych etapach analizy, przyjrano się istotności poszczególnych predyktorów, w jego obrębie. W związku z brakiem lepszego dopasowania żadnego z modeli w domenie strat, analizy efektów głównych i interakcji nie została przeprowadzona. Analizując efekty stałe

modelu 2 w domenie zysków, zaobserwowano istotny wpływ interwencji boost na liczbę wyborów optymalnych $b = 0,84$ ($t(491,87) = 4,08$, $p < 0,001$). Oznacza to, że osoby z grupy boost w pierwszym pomiarze miały średnio o 0,84 jednostki wyższe zyski w porównaniu do grupy kontrolnej. Grupa w warunku nudge'a dyskluzyjnego nie różniła się istotnie statystycznie względem grupy kontrolnej ($t(491,87) = -0,110$, $p = 0,912$), natomiast interwencja nudge'a normatywnego zmniejszała wyniki na skali zmiennej zależnej względem grupy kontrolnej $b = -0,41$ ($t(491,87) = -2,00$, $p = 0,046$). Efekt główny punktu pomiarowego wskazał również, że osoby badane notowały niższe wyniki w drugim pomiarze $b = -0,32$ ($t(296) = -2,18$, $p = 0,030$). Zaobserwowano także interakcję między punktem pomiarowym a interwencją nudge normatywny $b = 0,48$ ($t(296) = 2,23$, $p = 0,027$). Oznacza to, że różnice między grupą nudge normatywny a grupą kontrolną zmieniały się w zależności od punktu pomiarowego. Biorąc pod uwagę suboptymalny wynik tych osób w pierwszym pomiarze oraz brak interwencji w drugim pomiarze, jednymi z potencjalnych wytłumaczeń są efekt wprawy, lub pewna reaktancja względem deskryptywnej normy społecznej na której ta interwencja bazowała. Choć interakcja w grupie boost a drugim punktem pomiarowym wykazywała pewną tendencję, to nie różniła się istotnie od pozostałych grup w pomiarze drugim $b = -0,40$ ($t(296) = -1,85$, $p = 0,065$). Warto jednak wskazać, że w tej grupie zarówno w pierwszym jak i w drugim pomiarze osoby badane uzyskiwały najwyższe wyniki w zadaniu eksperymentalnym. W pewnym stopniu, uzyskane wyniki są zbliżone do tych odnotowanych w przypadku procedury FFR w badaniu 2A, gdzie również interwencja typu boost odnosiła najlepsze rezultaty w domenie zysków. Inaczej jest w przypadku domeny strat. W badaniu 2A nie tylko model uwzględniający zmienną grupującą był lepszy od bazowego, ale również wskazywał na lepsze działanie nudge'a dyskluzyjnego. Wizualizacja średniej liczby wyborów optymalnych w procedurze Franklina w badaniu 2B znajduje się na Rycinie 6.



Rycina 6. Liczba wyborów optymalnych w procedurze Franklin i in. (2019) w zależności od rodzaju interwencji, punktu pomiaru i w podziale na domenę zysków i strat. Kropki na rycinie przedstawiają średnie dla każdej z grup eksperymentalnych: grupy kontrolnej, grupy boost, grupy nudge’a dyskluzywnego i grupy nudge’a normatywnego.

Podobnie jak w przypadku badania 2A, w ramach dalszych kroków analizy wyników procedury Franklina, w ramach eksploracji, zbadano potencjalny związek pozostałych różnic indywidualnych ze zmienną zależną. Na podstawie najlepiej dopasowanego i jednocześnie najprostszego modelu liniowego uwzględniającego czynnik punktu pomiaru i interwencji behawioralnych, zbudowano sześć modeli oddzielnie dla domeny zysków:

- a) Model A - uwzględniający zmienne metryczkowe: wiek, status związku, posiadanie dzieci, poziom wykształcenia, status pracy

- b) Model B – uwzględniający zmienne osobowościowe zgodnie z modelem Wielkiej Piątki: ekstrawersja, otwartość na doświadczenie, sumienność, neurotyczność i ugodowość.
- c) Model C – uwzględniający cechy temperamentalne z kwestionariusza FCZ-KT (R): Żwawość, Perseweratywność, Wrażliwość sensoryczna, Wytrzymałość, Reaktywność emocjonalna i Aktywność, oraz Rytmiczność.
- d) Model D – uwzględniający wyniki na skali potrzeby poznania
- e) Model E - uwzględniający zdolności numeryczne

Żaden z modeli uwzględniających zmienne indywidualne, nie charakteryzował się lepszym dopasowaniem od modelu bazowego: model A (χ^2 (12) = 12,07, p = 0,440, AIC = 1926,10, BIC = 2022,80), model B (χ^2 (5) = 10,24, p = 0,069, AIC = 1887,70, BIC = 1953,30), model C (χ^2 (7) = 10,35, p = 0,169, AIC = 1755,20, BIC = 1828,50), model D (χ^2 (1) = 2,46, p = 0,117, AIC = 1724,4, BIC = 7635,3), model E, (χ^2 (1) = 2,68, p = 0,102, AIC = 1891,8, BIC = 1940,0).

Replikacja za pomocą procedury PDQ. Podobnie jak w badaniu 2A, w badaniu uwzględniono procedurę PDQ. Różnicą względem badania 2A był trzykrotny pomiar uzyskany za pomocą procedury: pomiar przed-interwencyjny, interwencyjny oraz po-interwencyjny oddalony tygodniowym odstępem od pozostałych z nich. Pozostałe aspekty pozostały tożsame z badaniem 2A. Analizy dokonano metodą mieszanych modeli liniowych z pomocą tych samych pakietów analitycznych do uprzednio. W ramach analiz zaplanowano serię mieszanych modeli liniowych, uwzględniających efekty losowe dla osób badanych. Analizy zostały wykonane oddzielnie dla domeny zysków i strat:

- 1) Model bazowy – model bazowy uwzględniał tylko stałą i losowe efekty dla osób badanych

- 2) Model 1 – dodawał do modelu zmienną zawierającą informacje o interwencji behawioralnej
- 3) Model 2 – dodawał do modelu zmienną zawierającą informacje punkcie pomiaru (trzy punkty) oraz interakcję między interwencją behawioralną a punktem pomiaru

W domenie zysków, w porównaniu z modelem bazowym, zarówno model 1 uwzględniający wpływ zmiennej interwencji przyniósł poprawę dopasowania jak i model 2, uwzględniający interakcję między interwencją behawioralną a punktem pomiarowym. Podobna sytuacja miała miejsce w domenie strat, zarówno model 1 jak i model 2 okazały się lepsze od modelu bazowego.

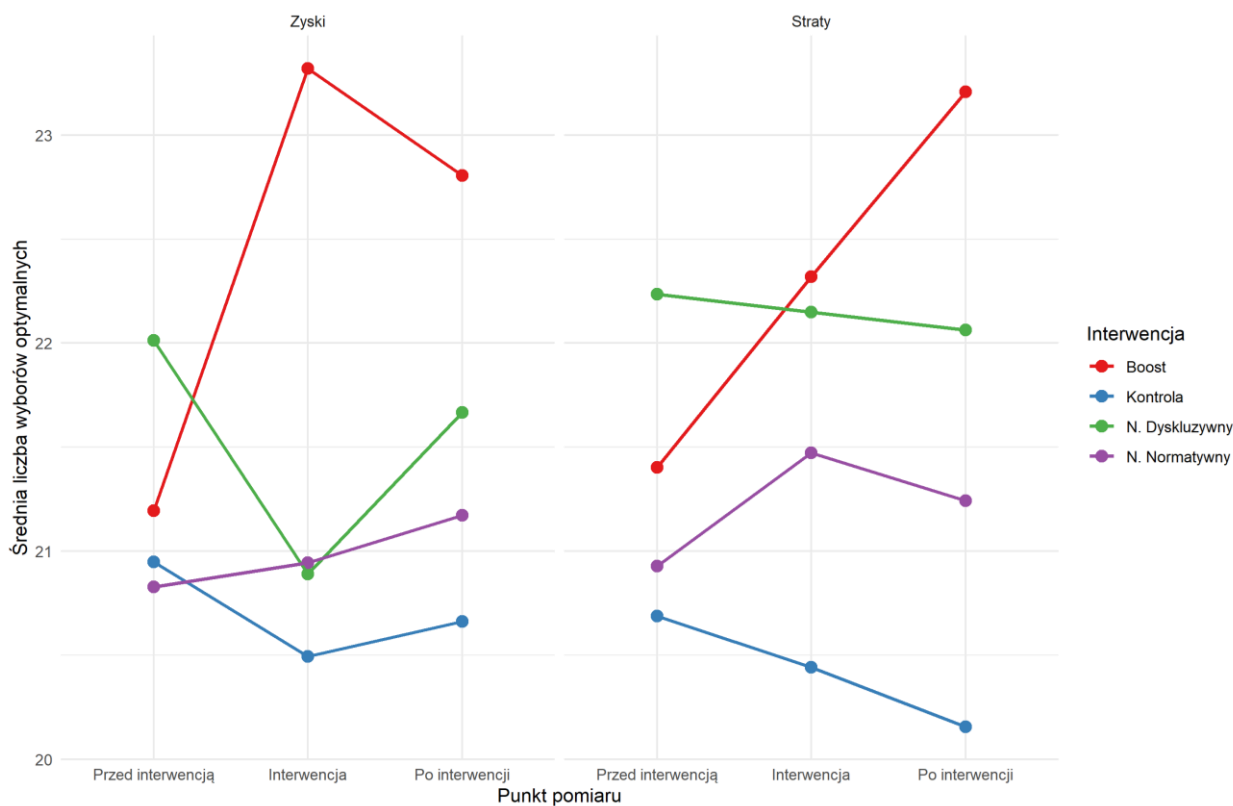
Tabela 8. Porównanie mieszanych modeli liniowych w domenie zysków i strat dla danych uzyskanych z procedury PDQ

Porównanie modeli	AIC	BIC	LogWorth	Statystyka X ²	Istotność (p)
Zyski					
Model bazowy	5009,13	5023,53	-2501,56	-	-
Model 1 – M bazowy	5005,30	5034,11	-2496,65	9,83	0,020
M2 – M1	4994,77	5062,01	-2483,39	26,52	0,001
Straty					
Model bazowy	4860,02	4874,43	-2427,01	-	-
Model 1 – M bazowy	4849,42	4878,23	-2418,71	16,61	0,001
M2 – M1	4847,89	4915,13	-2409,95	17,52	0,025

W związku z tym, że najlepsze dopasowanie osiągnęły modele 2 zarówno w domenie zysków i strat, w kolejnych etapach analizy, przyjrzano się istotności poszczególnych predyktorów, w obrębie każdego z modeli. W domenie zysków, na poziomie bazowym, czyli w pierwszym punkcie pomiarowym (przed-interwencyjnym) zarówno grupa boost ($t(572,24) = 0,34, p = 0,731$), nudge dyskluzyjny ($t(572,24) = 1,53, p = 0,126$) i nudge normatywny ($t(572,24) = -0,17, p = 0,868$) nie różniły się od grupy kontrolnej. Co oznacza, że dobór losowy pod kątem umiejętności i wiedzy mogąc wpływać na wyniki zapewnił brak różnic pomiędzy osobami. Jednakże, po uwzględnieniu interakcji z poziomem pomiaru, interwencja typu boost okazała się wyraźnie lepsza od grupy kontrolnej zarówno w pomiarze interwencyjnym $b = 2,58, (t(592) = 3,69, p < 0,001)$ jak i w trzecim pomiarze, po-interwencyjnym $b = 1,90, (t(592) = 2,72, p = 0,007)$. W ramach dalszej analizy, przeprowadzono porównania parami za pomocą pakietu „emmeans” (Lenth i in., 2024), które potwierdziło wynik dla pomiaru interwencji ($t(572) = -3,31, p = 0,006$), natomiast nie potwierdziła dla trzeciego pomiaru po-interwencyjnego ($t(572) = -2,989, p = 0,176$). Niejednoznaczny wynik zostanie przedyskutowany w podsumowaniu sekcji. Interpretując jedynie wyniki modelu, można byłoby stwierdzić, że również po usunięciu interwencji i po upływie tygodnia osoby badane wciąż były w stanie skutecznie zastosować wyuczoną heurystykę pozwalającą obliczać EV. Jednakże, po sprawdzeniu kontrastów, nie jest to w pełni możliwe. Żadna z pozostałych interwencji nie prowadziła do większej liczby optymalnych wyborów ani w drugim ani trzecim pomiarze (Rycina 7).

W domenie strat w pierwszym pomiarze, względem grupy kontrolnej zaobserwowano istotny efekt interwencji typu nudge'a dyskluzyjnego $b = 1,55 (t(598) = 2,45, p = 0,015)$. W drugim pomiarze można było zaobserwować pewną tendencję interwencji boost względem kontroli $b = 1,16 (t(592) = 1,78, p = 0,076)$, a w trzecim pomiarze interwencja boost okazała się

istotnie lepsza od grupy kontrolnej $b = 2,34$ ($t(592) = 3,57$, $p < 0,001$). Porównania parami za pomocą pakietu *emmeans* (Lenth i in., 2024) potwierdziły jedynie trzeci efekt ($t(592) = -4,70$, $p < 0,001$). Pozostałe interwencje nie wykazały istotnych różnic względem grupy kontrolnej. Różnicę pomiędzy grupą boost a grupą kontrolną w trzecim pomiarze można również zobaczyć na Rycinie 7. Obserwowany trend wyników może być spowodowany wzrostem umiejętności posługiwania się heurystyką wdrażaną przez interwencję typu boost pomiędzy domenami. Należy pamiętać, że osoby badane miały za zadanie dokonać wyboru zarówno w stratach jak i zyskach.



Rycina 7. Uśredniona liczba wyborów optymalnych w procedurze Franklin i in. (2019) w zależności od rodzaju interwencji, punktu pomiaru i w podziale na domenę zysków i strat. Punkty przedstawiają średnią liczbę wyborów optymalnych dla każdej z grup eksperymentalnych: grupy kontrolnej, grupy boost, grupy nudge’a dyskluzywnego i grupy nudge’a normatywnego. Skala osi Y zaczyna się od 20, gdyż wszystkie średnie w poszczególnych grupach były powyżej 20 na skali

średniej liczby wyborów optymalnych. Ustawienie początku osi Y na 0, utrudniało odczytanie wyników i relacji.

Podobnie jak w przypadku badania 2A, w ramach dalszych kroków analizy wyników procedury PDQ, zbadano potencjalny związek pozostałych różnic indywidualnych ze zmienną zależną. Na podstawie najlepiej dopasowanego i jednocześnie najprostszego modelu liniowego uwzględniającego czynnik punktu pomiaru i interwencji behawioralnych, zbudowano pięć modeli oddzielnie dla domen zysków i strat:

- a) Model A - uwzględniający zmienne metryczkowe: wiek, status związku, posiadanie dzieci, poziom wykształcenia, status pracy
- b) Model B – uwzględniający zmienne osobowościowe zgodnie z modelem Wielkiej Piątki: ekstrawersja, otwartość na doświadczenie, sumienność, neurotyczność i ugodowość.
- c) Model C – uwzględniający cechy temperamentalne z kwestionariusza FCZ-KT (R): Żwawość, Perseweratywność, Wrażliwość sensoryczna, Wytrzymałość, Reaktywność emocjonalna i Aktywność, oraz Rytmiczność.
- d) Model D – uwzględniający wyniki na skali potrzeby poznania
- e) Model E - uwzględniający zdolności numeryczne

Spośród pięciu typów modeli, modele A, B oraz C zarówno w domenie zysków jak i strat, nie charakteryzowały się lepszym dopasowaniem: Model A zyski ($\chi^2 (12) = 10,58$, $p = 0,565$, AIC = 5008,20, BIC = 5133,10), model A straty ($\chi^2 (12) = 19,99$, $p = 0,067$, AIC = 4851,9, BIC = 4976,8), model B zyski ($\chi^2 (5) = 8,18$, $p = 0,147$, AIC = 4908,9, BIC = 4999,8), model B straty ($\chi^2 (5) = 10,21$, $p = 0,069$, AIC = 4767,6, BIC = 4858,6), modelu C zyski ($\chi^2 (7) = 9,47$, $p = 0,2208$, AIC = 4560,9, BIC = 4659,9), model C straty ($\chi^2 (7) = 12,36$, $p = 0,089$, AIC = 4456,7, BIC = 4555,7).

Inaczej miała się sytuacja w przypadku modeli D gdzie zarówno w zyskach ($\chi^2(1) = 8,57, p = 0,003, AIC = 4514,1, BIC = 4584,6$) jak i stratach ($\chi^2(1) = 9,79, p = 0,002, AIC = 4374,2, BIC = 4444,6$) znaleziono lepsze dopasowanie do danych w porównaniu do modelu bazowego. Zarówno w zyskach $b = 1,53 (t(265)=2,924, p = 0,004)$ jak i w stratach $b = 1,44 (t(265)= 3,128, p = 0,002)$ wynik na skali potrzeby poznania okazał się być dobrym predyktorem zmiennej zależnej. Nie wchodził natomiast w interakcję z interwencjami behawioralnymi.

Podobne zależności zauważono w przypadku modeli E. Zarówno w zyskach ($\chi^2(1) = 5,06, p = 0,024, AIC = 4921, BIC = 4992,8$) jak i stratach ($\chi^2(1) = 6,003, p = 0,014, AIC = 4769,6, BIC = 4841,4$) znaleziono lepsze dopasowanie do danych w porównaniu do modelu bazowego. Zarówno w zyskach $b = 0,47 (t(291)=2,24, p = 0,026)$ jak i w stratach $b = 0,46 (t(291)= 2,44, p = 0,015)$ wynik zdolności numerycznych również okazał się być dobrym predyktorem zmiennej zależnej. Nie wchodził natomiast w interakcję z interwencjami behawioralnymi.

2.4.4 Podsumowanie

Celem badania 2B, była weryfikacja Hipotezy H2.4, dotyczącej trwałości efektów interwencji behawioralnych w czasie. Ważnym aspektem badania było również sprawdzenie, czy efekty oddziaływania interwencji będą utrzymywać się po jej usunięciu z sytuacji decyzyjnej. Przyjmuje się, że jedną z kluczowych charakterystyk badanych interwencji behawioralnych typu nudge, jest odwracalność efektu interwencji po usunięciu interwencji ze środowiska decyzyjnego (Thaler i Sunstein, 2008). Z kolei w przypadku boostingu, najczęściej, zakłada się utrzymywanie efektów po wycofaniu interwencji (Hertwig i Grüne-Yanoff, 2017). Schemat Badania 2B miał zapewnić możliwość sprawdzenia tych przewidywań. Zarówno w przypadku w procedury FFR

jak i PDQ, po pomiarze w którym występowała interwencja, następował tydzień przerwy i ponowny pomiar, prezentujący zadania wyboru już bez obecnej interwencji.

Analiza wykazała, że w przypadku procedury FFR, jedynie modele w domenie zysków przyniosły realną poprawę dopasowania względem modelu bazowego i to na nich skupiono się w analizie efektów głównych i interakcyjnych. Podobnie jak w Badaniu 2A, to interwencja boost jako jedyna okazała się prowadzić do istotnie większej liczby wyborów z wyższą EV niż w grupie kontrolnej. Analiza wykazała, że w drugim pomiarze następującym po tygodniu przerwy, osoby badane osiągały niższe wyniki pod kątem liczby optymalnych wyborów (z wyższą EV) niż w pierwszym pomiarze interwencyjnym. Niższych wyników w drugim pomiarze, nie zaobserwowano jedynie w grupie nudge'a normatywnego. Jeżeli chodzi o różnice indywidualne, włączenie ich jako współzmiennych do testowanych modeli, nie poprawiało w sposób istotny ich dopasowania.

Oprócz procedury FFR, w badaniach zastosowana została również procedura PDQ. Okazało się, że procedura PDQ jest bardziej czuła na wpływ interwencji behawioralnych. Zarówno w domenie strat jak i zysków uzyskano lepsze dopasowania uwzględniając informacji o interwencjach i ich interakcjach ze zmienną czasu pomiaru. W domenie zysków, analiza efektów interakcyjnych grupy boost, wykazała, że osoby badane z tej grupy uzyskały wyższe wyniki niż osoby w grupie kontrolnej zarówno w pomiarze interwencyjnym jak i w pomiarze po-interwencyjnym, tydzień po interwencji. Interpretując wyniki modelu uwzględniającego rodzaj interwencji i czas pomiaru, można byłoby stwierdzić, że po upływie tygodnia od interwencji, osoby badane wciąż były w stanie skutecznie zastosować wyuczoną heurystykę pozwalającą obliczać EV. Jednakże, w wyniku analizy kontrastów, czyli porównań między liczbą optymalnych wyborów w grupie boost a grupą kontrolną tydzień po interwencji,

okazało się, różnica nie była istotna. Wszystkie interwencje poza interwencją boost nie prowadziły do większej liczby optymalnych wyborów w drugim pomiarze. Z kolei w domenie strat, w pomiarze interwencyjnym grupa wystawiona na działanie nudge'a dyskluzyjnego charakteryzowała się większą liczbą wyborów z wyższą EV niż grupa kontrolna, co nie miało miejsca w pomiarze po-interwencyjnym. To wskazywałoby na „odwracalność” efektu interwencji, czyli brak utrzymywania się efektów oddziaływania na wybór po usunięciu interwencji. Należy również zwrócić uwagę na wzrost liczby wyborów o wyższej EV również w pomiarze po-interwencyjnym dla grupy boost. Analiza potencjalnych moderatorów, wykazała, że jedynie w przypadku potrzeby poznania i zdolności numerycznych zaobserwowano lepsze dopasowanie modelu względem modelu bazowego. Zarówno potrzeba poznania jak i zdolności numeryczne okazały się być istotnymi predyktorami liczby optymalnych wyborów, ale nie wchodziły w interakcję z interwencjami. **Wyniki badania 2B pozwoliły na odrzucenie hipotezy H2.4. Zarówno efekty interwencji boost jak i efekty interwencji nudge nie utrzymywały się w czasie.**

2.5 Dyskusja Badań 2A i 2B

Celem Badania 2A i 2B była ocena efektywności oddziaływania interwencji behawioralnych nudge i boost, mierzonej za pomocą liczby wyborów optymalnych rozumianych jako wybór opcji o wyższej wartości oczekiwanej (EV). Ocena objęła również elementy środowiska decyzyjnego i aspekty interwencji: efekt sformułowania wyniku wyboru (czy wybór był określony jako zysk czy strata), transparentność interwencji oraz utrzymywanie się efektów interwencji w czasie. W ramach eksploracji, uwzględniono również potencjalne zmienne różnic indywidualnych, mogących moderować efekty interwencji. Schemat Badania 2A i 2B wykorzystał loterie pieniężne, które można interpretować jako symulację środowiska „małego

świata” (Gigerenzer i in., 2022), gdzie prawdopodobieństwo wystąpienia wszystkich elementów sytuacji decyzyjnej są znane decydentowi.

Otrzymane wyniki w Badaniu 2A wskazują, że wybór konkretnej interwencji mającej skutkować zmianą zachowania, zależy od kontekstu (znaku) dokonywanych decyzji. Podobnie jak w badaniach Franklina i in. (2019) boosty okazały się skutecznym narzędziem do wspierania wyborów w domenie zysków, a nudge w domenie strat. Efekty uzyskano zarówno wykorzystując procedurę FFR zaczerpniętą z badań Franklina i in. (2019) jak i w przypadku wprowadzonej w tej pracy procedury PDQ. Tym samym, uzyskane efekty pozwalają na **przyjęcie Hipotez H2.1 i H2.2**. Efekt ramowania (ang. *framing*) nie jest jeszcze szeroko przebadanym czynnikiem w kontekście omawianych interwencji behawioralnych i wymaga dalszych badań. Do tej pory, wątek ten został poruszony tylko przez kilka badań (Folke i in., 2021; Ruggeri i in., 2021). Folke i in. (2021), który w swoich badaniach również wykorzystywał procedurę FFR, porównał oba typy interwencji, potwierdzając, że boosty są bardziej efektywne w kontekście decyzji ryzykownych, maksymalizując wybory z wyższą wartością oczekiwaną w domenie zysków, natomiast interwencje typu nudge prowadziły do większej liczby wyborów z wyższą wartością oczekiwaną w domenie strat (Folke i in., 2021). Z kolei w badaniach Ruggeriego i in. (2021) boosty okazały się być najbardziej efektywną formą interwencji, niezależnie od tego, czy decyzja dotyczyła domeny zysków czy strat. Rezultaty badania mogą prowadzić do wniosków, że boosty mające na celu umożliwienie ludziom podejmowania niezależnych decyzji zgodnie ze swoimi preferencjami, mogą wykazywać szczególną skuteczność w krajach o niskim poziomie ekonomicznym (Ruggeri i in., 2021).

Wyniki badania 2A pozwalają na **odrzućcie Hipotezy H2.3** dotyczącej wpływu transparentności na efektywność interwencji behawioralnych. Stwierdzono brak istotnego

wpływu transparentności na różnice pomiędzy oddziaływaniami interwencji nudge i boost. Zarówno w przypadku interwencji typu boost jak i nudge, liczba wyborów z wyższą wartością oczekiwaną nie była uwarunkowana świadomością osób badanych nt. interwencji. Wyniki są potwierdzeniem dotychczasowych rezultatów badań nad transparentnością w interwencji typu nudge, przeprowadzanych w większości na nudge'ach opcji domyślnej (de Ridder, 2023). Dodatkowo, rozszerzają te wyniki o inne, dotychczas pomijane interwencje typu nudge dyskluzywny i nudge normatywny, uwzględniając domenę wyborów pieniężnych. Ważne jest również rozszerzenie efektów na interwencje typu boost, które do tej pory, mimo założeń o braku roli transparentności w ich efektywności (Hertwig i Grüne-Yanoff, 2017), nie były zweryfikowane empirycznie.

W badaniu 2B również wykorzystano procedury FFR i PDQ. W przeciwieństwie do Badania 2A, w Badaniu 2B, modele zbudowane w oparciu o dane z procedury FFR, pozwoliły na analizę efektów głównych i interakcyjnych jedynie w domenie zysków. Interwencja typu boost prowadziła do istotnie wyższej liczby wyborów o wyższej wartości oczekiwanej (EV) w porównaniu z grupą kontrolną, co pokrywa się z wynikami z Badania 2A. Jednakże, pomimo początkowego zwiększenia liczby wyborów optymalnych (o wyższym EV) w pomiarze interwencyjnym, wyniki w drugim pomiarze po tygodniu ujawniły zmniejszenie liczby optymalnych wyborów, co jedynie nie dotyczyło grupy poddanej nudge'owi normatywnemu. Częściowo, uzyskane wyniki są zbliżone do tych odnotowanych w przypadku procedury FFR w badaniu 2A, gdzie również interwencja typu boost prowadziła do najwyższej liczby wyborów o wyższym EV w domenie zysków. Z kolei wyniki uzyskane w ramach procedury PDQ były bardziej wrażliwe, pokazując poprawę dopasowania względem modeli bazowych w obu domenach: zysków i strat. W domenie zysków, grupa boost charakteryzowała się większą liczbą

wyborów optymalnych (o wyższym EV) zarówno w pomiarze interwencyjnym, jak i po-
interwencyjnym, sugerując utrzymanie się efektu interwencji przez co najmniej tydzień.
Jednakże, analiza kontrastów wykazała, że istotność różnic między grupą boost a kontrolną
zanikała z czasem, co może wskazywać na ograniczoną trwałość efektu interwencji. W domenie
strat, jedynie grupa nudge'a dyskluzywnego wykazała wyższe wyniki wyborów optymalnych niż
grupa kontrolna w pomiarze interwencyjnym w porównaniu z grupą kontrolną, co nie utrzymało
się w pomiarze po-interwencyjnym. Wyniki sugerują, że szczególnie efekty interwencji typu
nudge mogą być odwracalne, a ich trwałość jest ograniczona po usunięciu bodźca
interwencyjnego. W przypadku interwencji boost można zaobserwować pewną tendencję do
utrzymywania się efektów interwencji, jednakże są one zbyt słabe i nieistotne.

Rezultaty Badania 2B pozwalają na **odrzućcie Hipotezy H2.4**, zakładającej
utrzymywanie się efektów interwencji w czasie. Choć w przypadku interwencji typu boost można
było zaobserwować pewną tendencję, zarówno w przypadku procedury Franklina jak i procedury
PDQ, oddziaływania interwencji typu nudge i boost zanikały po tygodniu od ekspozycji na
interwencję. Otrzymane wyniki są sprzeczne z założeniami interwencji boost (Hertwig i Grüne-
Yanoff, 2017), ale jednocześnie zgodne z „odwracalnością” efektów interwencji typu nudge
(Sunstein i Reisch, 2023). Efekty nudgingu wedle założeń autorów powinny zanikać w czasie. W
przypadku interwencji boost, w przyszłych badaniach nad utrzymywaniem się efektów w czasie,
warto byłoby wydłużyć czas treningu jaki osoby badane spędzały na przyswajanie heurystyki.
Niewykluczone też, że w przypadku innych sytuacji decyzyjnych i przekazywanych interwencji,
efekty boostów utrzymywałyby się dłużej. Wydaje się to tym bardziej zasadne, że jedne z
najnowszych badań porównujących efekty nudge'y i boostów wskazało na utrzymujące się
działanie boostów w czasie na przestrzeni siedmiu miesięcy (Paunov i Grüne-Yanoff, 2023).

W badaniach Paunova i Grüne-Yanoffa (2023) okazało się, że w warunku boost, osoby badane zużywały istotnie mniej energii elektrycznej i wody niż w warunku nudge. Jak podkreślają autorzy, efekt boostów pogłębiał się w czasie, co może wskazywać na dłuższy okres czasu niezbędny do rozwinięcia się pełnego efektu interwencji (Paunov i Grüne-Yanoff, 2023).

Badania 2A i 2B miały również ważny komponent eksploracyjny w postaci roli różnic indywidualnych jako potencjalnych moderatorów efektywności interwencji nudge i boost. Istotne efekty moderacyjne interwencji zostały zauważone jedynie w przypadku procedury PDQ. Może to oznaczać, że procedura PDQ była bardziej wrażliwa na obserwowane zmiany w zachowaniu. Procedura charakteryzowała się większą liczbą wyborów dokonywanych przez osoby badane oraz zróżnicowaniem wykorzystywanych poziomów prawdopodobieństwa i kwot wykorzystywanych – procedura FFR była krótsza i mniej zróżnicowana pod kątem poziomów ryzyka i kwot. Mimo to, warto zwrócić uwagę na te potencjalne efekty moderacyjne, gdyż wskazują, że mają podłoże bardziej osobowościowo-temperamentalne niż poznawcze. Efekty boostów i nudge’y w procedurze PDQ wchodziły w interakcję z takimi cechami jak sumienność, otwartość na doświadczenie czy ugodowość. Osoby o wyższych wynikach na tych skalach chętniej korzystały z rozwiązań proponowanych przez te interwencje. Podobne rezultaty w przypadku sumienności zanotowali Kamilçelebi i Songül (2020), którzy wskazywali, że osoby o wyższych wynikach na skali sumienności chętniej podążały za oddziaływaniami typu nudge opcji domyślnej. Interwencje behawioralne wchodziły również w interakcję z takimi cechami temperamentalnymi jak zwawość, wytrzymałość czy perseweratywność. Choć były to pierwsze badania uwzględniające cechy temperamentalne w badaniach nad interwencjami behawioralnymi, dotychczasowe projekty badawcze poruszające Regulacyjną Teorię Temperamentu w decyzjach finansowych, wskazywały już na istotną rolę perseweratywności

(Gąsiorowska, 2011). Interpretując wyniki na skalach osobowości i temperamentu, wyniki wskazują potencjalnie na profil osób, które w dążeniach do maksymalizacji korzyści płynących z decyzji finansowych otwarcie sięgają po wsparcie swoich procesów decyzyjnych. Brak interakcji ze zmiennymi potrzeby poznania czy umiejętności numerycznych może wskazywać jednak na podłoże bardziej emocjonalne czy motywacyjne niż chłodne analityczne, które potencjalnie warto byłoby eksplorować w kolejnych badaniach nad czynnikami mediującymi i moderującymi działanie interwencji. Takie rezultaty nie zostały zaobserwowane ani w procedurze FFR ani w Badaniu 2B, co wskazuje nie konieczność dalszej eksploracji tych czynników w badaniach nad interwencjami behawioralnymi. Wedle wiedzy autora, były to pierwsze badania uwzględniające czynniki osobowościowe i temperamentalne w ocenie efektywności interwencji nudge'y dyskluzyjnych i normatywnych oraz interwencjach boost.

2.6. Badania 3A i 3B – zarys ogólny

2.6.1 Metoda ogólna badań 3A i 3B

Badanie 3 składało się z dwóch części, których celem była ocena efektywności nudge'y i boostów za pomocą paradygmatu wyboru dyskretnego (ang. *discrete choice experiment*; DCE). Schemat Badania 3A uwzględniał pięć warunków badawczych: cztery interwencje (dwie interwencje typu nudge i dwie interwencje typu boost) oraz grupa kontrolna. Na podstawie Badania 3A, wyłoniono najbardziej efektywną interwencję typu boost i typu nudge, które zastosowano w badaniu 3B. Dodatkowo w badaniu 3B wprowadzono warunek badawczy w którym zaimplementowano je jednocześnie by zbadać efekt synergii. Zarówno w badaniu 3A jak i 3B, części behawioralnej (DCE) dokonano pomiaru różnic indywidualnych, w celu zbadania potencjalnych moderatorów oddziaływania interwencji.

Przeprowadzone badanie opierało się na schemacie wyboru dyskretnego (Discrete Choice Experiment, DCE) z wykorzystaniem bayesowskich schematów D- optymalnych (Kessels i in., 2011), zrealizowanych za pomocą pakietu *idefix* w języku R (Traets i in., 2020). Uczestnicy byli proszeni o dokonanie 35 wyborów, określając preferencje dotyczące jednego z dwóch programów oszczędnościowego, w których każda opcja była opisana przez dziewięć atrybutów decyzyjnych. Oprócz nich, w oddzielnych warunkach zostały również wprowadzone interwencje boost i nudge. Do badania zostały wybrane dwie interwencje typu boost: boost Reguła Franklina (Mata i in., 2010) i boost soczewkowy (Hertwig i Grüne-Yanoff, 2017) oraz dwie interwencje typu nudge: nudge normatywny (Franklin i in., 2019; Sunstein, 2014) oraz nudge dyskluzywny (Franklin i in., 2019; Sunstein, 2014). Piątym warunkiem była grupa kontrolna, w której osobom badanym zaprezentowano atrybuty decyzyjne bez dodatkowej interwencji.

Każda interwencja behawioralna opierała się na innym mechanizmie oddziaływania. Zasadą działania interwencji boost soczewkowy było wyposażenie uczestników badania w prostą heurystykę (ang. *Simple Heuristic*; Gigerenzer i Todd, 1999), która miała za zadanie ułatwić wybór programu zachęcając osoby badane do koncentrowania się na jednym, początkowo neutralnym atrybucie decyzyjnym. W tym celu do badania został wprowadzony atrybut, stworzonej na potrzeby badania „organizacji rekomendującej”. Atrybut ten odnosił się do nieistniejącej, prestiżowej organizacji specjalizującej się w ocenie programów oszczędnościowych na całym świecie. Z materiałów poprzedzających główną część procedury badawczej, uczestnicy zostali poinformowani, że podczas dokonywania wyboru, w celu optymalizacji wyboru, mogą ograniczyć się do podążania za rekomendacjami organizacji WBW. Interwencja została zaadaptowana ze współautorskich badań nad interwencjami behawioralnymi w sytuacji szczepień (Marcowski i in., w druku)

Podobny punkt odniesienia otrzymała interwencja nudge'a normatywnego. W tym warunku, w każdym wyborze uczestnicy mogli zobaczyć dopisek „*Większość osób wybierała tę opcję*”, który towarzyszył zawsze temu programowi w którym poziomem atrybutu była organizacja WBW. Mechanizm działania interwencji nawiązywał do deskryptywnej normy społecznej, opisującej zachowanie większości osób w danej sytuacji (Sunstein, 2014).

W przypadku kolejnych warunków badawczych, boost wg Reguły Franklina oraz nudge dyskluzyjnego wartość każdego z programów opierała się na wartości uzyskiwanej w rezultacie sumy dwóch atrybutów: racie dopłacanej przez pracodawcę oraz racie dopłacanej przez Państwo. Dodatkowo, atrybuty posiadały różne wagi. Pełny wzór służący do obliczenia wartości programu wyglądał następująco:

$$\text{Wartość programu: Rata pracodawcy} \times 100 + \text{Rata Państwa} \times 50$$

Wagi atrybutów zostały dobrane w oparciu o proporcjonalny udział wpłat Państwa i pracodawcy w realnych programach oszczędnościowych takich jak Pracownicze Plany Kapitałowe (np. Jedynak, 2018), imitujące realne ratalne proporcje pensji wpływające na konto oszczędzających. W warunku boost wg Reguły Franklina, uczestnikom prezentowano powyższą formułę obliczania wartości programu, rekomendowaną przez stworzony na potrzeby badania autorytet badawczy w dziedzinie oszczędzania. Dodatkowo, przed przystąpieniem do głównej części procedury uwzględniającej serię wyborów, osoby badane w warunku boostu wg Reguły Franklina miały za zadanie poprawnie wybrać program cechujący się wyższą wartością programu (zgodnie z powyższą formułą) w trzech wyborach treningowych. Podczas wszystkich wyborów tej grupie, uczestnicy mogli używać kalkulatora.

Mechanizm działania interwencji typu nudge dyskluzywnego był podobny do interwencji boost Reguły Franklina – kierunek interwencji wyznaczała wartość programu obliczona w oparciu o atrybuty raty Państwa i pracodawcy. Różnica polegała na tym, że nudge dyskluzywny wprost informował uczestników o wartości programu. W tym warunku badawczym, osoby badane nie dowiadywały się o formule obliczania wartości programu, a przy okazji każdego wyboru prezentowano im wartość programu A i B. Zasady działania opisanych interwencji oraz ich wykorzystanie w Badaniu 3A i Badaniu 3B zostały podsumowane w Tabeli 9.

Badania przeprowadzono online za pomocą aplikacji RShiny (Chang i in., 2015) i platformy Qualtrics (Qualtrics, Provo, UT; qualtrics.com). Uczestnicy wyrazili świadomą zgodę, a badanie przestrzegało wytycznych etycznych, zatwierdzonych przez Komisję Etyki Wydziału Psychologii w Warszawie Uniwersytetu SWPS. Rekrutacja osób badanych odbywała się za pośrednictwem systemu SONA Uniwersytetu SWPS, a osoby badane otrzymywały za udział w badaniu uczelniane punkty za aktywność.

Tabela 9. Porównanie interwencji stosowanych w badaniu 3A i 3B

Rodzaj interwencji	Charakterystyka działania	Sposób prezentacji interwencji	Nośnik interwencji/ kierunek interwencji	Obecna w bad. 3A	Obecna w bad. 3B
Boost Soczewkowy (BS)	Osobom badanym przekazywano prostą heurystykę opierającą się na wyborze programu w którym pojawiała się rekomendacja prestiżowej agencji WBW	Nazwa agencji WBW pojawiająca się jako poziom atrybutu „Organizacja rekomendująca”; instrukcja informująca o organizacji WBW pojawiała się przed częścią behawioralną	Organizacja WBW – poziom atrybutu org. rekomendującej	X	-
Boost wg. Reguły Franklina (BRF)	Osobom badanym przekazywano prostą heurystykę opierającą się na stosowaniu reguły obliczania wartości programu uwzględniającą dwa atrybuty: ratę Państwa i ratę pracodawcy	Dwa atrybuty decyzyjne mogły przyjąć różne poziomy; instrukcja informująca o wzorze na obliczenie wartości programu pojawiała się przed częścią behawioralną	Reguła Franklina – wartość programu składająca się z dwóch atrybutów	X	X
Nudge Dyskluzywny (ND)	Podanie do wiadomości osób badanych wartości programów bez wprowadzenia do reguły obliczania	Prezentacja wartości programów poniżej atrybutów decyzyjnych w każdym wyborze	Reguła Franklina – wartość programu składająca się z dwóch atrybutów	X	-
Nudge Normatywny (NN)	Wprowadzenie wiadomości o preferowanym programie przez większość osób badanych; programem podkreślanym przez wiadomość był ten zawierający organizację WBW w atrybutcie org. rekomendującej; wiadomość odwoływała się do deskryptywnej normy społecznej	Prezentacja wiadomości „Większość osób wybierała tę opcję” poniżej atrybutów decyzyjnych	Organizacja WBW – poziom atrybutu org. rekomendującej	X	X
B. wg. Reguły Franklina x Nudge Normatywny	Połączenie interwencji BRF i NN; wiadomość nawiązująca do normy społecznej podkreślała program z większą wartością według reguły Franklina	Instrukcja przed częścią behawioralną informująca o regule Franklina i wiadomość „Większość osób wybierała tę opcję” poniżej atrybutów decyzyjnych	„Większość osób wybierała tę opcję” poniżej atrybutów decyzyjnych	-	X

2.7 Badanie 3A *Efektywność interwencji nudge i boost w zmianie zachowań oszczędnościowych*

2.7.1 Osoby badane

W Badaniu 3A, wzięło udział 672 osoby które zostały losowo przydzielone do jednego z pięciu warunków badawczych: 1) grupa nudge dyskluzywny (129 osób), 2) grupa nudge normatywny (133 osoby), 3) grupa boost Soczewkowy (134 osób), 4) grupa boost reguła Franklina (136 osoby), 5) grupa kontrolna (140 osób). Zgodnie z wyliczeniami analizy mocy według Johnson i Orme (2010) dla schematu decyzji wieloatrybutowych, minimalna liczba osób dla warunku badania wynosiła 114 osób (w sumie 570 dla pięciu warunków decyzyjnych). Liczebność ta została zwiększona o 15% ze względu na potencjalną konieczność wykluczenia przypadków odstających, z uwagi na zbyt krótki czas odpowiedzi.

2.7.2 Narzędzia/Procedura

W każdej sytuacji decyzyjnej, osobom badanym prezentowano wybór pomiędzy dwoma programami oszczędnościowymi (Rycina 8), różniącymi się na poziomie dziewięciu atrybutów decyzyjnych (Tabela 10) Dodatkowo, w warunkach nudge'a normatywnego (Rycina 8, Panel B) i nudge'a dyskluzywnego (Rycina 8, Panel C), bezpośrednio pod oferowanymi programami pojawiały się informacje będące częścią działania interwencji. W sumie, każda osoba uczestnicząca w badaniu, dokonywała 35 wyborów. W każdym wyborze prezentowany był inny zestaw poziomów atrybutów, a każda osoba otrzymywała losową kolejność atrybutów w celu przeciwdziałania potencjalnym artefaktom związanym z kolejnością prezentacji bodźców

Panel A

Wybór: 1 / 35

	Program A	Program B
Liczba Twoich znajomych którzy uczestniczą w programie oszczędnościowym	30% z Twoich znajomych obecnie korzysta z programu oszczędnościowego	10% z Twoich znajomych obecnie korzysta z programu oszczędnościowego
Otrzymałeś/aś rekomendację do uczestniczenia w tym programie od	Bliski znajomy/Ktoś z rodziny	Doradca finansowy z banku
Czy w umowie z pracodawcą jest przewidziana podwyżka?	Zaplanowana podwyżka w wysokości 10% pensji po 6 miesięcznym okresie próbnym	Zaplanowana podwyżka w wysokości 10% pensji po 6 miesięcznym okresie próbnym
Organizacja rekomendująca	WBW	DEOC
Jaką kwotę dokładałby pracodawca?	1% miesięcznej pensji	2% miesięcznej pensji
Ile odkładał(a)bym miesięcznie od swojej pensji?	5%	1%
Jaką kwotę dokładałoby Państwo?	2% miesięcznej pensji	1% miesięcznej pensji
Moment rozpoczęcia oszczędzania	Natychmiast w momencie rozpoczęcia pracy	3 miesiące od rozpoczęcia pracy w danej firmie
Aktualne oszczędności na Twoim koncie	Ponad 10 tys. złotych	Mniej niż 1000 zł

W której sytuacji jest bardziej prawdopodobne, że zapisał(a)byś się do programu oszczędnościowego? Wybór należy do Ciebie:

Program A Program B

OK

Panel B

Wybór: 1 / 35

	Program A	Program B
Liczba Twoich znajomych którzy uczestniczą w programie oszczędnościowym	30% z Twoich znajomych obecnie korzysta z programu oszczędnościowego	10% z Twoich znajomych obecnie korzysta z programu oszczędnościowego
Aktualne oszczędności na Twoim koncie	Ponad 10 tys. złotych	Mniej niż 1000 zł
Organizacja rekomendująca	WBW	DEOC
Otrzymałeś/aś rekomendację do uczestniczenia w tym programie od	Bliski znajomy/Ktoś z rodziny	Doradca finansowy z banku
Moment rozpoczęcia oszczędzania	Natychmiast w momencie rozpoczęcia pracy	3 miesiące od rozpoczęcia pracy w danej firmie
Czy w umowie z pracodawcą jest przewidziana podwyżka?	Zaplanowana podwyżka w wysokości 10% pensji po 6 miesięcznym okresie próbnym	Zaplanowana podwyżka w wysokości 10% pensji po 6 miesięcznym okresie próbnym
Ile odkładał(a)bym miesięcznie od swojej pensji?	5%	1%
Jaką kwotę dokładałby pracodawca?	1% miesięcznej pensji	2% miesięcznej pensji
Jaką kwotę dokładałoby Państwo?	2% miesięcznej pensji	1% miesięcznej pensji

Większość osób wybierała tę opcję

W której sytuacji jest bardziej prawdopodobne, że zapisał(a)byś się do programu oszczędnościowego? Wybór należy do Ciebie:

Program A Program B

OK

Panel C

Wybór: 1 / 35

	Program A	Program B
Organizacja rekomendująca	WBW	DEOC
Czy w umowie z pracodawcą jest przewidziana podwyżka?	Zaplanowana podwyżka w wysokości 10% pensji po 6 miesięcznym okresie próbnym	Zaplanowana podwyżka w wysokości 10% pensji po 6 miesięcznym okresie próbnym
Liczba Twoich znajomych którzy uczestniczą w programie oszczędnościowym	30% z Twoich znajomych obecnie korzysta z programu oszczędnościowego	10% z Twoich znajomych obecnie korzysta z programu oszczędnościowego
Moment rozpoczęcia oszczędzania	Natychmiast w momencie rozpoczęcia pracy	3 miesiące od rozpoczęcia pracy w danej firmie
Jaką kwotę dokładałby pracodawca?	1% miesięcznej pensji	2% miesięcznej pensji
Ile odkładał(a)bym miesięcznie od swojej pensji?	5%	1%
Aktualne oszczędności na Twoim koncie	Ponad 10 tys. złotych	Mniej niż 1000 zł
Jaką kwotę dokładałoby Państwo?	2% miesięcznej pensji	1% miesięcznej pensji
Otrzymałeś/aś rekomendację do uczestniczenia w tym programie od	Bliski znajomy/Ktoś z rodziny	Doradca finansowy z banku

Wartość tego programu to: 200

Wartość tego programu to: 250

W której sytuacji jest bardziej prawdopodobne, że zapisał(a)byś się do programu oszczędnościowego? Wybór należy do Ciebie:

Program A Program B

OK

Rycina 8. Przykładowe sytuacje wyboru w poszczególnych warunkach badawczych: Panel A)

Grupa kontrolna, boost Reguła Franklina i boost soczewkowy; Panel B) Nudge'a Normatywny;

Panel C) Nudge Dyskluzywny. Dodatkowe instrukcje poprzedzające wybory decyzyjne w warunkach boostów zostały zawarte w suplemencie online na stronie OSF.io: <https://osf.io/n2hcj/>

Tabela 10. Atrybuty decyzji wraz ze swoimi poziomami wykorzystane w badaniach 3A i 3B

Nazwa atrybutu	Poziomy atrybutu
1) Jaką kwotę dokładałoby Państwo?	0% (Brak udziału Państwa); 1%, 2%, 3% miesięcznej pensji
2) Jaką kwotę dokładałby pracodawca?	1%; 2%, 5% miesięcznej pensji
3) Ile odkładał(a)bym miesięcznie od swojej pensji?	1%; 2%, 5% miesięcznej pensji
4) Liczba Twoich znajomych którzy uczestniczą w programie oszczędnościowym	Nikt z Twoich znajomych obecnie nie korzysta z programu oszczędnościowego; 10% z Twoich znajomych obecnie korzysta z programu oszczędnościowego; 30% z Twoich znajomych obecnie korzysta z programu oszczędnościowego
5) Moment rozpoczęcia oszczędzania	Natychmiast w momencie rozpoczęcia pracy; 3 miesiące od rozpoczęcia pracy w danej firmie; 6 miesięcy od rozpoczęcia pracy w danej firmie
6) Czy w umowie z pracodawcą jest przewidziana podwyżka?	Brak planowanej podwyżki; Zaplanowana podwyżka w wysokości 10% pensji po 3 miesięcznym okresie próbnym; Zaplanowana podwyżka w wysokości 10% pensji po 6 miesięcznym okresie próbnym
7) Aktualne oszczędności na Twoim koncie	Mniej niż 1000 zł; 1000-5000 zł; 5000-10 tys zł; Ponad 10 tys. złotych
8) Otrzymałeś/aś rekomendację do uczestniczenia w tym programie od	Bliski znajomy/Ktoś z rodziny; Doradca finansowy z banku; Pracownik ministerstwa finansów
9) Organizacja rekomendująca	DEOC; WBW

Dodatkowo, w badaniu uwzględniono serię narzędzi mierzących różnice indywidualne, potencjalnie różnicujące osoby badane podczas wyboru programów oszczędnościowych. W celu potwierdzenia dotychczas opisywanych zależności między perspektywami czasowymi a zachowaniami oszczędnościowymi (Jacobs-Lawson i Hershey, 2005) wprowadzono do badania Kwestionariusz Perspektywy Czasowej Zimbardo w polskiej adaptacji (Przepiorka i in., 2016).

Ważnym aspektem była również weryfikacja związku skłonności do ryzyka z efektywnością interwencji behawioralnych. Ze względu na podskale dot. zachowań finansowych, wybrano polską adaptację kwestionariusza DOSPERT mierzącego ryzyko domenowe (Blais i Weber, 2006). Ostatnią procedurą mierzącą różnice indywidualne była procedura pomiaru tempa dyskontowania w odroczeniu dla strat i zysków (*5-trial discounting task*; Koffarnus i Bickel, 2014), które jest wskazywane jako ważny element zachowań oszczędnościowych (Ersner-Hershfield i in., 2009). Dodatkowo, włączono zestaw pytań demograficznych dotyczących płci, wieku, miesięcznych zarobków i posiadanych oszczędności.

W celu zbadania preferencji względem atrybutów decyzyjnych oraz wpływu interwencji na wybór programu oszczędnościowego dane w badaniu zostały przeanalizowane w środowisku R (R Core Team, 2022), wykorzystując uogólnione liniowe modele mieszane o rozkładzie dwumianowym z efektem losowym na poziomie uczestnika (Hoogink i in., 2020). Do obliczeń modeli mieszanych został wykorzystany pakiet glmmTMB (Brooks i in., 2023). Zmienna zależna reprezentowała wybór między scenariuszem A i B, a predyktorami były atrybuty scenariuszy (Tabela 10) oraz zmienne warunków eksperymentalnych (cztery interwencje behawioralne i grupa kontrolna). Aby zapewnić jakość danych, usunięto wybory o średnim czasie reakcji mniejszym niż pięć sekund na jeden wybór. Punkt odcięcia został przyjęty w oparciu o literaturę opisującą przeciętną prędkość czytania ze zrozumieniem (Rayner i in., 2016) oraz minimalną liczbę słów jaką osoba badana musiała świadomie przeczytać, by móc podjąć świadomy, poinformowany wybór. Według badań m.in. Just i Carpenter (1987; za Rayner i in., 2016) przeciętna prędkość czytania ze zrozumieniem wynosi około 250 słów na minutę, co można sprowadzić do czterech słów na sekundę. Sugerując się literaturą przedmiotu oraz mając na uwadze, że trzy prezentowane atrybuty wraz z poziomami składają się przynajmniej z dwudziestu

słów, pięć sekund zostało przyjęte jako minimalny czas wystarczający by zapoznać się z niezbędnymi informacjami do podjęcia decyzji w warunkach proponujących najkrótszy proces wyboru docelowego programu: nudge'a normatywnego i boost'a soczewkowego oraz kliknięcie w przycisk zatwierdzający wybór. Choć w tych dwóch warunkach potencjalnie można byłoby sprowadzić zapoznanie się z sytuacją decyzyjną tylko do jednego atrybutu (Organizacji rekomendującej w warunku boost soczewkowy lub obserwację za czym optowała większość w warunku nudge'a normatywnego), oznaczałoby to wówczas nie zapoznanie się wcale z innymi atrybutami i zgodnie z literaturą przedmiotu (np. Gigerenzer i Todd, 1999) nie byłoby to postrzegane jako adaptacyjne. Biorąc to pod uwagę, okrojenie średniego minimalnego czasu do 5 sekund na wybór można rozpatrywać jako konserwatywny i arbitralny wybór metodologiczny, miało natomiast na celu przede wszystkim wykluczenie osób, które bez czytania i zapoznania się z atrybutami wybierały jakąkolwiek opcję, by jak najszybciej zakończyć badanie i otrzymać punkty. Tendencja do czytania poleceń bez zrozumienia, jest zauważalną tendencją wśród osób badanych rekrutowanych poprzez platformy internetowe, co może stanowić zagrożenie dla jakości danych (Douglas i in., 2023). Wykluczenie osób poniżej średniego czasu pięciu sekund oznaczało zmniejszenie próby badawczej do 579 osób badanych, będąc wciąż powyżej minimalnej próby uzyskanej w ramach analizy mocy (570).

Analiza odbywała się w dwóch etapach: najpierw dopasowano model do danych oddzielnie dla każdego warunku, aby obliczyć względną ważność atrybutów dla podejmowanej decyzji. Obliczono ją na podstawie testów ilorazu wiarygodności, a względną ważność każdego atrybutu za pomocą znormalizowanej statystyki LogWorth ($-\log_{10}(p)$), gdzie p oznacza poziom istotności otrzymany z testów ilorazu wiarygodności. Następnie model został dopasowany ponownie, obejmując wszystkie atrybuty i warunki. Dodatkowo, dodano analizę wpływu interwencji

behawioralnych, a atrybuty interwencji zakodowano jako zmienną binarną. W oddzielnej sekcji, uwzględniono serię modeli uwzględniających różnice indywidualne.

2.7.3 Wyniki

Pierwszą przeprowadzoną analizą była analiza relatywnej ważności atrybutów dla podejmowanej decyzji. Relatywna ważność poszczególnych atrybutów różniła się pomiędzy warunkami badawczymi co zostało zaprezentowane na Rycinie 9. Należy podkreślić, że statystyka LogWorth została znormalizowana, co oznacza, że najważniejszy atrybut w każdym warunku osiągnął po normalizacji wartość 1, a wartość każdego następnego atrybutu została ujęta relatywnie w odniesieniu do najważniejszego atrybutu w hierarchi. W warunku kontroli, wśród trzech najważniejszych atrybutów pojawiła się planowana podwyżka (RI=1,00), aktualne środki na koncie (RI=0,54) oraz rata Państwa (RI=0,52). Tuż za ratą Państwa znalazła się rata pracodawcy (RI=0,51). Warto podkreślić, że wyniki zostały uzyskane oddzielnie dla danych z warunku kontroli, czyli przedstawiona hierarchia atrybutów funkcjonowała przy braku oddziaływania jakiegokolwiek interwencji. Oznacza to, że przy braku oddziaływania interwencji, osoby badane w pierwszej kolejności preferowały opierać się na informacjach zawartych w atrybucie planowanej podwyżki, a na ich wyborze informacje zawarte w atrybutach aktualnych środków na koncie, oraz rat Państwa i pracodawcy ważyły prawie dwukrotnie mniej.

W porównaniu do grupy kontrolnej, w warunkach, które informowały o wartości programów oszczędnościowych (boost wg Reguły Franklina oraz nudge dyskluzyny) hierarchia atrybutów uległa zmianie, w czołówce atrybutów plasowały się raty dopłacane przez pracodawcę i Państwo oraz jako trzecia pozycjonowała się planowana podwyżka. W warunku boost wg Reguły Franklina (BRF) najważniejsza była rata pracodawcy (RI=1,00), tuż za nią uplasowała się rata Państwa (RI=0,98), a trzecia w kolejności znalazła się planowana podwyżka (RI=0,45). W

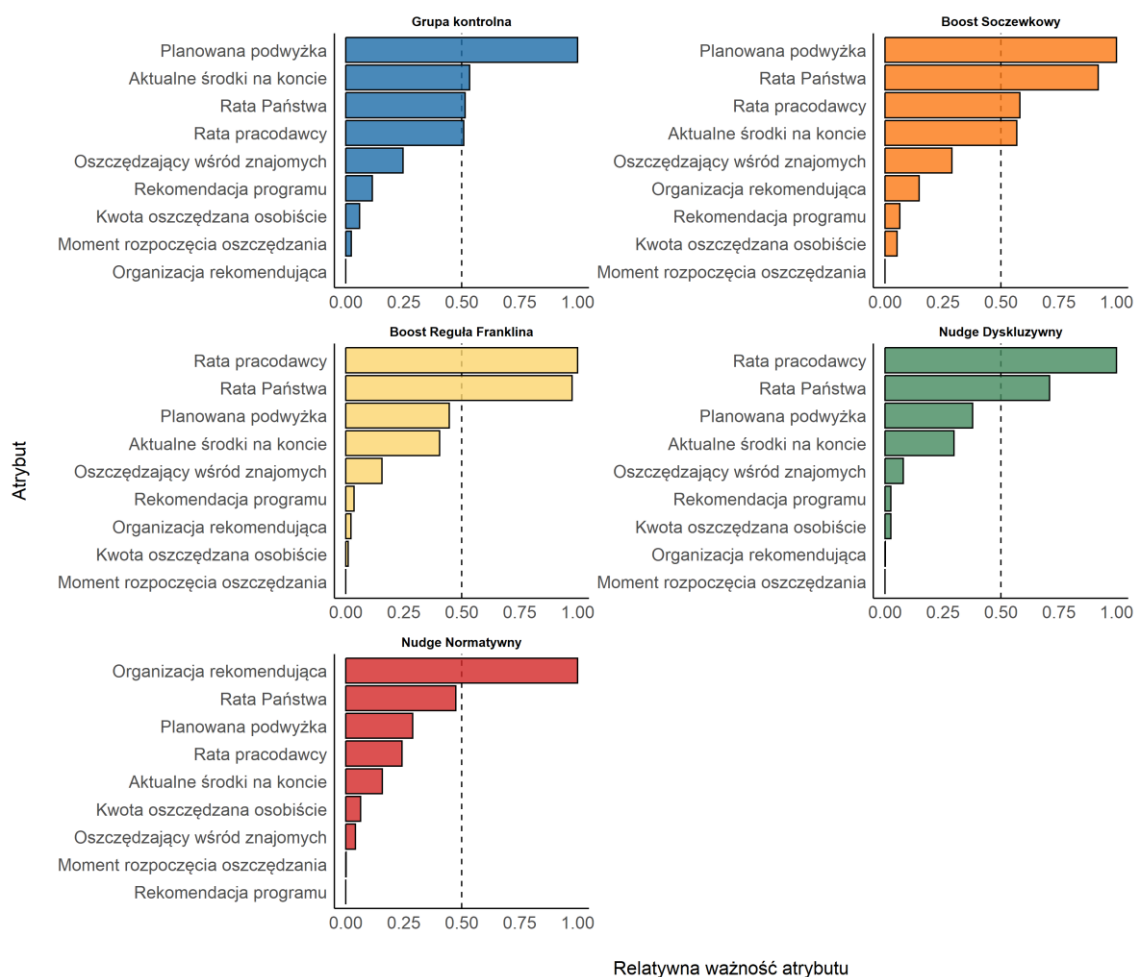
warunku nudge'a dyskluzyjnego (ND) największą rolę odgrywała również rata pracodawcy (RI=1,00), a kolejnymi była rata Państwa (RI= 0,71) oraz planowana podwyżka (RI=0,38).

Innymi słowy, dwa atrybuty wskazywane przez interwencje BRF i NN jako najważniejsze, czyli rata pracodawcy oraz rata Państwa, mające kolejno największą wagę w obliczeniu wartości oczekiwanej programu, znalazły odzwierciedlenie swoich wag w hierarchi atrybutów w tych warunkach. Jako trzeci atrybut plasowała się planowana podwyżka, czyli atrybut preferowany najmocniej, przy braku występowania jakichkolwiek innych interwencji w grupie kontrolnej.

W warunku boost Soczewkowy, wśród trzech pierwszych atrybutów znalazła się planowana podwyżka (RI=1,00), rata Państwa (RI=0,92) oraz rata pracodawcy (RI=0,58). Choć atrybut organizacji rekomendującej (RI=0,15) był w tym warunku przedmiotem interwencji, nie znajdował się wśród najważniejszych atrybutów, jak w warunku nudge'a normatywnego. Mimo to, warto zwrócić uwagę, że znalazł się zauważalnie wyżej w hierarchi atrybutów niż w grupie kontrolnej, boostu Reguły Franklina i nudge'a dyskluzyjnego. Dodatkowo, jako jedyny spośród interwencji, boost soczewkowy nie zmienił dominacji atrybutu planowanej podwyżki, zaobserwowanego na pierwszym miejscu również w grupie kontrolnej.

W warunku nudge'a normatywnego, pierwsze trzy miejsca w hierarchi atrybutów zajęły następujące cechy: organizacja rekomendująca (RI=1,00), rata Państwa (RI=0,48) oraz planowana podwyżka (RI=0,29). Jedynie w tym warunku została zaobserwowana znacząca waga atrybutu organizacji rekomendującej. Jest to związane z tym, że informacja przekazywana przez interwencję nudge'a normatywnego współwystępowała zawsze z pojawianiem się organizacji WBW w atrybucie organizacji rekomendującej. W porównaniu do grupy kontrolnej, jest to znaczącą zmianą hierarchi – atrybut planowanej podwyżki waży jedynie 30% atrybutu organizacji rekomendującej.

Relatywna ważność atrybutów dla podejmowanej decyzji (Badanie 3a)



Rycina 9. Relatywna ważność poszczególnych atrybutów dla podejmowanej decyzji o przystąpieniu do programu oszczędnościowego

W celu identyfikacji wpływu użytych w badaniu atrybutów na wybór programu oszczędnościowego, ocenie poddano parametry modeli oddzielnie dla danych z grupy kontrolnej (Tabela 11) oraz całych danych uwzględniających grupę kontrolną i cztery warunki zawierające interwencje (Tabela 12). Jak można zauważyć w Tabeli 11, w grupie kontrolnej prawie wszystkie atrybuty decyzji dot. programów oszczędnościowych okazały się istotnie wpływać na wybory

osób badanych. Jedynie nieistotny okazał się atrybut organizacji rekomendującej ($\chi^2(1) = 2,07$, $p=0,150$). Z kolei w modelu obliczonym na całych danych (Tabela 12) wszystkie atrybuty decyzyjne okazały się istotne włącznie z atrybutem organizacji rekomendującej ($\chi^2(1) = 176,69$, $p < 0,001$)

Tabela 11. Współczynniki atrybutów (test χ^2 Walda) wraz oszacowanymi parametrami modelu na danych grupy kontrolnej. Parametry modelu zostały przedstawione wraz z odpowiadającym im ilorazem szans (OR) oraz przedziałami ufności (CI), które reprezentują wpływ różnych atrybutów na decyzje dotyczące wyboru programu oszczędnościowego.

Parametr	χ^2	df	OR	95% CI	z	p
Model bazowy (Przecięcie)			0,15	[0,12; 0,19]	-16,96	< 0,001
Rata Państwa	121,26	3	-	-	-	< 0,001
0% (Brak udziału Państwa; pkt odniesienia)	-	-	-	-	-	-
1% miesięcznej pensji			1,35	[1,19; 1,53]	4,61	0,005
2% miesięcznej pensji			1,79	[1,59; 2,02]	9,41	< 0,001
3% miesięcznej pensji			2,04	[1,74; 2,40]	8,66	< 0,001
Rata pracodawcy	115,93	2	-	-	-	< 0,001
1% miesięcznej pensji (pkt odniesienia)	-	-	-	-	-	-
2% miesięcznej pensji			1,47	[1,31; 1,65]	6,48	< 0,001
5% miesięcznej pensji			1,92	[1,71; 2,17]	10,59	< 0,001
Rata własna	16,94	2				< 0,001
1% miesięcznej pensji (pkt odniesienia)	-	-	-	-	-	-
2% miesięcznej pensji			1,17	[1,05; 1,31]	2,79	0,005
5% miesięcznej pensji			1,27	[1,13; 1,43]	3,89	0,002
Procent oszczędzających znajomych	58,59	2	-	-	-	< 0,001
Nikt z Twoich znajomych obecnie nie korzysta z programu oszczędnościowego (pkt odniesienia)	-	-	-	-	-	-
10% z Twoich znajomych oszczędza			1,24	[1,11; 1,39]	3,71	< 0,001
30% z Twoich znajomych oszczędza			1,57	[1,40; 1,76]	7,61	< 0,001

Moment rozpoczęcia oszczędzania	9,14	2	-	-	-	0,010
Natychmiast w momencie rozpoczęcia pracy (pkt odniesienia)	-	-	-	-	-	-
3 miesiące od rozpoczęcia pracy			0,92	[0,82; 1,04]	-1,39	0,164
6 miesięcy od rozpoczęcia pracy			1,11	[0,98; 1,25]	1,65	0,100
Planowana podwyżka	221,24	2	-	-	-	< 0,001
Brak planowanej podwyżki (pkt odniesienia)	-	-	-	-	-	-
Zaplanowana podwyżka w wys. 10% po 3 mies. okresie próbnym			2,29	[2,05; 2,57]	14,31	< 0,001
Zaplanowana podwyżka w wys. 10% po 6 mies. okresie próbnym			2,03	[1,78; 2,30]	10,82	< 0,001
Aktualne oszczędności	125,85	3	-	-	-	< 0,001
Mniej niż 1000 zł (pkt odniesienia)	-	-	-	-	-	-
1000-5000 zł			1,34	[1,18; 1,52]	4,47	< 0,001
5000-10 000 zł			2,03	[1,78; 2,30]	10,72	< 0,001
Ponad 10 tys. zł			1,67	[1,45; 1,93]	7,07	< 0,001
Pochodzenie rekomendacji	29,12	2				< 0,001
Bliski znajomy/Ktoś z rodziny (pkt odniesienia)	-	-	-	-	-	-
Doradca finansowy z banku			1,00	[0,86; 1,04]	-0,09	0,927
Pracownik ministerstwa finansów			0,68	[0,63; 0,82]	-5,00	< 0,001
Organizacja rekomendująca	2,07	1				0,150
DEOC (pkt odniesienia)	-	-	-	-	-	-
WBW			1,07	[0,98; 1,17]	1,44	0,150

Adnotacja. „pkt. odniesienia” oznacza poziom bazowy atrybutu w porównaniu do którego analizowano zmianę ilorazu szans pozostałych poziomów atrybutu

W celu weryfikacji wpływu atrybutów decyzyjnych na decyzję wyboru programu oszczędnościowego, poddano ocenie oszacowania parametrów modelu uwzględniającego całe dostępne dane (Tabela 12). Wśród dziewięciu atrybutów decyzyjnych, największym ilorazem szans odznaczały się dwa poziomy atrybutu planowanej podwyżki, czyli zaplanowana podwyżka po 3- i 6-miesięcznym okresie próbnym, w porównaniu do braku podwyżki czyli poziomu

pierwszego atrybutu (3 miesięczny okres: OR = 2,45, 95% CI [2,18; 2,76]; z = 14,85, p < 0,001; OR = 2,12, 95% CI [1,86; 2,42]; z = 11,09, p < 0,001). Trzeci pod kątem ilorazu szans okazał się poziom raty Państwa w wysokości 3% miesięcznej pensji. W porównaniu do braku wsparcia Państwa zwiększał on szansę wyboru danego programu dwukrotnie (OR = 2,00, 95% CI [1,69; 2,37]; z = 8,13, p < 0,001). Na czwartym miejscu znalazł się poziom aktualnych oszczędności 5-10 tys. złotych (OR = 1,99, 95% CI [1,74; 2,27]; z = 10,04, p < 0,001) a na piątym rata pracodawcy w wysokości 5% miesięcznej pensji (OR = 1,97, 95% CI [1,74; 2,22]; z = 10,56, p < 0,001). Wśród poziomów atrybutów, które zwiększały szanse na wybór programu oszczędnościowego znalazły się również rata Państwa w wysokości 1% i 2%, aktualne oszczędności większe niż 10 tys. złotych i oscylujące między 1 a 5 tys. złotych, 30% oszczędzających znajomych i 10% oszczędzających znajomych, rata pracodawcy w wysokości 2% miesięcznej pensji i rata własna w wysokości 2% i 5% miesięcznej pensji. Szczegółowe oszacowania dla tych poziomów atrybutów zostały przedstawione w tabeli 12. Jednakże, nie wszystkie poziomy atrybutów zwiększały szanse na wybór danej opcji. Rekomendacja ze strony pracownika ministerstwa, zmniejszała szanse na wybór programu w porównaniu do sytuacji gdy program był polecany przez bliskiego znajomego/kogoś z rodziny (OR = 0,66, 95% CI [0,57; 0,77]; z = -5,27, p < 0,001).

Kolejnym krokiem była analiza interakcji pierwszego stopnia oznaczających wystąpienie interwencji. Ze względu na zagnieżdżenie danych w użytkownikach oraz posiadanie informacji o strukturze atrybutów w każdej opcji, w warunku BRF i NN, wystąpienie interwencji w kolejnych sytuacjach decyzyjnych zostało zakodowane w formie dychotomicznej w ramach dodatkowej zmiennej „Interwencja” – programy oszczędnościowe zawierające interwencje zostały zakodowane jako jeden, a programy niezawierające interwencji zostały zakodowane jako zero.

Było to szczególnie ważne dla adekwatnego zbadania czy współwystępowanie interwencji w danym programie będzie wpływać na jego wybór inaczej w zależności od warunku badawczego. Z kolei w warunkach BS i ND, interwencja była obecna w programach gdzie był obecny poziom WBW atrybutu organizacji rekomendującej. W związku z tym, interakcja pomiędzy warunkiem BS lub ND a poziomem WBW, oznaczała program z oddziaływującą interwencją. W konsekwencji powyższego kodowania i struktury danych, efekty oddziaływania interwencji na wybór programu oszczędnościowego zostały uwidocznione w Tabeli 12 jako:

- Dla warunku boost soczewkowy: interakcja pomiędzy warunkiem badawczym BS a atrybutem organizacji rekomendującej.
- Dla warunku boost wg Reguły Franklina: interakcja pomiędzy warunkiem badawczym BRF a zmienną Interwencja.
- Dla warunku nudge dyskluzywny: interakcja pomiędzy warunkiem badawczym ND a atrybutem organizacji rekomendującej
- Dla warunku nudge normatywny: interakcja pomiędzy warunkiem badawczym NN a zmienną Interwencja

Analiza interakcji wykazała, że wystąpienie każdej z interwencji zwiększało szansę wyboru programu na który oddziaływała. Największy iloraz szans został zanotowany w przypadku wystąpienia interwencji nudge'a normatywnego (OR = 2,46, 95% CI [2,15, 2,82]; $z = 13,01$, $p < 0,001$). Kolejno uplasował się nudge dyskluzywny (OR = 2,32, 95% CI [2,02, 2,66]; $z = 11,94$, $p < 0,001$), a następnie boost Reguła Franklina (OR = 1,93, 95% CI [1,64, 2,28]; $z = 7,83$, $p < 0,001$). Spośród interwencji, najniższy iloraz szans został odnotowany w grupie boostu Soczewkowego, który był na granicy istotności (OR = 1,13, 95% CI [1,03, 1,34]; $z = 2,36$, $p = 0,018$). Oprócz analizy ilorazu szans, w celu wyłonienia najefektywniejszych interwencji,

przeprowadzono porównania parami statystyki Z. W porównaniach uwzględniono poprawkę Bonferroniego. Stwierdzono istotność statystyczną między interwencją BRF a boostem soczewkowym ($p < 0,001$), natomiast nie stwierdzono istotności statystycznej pomiędzy interwencją NN a interwencją ND ($p=0,285$). Interwencja BRF różniła się również statystycznie zarówno od interwencji NN ($p < 0,001$) jak i interwencji ND ($p < 0,001$). Wyniki oznaczają, że w parze interwencji typu boost lepszą okazała się interwencja BRF, która była słabsza od obydwu interwencji typu nudge. Porównanie statystyk Z nie wyróżniło istotnych różnic pomiędzy interwencjami typu nudge, dlatego selekcją nudge'a normatywnego do badania 3B sugerowano się wyższym ilorazem szans ($OR = 2,46$).

Tabela 12. Współczynniki atrybutów (test χ^2 Walda) oraz interwencji behawioralnych wraz oszacowanymi parametrami modelu odpowiadające ilorazowi szans (OR) wraz z przedziałami ufności (CI), które reprezentują wpływ różnych atrybutów na decyzje dotyczące wyboru programu oszczędnościowego.

Parametr	χ^2	df	OR	95% CI	z	p
Model bazowy (Przecięcie)			0,15	[0,12; 0,18]	-14,63	< 0,001
Rata Państwa	584,92	3	-	-	-	< 0,001
0% (Brak udziału Państwa; punkt odniesienia)	-	-	-	-	-	-
1% miesięcznej pensji			1,37	[1,20; 1,56]	4,66	< 0,001
2% miesięcznej pensji			1,84	[1,62; 2,08]	9,44	< 0,001
3% miesięcznej pensji			2,00	[1,69; 2,37]	8,13	< 0,001
Rata pracodawcy	339,92	2	-	-	-	< 0,001
1% miesięcznej pensji (punkt odniesienia)	-	-	-	-	-	-
2% miesięcznej pensji			1,49	[1,32; 1,68]	6,46	< 0,001
5% miesięcznej pensji			1,97	[1,74; 2,23]	10,56	< 0,001
Rata własna	76,95	2	-	-	-	< 0,001

1% miesięcznej pensji (punkt odniesienia)	-	-	-	-	-	-
2% miesięcznej pensji			1,15	[1,02; 1,28]	2,34	0,019
5% miesięcznej pensji			1,26	[1,11; 1,43]	3,65	0,002
Procent oszczędzających znajomych	183,83	2	-	-	-	< 0,001
Nikt z Twoich znajomych obecnie nie korzysta z programu oszczędnościowego (punkt odniesienia)	-	-	-	-	-	-
10% z Twoich znajomych oszczędza			1,22	[1,08; 1,37]	3,31	0,001
30% z Twoich znajomych oszczędza			1,52	[1,34; 1,71]	6,77	< 0,001
Moment rozpoczęcia oszczędzania	19,20	2	-	-	-	< 0,001
Natychmiast w momencie rozpoczęcia pracy (punkt odniesienia)	-	-	-	-	-	-
3 miesiące od rozpoczęcia pracy			0,93	[0,82; 1,06]	-1,10	0,273
6 miesięcy od rozpoczęcia pracy			1,08	[0,96; 1,23]	1,28	0,200
Planowana podwyżka	790,64	2	-	-	-	< 0,001
Brak planowanej podwyżki (punkt odniesienia)	-	-	-	-	-	-
Zaplanowana podwyżka w wys. 10% po 3 mies. okresie próbnym			2,45	[2,18; 2,76]	14,85	< 0,001
Zaplanowana podwyżka w wys. 10% po 6 mies. okresie próbnym			2,12	[1,86; 2,42]	11,09	< 0,001
Aktualne oszczędności	460,15	3	-	-	-	< 0,001
Mniej niż 1000 zł (punkt odniesienia)	-	-	-	-	-	-
1000-5000 zł			1,35	[1,19; 1,55]	4,49	< 0,001
5000-10 000 zł			1,99	[1,74; 2,27]	10,04	< 0,001
Ponad 10 tys. zł			1,70	[1,46; 1,97]	7,02	< 0,001
Pochodzenie rekomendacji	61,39	2				< 0,001
Bliski znajomy/Ktoś z rodziny (punkt odniesienia)	-	-	-	-	-	-
Doradca finansowy z banku			0,98	[0,88; 1,09]	-0,31	0,758
Pracownik ministerstwa finansów			0,66	[0,57; 0,77]	-5,27	< 0,001
Organizacja rekomendująca	176,69	1				< 0,001
DEOC (punkt odniesienia)	-	-	-	-	-	-

WBW			1,10	[1,00; 1,21]	1,94	0,053
Interwencje						
Boost Soczewkowy * Organizacja rekomendująca WBW	5,56	1	-	-	-	0,020
				1,17	[1,03; 1,43]	2,36
Nudge Normatywny * Organizacja rekomendująca WBW	169,35	1				< 0,001
				2,46	[2,15; 2,82]	13,01
Boost R. Franklina * Interwencja	61,38	1	-	-	-	< 0,001
				1,93	[1,64; 2,28]	7,83
Nudge Dyskluzywny * Interwencja	142,60	1				< 0,001
				2,32	[2,02; 2,66]	11,94

Adnotacja. „pkt. odniesienia” oznacza poziom bazowy atrybutu w porównaniu do którego analizowano zmianę ilorazu szans pozostałych poziomów atrybutu. Zostały przedstawione wybrane efekty, które okazały się istotne lub wynikające z przyjętego schematu badawczego. Pełna lista efektów dostępna jest w suplemencie online na OSF.io: <https://osf.io/n2hcj/>

Rola różnic indywidualnych w wyborze programów oszczędnościowych. Kolejnym krokiem była analiza uzyskanych efektów włączająca czynniki różnic indywidualnych i danych demograficznych. W ramach analizy o charakterze eksploracyjnym, zbadano potencjalny związek różnic indywidualnych ze zmienną zależną oraz interakcji z interwencjami behawioralnymi. Na podstawie modelu uwzględniającego atrybuty sytuacji decyzyjnej, warunków badawczych i efektów interwencji, zbudowano pięć modeli oddzielnie dla każdego narzędzia badawczego:

- a) Model 0 (bazowy) – uwzględniający atrybuty sytuacji decyzyjnej, warunek badawczy i efekty interwencji
- b) Model 1 – uwzględniający zlogarytmizowane parametry dyskontowe (k) w warunku strat i zysków

- c) Model 2 – uwzględniający cztery skale z Kwestionariusza Postrzegania Czasu Zimbardo: perspektywę Przeszłą-Pozytywną, perspektywę Przeszłą-Negatywną, perspektywę Teraźniejszą-Hedonistyczną i perspektywę Przyszłą
- d) Model 3 – uwzględniający dane demograficzne: płeć, wiek, miesięczne zarobki, przybliżoną kwotę oszczędności
- e) Model 4 – uwzględniający 6 skal z kwestionariusza DOSPERT mierzącego ryzyko domenowe: ryzyko etyczne, ryzyko finansowe-inwestycyjne, ryzyko finansowe-gier hazardowych, ryzyko zdrowotne-ubezpieczeniowe, ryzyko rekreacyjne, ryzyko społeczne

W analizie porównano pięć modeli, w tym model bazowy (M0) i cztery kolejne modele (M1-M4), oceniając ich dopasowanie przy użyciu kryteriów AIC, BIC, LogWorth oraz statystyki χ^2 . Każdy z modeli M1-M4 wykazał istotne statystycznie ulepszenie w porównaniu z modelem bazowym, co jest widoczne w niższych wartościach AIC i BIC oraz wyższej LogWorth. Pod kątem AIC, statystyki LogWorth i χ^2 najlepiej dopasowany był model 4, zawierający jako predyktory skale DOSPERT. Jednakże, w związku z licznymi predyktorami i interakcjami, został negatywnie oceniony przez wskaźnik BIC, który penalizuje modele z większą liczbą parametrów, choć BIC wciąż lepiej oceniał model 4 niż model bazowy.

Tabela 13. Porównanie mieszanych modeli logitowych z dodanymi efektami losowymi na poziomie uczestników

Porównanie modeli	Stopnie swobody (DF)	AIC	BIC	LogWorth	Statystyka χ^2	Istotność (p)
Model bazowy (M0)	103	56710,51	57604,53	-28252,25	-	-
M1 – M0	159	48743,84	50101,95	-24212,42	120,06	< 0,001
M2-M0	199	48767,95	50468,16	-24184,97	174,86	< 0,001
M3-M0	223	48720,47	50625,73	-24137,24	270,33	< 0,001
M4-M0	247	48696,96	50807,26	-24101,48	341,85	< 0,001

W związku z obiecującymi wynikami, w kolejnych krokach skupiono się na mierzonych cechach psychologicznych w każdym z modeli jako potencjalnych moderatorach oddziaływania interwencji behawioralnych (Tabela 13). W tym celu szczegółowej analizie poddano interakcje II-stopnia obejmujące interakcje dotyczące interwencji behawioralnych w poszczególnych warunkach oraz cechy psychologiczne zawarte w poszczególnych modelach. W związku ze znaczną liczbą porównań, w Tabeli 14 skupiono się na istotnych parametrach.

Tabela 14. Współczynniki atrybutów (test χ^2 Walda) wraz oszacowanymi parametrami dla interakcji II stopnia modeli 1-4. Wybrane parametry modeli zostały przedstawione wraz odpowiadającymi im ilorazami szans (OR) oraz przedziałami ufności (CI), które reprezentują wpływ różnych interwencji na decyzje dotyczące wyboru programu oszczędnościowego.

Parametr	χ^2	df	OR	95% CI	z	p
Model 1 – parametry dyskontowe k						
			-	-	-	0,002
Nudge Dyskluzywny * Interwencja * ln(k) w warunku zysków	12,51	2	0,95	[0,91; 0,99]	-2,23	0,019
			-	-	-	0,001
Nudge Normatywny * Organizacja rekomendująca WBW * ln(k) w warunku zysków	14,79	2	0,95	[0,91; 0,99]	-2,36	0,018
Model 2 – kwestionariusz Postrzegania Czasu						
Boost Reguła Franklina * Interwencja * perspektywa przeszła negatywna	12,29	2	-	-	-	0,002

			0,88	[0,8; 0,98]	-2,28	0,023
Nudge Normatywny * Organizacja rekomendująca WBW * perspektywa przeszła pozytywna	20,40	2	-	-	-	< 0,001
			1,20	[1,06; 1,36]	2,89	0,001

Model 3 – zmienne demograficzne

			-	-	-	< 0,001
Nudge Normatywny * Organizacja rekomendująca WBW * zarobki	24,07	2				
			1,51	[1,15; 1,99]	2,94	0,003

Model 4 – kwestionariusz DOSPERT

Boost Soczewkowy * Organizacja rekomendująca WBW * R. Finansowe- Inwestycyjne	22,08	2	-	-	-	< 0,001
			0,96	[0,94; 0,99]	-3,00	0,003
Nudge Dyskluzywny * Interwencja * R. Finansowe-Inwestycyjne	25,68	2				< 0,001
			1,05	[1,02; 1,08]	3,37	0,001
Nudge Normatywny * Organizacja rekomendująca WBW * R. Społeczne	15,59	2	-	-	-	< 0,001
			0,98	[0,96; 0,99]	-2,60	0,009
Nudge Dyskluzywny * Interwencja * R. społeczne	10,09	2	-	-	-	0,006
			1,02	[1,00; 1,03]	2,11	0,035

Adnotacja. Zostały przedstawione wybrane efekty, które okazały się istotne lub wynikające z przyjętego schematu badawczego. Pełna lista efektów dostępna jest w suplemencie online na OSF.io: <https://osf.io/n2hcj/>. Znak „*” oznacza interakcję.

W modelu 1, uwzględniającym parametry dyskontowe $\ln(k)$, zanotowano dwie istotne interakcje moderujące oddziaływanie interwencji behawioralnych. Wśród osób charakteryzujących się wyższym tempem dyskontowania w warunku zysków, szansa na wybór programu oszczędnościowego z interwencją ND zmniejszała się w porównaniu do programu bez interwencji ND (OR = 0,95, 95% CI [0,91; 0,99], $z = -2,23$, $p = 0,019$). Podobna zależność została zaobserwowana w przypadku oddziaływania NN, w miarę wzrastania tempa dyskontowania w warunku zysków, było mniej prawdopodobne, że osoby badane wybiorą program wspierany przez interwencję (OR = 0,95, 95% CI [0,91; 0,99], $z = -2,26$, $p = 0,018$).

Wśród czynników analizowanych w Modelu 2, zanotowano dwie istotne interakcje z interwencjami. Pierwsza z nich dotyczyła podskali mierzącej perspektywę przeszłą negatywną. W miarę wzrastania wyników na skali perspektywy przeszłej negatywnej, szansa na wybór programu wspieranego przez interwencję BRF była mniejsza niż wybór programu bez interwencji BRF (OR = 0,88, 95% CI [0,8; 0,98], $z = -2,28$, $p = 0,023$). Inny kierunek relacji został zanotowany w przypadku drugiej interakcji z interwencją NN. W przypadku gdy wzrastały wyniki na skali perspektywy przeszłej pozytywnej, prawdopodobieństwo wyboru programu wskazywanego przez interwencję NN, było większe niż wyboru opcji bez interwencji (OR = 1,20, 95% CI [1,06; 1,36], $z = 2,89$, $p = 0,001$). W trzecim modelu uwzględniającym zmienne demograficzne, zanotowano również dwie interakcje wskazujące na potencjalne moderatory interwencji wśród cech socjo-demograficznych. Jedna z nich dotyczyła zmiennej płci a druga zmiennej zarobki. W związku z tym, że wśród uczestników była znaczna rozbieżność w

liczebności pomiędzy kobietami (N=516) i mężczyznami (N=95), nie uwzględniono różnic międzypłciowych ponieważ wykorzystywana analiza zakłada zbliżoną reprezentację poziomów uwzględnionego czynnika. Natomiast w miarę zwiększania się zarobków osób badanych, zwiększało się prawdopodobieństwo wyboru programu wspieranego przez interwencję NN w porównaniu do prawdopodobieństwa wyboru programu bez interwencji NN.

W modelu 4, zanotowano cztery istotne interakcje drugiego stopnia wskazujące opisujące moderacje interwencji behawioralnych. W miarę wzrastania ryzyka finansowo-inwestycyjnego, w porównaniu do programu bez oddziaływania interwencji, malała szansa, że zostanie wybrana opcja którą wskazywała interwencja BS (OR = 0,96, 95% CI [0,94; 0,99], $z = -3,00$, $p = 0,003$). Innymi słowy, interwencja BS okazywała się przeciwnie skuteczna w momencie gdy osoby badane charakteryzowały się wyższym ryzykiem finansowo-inwestycyjnym. Ryzyko finansowo-inwestycyjne okazało się również wchodzić w istotną interakcję z interwencją ND. W tym przypadku wynik był odwrotny w stosunku do interwencji BS. W porównaniu do programu bez występującej interwencji, ND zwiększała szansę na wybór programu, wśród osób o wyższej skłonności do ryzyka finansowo-inwestycyjnego (OR = 1,05, 95% CI [1,02; 1,08], $z = 3,37$, $p = 0,001$). Podobnie w przypadku ryzyka społecznego - wysokie wyniki na skali ryzyka społecznego, również zwiększały szansę na wybór programu wspieranego przez interwencję ND, w porównaniu do programu nie wspieranego (OR = 1,02, 95% CI [1,00; 1,04], $z = 2,11$, $p = 0,035$). Z kolei w przypadku interwencji NN, skłonność do ryzyka społecznego miała odwrotny skutek. Osoby charakteryzujące się wyższymi wynikami na skali skłonności do ryzyka społecznego, mniej chętnie wybierali opcje, wskazywane przez interwencję NN (OR = 0,98, 95% CI [0,96; 0,99], $z = -2,60$, $p = 0,009$).

2.7.4 Podsumowanie

Głównym celem Badania 3A była ocena efektywności interwencji behawioralnych typu nudge i boost za pomocą paradygmatu wyboru dyskretnego. Uwzględniono w planie badawczym cztery rodzaje interwencji: boost soczewkowy, boost wg Reguły Franklina, nudge dyskluzyjny oraz nudge normatywny. Dodatkowo, w badaniu wzięto pod uwagę potencjalne moderatory efektywności interwencji w postaci różnic indywidualnych, których pomiar został przeprowadzony w oparciu o Kwestionariusz Perspektywy Czasowej Zimbardo, kwestionariusz postrzegania ryzyka domenowego DOSPERT, procedurę pomiaru tempa dyskontowania w odroczeniu dla strat i zysków, wiek i płeć. Uwzględniono również dane demograficzne takie jak miesięczne zarobki i oszczędności.

Przeprowadzone analizy wykazały, że najefektywniejszą pod kątem zwiększania prawdopodobieństwa wyboru programu oszczędnościowego, była interwencja nudge'a normatywnego. Okazała się bardziej oddziaływać na wybór osób badanych niż boost wg Reguły Franklina jak i boost soczewkowy. Porównanie parami, nie wskazało na istotne różnice pomiędzy interwencjami nudge normatywny oraz nudge dyskluzyjny, chociaż nudge normatywny charakteryzował się wyższym ilorazem szans. W oparciu o uzyskane wyniki wybrano interwencje do kolejnego badania: jako najlepszy pod kątem zmiany zachowania z kategorii nudge wybrano nudge normatywny, a z kategorii boostów interwencję boost wg Reguły Franklina.

Analizy wskazały również na różnice w hierarchii atrybutów w poszczególnych grupach. W grupie kontrolnej, w której nie występowała żadna interwencja behawioralna najwyżej w hierarchii znalazły się atrybuty planowanej podwyżki, aktualnych środków na koncie, raty Państwa i raty pracodawcy. W porównaniu do grupy kontrolnej, inaczej wyglądała hierarchia

atrybutów w warunkach badawczych, które informowały o wartości programów oszczędnościowych (boost wg Reguły Franklina oraz nudge dyskluzyny), trzema najważniejszymi atrybutami okazały się raty dopłacane przez pracodawcę i Państwo, oraz informacja o planowanej podwyżce. Innymi słowy, dwa atrybuty wskazywane przez interwencje BRF i NN jako najważniejsze, czyli rata pracodawcy oraz rata Państwa, mające kolejno największą wagę w obliczeniu wartości oczekiwanej programu, znalazły odzwierciedlenie swoich wag w hierarchii atrybutów w tych warunkach. Jako trzeci atrybut plasowała się planowana podwyżka, czyli atrybut preferowany najmocniej, przy braku występowania jakichkolwiek innych interwencji w grupie kontrolnej. W warunku boost Soczewkowy, wśród trzech pierwszych atrybutów znalazła się planowana podwyżka, rata Państwa oraz rata pracodawcy. Mimo, że atrybut organizacji rekomendującej był w tym warunku przedmiotem interwencji, nie znajdował się w czołówce, jak w warunku nudge'a normatywnego. Jedynie w warunku nudge'a normatywnego atrybut organizacji rekomendującej znalazł się na pierwszym miejscu. Było to związane z tym, że informacja przekazywana przez interwencję nudge'a normatywnego współwystępowała zawsze z pojawianiem się organizacji WBW w atrybucie organizacji rekomendującej. Podsumowując zmiany w hierarchii atrybutów, największe różnice względem ułożenia atrybutów w grupie kontrolnej, zaobserwowano kolejno w hierarchii atrybutów grupy z nudgem normatywnym, nudgem dyskluzynnym i boostem wg Reguły Franklina.

Ważnym elementem Badania 3a była również analiza potencjalnych cech moderujących efektywność interwencji behawioralnych. Wyniki wskazały, że czynniki takie jak tempo dyskontowania, perspektywa czasowa, czynniki demograficzne oraz skłonność do ryzyka, moderują efektywność interwencji. Zaobserwowano, że interwencje są różnie odbierane w

zależności od indywidualnych cech psychologicznych i demograficznych uczestników, co wskazuje na konieczność dostosowania strategii interwencyjnych do specyficznych grup docelowych. Najwyższym ilorazem szans odznaczała się interakcja interwencji nudge'a normatywnego ze zmienną zarobki. Analiza wykazała, że w miarę zwiększania się zarobków, osoby badane chętniej podążały za kierunkiem wskazywanym przez nudge normatywny. Podobna relacja została zaobserwowana w przypadku skali perspektywy przeszłej pozytywnej. Osoby badane uzyskujące wyższe wyniki na skali przeszłej pozytywnej chętniej podążały za wskazówkami przekazywanymi poprzez interwencję nudge'a normatywnego. Z kolei odwrotny kierunek zależności został zaobserwowany w przypadku interwencji typu boost wg Reguły Franklina i perspektywy przeszłej negatywnej. Im wyższe wyniki osoby badane uzyskiwały na skali perspektywy przeszłej negatywnej, tym mniejsza była szansa, że podążą za opcją wskazywaną przez interwencję boost wg Reguły Franklina.

2.8 Badanie 3B *Efekt synergii interwencji nudge i boost w zmianie zachowań oszczędnościowych*

Celem Badania 3B była ocena potencjału synergii między interwencjami typu boost i nudge za pomocą paradygmatu wyboru dyskretnego (ang. *discrete choice experiment*). Bazując na wynikach Badania 3A, jako interwencje o największym efekcie z grupy nudge wybrano nudge normatywny a jako najlepszy z grupy boost, boost Reguła Franklina.

2.8.1 Osoby badane

W badaniu 3B, wzięło udział 502 osoby rekrutowanych poprzez platformę SONA Uniwersytetu SWPS. Osoby badane zostały losowo przydzielone do jednego z czterech warunków badawczych: grupa nudge normatywny (1), grupa boost wg Reguły Franklina (2), grupy łączącej boost reguła Franklina z nudgem normatywnym (4) oraz grupy kontrolnej .

Zgodnie z wyliczeniami analizy mocy według Johnson i Orme (2010) dla schematu decyzji wieloatrybutowej, minimalna liczba osób dla badania wynosiła 456 osób.

2.8.2 Procedura

Podobnie jak w Badaniu 3A, w każdej sytuacji decyzyjnej, osobom badanym prezentowano wybór pomiędzy dwoma programami oszczędnościowymi (Rycina 8), różniącymi się pod kątem dziewięciu atrybutów decyzyjnych (Tabela 10). Dodatkowo, w warunku nudge'a normatywnego (Rycina 8, Panel B) i warunku synergii łączącym boost wg Reguły Franklina i nudge normatywny (Rycina 8, Panel B), bezpośrednio pod oferowanymi programami pojawiała się informacja będąca częścią działania interwencji brzmiąca „Większość osób wybierała tę opcję”, która nawiązywała do deskryptywnej normy społecznej. W sumie, każda osoba uczestnicząca w badaniu, dokonywała 35 wyborów. W każdym wyborze prezentowany był inny zestaw poziomów atrybutów, a każda osoba otrzymywała losową kolejność atrybutów w celu przeciwdziałania potencjalnym artefaktom związanym z kolejnością prezentacji bodźców. Ważną różnicą względem badania 3A był mechanizm działa nudge'a normatywnego. O ile podobnie jak wcześniej działał poprzez wyświetlanie informacji nawiązującej do deskryptywnej normy społecznej, to w badaniu 3B, bodziec wyświetlał się pod opcją z większą wartością oczekiwaną obliczoną w oparciu o Regułę Franklina. W niniejszy sposób, zapewniono, że we wszystkich trzech grupach: grupie boost RF, grupie NN oraz grupie łączącej te dwie interwencje, kierunek wskazywany przez interwencje będzie identyczny, co umożliwi zbadanie efektu synergii.

Tak jak w Badaniu 3A, w celu wyłonienia najlepszych interwencji behawioralnych, zebrane dane zostały przeanalizowane przy użyciu uogólnionego liniowego modelu mieszanego o rozkładzie dwumianowym z efektem losowym na poziomie uczestnika (Hoogink i in., 2020). Relatywną ważność atrybutów dla podejmowanej decyzji obliczono na podstawie testów ilorazu

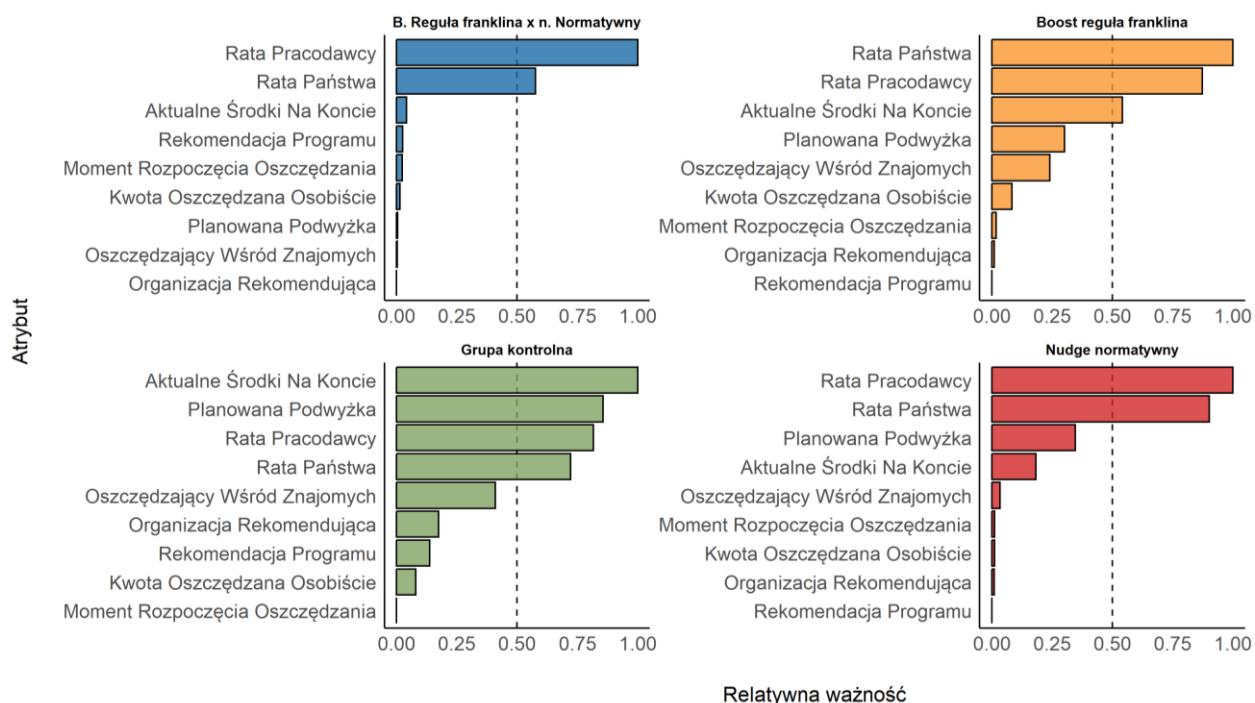
wiarygodności, a względną ważność każdego atrybutu za pomocą znormalizowanej statystyki LogWorth ($-\log_{10}(p)$), gdzie poziom istotności uzyskano z testów ilorazu wiarygodności.

2.8.3 Wyniki

Rycina 10 prezentuje, jak relatywna ważność poszczególnych atrybutów różniła się w zależności od warunków badawczych. W grupie kontrolnej dominującym atrybutem były aktualne środki na koncie (WA = 1), zaś planowana podwyżka i rata pracodawcy otrzymały nieco mniejszą wagę, odpowiednio 0,86 i 0,82. W warunku Nudge Normatywnym (NN) największą relatywną ważność miał atrybut rata pracodawcy (waga = 1), za nim plasowały się rata Państwa (waga = 0,90) i planowana podwyżka (waga = 0,35). W grupie boost wg Reguły Franklina (RF) rata Państwa była ponownie uznawana za atrybut najważniejszy (waga = 1), a rata pracodawcy oraz aktualne środki na koncie posiadały kolejno wagi 0,87 i 0,54.

W warunku połączenia Boost Reguły Franklina z Nudge Normatywnym, najwyższą wagę przypisano racie pracodawcy (waga = 1), natomiast rata Państwa i aktualne środki na koncie uzyskały mniejszą relatywną ważność, wynoszącą odpowiednio 0,58 oraz 0,04. Jest to odzwierciedlenie wpływu interwencji na hierarchię atrybutów – rata pracodawcy i rata Państwa, podkreślone jako najistotniejsze przez interwencje, zyskały na znaczeniu w preferencjach respondentów w tych warunkach. Natomiast, wyniki najbliższe przewidywaniom reguły Franklina można zauważyć w warunku synergii interwencji RF i NN. Sugerowana waga atrybutu raty pracodawcy we wzorze przekazywanym w ramach interwencji RF wynosiła 100, a dla atrybutu raty Państwa 50. W warunku synergii, preferowane wagi atrybutów są najbardziej zbliżone do wartości oczekiwanej określonej przez wzór RF.

Waga atrybutów dla podejmowanej decyzji (Badanie 3b)



Rycina 10. Relatywna ważność poszczególnych atrybutów sytuacji decyzyjnej dla

podejmowanej decyzji o przystąpieniu do programu oszczędnościowego w badaniu 3B

W celu identyfikacji efektów różnych atrybutów na wybór programu oszczędnościowego, ocenie poddano parametry modeli oddzielnie dla danych z grupy kontrolnej (Tabela 15) oraz całych danych uwzględniających grupę kontrolną i trzy warunki zawierające interwencje (Tabela 16). Jak można zauważyć w Tabeli 15, w grupie kontrolnej wszystkie atrybuty decyzji dot. programów oszczędnościowych okazały się istotnie wpływać na wybory osób badanych. Jedynie nieistotny okazał się atrybut momentu rozpoczęcia oszczędzania ($\chi^2(2) = 2,28, p = 0,264$). Z kolei w modelu obliczonym na całych danych (Tabela 16) wszystkie atrybuty decyzyjne okazały się istotne włącznie z atrybutem momentu rozpoczęcia oszczędzania ($\chi^2(2) = 17,34, p < 0,001$).

Tabela 15. Współczynniki atrybutów (test χ^2 Walda) wraz oszacowanymi parametrami modelu na danych grupy kontrolnej. Parametry modelu zostały przedstawione wraz odpowiadającym im ilorazowi szans (OR) oraz przedziałami ufności (CI), które reprezentują wpływ różnych atrybutów na decyzje dotyczące wyboru programu oszczędnościowego.

Parametr	χ^2	df	OR	95% CI	z	p
Model bazowy (Przecięcie)			0,17	[0,14; 0,21]	-17,17	< 0,001
Rata Państwa	118,39	3	-	-	-	< 0,001
0% (Brak udziału Państwa; pkt odniesienia)	-	-	-	-	-	-
1% miesięcznej pensji			1,40	[1,24; 1,58]	5,56	< 0,001
2% miesięcznej pensji			1,58	[1,41; 1,77]	7,96	< 0,001
3% miesięcznej pensji			1,99	[1,71; 2,31]	8,93	< 0,001
Rata pracodawcy	97,39	2	-	-	-	< 0,001
1% miesięcznej pensji (pkt odniesienia)	-	-	-	-	-	-
2% miesięcznej pensji			1,40	[1,26; 1,56]	6,10	< 0,001
5% miesięcznej pensji			1,84	[1,64; 2,06]	10,61	< 0,001
Rata własna	8,80	2				0,001
1% miesięcznej pensji (pkt odniesienia)	-	-	-	-	-	-
2% miesięcznej pensji			1,14	[1,03; 1,26]	2,54	0,011
5% miesięcznej pensji			1,22	[1,09; 1,37]	3,48	< 0,001
Procent oszczędzających znajomych	58,03	2	-	-	-	< 0,001
Nikt z Twoich znajomych obecnie nie korzysta z programu oszczędnościowego (pkt odniesienia)	-	-	-	-	-	-
10% z Twoich znajomych oszczędza			1,31	[1,18; 1,18]	4,98	< 0,001
30% z Twoich znajomych oszczędza			1,50	[1,35; 1,45]	7,36	< 0,001
Moment rozpoczęcia oszczędzania	2,28	2	-	-	-	0,264
Natychmiast w momencie rozpoczęcia pracy (pkt odniesienia)	-	-	-	-	-	-
3 miesiące od rozpoczęcia pracy			0,95	[0,85; 1,07]	-0,83	0,405
6 miesięcy od rozpoczęcia pracy			1,05	[0,94; 1,17]	0,81	0,420

Planowana podwyżka	103,23	2	-	-	-	< 0,001
Brak planowanej podwyżki (pkt odniesienia)	-	-	-	-	-	-
Zaplanowana podwyżka w wys. 10% po 3 mies. okresie próbnym			1,79	[1,61; 1,99]	10,83	< 0,001
Zaplanowana podwyżka w wys. 10% po 6 mies. okresie próbnym			1,53	[1,36; 1,72]	6,99	< 0,001
Aktualne oszczędności	97,39	3	-	-	-	< 0,001
Mniej niż 1000 zł (pkt odniesienia)	-	-	-	-	-	-
1000-5000 zł			1,33	[1,18; 1,50]	4,69	< 0,001
5000-10 000 zł			2,01	[1,79; 2,27]	11,41	< 0,001
Ponad 10 tys. zł			1,68	[1,47; 1,92]	7,66	< 0,001
Pochodzenie rekomendacji	7,28	2				0,026
Bliski znajomy/Ktoś z rodziny (pkt odniesienia)	-	-	-	-	-	-
Doradca finansowy z banku			1,11	[1,01; 1,23]	2,19	0,029
Pracownik ministerstwa finansów			0,82	[0,71; 0,94]	-2,91	0,004
Organizacja rekomendująca	6,32	1				0,012
DEOC (pkt odniesienia)	-	-	-	-	-	-
WBW			1,23	[1,02; 1,34]	1,44	0,001

Adnotacja. „pkt odniesienia” oznacza poziom bazowy atrybutu w porównaniu do którego analizowano zmianę ilorazu szans pozostałych poziomów atrybutu

W celu sprawdzenia, jak różne atrybuty decyzyjne wpływają na decyzję o wyborze programu oszczędnościowego, poddano ocenie oszacowania parametrów modelu uwzględniającego całe dostępne dane (Tabela 16). Wśród dziewięciu atrybutów decyzyjnych, największym ilorazem szans odznaczał się poziom atrybutu aktualnych oszczędności w wysokości 5-10 tysięcy złotych oszczędności, w porównaniu do oszczędności mniejszych 1000 zł czyli poziomu pierwszego atrybutu (OR = 2,01, 95% CI [1,79; 2,27]; $z = 11,41$, $p < 0,001$). Drugi pod kątem ilorazu szans okazał się poziom raty Państwa w wysokości 3% miesięcznej pensji. W porównaniu do braku wsparcia Państwa zwiększał on szansę wyboru danego programu prawie

dwukrotnie (OR = 1,99, 95% CI [1,71; 2,31]; $z = 8,93$, $p < 0,001$). Na trzecim miejscu znalazł się poziom raty pracodawcy w wysokości 5% miesięcznej pensji (OR = 1,84, 95% CI [1,64; 2,06]; $z = 10,61$, $p < 0,001$) a na czwartym planowana podwyżka po 3 miesiącach okresu próbnego (OR = 1,79, 95% CI [1,61; 1,99]; $z = 10,83$, $p < 0,001$). Wśród poziomów atrybutów, które zwiększały szanse na wybór programu oszczędnościowego znalazły się również rata Państwa w wysokości 1% i 2%, aktualne oszczędności większe niż 10 tys. złotych i oscylujące między 1 a 5 tys. złotych, 30% oszczędzających znajomych i 10% oszczędzających znajomych, rata pracodawcy w wysokości 2% miesięcznej pensji i rata własna w wysokości 2% i 5% miesięcznej pensji. Szczegółowe oszacowania dla tych poziomów atrybutów zostały przedstawione w tabeli 16. Jednakże, nie wszystkie poziomy atrybutów zwiększały szanse na wybór danej opcji. Rekomendacja ze strony pracownika ministerstwa, zmniejszała szanse na wybór programu w porównaniu do sytuacji gdy program był polecany przez bliskiego znajomego/kogoś z rodziny (OR = 0,82, 95% CI [0,71; 0,94]; $z = -2,91$, $p = 0,004$).

Kolejnym krokiem była analiza interakcji pierwszego stopnia pomiędzy zmienną „Interwencja” a dychotomicznymi zmiennymi grupującymi (BRF, NN i BRF x NN), oznaczającymi wystąpienie interwencji. Podobnie jak w badaniu 3A, ze względu na zagnieżdżenie danych w użytkownikach oraz posiadanie informacji o strukturze atrybutów w każdej opcji, we wszystkich warunkach z oddziaływanymi interwencjami, wystąpienie interwencji w kolejnych sytuacjach decyzyjnych zostało zakodowane w formie dychotomicznej w ramach dodatkowej zmiennej „Interwencja” – programy oszczędnościowe zawierające interwencje zostały zakodowane jako jeden, a programy niezawierające interwencji zostały zakodowane jako zero. Było to szczególnie ważne dla adekwatnego zbadania czy współwystępowanie interwencji w danym programie będzie wpływać na jego wybór inaczej w

zależności od warunku badawczego. W konsekwencji powyższego kodowania i struktury danych, efekty oddziaływania interwencji na wybór programu oszczędnościowego zostały zaprezentowane w Tabeli 16 jako:

- Dla warunku boost wg Reguły Franklina: interakcja pomiędzy warunkiem badawczym BRF a zmienną Interwencja.
- Dla warunku nudge normatywny: interakcja pomiędzy warunkiem badawczym NN a zmienną Interwencja
- Dla warunku synergii BRF i NN: : interakcja pomiędzy warunkiem badawczym BRF x NN a zmienną Interwencja

Analiza interakcji wykazała, że wystąpienie każdej z interwencji zwiększało szansę wyboru programu na który oddziaływała. Największy iloraz szans został zanotowany w przypadku wystąpienia dwóch interwencji razem w warunku synergii BRF i NN (OR = 4,92, 95% CI [4,28; 5,67]; $z = 22,13$, $p < 0,001$). Kolejną interwencją pod kątem ilorazu szans był nudge normatywny (OR = 2,41, 95% CI [2,10; 2,77]; $z = 12,48$, $p < 0,001$), a następnie boost Reguła Franklina (OR = 1,49, 95% CI [1,31; 1,71]; $z = 5,91$, $p < 0,001$). Oprócz analizy ilorazu szans, w celu wyłonienia najefektywniejszych interwencji, przeprowadzono porównania parami statystyki Z. W porównaniach uwzględniono poprawkę Bonferroniego. Stwierdzono istotność statystyczną zarówno między grupą łączącą interwencje BRF i NN a grupą z samą interwencją NN ($p < 0,001$) i samą interwencją BRF ($p < 0,001$). Interwencja BRF różniła się również statystycznie od interwencji NN ($p < 0,001$). Wyniki wskazują na efekt synergii pomiędzy interwencją nudge i boost – istotnie najwięcej decyzji o wyższej wartości oczekiwanej podejmowały osoby z grupy łączącej interwencje.

Tabela 16. Współczynniki atrybutów (test χ^2 Walda) oraz interwencji behawioralnych wraz oszacowanymi parametrami modelu odpowiadające ilorazowi szans (OR) wraz z przedziałami ufności (CI), które reprezentują wpływ różnych atrybutów na decyzje dotyczące wyboru programu oszczędnościowego.

Parametr	χ^2	df	OR	95% CI	z	p
Model bazowy (Przecięcie)			0,17	[0,14; 0,21]	-17,17	< 0,001
Rata Państwa	314,53	3	-	-	-	< 0,001
0% (Brak udziału Państwa; punkt odniesienia)	-	-	-	-	-	-
1% miesięcznej pensji			1,40	[1,24; 1,58]	5,56	< 0,001
2% miesięcznej pensji			1,58	[1,41; 1,77]	7,96	< 0,001
3% miesięcznej pensji			1,99	[1,71; 2,31]	8,93	< 0,001
Rata pracodawcy	140,93	2	-	-	-	< 0,001
1% miesięcznej pensji (punkt odniesienia)	-	-	-	-	-	-
2% miesięcznej pensji			1,40	[1,26; 1,56]	6,10	< 0,001
5% miesięcznej pensji			1,84	[1,64; 2,06]	10,61	< 0,001
Rata własna	44,94	2	-	-	-	< 0,001
1% miesięcznej pensji (punkt odniesienia)	-	-	-	-	-	-
2% miesięcznej pensji			1,14	[1,03; 1,26]	2,54	0,011
5% miesięcznej pensji			1,22	[1,09; 1,37]	3,48	< 0,001
Procent oszczędzających znajomych	107,35	2	-	-	-	< 0,001
Nikt z Twoich znajomych obecnie nie korzysta z programu oszczędnościowego (punkt odniesienia)	-	-	-	-	-	-
10% z Twoich znajomych oszczędza			1,31	[1,18; 1,45]	4,98	< 0,001
30% z Twoich znajomych oszczędza			1,50	[1,35; 1,67]	7,36	< 0,001
Moment rozpoczęcia oszczędzania	17,34	2	-	-	-	< 0,001
Natychmiast w momencie rozpoczęcia pracy (punkt odniesienia)	-	-	-	-	-	-
3 miesiące od rozpoczęcia pracy			0,95	[0,85; 1,07]	-0,83	0,405
6 miesięcy od rozpoczęcia pracy			1,05	[0,94; 1,17]	0,81	0,420

Planowana podwyżka	305,52	2	-	-	-	< 0,001
Brak planowanej podwyżki (punkt odniesienia)	-	-	-	-	-	-
Zaplanowana podwyżka w wys. 10% po 3 mies. okresie próbnym			1,79	[1,61; 1,99]	10,83	< 0,001
Zaplanowana podwyżka w wys. 10% po 6 mies. okresie próbnym			1,53	[1,36; 1,72]	6,99	< 0,001
Aktualne oszczędności	329,99	3	-	-	-	< 0,001
Mniej niż 1000 zł (punkt odniesienia)	-	-	-	-	-	-
1000-5000 zł			1,33	[1,18; 1,50]	4,69	< 0,001
5000-10 000 zł			2,01	[1,79; 2,27]	11,41	< 0,001
Ponad 10 tys. zł			1,68	[1,47; 1,92]	7,66	< 0,001
Pochodzenie rekomendacji	29,82	2				< 0,001
Bliski znajomy/Ktoś z rodziny (punkt odniesienia)	-	-	-	-	-	-
Doradca finansowy z banku			1,11	[1,01; 1,23]	2,19	0,029
Pracownik ministerstwa finansów			0,82	[0,71; 0,94]	-2,91	0,004
Organizacja rekomendująca	11,96	1				< 0,001
DEOC (punkt odniesienia)	-	-	-	-	-	-
WBW			1,23	[1,13; 1,34]	4,82	< 0,001
Interwencje						
Boost R. Franklina * Interwencja	34,91	1	-	-	-	< 0,001
			1,49	[1,31; 1,71]	5,91	< 0,001
Nudge Normatywny * Interwencja	155,66	1				< 0,001
			2,41	[2,10; 2,77]	12,48	< 0,001
(Boost RF x NN) * Interwencja	489,94	1	-	-	-	< 0,001
			4,92	[4,28; 5,67]	22,13	< 0,001

Adnotacja. Zostały przedstawione wybrane efekty, które okazały się istotne lub wynikające z przyjętego schematu badawczego. Pełna lista efektów dostępna jest w suplemencie online na OSF.io: <https://osf.io/n2hcj/>

Rola różnic indywidualnych w wyborze programów oszczędnościowych. Kolejnym krokiem była analiza uzyskanych efektów włączająca czynniki różnic indywidualnych i danych

demograficznych. W ramach analizy o charakterze eksploracyjnym, zbadano potencjalny związek różnic indywidualnych ze zmienną zależną oraz interakcji z interwencjami behawioralnymi. Na podstawie modelu uwzględniającego atrybuty sytuacji decyzyjnej, warunków badawczych i efektów interwencji, zbudowano trzy modele oddzielnie dla każdego narzędzia badawczego:

- a) Model 0 (bazowy) – uwzględniający atrybuty sytuacji decyzyjnej, warunek badawczy i efekty interwencji
- b) Model 1 – uwzględniający cztery skale z Kwestionariusza Postrzegania Czasu Zimbardo: perspektywę Przeszłą-Pozytywną, perspektywę Przeszłą-Negatywną, perspektywę Teraźniejszą-Hedonistyczną i perspektywę Przyszłą
- c) Model 2 – uwzględniający dane demograficzne: płeć, wiek, miesięczne zarobki, przybliżoną kwotę oszczędności
- d) Model 3 – uwzględniający 6 skal z kwestionariusza DOSPERT mierzącego ryzyko domenowe: ryzyko etyczne, ryzyko finansowe-inwestycyjne, ryzyko finansowe-gier hazardowych, ryzyko zdrowotne-ubezpieczeniowe, ryzyko rekreacyjne, ryzyko społeczne

Tabela 17. Porównanie mieszanych modeli logitowych z dodanymi efektami losowymi na poziomie uczestników

Porównanie modeli	Stopnie swobody (DF)	AIC	BIC	LogWorth	Statystyka χ^2	Istotność (p)
Model bazowy (M0)	84	45156	45868	-22494	-	-
M1-M0	188	45150	46742	-22387	214,86	< 0,001
M2-M0	202	45322	47217	-22414	122	0,2398
M3-M0	213	45110	46914	-22342	304.38	< 0,001

W analizie porównano cztery modele, w tym model bazowy (M0) i trzy kolejne modele (M1-M3), oceniając ich dopasowanie przy użyciu kryteriów AIC, BIC, LogWorth oraz statystyki χ^2 . Modele M1 i M3 wykazały istotne statystycznie ulepszenie w porównaniu z modelem bazowym, co jest widoczne w niższych wartościach AIC i BIC oraz wyższej LogWorth. Analizując wskaźnik AIC i statystyki LogWorth najlepiej dopasowany był model 1, zawierający jako predyktory z kwestionariusza Postrzegania Czasu. W kolejnych krokach skupiono się na mierzonych cechach psychologicznych w każdym z modeli jako potencjalnych moderatorach oddziaływania interwencji behawioralnych (Tabela 17). W tym celu analizie poddano interakcje II-stopnia obejmujące interakcje dotyczące interwencji behawioralnych w poszczególnych warunkach oraz cechy psychologiczne zawarte w poszczególnych modelach. W związku ze znaczną liczbą porównań, w Tabeli 18 skupiono się na istotnych parametrach.

Tabela 18. Współczynniki atrybutów (test χ^2 Walda) wraz oszacowanymi parametrami dla interakcji drugiego stopnia modeli 1 i 3. Wybrane parametry modeli zostały przedstawione wraz odpowiadającymi im ilorazami szans (OR) oraz przedziałami ufności (CI), które reprezentują wpływ różnych interwencji na decyzje dotyczące wyboru programu oszczędnościowego.

Parametr	χ^2	df	OR	95% CI	z	p
Model 1 – kwestionariusz Postrzegania Czasu						
			-	-	-	0,008
Boost Reguła Franklina * Interwencja * perspektywa przeszła negatywna	7,15	1	0,85	[0,75; 0,96]	-2,67	0,008

Nudge Normatywny * Interwencja * perspektywa przeszła negatywna	13,55	1	-	-	-	< 0,001
			1,25	[1,11; 1,41]	3,68	< 0,001
(Boost RF x NN) * Interwencja * perspektywa przeszła negatywna	5,51	1	-	-	-	0,002
			1,16	[1,03; 1,32]	2,35	0,019
(Boost RF x NN) * Interwencja * perspektywa terażniejsza hedonistyczna	39,47	1	-	-	-	< 0,001
			0,62	[0,54; 0,72]	-6,28	0,001

Model 3 – kwestionariusz DOSPERT

Nudge Normatywny * Interwencja * R. Społeczne	31,28	2	-	-	-	< 0,001
			1,01	[1,00; 1,02]	2,33	0,020
(Boost RF x NN) * Interwencja * R. Społeczne	169,40	2	-	-	-	< 0,001
			1,03	[1,02; 1,04]	5,65	< 0,001

Adnotacja. Zostały przedstawione wybrane efekty, które okazały się istotne lub wynikające z przyjętego schematu badawczego. Pełna lista efektów dostępna jest w suplemencie online na OSF.io: <https://osf.io/n2hcj/>

Wśród czynników analizowanych w modelu 1, zaobserwowano cztery istotne interakcje z interwencjami. Trzy z nich dotyczyły podskali mierzącej perspektywę przeszłą negatywną, z którą w interakcje wchodziły wszystkie trzy interwencje. W miarę wzrastania wyników na skali perspektywy przeszłej negatywnej, szansa na wybór programu wspieranego przez interwencję BRF była mniejsza niż wybór programu bez interwencji BRF (OR = 0,85, 95% CI [0,75; 0,96], z

= -2,67, $p = 0,008$). Inny kierunek relacji został zanotowany w przypadku interakcji z interwencją NN i z łączonymi interwencjami. W przypadku gdy wzrastały wyniki na skali perspektywy przeszłej pozytywnej, prawdopodobieństwo wyboru programu wskazywanego przez interwencję NN, było większe niż wyboru opcji bez interwencji (OR = 1,25, 95% CI [1,11; 1,41], $z = 3,68$, $p < 0,001$). Podobny kierunek relacji został odnotowany w przypadku grupy synergicznej: (OR = 1,16, 95% CI [1,03; 1,32], $z = 2,35$, $p < 0,019$). Czwarty efekt interakcyjny został zanotowany w przypadku grupy łączącej dwie interwencje i perspektywę terażniejszo-hedonistyczną. W miarę jak wyniki na skali perspektywy terażniejszo-hedonistycznej wzrastały, szansa na wybór opcji wspieranej przez dwie interwencje malała (OR = 0,62, 95% CI [0,54; 0,72], $z = -6,28$, $p < 0,001$)

W modelu 3, zaobserwowano dwie istotne interakcje II-stopnia opisujące moderację interwencji behawioralnych. W miarę wzrastania wyników na skali ryzyka społecznego, w porównaniu do programu bez oddziaływania interwencji, wzrastała szansa, że zostanie wybrana opcja którą wskazywała interwencja NN (OR = 1,01, 95% CI [1,00; 1,02], $z = -2,33$, $p = 0,020$). Podobną relację odnotowano w przypadku grupy łączącej dwie interwencje, również w miarę wzrastania wyników na skali ryzyka społecznego, zwiększała się szansa na wybór opcji wspieranej przez NN i BRF (OR = 1,03, 95% CI [1,02; 1,04], $z = 5,65$, $p < 0,001$).

2.8.4 Podsumowanie

Celem Badania 3B, była analiza potencjalnego efektu synergii pomiędzy interwencjami nudge i boost. Efekt synergii w przyjętym podejściu odnosi się do zjawiska, gdzie połączenie dwóch interwencji przynosi większe efekty, niż można by oczekiwać na podstawie indywidualnych efektów. W oparciu o wyniki badania 3A, do badania 3B zakwalifikowano dwa rodzaje interwencji: nudge normatywny i boost wg Reguły Franklina. Dodatkowo, w celu zbadania efektu synergii, utworzono trzecią grupę w której dwie interwencje wspólnie

oddziaływały na wybory osób badanych. Czwarta grupa była grupą kontrolną, gdzie osoby badane podejmowały decyzje bez oddziaływania żadnej interwencji behawioralnej.

W pierwszej części analizy, skupiono się na hierarchii atrybutów decyzji w poszczególnych warunkach badawczych. W trzech warunkach interwencyjnych, dwa pierwsze miejsca, były zajmowane przez atrybuty wspierane przez interwencje nudge i boost: rata wpłacana przez pracodawcę i rata wpłacana przez Państwo. Proporcja wagi atrybutów najbardziej zbliżona do wag sugerowanych przez regułę Franklina została odnotowana w grupie łączącej interwencje. Innymi słowy, to właśnie w hierarchii atrybutów w grupie łączącej interwencje, wybory osób badanych najbardziej odzwierciedlały wagi atrybutów zakodowane we wzorze obliczania wartości oczekiwanej planów oszczędnościowych, wzoru *implicite* przekazywanego przez wszystkie trzy interwencje. Kierunek oddziaływania każdej z interwencji, określała wartość obliczana w oparciu o Regułę Franklina – opcja, która w danym wyborze charakteryzowała się wyższą wartością była wskazywana przez nudge, boost lub obie interwencje jednocześnie (w warunku łączonym).

Analiza oszacowań parametrów modelu wraz z ilorazami szans, wykazała, że łączenie interwencji behawioralnych nudge i boost bardziej zwiększało prawdopodobieństwo wyboru programu oszczędnościowego niż stosowanie pojedynczych interwencji. Pojedyncze interwencje również okazały się istotnie zwiększać szansę wyboru wspieranych opcji, jednakże ich iloraz szans był istotnie mniejszy niż w przypadku grupy łączącej interwencje. Wyniki wskazują na istnienie efektu synergii pomiędzy interwencjami. Tym samym, **wyniki analizy Badania 3B, pozwoliły na przyjęcie Hipotezę 3.**

Ostatnią część analizy dotyczyła potencjalnych moderatorów. Zaobserwowano istotne interakcje z wynikami na skalach perspektywy przeszłej negatywnej, terażniejszej-hedonistycznej

i ryzyka społecznego. W miarę jak wyniki na skali przeszłej negatywnej z Kwestionariusza Postrzegania Czasu Zimbardo oraz wyniki na skali ryzyka społecznego z kwestionariusza DOSPERT wzrastały, rosła również szansa na wybór opcji wspieranej przez interwencje nudge'a normatywnego i łączonych interwencji. Zaobserwowano również dwie odwrotne relacje. W miarę jak wyniki na skali przeszłej negatywnej i terażniejszej-hedonistycznej z Kwestionariusza Postrzegania Czasu Zimbardo wzrastały, kolejno malała szansa na wybór opcji wspieranej przez boost Regułę Franklina i przez łączone interwencje.

2.9 Dyskusja Badań 3A i 3B

Celem Badań 3A i 3B była ocena efektywności oddziaływania interwencji behawioralnych nudge i boost na zmianę zachowań oszczędnościowych, w postaci wyboru programu oszczędnościowego opisanego za pomocą atrybutów decyzyjnych. Dodatkowo, ważnym elementem była ocena potencjalnego efektu synergii pomiędzy dwoma kategoriami interwencji. W ramach eksploracji, uwzględniono również zmienne z zakresu różnic indywidualnych, mogących moderować efekty interwencji. Schemat Badania 3A i 3B został osadzony w paradygmacie wyboru dyskretnego, który można interpretować jako symulację środowiska „dużego świata” (Gigerenzer i in., 2022), ponieważ prawdopodobieństwo wystąpienia elementów sytuacji decyzyjnej jest nieznane decydentowi.

W oparciu o wyniki badania 3A wybrano dwie interwencje, po jednej z każdej kategorii interwencji, które w swojej kategorii najbardziej zmieniły szansę na wybór programu oszczędnościowego. W grupie nudge, była to interwencja typu nudge normatywny, a w grupie boost, była to interwencja typu boost wg Reguły Franklina. Następnie, wybrane interwencje zastosowano w badaniu 3B, w którym oprócz warunków z pojedynczymi interwencjami,

dotatkowo utworzono warunek badawczy, które łączył je obydwie, aby zbadać efekt połączenia – synergii interwencji.

Zarówno w badaniu 3A jak i 3B, analizy zostały wykonane w dwóch etapach: analizę, której efektem była hierarchia atrybutów w poszczególnych warunkach (oparta o statystyki LogWorth) oraz analizę, której rezultatem były oszacowania parametrów wraz z ilorazem szans poszczególnych atrybutów i interwencji (opartą o uogólniony liniowy model mieszany o rozkładzie dwumianowym). W warunkach kontrolnych w obydwu badaniach wśród atrybutów o najwyższej wadze znalazły się następujące atrybuty: planowana podwyżka, aktualne oszczędności i na dalszych miejscach informacje o miesięcznej racie Państwa i pracodawcy. Występowanie prawie wszystkich interwencji behawioralnych w obydwu badaniach zauważalnie zmieniało hierarchię atrybutów. Jedynie w warunku interwencji boostu Soczewkowego w Badaniu 2A, atrybut planowanej podwyżki pozostawał niezmiennie jako najważniejszy atrybut w hierarchii. W przypadku pozostałych interwencji, główne atrybuty będące przedmiotem oddziaływania, tj. atrybut raty pracodawcy, raty Państwa oraz organizacji rekomendującej stawały się kluczowymi atrybutami dla podejmowania decyzji i ważyły najwięcej.

Drugi etap analiz wykazał, że prawdopodobieństwo wyboru programów oszczędnościowych najbardziej zwiększały kolejno interwencja nudge'a normatywnego, nudge'a dyskluzywnego, boostu wg Reguły Franklina oraz boostu Soczewkowego. Obydwie interwencje typu nudge wywierały silniejszy wpływ na wybór desygnowanych programów oszczędnościowych w porównaniu do interwencji typu boost. Jest to niezgodne z założeniami o prostych heurystykach Gigerenzera, Hertwiga i Grune-Yanoffa, wedle których, w sytuacjach „dużych światów”, gdzie ryzyko jest nieznane, lepiej sprawdzają się heurystyki „upraszczające” wybór i pozwalające na wybór zgodnie ze swoimi preferencjami (Gigerenzer i in., 2022; Hertwig

i Grüne-Yanoff, 2017). W świetle otrzymanych wyników, można zadać pytanie, czy „optymalność” prostych heurystyk jest zawsze czymś docenianym i preferowanym przez decydentów. Być może wybór programów oszczędnościowych dla decydentów, był na tyle skomplikowany, że osoby badane będąc przytłoczone poznawczo chętniej wybierały drogę oferowaną przez interwencję nudge. Zgodnie z wiedzą autora, jest to pierwszy projekt badawczy porównujący ze sobą interwencje typu nudge i boost za pomocą paradygmatu wyborów dyskretnych.

Badanie 3B dostarczyło informacji w kontekście Pytania badawczego nr 3, dotyczącego tego czy łączenie interwencji typu nudge i boost, będzie charakteryzowało się wyższą skutecznością w zmianie zachowania niż efekty osiągnęte za pomocą osobnych, pojedynczych interwencji. Wnioski płynące z przeprowadzonego badania sugerują, że łączenie interwencji nudge i boost istotnie zwiększało szanse na wybór programu niż interwencje typu nudge i boost stosowane rozłącznie. Oznacza to, że wystąpił efekt synergii pomiędzy dwoma interwencjami.

W oparciu o uzyskane rezultaty, można przyjąć hipotezę H3. Wyniki mogą mieć zastosowanie dla praktyków polityk publicznych, rozszerzając zestaw efektywnych technik nie tylko o interwencje typu nudge i boost, także o oddziaływania bardziej złożone, składające się z kilku metod wpływu na zmianę zachowania.

Wyniki Badań 3A i 3B podkreślają znaczenie cech moderujących efektywność interwencji behawioralnych, ujawniając złożoność zależności między indywidualnymi predyspozycjami a odbiorem interwencji. W badaniu 3A zaobserwowano, że tempo dyskontowania, perspektywa czasowa, czynniki demograficzne, oraz skłonność do ryzyka istotnie są związane z reakcjami uczestników na interwencje, z największym efektem moderującym obserwowanym dla interakcji między zarobkami a nudge'em normatywnym.

Wyniki sugerują, że osoby z wyższymi zarobkami oraz te z pozytywną perspektywą przeszłą są bardziej skłonne do podążania za wskazówkami nudge'a normatywnego, podczas gdy osoby z negatywną perspektywą przeszłą są mniej skłonne do reagowania na interwencje typu boost. W badaniu 3B, interakcje z perspektywą przeszłą negatywną, terażniejszą-hedonistyczną, i ryzykiem społecznym pokazały, że wzrost wyników w tych skalach korelował z większą skłonnością do wyboru opcji wspieranych przez nudge normatywny i łączone interwencje, natomiast wzrost na skali perspektywy przeszłej negatywnej i terażniejszej-hedonistycznej zmniejszał szansę na wybór opcji wspieranej przez boost wg Reguły Franklina. Takie wyniki wskazują na potrzebę indywidualizacji strategii interwencyjnych, aby zwiększyć ich skuteczność, biorąc pod uwagę zróżnicowane predyspozycje psychologiczne i demograficzne odbiorców.

2.10 Dyskusja ogólna

Interwencje behawioralne nudge i boost są coraz chętniej wykorzystywane przez architektów decyzji będących praktykami polityk publicznych i biznesu (Reijula i Hertwig, 2022). Promowane jako rozwiązania „tanie, proste i efektywne” w implementacji (Benartzi i in., 2017), coraz częściej znajdują zastosowanie w programach oszczędnościowych, akcjach zwiększających liczbę szczepień czy kampaniach charytatywnych (Sunstein i Reisch, 2023). Mimo wielu różnic pomiędzy podejściami nudge i boost, warto zauważyć, że to co je łączy, to nie tylko wspólne fundamenty w postaci teorii „ograniczonej racjonalności” Herberta Simona (1997), ale również cel w którym mogą być wykorzystane przez architektów decyzji. Zarówno nudging jak i boosting, dąży do zmiany zachowania poprzez oddziaływanie na architekturę sytuacji decyzyjnej (Hertwig i Grüne-Yanoff, 2017), bez zmiany przekonań i postaw decydentów ani wykorzystywania pieniężnych kar i nagród. W związku ze wzrostem zainteresowania

wspomnianymi narzędziami w politykach publicznych, pojawia się rosnąca potrzeba wytworzenia praktyk ewaluacyjnych, które niskim kosztem i warunkach laboratoryjnych będą w stanie ocenić potencjał danych interwencji i przewidzieć w jakich warunkach i domenach najlepiej je implementować (Olejniczak i in., 2020; Ruggeri i in., 2021). Niniejszy projekt jest próbą odpowiedzi na tę potrzebę, wykorzystując do tego metodologię i teorię badań psychologii decyzji. W projekcie wykorzystano procedury badań behawioralnych bazujących na loteriach oraz nowoczesnych paradygmatach wyboru wieloatrybutowego. W pracy wprowadzone zostały także elementy dotyczące w znacznym stopniu pomijane w badaniach nad wykorzystaniem interwencji behawioralnych w politykach publicznych: różnice indywidualne oraz ocenę efektu synergii pomiędzy dwoma kategoriami interwencji. Ważnym komponentem projektu, jest również rozróżnienie potocznie rozumianego ryzyka na decyzje w warunkach ryzyka, gdzie prawdopodobieństwo zdarzeń jest znane i decyzje w warunkach niepewności, gdzie prawdopodobieństwo wystąpienia poszczególnych elementów decyzji jest nieznane (Knight, 1921). Rozróżnienie na decyzje w ryzyku i niepewności, pozwoliło na bardziej trafne odwzorowanie sytuacji decyzyjnych w których interwencje mogą działać zgodnie ze swoimi założeniami (Grüne-Yanoff i Hertwig, 2016).

Głównym celem przeprowadzonych badań była identyfikacja czynników decydujących o skuteczności interwencji behawioralnych typu nudge i boost w zmianie zachowania w sytuacjach ryzykownych i niepewnych. Podczas pierwszego etapu projektu (Badanie 1), w celu syntezy dowodów i opisu skuteczności dotychczas przeprowadzonych badań, została przeprowadzona metaanaliza porównująca interwencje nudge i boost. W drugim etapie (Badania 2A i 2B), ocenie poddano szereg czynników sytuacyjnych i osobowościowych, które potencjalnie mogą moderować efektywność interwencji behawioralnych. Zgodnie z literaturą przedmiotu wśród

czynników sytuacyjnych należy uwzględnić szczególnie transparentność interwencji oraz utrzymywanie się efektów interwencji w czasie, również po wycofaniu oddziaływania ze środowiska decyzyjnego (Hertwig i Grüne-Yanoff, 2017; de Ridder, 2023). Ważnym elementem projektu było zbadanie roli różnic indywidualnych w efektywności interwencji behawioralnych, których potencjalny efekt moderacyjny, nie został do tej pory należycie zbadany empirycznie (Gillebaart i in., 2023; Mertens i in., 2022). Trzecim elementem projektu (Badania 3A i 3B), była selekcja najbardziej efektywnych dla zmiany zachowania interwencji z kategorii nudge i boost oraz sprawdzenie potencjalnego efektu synergii. Dotychczas, nie badano synergii pomiędzy interwencjami nudge i boost, a jedyne zainteresowanie badaczy w kontekście synergii, dotyczyło łączenia interwencji typu nudge z klasycznymi instrumentami polityk publicznych (Alt i in., 2024).

Pierwszym elementem projektu była metaanaliza dotychczasowych badań porównujących interwencje nudge i boost. Badanie 1, dostarczyło odpowiedzi na pytanie dotyczące różnic pomiędzy dwoma kategoriami opisywanych interwencji. Otrzymane wyniki, wskazały na statystycznie istotną, ale niską skuteczność interwencji behawioralnych typu nudge i boost, ujętych jako jedną, homogeniczną kategorię. Choć stosunkowo mały efekt, jak wskazują Thaler i Sunstein (2008) oraz Hertwig i Grune-Yanoff (2017), niewielkie istotne efekty mogą mieć znaczące znaczenie w przypadku stosowania interwencji w politykach publicznych. Według Thalera i Sunsteina (2008), nawet niewielki efekt może oznaczać znaczącą zmianę zachowania wśród kilkuset tysięcy osób, jeśli interwencja zostanie zaimplementowana np. w populacji całego kraju.

Uwzględnienie w metaanalizie rodzaju interwencji jako czynnika moderacyjnego, pozwoliło na wykazanie dwukrotnie większego efektu interwencji typu boost w porównaniu do

interwencji typu nudge. Efekt ten może uzasadniać potrzebę ponoszenia wyższych kosztów wdrożeniowych interwencji typu boost. Zazwyczaj, potencjalny koszt interwencji boost jest związany z wyższym wysiłkiem poznawczym, pojawiającym się podczas zdobywania nowych kompetencji (Kozyreva i in., 2020). Wyniki są szczególnie obiecujące w porównaniu do dotychczasowych metaanaliz uwzględniających jedynie interwencje nudge (Hummel i Maedche, 2019; Mertens i in., 2022) lub interwencje behawioralne jako ogólną, homogeniczną kategorię (Nisa i in., 2019). Choć oryginalne badania Mertens i in. (2022) wskazywały na średnią siłę efektu, to po nałożeniu na wyniki poprawki wynikającej z tendencji publikacyjnej, efekt się zmniejszył (Szasi i in., 2022), wahając się od braku efektu, przy założeniu bardzo konserwatywnego podejścia (Maier i in., 2022) do średniej siły efektu, przy założeniu najmniej konserwatywnego podejścia (Szasi i in., 2022). Z kolei badania Nisy i in. (2019), poświęcone metaanalizie nudge'y i innych interwencji behawioralnych w domenie próśrodowiskowej, również wskazały na istotną, ale stosunkowo małą siłę efektu podobną do badań Szasiego i in. (2022). Wyniki Badania 1 w połączeniu z wynikami Mertens i in. (2022) i Nisy i in. (2019) wskazują raczej na małą siłę efektu interwencji typu nudge. Innym wspólnym wynikiem uzyskanym zarówno jako rezultat Badania 1 jak i przywołanych badań (Mertens i in., 2022; Nisa i in., 2019) jest heterogeniczność oddziaływania interwencji behawioralnych. O ile badania wcześniejsze wskazywały na różne średnie efektów w zależności od stosowanych rodzajów nudge'y czy interwencji behawioralnych (Mertens i in., 2022; Nisa i in., 2019), to w przypadku Badania 1, zauważono różne efekty interwencji nudge i boost. Co ważne, to interwencje boost, zanotowały dwukrotnie większy efekt niż interwencje nudge. Dla zastosowań praktycznych, oznacza to, że w zależności od wyboru interwencji i jej odpowiedniego dopasowania do środowiska decyzyjnego, architekt decyzji może uzyskać wyniki wahające się od braku efektu do nawet średniej siły efektu stosowanego oddziaływania. Jest to kluczowy

wynik dla tworzenia polityk publicznych i wymaga dalszych badań nie tylko nad oceną poszczególnych narzędzi, ale również ich oceny w różnych sytuacjach decyzyjnych.

Wyniki Badań 2A i 2B dostarczyły odpowiedzi na pytanie badawcze nr 2: „**Jakie czynniki moderują różnice w oddziaływaniach typu boost i nudge?**”. Wyniki badania 2A pozwalają na **odrzućenie hipotezy H2.3**, dotyczącej wpływu transparentności na efektywność interwencji behawioralnych. Zarówno w przypadku interwencji typu boost jak i nudge, liczba wyborów z wyższą wartością oczekiwaną nie była uwarunkowana świadomością osób badanych nt. interwencji. Wyniki są potwierdzeniem dotychczasowych rezultatów badań nad transparentnością w interwencji typu nudge, przeprowadzanych w większości na nudge'ach opcji domyślnej (de Ridder, 2023). Dodatkowo, rozszerzają te wyniki o inne, dotychczas pomijane interwencje typu nudge dyskluzyjny i nudge normatywny, uwzględniając domenę wyborów pieniężnych. W poprzednich badaniach koncentrujących się na aspektach transparentności, nie badano również interwencji boost (Hertwig i Grüne-Yanoff, 2017). Jest to ważny wynik z perspektywy polityk publicznych ponieważ może pełnić rolę w utwierdzeniu architektów decyzji, że pełne poinformowanie o funkcjonujących interwencjach behawioralnych w sytuacji decyzyjnej, nie powinno wpłynąć na efektywność działania tych interwencji (Bruns i in., 2023). Wprowadzenie powszechnego informowania o stosowanych praktykach interwencji behawioralnych mogłoby pozwolić na polepszenie wizerunku architektów polityk społecznych i rozwiązań biznesowych w oczach pojedynczych decydentów i potencjalnie zmniejszyć odczucie bycia „manipulowanym” (Sunstein i in., 2019).

Rezultaty Badania 2 pozwalają również na **odrzućenie hipotezy badawczej H2.4**, zakładającej utrzymywanie się efektów interwencji w czasie. Choć w przypadku interwencji typu boost można było zaobserwować pewną tendencję, zarówno w przypadku procedury FFR

(Franklin, Folke i Ruggeri; Franklin i in., 2019) jak i procedury PDQ (Madden i in., 2009), oddziaływania interwencji typu nudge i boost zanikały po tygodniu od ekspozycji na interwencję. O ile jest to sprzeczne z założeniami interwencji boost (Hertwig i Grüne-Yanoff, 2017), to zgodne z „odwracalnością” efektów interwencji typu nudge (Thaler i Sunstein, 2008; Sunstein i Reisch, 2023). Efekty nudgingu wedle założeń autorów powinny zanikać w czasie, po wycofaniu interwencji. W przypadku interwencji boost, w przyszłych badaniach nad utrzymywaniem się efektów w czasie, potencjalnie dobrym krokiem byłoby wydłużenie czasu fazy treningowej w trakcie której osoby badane spędzały na przyzwajaniu heurystyki. Wydaje się to tym bardziej zasadne, że jedne z najnowszych badań porównujących efekty nudge’y i boostów wskazało na utrzymujące się działanie boostów w czasie na przestrzeni siedmiu miesięcy (Paunov i Grüne-Yanoff, 2023). W badaniach Paunova i Grüne-Yanoffa (2023) okazało się, że w warunku boost, osoby badane zużywały istotnie mniej energii elektrycznej i wody niż w warunku nudge. Jak podkreślają autorzy, efekt boostów pogłębiał się w czasie, co może wskazywać na dłuższy okres czasu niezbędny do rozwinięcia się pełnego efektu interwencji (Paunov i Grüne-Yanoff, 2023). Mniej jednoznaczne wyniki badań uzyskali Caballero i Ploner (2022), którzy również badając nudging i boosting w domenie środowiskowej, zauważyli wzrastanie efektów tych interwencji na przestrzeni kilku dni trwania eksperymentu. Co ciekawe, zauważyli jednak, że w przypadku osób o niższym dochodzie, boosty negatywnie wpływały na oszczędzanie energii, co nie miało miejsca w przypadku osób o wyższym dochodzie. Autorzy tłumaczą to związkiem między niższymi zdolnościami poznawczymi a niskim poziomem zamożności (Caballero i Ploner, 2022). Przedstawione autorskie badania są dotychczas jedynymi, które porównywały interwencje nudge i boost w domenie finansów, uwzględniając czynnik zmiany efektywności w czasie. Rezultaty Badania 2B oraz cytowanych badań Paunova i Grüne-Yanoffa (2023) i Caballero i Plonera (2022), wskazują na złożoność wpływu upływającego czasu na efektywność interwencji, ale

również na konieczność uwzględniania w tym wątku badawczym różnic indywidualnych oraz domeny podejmowania decyzji. Mimo, że Badanie 2B utrzymane w domenie decyzji finansowych, nie pokrywa się z wynikami z badaniami przeprowadzonymi w domenie środowiskowej, w warunkach dłuższej ekspozycji na interwencję np. kilku miesięcy jak w przypadku badania Paunova i Grüne-Yanoffa (2023) być może mogłoby się to zmienić również w domenie decyzji finansowych. Jak wskazuje Battaglio i in. (2018), eksperci polityk publicznych coraz chętniej rozważają nie tylko interwencje, które działają natychmiastowo jak przypadku klasycznych prawnych i ekonomicznych rozwiązań, ale również narzędzia, które potrzebują więcej czasu na uaktywnienie swojego pełnego potencjału, szczególnie jeśli działają z poszanowaniem wolnego wyboru jednostki (Battaglio i in., 2018).

Innymi hipotezami na które odpowiadały Badania 2A i 2B były hipotezy H2.1 i H2.2, dotyczące występowania różnic między interwencjami typu nudge i boost w zależności od znaku konsekwencji wyboru, tj. czy interwencje będą inaczej oddziaływać w domenie zysków i strat. Wyniki badania 2A w całości zreplikowały wyniki badania Franklina i in. (2019) wskazujące na to, że interwencje typu boost lepiej sprawdzały się w domenie zysków, a interwencje typu nudge w domenie strat (Franklin i in., 2019). Choć rezultaty badania 2A pozwalają na **przyjęcie hipotez H2.1 i H2.2**, nie jest to tak jednoznaczne w przypadku wyników badania 2B, gdzie osoby z grupy boost w domenie zysków ponownie wybierały więcej opcji o wyższej wartości oczekiwanej, natomiast w domenie strat, interwencje behawioralne nie różniły się istotnie od grupy kontrolnej. Efekt ramowania (ang. *framing*) nie jest jeszcze szeroko przebadanym czynnikiem w kontekście omawianych interwencji behawioralnych i wymaga dalszych badań. Należy natomiast wspomnieć o dwóch publikacjach, które poruszają ten wątek. Folke i in. (2021), który w badaniach wykorzystywał procedurę FFR, porównał nudge i boosty, potwierdzając, że boosty są

bardziej efektywne w kontekście decyzji ryzykownych, maksymalizując wybory z wyższą wartością oczekiwaną w domenie zysków, natomiast interwencje typu nudge prowadziły do większej liczby wyborów z wyższą wartością oczekiwaną w domenie strat (Folke i in., 2021). Dodatkowo, badacze zaobserwowali, że boosty efektywnie promują podejmowanie decyzji ryzykownych, tylko w domenie zysków, w przeciwieństwie do nudgingu, które wywiera mniejszy wpływ, niezależnie od rodzaju domeny. Jeszcze bardziej optymistyczne wyniki szczególnie w kontekście interwencji typu boost zostały uzyskane na ogólnej populacji libańskich uchodźców (Ruggeri i in., 2021). Boosty okazały się być najbardziej efektywną formą interwencji, niezależnie od tego, czy decyzja dotyczyła domeny zysków czy strat. Rezultaty badania mogą prowadzić do wniosków, że boosty mające na celu umożliwienie ludziom podejmowania niezależnych decyzji zgodnie ze swoimi preferencjami, mogą wykazywać szczególną skuteczność w krajach o niskim poziomie ekonomicznym (Ruggeri i in., 2021). Wyniki badań autorskich i omówionych w tym paragrafie, mogą być szczególnie przydatne w tych obszarach polityk publicznych gdzie ryzykowne decyzje mogą prowadzić do lepszych rezultatów w przyszłości, jak np. decyzje w domenie zdrowia o poddawaniu się zabiegom medycznym. Choć wybór opcji, o zdrowotnie mniejszej wartości oczekiwanej w danej chwili może być postrzegany bardziej awersyjnie, to w efekcie może prowadzić do lepszego stanu zdrowia w przyszłości (Krawiec i in., 2022). Właśnie w sytuacjach gdzie bardziej korzystna opcja oznacza większe ryzyko, a decydenci nie posiadają wystarczającej wiedzy i zdolności poznawczych, Ruggeri i Cafarelli (2023) proponują interwencje behawioralne typu nudge i boost jako te najbardziej obiecujące.

Badanie 3A i 3B dostarczyło informacji w kontekście pytania badawczego nr 3, dotyczącego tego czy łączenie interwencji typu nudge i boost, będzie charakteryzowało się

wyższą skutecznością niż efekty osiągane za pomocą osobnych interwencji. Analiza danych dostarczyła jednoznacznych wyników, wskazujących na to, że łączenie interwencji nudge i boost istotnie bardziej zwiększało szanse na wybór programu niż interwencje typu nudge i boost stosowane oddzielnie. Oznacza to, że istotnie wystąpił efekt synergii pomiędzy dwoma interwencjami, **co pozwala na przyjęcie Hipotezy 3**. Wyniki mogą mieć praktyczne zastosowanie dla praktyków polityk publicznych, rozszerzając zestaw efektywnych technik nie tylko o interwencje typu nudge i boost, ale również o warianty łączące je ze sobą (Alt i in., 2024; Banerjee i John, 2024). Dotychczas, w badaniach empirycznych koncentrowano się głównie na łączeniu interwencji behawioralnych z klasycznymi narzędziami polityk publicznych, a nie na łączeniu ze sobą różnych rodzajów interwencji behawioralnych (Alt i in., 2024). Warto zaznaczyć, że wedle wiedzy autora Badanie 3A i 3B nie tylko jako pierwsze porównywało za pomocą paradygmatu wyborów dyskretnych interwencje nudge i boost, ale również jako pierwsze testowało atrybuty decyzyjne wywodzące się z programu Save More Tomorrow. Analiza wagi atrybutów w warunkach kontroli wskazała, że ważny element programu SMT, czyli informacja o planowanej podwyżce był najważniejszym atrybutem oddziałującym na wybór osób badanych. Jednakże, trzy interwencje behawioralne: nudge normy społecznej, nudge dyskluzyny oraz boost wg Reguły Franklina były w stanie zmienić dominację planowanej podwyżki jako dominującego atrybutu. Wyniki wskazują na potencjalne znaczenie innych elementów sytuacji decyzyjnej, które mogą zaważyć na skuteczności programu, co było już podkreślane jako ważny kierunek badawczy przez wcześniejszych autorów (Reiff i in., 2023). Zgodnie z wiedzą autora, jest to pierwsze badanie testujące interwencje behawioralne typu nudge i boost w dziedzinie oszczędzania za pomocą paradygmatu decyzji wieloatrybutowej. Dotychczasowe, nieliczne przedsięwzięcia uwzględniające paradygmat wyboru dyskretnego w badaniach nad interwencjami behawioralnymi, były poświęcone interwencjom typu nudge w

domenie leczenia chorób reumatoidalnych (Hazlewood i in., 2020), zrównoważonego żywienia (Gottselig i in., 2023) czy zachowań prośrodowiskowych (Ouvrard i in., 2020). Hazlewood i in. (2020) w pilotażowych badaniach wprowadzili interwencję, która okazała się skutecznie prowadzić do zwiększania przejrzystości sytuacji decyzyjnej. Pacjenci wskazywali, że lepiej rozumieją sytuacje leczenia i przywiązywali większą wagę do atrybutów bardziej pożądanym z perspektywy lekarzy, jak wykorzystanie mniej znanych, ale bardziej obiecujących metod terapii (Hazlewood i in., 2020). Z kolei badania Gottselig i in. (2023) wskazały na brak oczekiwanych rezultatów nudge'y normatywnych dla zmniejszenia konsumpcji produktów mięsnych. W ich badaniach, nudge nawiązujący do normy społecznej sprawiał, że osoby badane były skłonne płacić nawet więcej za produkty mięsne niż grupa kontrolna. Jest to jeden z przykładów przeciwnego działania interwencji behawioralnej (ang. *backfiring*), który jest raportowany przy okazji niektórych zachowań i domen, najczęściej w postaci zwiększonej reaktancji względem doświadczanego oddziaływania behawioralnego (Banerjee i in., 2023; Bolton i in., 2020). Z kolei badania Ouvrarda i in. (2020) wskazały na pozytywne rezultaty oddziaływania nudge'y na akceptację polityk proklimatycznych. Na tle wymienionych badań, uzyskane wyniki wydają się szczególnie obiecujące, co może być związane z odpowiednim dopasowaniem określonych rodzajów interwencji do domeny podejmowania decyzji (Marchiori i in., 2017). Podsumowując, Badania 3A i 3B odpowiadają na rosnące zainteresowanie wykorzystaniem paradygmatu decyzji wieloatrybutowych w testowaniu interwencji behawioralnych, co może pozwolić na dokładniejsze zbadanie charakterystyki i potencjału rozwiązań nudge i boost przed ich wdrożeniem. Obiecujące wyniki badań własnych, są jednakże pierwszymi przeprowadzonymi w domenie oszczędzania i wymagają dalszych pogłębionych badań. Wyniki badań z innych domen decyzyjnych wskazują, że należy zachować ostrożność podczas generalizacji uzyskanych wyników nad synergią interwencji na inne obszary decyzji prośrodowiskowych czy zdrowotnych.

Przedstawiona praca miała również ważny komponent eksploracyjny w postaci zbadania roli różnic indywidualnych dla efektywności interwencji. Różnice indywidualne zostały ujęte we wszystkich czterech badaniach eksperymentalnych: 2A, 2B, 3A i 3B. W zależności od procedury, analizy o charakterze eksploracyjnym dostarczyły zróżnicowanych wyników. W przypadku badań 2A i 2B jedynie w ramach procedury PDQ w Badaniu 2A, efekty boostów i nudge'y wchodziły w interakcję z takimi cechami jak sumienność, otwartość na doświadczenie czy ugodowość. Osoby o wyższych wynikach na tych skalach chętniej korzystały z rozwiązań proponowanych przez interwencje. Podobne rezultaty w przypadku sumienności zanotowali Kamilçelebi i Songül (2020), którzy wskazywali, że osoby o wyższych wynikach na skali sumienności chętniej podążały za oddziaływaniami typu nudge opcji domyślnej. Wyniki Badania 2A wskazały również na interakcje nudge'y i boostów z niektórymi cechami temperamentalnymi: żwawością, wytrzymałością i perseweratywnością. Choć były to pierwsze badania uwzględniające cechy temperamentalne w badaniach nad interwencjami behawioralnymi, dotychczasowe projekty badawcze poruszające Regulacyjną Teorię Temperamentu w decyzjach finansowych, wskazywały już na istotną rolę perseweratywności (Gąsiorowska, 2011). Interpretując wyniki na skalach osobowości i temperamentu, wyniki wskazują potencjalnie na profil osób, które w dążeniach do maksymalizacji korzyści płynących z decyzji finansowych, chętnie sięgają po wsparcie swoich procesów decyzyjnych, co pokrywa się z założeniami autorów rekomendujących wykorzystywanie interwencji behawioralnych w politykach publicznych (Hertwig i Ryall, 2020; Ruggeri i in., 2021). Brak interakcji z potrzebą poznania czy umiejętnościami numerycznymi może wskazywać bardziej na podłoże emocjonalne czy motywacyjne niż chłodne analityczne, które potencjalnie warto byłoby eksplorować w kolejnych badaniach nad czynnikami mediującymi i moderującymi działanie interwencji. Waga emocji i afektu w przetwarzaniu informacji jest wielokrotnie poruszana przez proponentów podejścia

nudge (Sunstein i Reisch, 2023; Zhang i Xu, 2016). Podkreślane jest znaczenie regulacji emocji i podatności na afekt w działaniu interwencji behawioralnych, szczególnie powołując się na przykłady decyzji podejmowanych w internecie, w kontekście dezinformacji czy prywatności (Pennycook i in., 2020; Sunstein i Reisch, 2023; Zhang i Xu, 2016). Z kolei brak interakcji interwencji behawioralnych z potrzebą poznania czy umiejętnościami numerycznymi jest szczególnie ważny w kontekście podejścia boost, gdzie dotychczasowe wyniki informowały o mniejszym zastosowaniu tych interwencji u osób o mniejszych zdolnościach numerycznych czy poznawczych (Caballero i Ploner, 2022). Przedstawione wyniki mogą wskazywać, że jeśli prosta heurystyka jest wystarczająco przystępna poznawczo, jej efektywne wykorzystanie może być w zasięgu większości decydentów, tak jak zakładają proponenci boostów (Folke i in., 2021; Paunov i Grüne-Yanoff, 2023). Wedle wiedzy autora, były to pierwsze badania uwzględniające czynniki osobowościowe i temperamentalne w ocenie efektywności interwencji nudge typu dyskluzyjnego oraz interwencjach boost. Dotychczasowe badania uwzględniające czynniki osobowościowe skupiały się przede wszystkim na nudge'ach opcji domyślnej (Ingendahl i in., 2021; Kamilçelebi i Songül, 2020) i nudge'ach normatywnych (Ingendahl i in., 2021).

Wyniki badań 3A i 3B podkreślają znaczenie różnic indywidualnych dla efektywności interwencji behawioralnych, ujawniając złożoną zależność między indywidualnymi predyspozycjami a rodzajem stosowanej interwencji. W badaniu 3A zaobserwowano, że tempo dyskontowania, perspektywa czasowa, czynniki demograficzne, oraz skłonność do ryzyka istotnie wpływają na reakcje uczestników na interwencje, z największym efektem moderującym obserwowanym dla interakcji między zarobkami a nudge'em normatywnym. Wyniki sugerują, że osoby z wyższymi zarobkami oraz te z pozytywną perspektywą przeszłą są bardziej skłonne do podążania za wskazówkami nudge'a normatywnego, podczas gdy osoby z negatywną

perspektywą przeszłą są mniej skłonne do reagowania na interwencje typu boost. Uzyskane wyniki pokrywają się z dotychczasowymi badaniami Maison i in. (2019) badającymi postawy oszczędnościowe na polskiej ogólnopolskiej próbie, gdzie obiektywnie i subiektywnie lepsza sytuacja finansowa, sprzyjała posiadaniu oszczędności. Należy zwrócić uwagę, że wyniki są zbieżne mimo, że badania Maison i in. (2019) były badaniami deklaracyjnymi, korzystającymi z kwestionariuszy samoopisowych. W kontekście wyników na skalach perspektyw czasowych, dotychczasowe badania na zachowaniach inwestycyjnych, podkreślały rolę perspektywy terażniejszej hedonistycznej i perspektywy przyszłościowej (Sekścińska i in., 2017; 2018). Wskazuje to na zróżnicowany związek perspektyw czasowych z zachowaniami należącymi do domeny zachowań finansowych, szczególnie po uwzględnieniu interwencji. Mimo, że w Badaniu 3A efekty interwencji osadzonych w zachowaniach oszczędnościowych były moderowane przez przeszłe perspektywy czasowe, nie musi być to prawdziwe dla innych zachowań finansowych. Z kolei w badaniu 3B, interakcje z perspektywą przeszłą negatywną, terażniejszą-hedonistyczną, i ryzykiem społecznym pokazały, że wzrost wyników w tych skalach korelował z większą skłonnością do wyboru opcji wspieranych przez nudge normatywny i łączone interwencje, natomiast wzrost na skali perspektywy przeszłej negatywnej i terażniejszej-hedonistycznej zmniejszał szansę na wybór opcji wspieranej przez boost wg Reguły Franklina. Wyniki dla interwencji wg Reguły Franklina w Badaniu 3B bardziej pokrywają się z wynikami badań Sekścińskiej i in. (2017) nad zachowaniami inwestycyjnymi. Takie rezultaty wskazują na potrzebę indywidualizacji strategii interwencyjnych, aby zwiększyć ich skuteczność, biorąc pod uwagę zróżnicowane predyspozycje psychologiczne i demograficzne odbiorców.

Podsumowując wyniki badań, warto również ocenić interwencje z poziomu „małych” i „dużych światów”. Gigerenzer i in. (2022), zakłada, że to proste heurystyki, będą lepiej

sprawdzać się w sytuacjach niepewnych, gdzie prawdopodobieństwo nie jest znane. Twierdzi, że proste heurystyki mogą umożliwić „odsianie” niepotrzebnych informacji i w konsekwencji skoncentrowanie się na najważniejszych aspektach decyzji (Gigerenzer & Todd, 1999a). Wyniki badań 3A i 3B, które odzwierciedlały sytuacje niepewne, wskazują, że mimo iż boosty wprowadzały proste heurystyki w swoich warunkach badawczych, to interwencje typu nudge bardziej wpływały na zmianę zachowania. Z kolei w badaniach 2A i 2B, które odzwierciedlały sytuacje ryzykowne w rozumieniu Knighta (1921), wyniki wskazywały, że efektywność interwencji była różna w zależności od sformułowania sytuacji: boosty okazywały się bardziej wpływać na decyzje w domenie zysków, a nudge w domenie strat. Gdyby interwencje odnosiły rezultaty lub były przyjmowane zgodnie z założeniami Gigerenzera i in. (2022), można byłoby oczekiwać odwrotnych wyników: boostów mocniej oddziaływujących w warunkach niepewności („dużych światach”), a nudge’y mocniej oddziaływujących w warunkach ryzyka („małych światach”). Może to oznaczać, że proste heurystyki nie były komunikowane w przedstawionych badaniach w prosty sposób za pomocą interwencji boost, lub reguły postępowania proponowane przez interwencje typu nudge były bardziej przystępne. Może to być również kwestia świadomości osób badanych odnośnie tego, która reguła „powinna” zgodnie z założeniami ekspertów, lepiej się sprawdzać w danych warunkach (Ruggeri i in., 2021). W kolejnych badaniach, potencjalnie warto byłoby zwrócić uwagę na dopasowanie rodzaju interwencji do środowiska decyzji w ryzyku i niepewności.

Wśród ograniczeń autorskich badań eksperymentalnych należy wymienić przede wszystkim przeprowadzenie badań w formie online oraz opieranie wyników na próbie studentów Uniwersytetu SWPS rekrutowanych za pośrednictwem systemu SONA. Choć Badanie 2A zreplikowało wyniki badań Franklina i in. (2019), które badacze przeprowadzili na ogólnej

populacji w USA i w Serbii, uzyskane wyniki, warto byłoby zreplikować na próbie ogólnopolskiej. Byłoby to tym bardziej istotne w razie planowanej implementacji efektów interwencji w realnych sytuacjach decyzyjnych. Badania porównujące populacje internetowych paneli badawczych takich jak Prolific, Amazon Turk i SONA, wskazują na niskie motywacje osób rekrutowanych za pośrednictwem SONA, mogące skutkować bardzo krótkim czasem reakcji, charakteryzującym się brakiem czytania ze zrozumieniem (Douglas i in., 2023). Kolejnym ograniczeniem badań jest przestrzeń ich przeprowadzenia. Wszystkie badania były badaniami online, co szczególnie w przypadku Badania 2A i 2B było uwarunkowane utrudnieniami w przeprowadzaniu eksperymentów spowodowanych pandemią COVID-19. Być może, szczególnie w przypadku interwencji boost, ograniczenie nieporządkanych bodźców zewnętrznych w laboratorium, sprzyjałoby lepszemu przyswojeniu prostej heurystyki decyzyjnej (Gigerenzer i Todd, 1999b). Z drugiej jednak strony, interwencje boost oparte na prostych heurystykach mają z założenia działać w trudnych warunkach natłoku informacji i rozwijać np. częściową ignorację umożliwiając filtrowanie tylko najważniejszych bodźców (Kozyreva i in., 2020). W związku z tym, ich efektywność powinna pojawiać się również w trudnych warunkach, co potwierdziły badania Franklina i in. (2019) oraz badania własne.

W kontekście badań 3A i 3B, pewnym ograniczeniem był arbitralny wybór atrybutów sytuacji wyboru programu oszczędnościowego, które wcale nie muszą odzwierciedlać atrybutów rzeczywiście najważniejszych dla decydentów w tej sytuacji. Atrybuty były głównie wybrane w oparciu o literaturę nt. programu Save More Tomorrow i zjawiska oszczędzania w USA i w Polsce (Findley i Cottle Hunt, 2019; Jedynek, 2018; Thaler i Benartzi, 2004), co w pewnym stopniu ogranicza możliwość generalizowania wyników. Dodatkowo, ze względu na chęć uwzględnienia możliwie jak największej liczby atrybutów umotywowanych w literaturze,

końcowa liczba atrybutów decyzji była relatywnie wysoka (9 atrybutów) w porównaniu do liczby atrybutów zazwyczaj wykorzystywanych w badaniach nad paradygmatem wyborów dyskretnych, która zazwyczaj oscyluje między 5 a 7 atrybutami (Johnson i Orme, 2010). Ze względu na potencjalne obciążenie poznawcze podczas podejmowania decyzji, taka liczba atrybutów mogła zwiększać chęć do podążania za interwencjami behawioralnymi, szczególnie za interwencjami typu nudge, oferującymi proste rozwiązanie sytuacji decyzyjnej.

Potencjalnym ograniczeniem i kierunkiem dalszych badań, jest również zakres Badania 1, metaanalizy. Badanie obejmowało swoim zakresem publikacje do końca października 2021 roku, które oprócz dostarczenia odpowiedzi na pierwsze pytanie badawcze, posłużyło również do oszacowania wielkości próby Badań 2A i 2B. Było to jeszcze zanim Mertens i in. (2022) opublikowali największą do tej pory metaanalizę interwencji behawioralnych typu nudge. W kolejnych badaniach należałoby uwzględnić kolejne publikacje z lat 2022 i 2023, w których pojawiło się kilka nowych badań porównujących ze sobą te dwie kategorie interwencji, mogąc zmieniać opisane wyniki.

Przeprowadzone badania, pogłębiają rozumienie szeregu czynników sytuacyjnych i osobowościowych dla efektywności interwencji behawioralnych w decyzjach ryzykownych i niepewnych. Dodatkowo, jako pierwsze oferują ocenę efektywności łączonych interwencji nudge i boost. Wnioski z tych badań podkreślają złożoność wpływu interwencji behawioralnych na proces decyzyjny, ilustrując, jak różne aspekty proceduralne, transparentność interwencji, oraz różnice indywidualne wpływają na skuteczność nudge i boost. Mimo że nie wszystkie hipotezy znalazły potwierdzenie, uzyskane wyniki wskazują na wartość dalszej eksploracji interwencji behawioralnych w różnych kontekstach decyzyjnych oraz na potrzebę uwzględniania szerokiego zakresu czynników moderujących ich efektywność. Wyniki te poszerzają rozumienie dynamiki

oddziaływań behawioralnych, oferując wskazówki dla praktyków polityk publicznych oraz badaczy zainteresowanych optymalizacją strategii interwencyjnych. Dalsze badania powinny skupić się na ocenie poszczególnych rodzajów interwencji zawartych w systematykach nudge'y i boostów w różnych domenach decyzyjnych oraz kontynuować wykorzystanie paradygmatu decyzji wieloatrybutowych. Kolejne wysiłki prowadzone w tym zakresie mogą umożliwić oszczędne wykorzystanie środków publicznych podczas badań przed-wdrożeniowych nad interwencjami oraz możliwie dokładne i trafne określenie skutków interwencji behawioralnych.

3. PODSUMOWANIE

Prezentowana praca podjęła temat identyfikacji czynników decydujących o efektywności interwencji behawioralnych w decyzjach ryzykownych, koncentrując się przede wszystkim na dwóch kategoriach interwencji: nudgingu i boostingu. Dotychczasowe badania poświęcone wykorzystaniu interwencji behawioralnych do zmiany zachowania z perspektywy polityk publicznych, przede wszystkim koncentrowały się na nudgingu (Thaler i Sunstein, 2008), podczas gdy istnieje również alternatywne podejście, sprzyjające bardziej świadomemu przetwarzaniu informacji i wyborom zgodnie z własnymi preferencjami – boosting (Hertwig i Grüne-Yanoff, 2017). Jak zauważa Marchiori i in. (2017) czy Ingendahl i in. (2021), badania realizowane do tej pory, koncentrowały się przede wszystkim na wybranych interwencjach typu nudge (szczególnie interwencji nudge opcji domyślnej) oraz poza nielicznymi wyjątkami (Caballero i Ploner, 2022; Ingendahl i in., 2021; Kamilçelebi i Songül, 2020) nie uwzględniały różnic indywidualnych jako moderatorów efektywności interwencji. Dodatkowo, w dotychczasowych badaniach, nie testowano interwencji jako komplementarnych, które łącznie prowadziłyby do efektu synergii. Ważnym komponentem przeprowadzonego projektu, było również rozróżnienie potocznie rozumianego ryzyka na decyzje w warunkach ryzyka, gdzie

prawdopodobieństwo zdarzeń jest znane i decyzje w warunkach niepewności, gdzie prawdopodobieństwo wystąpienia poszczególnych elementów decyzji jest nieznane (Knight, 1921). Rozróżnienie na decyzje w ryzyku i niepewności, pozwoliło na bardziej trafne odwzorowanie sytuacji decyzyjnych w których mogą działać opisywane interwencje zgodnie ze swoimi założeniami (Grüne-Yanoff i Hertwig, 2016). Przeprowadzone badania dostarczyły odpowiedzi na trzy główne pytania badawcze, które dotyczyły różnic pomiędzy interwencjami nudge i boost, występowania potencjalnych moderatorów w postaci różnic indywidualnych i czynników sytuacyjnych oraz potencjalnego efektu synergii. Podsumowując wyniki wszystkich przeprowadzonych badań, a jednocześnie biorąc pod uwagę główne pytania badawcze, można stwierdzić, co następuje:

- 1) Metaanaliza badań porównujących interwencje nudge i boost wykazała, że boosty charakteryzowały się prawie dwukrotnie większą siłą efektu niż nudge. Z kolei interwencje behawioralne traktowane jako pojedyncza kategoria interwencji odnotowały małą siłę efektu. Choć uzyskano stosunkowo mały efekt ogólny, jak wskazują Thaler i Sunstein (2008) oraz Hertwig i Grune-Yanoff (2017), niewielkie istotne efekty mogą mieć znaczące znaczenie w przypadku stosowania interwencji w politykach publicznych.
- 2) Interwencje typu boost prowadziły do większej liczby decyzji o wyższej wartości oczekiwanej w domenie zysków, a interwencje typu nudge w domenie strat. Wyniki potwierdziły rezultaty badań (Franklina i in., 2019) przeprowadzone na populacji ogólnej w USA i w Serbii. Dodatkowo, ani transparentność interwencji ani upływ czasu nie wpływały na efektywności interwencji, co potwierdza dotychczasowe badania nad transparentnością (de Ridder, 2023) i dostarcza nowych danych w kontekście utrzymania efektywności interwencji w czasie w domenie decyzji finansowych.

- 3) Nowatorskie wykorzystanie paradygmatu wyborów dyskretnych w badaniach nad nudge'ami i boostami, umożliwiło porównanie interwencji w obszarze zachowań oszczędnościowych. Okazało się, że interwencje typu nudge bardziej zwiększały prawdopodobieństwo wyboru programu oszczędnościowego niż interwencje boost, a najlepszą z grupy nudge był nudge normy społecznej. Dodatkowo, połączenie interwencji nudge i interwencji boost prowadziło do uzyskania większych efektów niż w przypadku pojedynczych interwencji, potwierdzając hipotezę o efektach synergii w wyniku łączenia interwencji oraz podkreślając komplementarny charakter nudge'y i boostów.
- 4) Wśród różnic indywidualnych, istotnymi moderatorami efektywności interwencji okazały się niektóre czynniki osobowościowe (sumienność, otwartość na doświadczenie i ugodowość) oraz niektóre czynniki temperamentalne (żwawość, wytrzymałość, perseweratywność), potwierdzając dotychczasowe wyniki badań nad konstruktami osobowości i temperamentu w obszarze interwencji behawioralnych i zachowaniach finansowych (Gąsiorowska, 2011; Kamilçelebi i Songül, 2020).
- 5) Ważnymi różnicami indywidualnymi uwzględnionymi w badaniach były również perspektywy czasowe (Kwestionariusz Postrzegania Czasu Zimbardo), domenowa skłonność do podejmowania ryzyka (Kwestionariusz DOSPERT) i czynniki demograficzne jak zarobki. Okazało się, że osoby z wyższymi zarobkami oraz te z wyższymi wynikami na skali perspektywy przeszłej pozytywnej były bardziej skłonne do podążania za wskazówkami nudge'a normatywnego, podczas gdy osoby z wyższymi wynikami na skali perspektywy przeszłej negatywnej były mniej skłonne do reagowania na interwencje typu boost. Z kolei wyższe prawdopodobieństwo podążania za łączonymi interwencjami nudge'a normatywnego i boostu wg Reguły Franklina było skorelowane z

wyższymi wynikami na skali perspektywy przeszłej negatywnej i niższymi wynikami na skali terażniejszej hedonistycznej.

Przeprowadzone badania przybliżają do lepszego zrozumienia czynników sytuacyjnych i osobowościowych leżących u podstaw efektywności interwencji behawioralnych typu nudge i boost oraz zachęcają do postrzegania tych dwóch kategorii interwencji jako komplementarnych, a nie konkurencyjnych. Wyniki dostarczają informacji nt. możliwych skutków wprowadzania interwencji w politykach publicznych oraz oferują nowatorskie sposoby ich testowania w warunkach laboratoryjnych.

4. BIBLIOGRAFIA

- Alt, M., Bruns, H., DellaValle, N., & Murauskaite-Bull, I. (2024). Synergies of interventions to promote pro-environmental behaviors – A meta-analysis of experimental studies. *Global Environmental Change*, *84*, 102776. <https://doi.org/10.1016/j.gloenvcha.2023.102776>
- APA Dictionary of Psychology. (2024). *APA Dictionary of Psychology*.
<https://dictionary.apa.org/>
- Banerjee, S., Galizzi, M. M., John, P., & Mourato, S. (2023). Immediate backfire? Nudging sustainable food choices and psychological reactance. *Food Quality and Preference*, 104923.
- Banerjee, S., & John, P. (2024). Nudge plus: Incorporating reflection into behavioral public policy. *Behavioural Public Policy*, *8*(1), 69–84.
- Basol, M., Roozenbeek, J., & van der Linden, S. (2020). Good news about bad news: Gamified inoculation boosts confidence and cognitive immunity against fake news. *Journal of cognition*, *3*(1).
- Battaglio, R. P., Belardinelli, P., Bellé, N., & Cantarelli, P. (2018). Behavioral Public Administration ad fontes: A Synthesis of Research on Bounded Rationality, Cognitive Biases, and Nudging in Public Organizations. *Public Administration Review*, *79*(3), 304–320.
- 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>
- Blais, A.-R., & Weber, E. U. (2006). A domain-specific risk-taking (DOSPERT) scale for adult populations. *Judgment and Decision making*, *1*(1), 33–47.

- Bolton, G., Dimant, E., & Schmidt, U. (2020). *When a nudge backfires: Combining (im) plausible deniability with social and economic incentives to promote behavioral change*. https://papers.ssrn.com/sol3/papers.cfm?abstract_id=3531419
- Bovens, L. (2009). The Ethics of Nudge. W T. Grüne-Yanoff & S. O. Hansson (Red.), *Preference Change* (s. 207–219). Springer Netherlands. https://doi.org/10.1007/978-90-481-2593-7_10
- Bradt, J. (2019). Comparing the effects of behaviorally informed interventions on flood insurance demand: An experimental analysis of ‘boosts’ and ‘nudges’. *Behavioural Public Policy*, 1–31.
- Brooks, M., Bolker, B., Kristensen, K., Maechler, M., Magnusson, A., McGillicuddy, M., Skaug, H., Nielsen, A., Berg, C., Benth, K. van, Sadat, N., Lüdtke, D., Lenth, R., O’Brien, J., Geyer, C. J., Jagan, M., Wiernik, B., & Stouffer, D. B. (2023). *glmmTMB: Generalized Linear Mixed Models using Template Model Builder* (1.1.8) [Software]. <https://cran.r-project.org/web/packages/glmmTMB/index.html>
- Bruns, H., Fillon, A., Maniadis, Z., & Paunov, Y. (2023). ‘Fear of the Light’? Transparency does not reduce the effectiveness of nudges. A data-driven review. *University of Cyprus Working Papers in Economics, 04–2023*. <https://www.ucy.ac.cy/econ/wp-content/uploads/sites/48/2023/08/04-2023.pdf>
- Bruns, H., Kantorowicz-Reznichenko, E., Klement, K., Jonsson, M. L., & Rahali, B. (2018). Can nudges be transparent and yet effective? *Journal of Economic Psychology*, 65, 41–59.
- Caballero, N., & Ploner, M. (2022). Boosting or nudging energy consumption? The importance of cognitive aspects when adopting non-monetary interventions. *Energy Research & Social Science*, 91, 102734. <https://doi.org/10.1016/j.erss.2022.102734>

- Chaiken, S. (1980). Heuristic versus systematic information processing and the use of source versus message cues in persuasion. *Journal of personality and social psychology*, 39(5), 752.
- Chang, W., Cheng, J., Allaire, J., Xie, Y., & McPherson, J. (2015). Package 'shiny'. See <http://citeseerx.ist.psu.edu/viewdoc/download>.
- Cohen, J. (1992). A power primer. *Psychological bulletin*, 112(1), 155.
- Cohen, J. (2013). *Statistical power analysis for the behavioral sciences*. Academic press.
- Cokely, E. T., Galesic, M., Schulz, E., Ghazal, S., & Garcia-Retamero, R. (2016). Berlin Numeracy Test. *Judgment and Decision Making*.
- Constantino, S. M., Sparkman, G., Kraft-Todd, G. T., Bicchieri, C., Centola, D., Shell-Duncan, B., Vogt, S., & Weber, E. U. (2022). Scaling up change: A critical review and practical guide to harnessing social norms for climate action. *Psychological Science in the Public Interest*, 23(2), 50–97.
- Cyniak-Cieciura, M. (2017). *Temperament a nasilenie objawów pourazowego zaburzenia stresowego*.
- Cyniak-Cieciura, M., Zawadzki, B., & Strelau, J. (2018). The development of the revised version of the Formal Characteristic of Behaviour–Temperament Inventory FCB-TI (R). *Personality and Individual Differences*, 127, 117–126.
- de Ridder, D. (2023). Nudgeability and beyond: Affording people with opportunities to make the right choice. W *Research Handbook on Nudges and Society* (s. 17–33). Edward Elgar Publishing.
- De Ridder, D., Kroese, F., & Van Gestel, L. (2022). Nudgeability: Mapping Conditions of Susceptibility to Nudge Influence. *Perspectives on Psychological Science*, 17(2), 346–359. <https://doi.org/10.1177/1745691621995183>

- 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>
- Douglas, B. D., Ewell, P. J., & Brauer, M. (2023). Data quality in online human-subjects research: Comparisons between MTurk, Prolific, CloudResearch, Qualtrics, and SONA. *Plos one*, 18(3), e0279720.
- Drexler, A., Fischer, G., & Schoar, A. (2014). Keeping it simple: Financial literacy and rules of thumb. *American Economic Journal: Applied Economics*, 6(2), 1–31.
- Epstein, S. (1994). Integration of the cognitive and the psychodynamic unconscious. *American psychologist*, 49(8), 709.
- Ersner-Hershfield, H., Wimmer, G. E., & Knutson, B. (2009). Saving for the future self: Neural measures of future self-continuity predict temporal discounting. *Social cognitive and affective neuroscience*, 4(1), 85–92.
- Findley, T. S., & Cottle Hunt, E. (2019). The Save More Tomorrow Program and the Household Balance Sheet: A Theoretical Investigation. Available at SSRN 3315808.
https://papers.ssrn.com/sol3/papers.cfm?abstract_id=3315808
- Folke, T., Bertoldo, G., D’Souza, D., Ali, S., Stablum, F., & Ruggeri, K. (2021). Boosting promotes advantageous risk-taking. *Humanities and Social Sciences Communications*, 8(1), 1–10.
- Franklin, M., Folke, T., & Ruggeri, K. (2019). Optimising nudges and boosts for financial decisions under uncertainty. *Palgrave Communications*, 5(1), Article 1.
<https://doi.org/10.1057/s41599-019-0321-y>
- Gąsiorowska, A. (2011). Gender as a moderator of temperamental causes of impulse buying tendency. *Journal of Customer Behaviour*, 10(2), 119–142.

- Gigerenzer, G., & Goldstein, D. G. (1999). Betting on one good reason: The take the best heuristic. W *Simple heuristics that make us smart* (s. 75–95). Oxford University Press.
https://pure.mpg.de/rest/items/item_2102907/component/file_2102906/content
- Gigerenzer, G., Reb, J., & Luan, S. (2022). Smart Heuristics for Individuals, Teams, and Organizations. *Annual Review of Organizational Psychology and Organizational Behavior*, 9(1), 171–198. <https://doi.org/10.1146/annurev-orgpsych-012420-090506>
- Gigerenzer, G., & Todd, P. M. (1999a). Fast and frugal heuristics: The adaptive toolbox. W *Simple heuristics that make us smart* (s. 3–34). Oxford University Press.
https://pure.mpg.de/rest/items/item_2102905/component/file_2102904/content
- Gigerenzer, G., & Todd, P. M. (1999b). *Simple heuristics that make us smart*. Oxford University Press, USA.
- Gillebaart, M., Blom, S. S., Benjamins, J. S., De Boer, F., & De Ridder, D. T. (2023). The role of attention and health goals in nudging healthy food choice. *Frontiers in Psychology*, 14.
<https://www.ncbi.nlm.nih.gov/pmc/articles/PMC10620715/>
- Goldstein, D. G., & Gigerenzer, G. (1999). The recognition heuristic: How ignorance makes us smart. W *Simple heuristics that make us smart* (s. 37–58). Oxford University Press.
https://pure.mpg.de/rest/items/item_2102900/component/file_2102899/content
- Gollwitzer, P. M. (1999). Implementation intentions: Strong effects of simple plans. *American psychologist*, 54(7), 493.
- Gottselig, V., Wuppermann, A., & Herrmann, C. (2023). Effects of green nudges on consumer valuation of sustainable food: A discrete choice experiment. *GAIA-Ecological Perspectives for Science and Society*, 32(2), 233–240.
- Grüne-Yanoff, T., & Hertwig, R. (2016). Nudge versus boost: How coherent are policy and theory? *Minds and Machines*, 26(1), 149–183.

- Hallsworth, M., Egan, M., Rutter, J., & McCrae, J. (2018). *Behavioural government: Using behavioural science to improve how governments make decisions*.
<https://policycommons.net/artifacts/10773962/behavioural-government/11651566/>
- Hallsworth, M., List, J. A., Metcalfe, R. D., & Vlaev, I. (2017). The behavioralist as tax collector: Using natural field experiments to enhance tax compliance. *Journal of public economics*, *148*, 14–31.
- Hallsworth, M., Snijders, V., Burd, H., Prestt, J., Judah, G., Huf, S., & Halpern, D. (2016). Applying behavioral insights: Simple ways to improve health outcomes. *Doha, Qatar: World Innovation Summit for Health*, 29–30.
- Hazlewood, G. S., Marshall, D. A., Barber, C. E., Li, L. C., Barnabe, C., Bykerk, V. P., Tugwell, P., Hull, P. M., & Bansback, N. (2020). Using a Discrete-Choice Experiment in a Decision Aid to Nudge Patients Towards Value-Concordant Treatment Choices in Rheumatoid Arthritis: A Proof-of-Concept Study. *Patient Preference and Adherence*, *Volume 14*, 829–838. <https://doi.org/10.2147/PPA.S221897>
- Hedges, L. V., & Vevea, J. L. (1998). Fixed-and random-effects models in meta-analysis. *Psychological methods*, *3*(4), 486.
- Hertwig, R., & Grüne-Yanoff, T. (2017). Nudging and boosting: Steering or empowering good decisions. *Perspectives on Psychological Science*, *12*(6), 973–986.
- Hertwig, R., Pleskac, T. J., Pachur, T., & Rationality, C. for A. (2019). Reckoning with uncertainty: Our research program. *Taming Uncertainty*, 3–28.
- Hertwig, R., & Ryall, M. D. (2020). Nudge versus boost: Agency dynamics under libertarian paternalism. *The Economic Journal*, *130*(629), 1384–1415.
- Higgins, J. P. T., & Thompson, S. G. (2002). Quantifying heterogeneity in a meta-analysis. *Statistics in Medicine*, *21*(11), 1539–1558. <https://doi.org/10.1002/sim.1186>

- Hoogink, J., Verelst, F., Kessels, R., Van Hoek, A. J., Timen, A., Willem, L., Beutels, P., Wallinga, J., & De Wit, G. A. (2020). Preferential differences in vaccination decision-making for oneself or one's child in The Netherlands: A discrete choice experiment. *BMC Public Health*, *20*(1), 1–14.
- 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.
- Ingendahl, M., Hummel, D., Maedche, A., & Vogel, T. (2021). Who can be nudged? Examining nudging effectiveness in the context of need for cognition and need for uniqueness. *Journal of Consumer Behaviour*, *20*(2), 324–336. <https://doi.org/10.1002/cb.1861>
- Jachimowicz, J. M., Duncan, S., Weber, E. U., & Johnson, E. J. (2019). When and why defaults influence decisions: A meta-analysis of default effects. *Behavioural Public Policy*, *3*(2), 159–186.
- Jackson, D., & Turner, R. (2017). Power analysis for random-effects meta-analysis. *Research synthesis methods*, *8*(3), 290–302.
- Jacobs-Lawson, J. M., & Hershey, D. A. (2005). Influence of future time perspective, financial knowledge, and financial risk tolerance on retirement saving behaviors. *Financial Services Review-greenwich-*, *14*(4), 331.
- Jedynak, T. (2018). Wpływ wprowadzenia pracowniczych planów kapitałowych na wysokość przyszłych świadczeń emerytalnych w Polsce. *Ubezpieczenia Społeczne. Teoria i praktyka*, *2/2018*, 33–55.
- Johnson, R., & Orme, B. (2010). Sample size issues for conjoint analysis. *Getting started with conjoint analysis: strategies for product design and pricing research*. Madison: Research Publishers LLC, 57–66.

- Just, M. A., & Carpenter, P. A. (1987). *The psychology of reading and language comprehension*. Allyn & Bacon. <https://psycnet.apa.org/record/1986-98384-000>
- Kahneman, D. (2003). Maps of Bounded Rationality: Psychology for Behavioral Economics. *American Economic Review*, 93(5), 1449–1475.
<https://doi.org/10.1257/000282803322655392>
- Kahneman, D. (2011). *Thinking, fast and slow*. macmillan.
- Kamilçelebi, H., & Songül, G. (2020). Nudging in behavioral public policies in Turkey: A research on the effects of the big-5 personality traits on the automatic enrollment system. *Istanbul Ticaret Üniversitesi Sosyal Bilimler Dergisi*, 19(Temmuz 2020 (Özel Ek)), 275–294.
- Kessels, R., Jones, B., & Goos, P. (2011). Bayesian optimal designs for discrete choice experiments with partial profiles. *Journal of Choice Modelling*, 4(3), 52–74.
[https://doi.org/10.1016/S1755-5345\(13\)70042-3](https://doi.org/10.1016/S1755-5345(13)70042-3)
- Knight, F. (1921). Risk, uncertainty and profit. *Vernon Press Titles in Economics*.
<https://ideas.repec.org/b/vpr/ecbook/14.html>
- Knudsen, H. K., & Havens, J. R. (2021). Using conjoint analysis to study health policy changes: An example from a cohort of persons who use drugs. *International Journal of Drug Policy*, 98, 103425. <https://doi.org/10.1016/j.drugpo.2021.103425>
- Koffarnus, M. N., & Bickel, W. K. (2014). A 5-trial adjusting delay discounting task: Accurate discount rates in less than one minute. *Experimental and clinical psychopharmacology*, 22(3), 222.
- Kozyreva, A., Lewandowsky, S., & Hertwig, R. (2020). Citizens Versus the Internet: Confronting Digital Challenges With Cognitive Tools. *Psychological Science in the Public Interest*, 21(3), 103–156. <https://doi.org/10.1177/1529100620946707>

- Kozyreva, A., Pleskac, T. J., Pachur, T., & Hertwig, R. (2019). 18 Interpreting Uncertainty: A Brief History of Not Knowing. *Taming uncertainty*, 343.
- Krawiec, J. M., Mizak, S., Tagliabue, M., & Białaszek, W. (2022). Delay discounting of money and health outcomes, and adherence to policy guidelines during the COVID-19 pandemic. *Frontiers in Public Health*, 10, 953743.
- 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>
- Kroese, F. M., Marchiori, D. R., & de Ridder, D. T. D. (2016). Nudging healthy food choices: A field experiment at the train station. *Journal of Public Health*, 38(2), e133–e137.
<https://doi.org/10.1093/pubmed/fdv096>
- Kuznetsova, A., Brockhoff, P. B., & Christensen, R. H. B. (2015). Package ‘lmerTest’. *R package version*, 2(0), 734.
- Lejarraga, T., & Hertwig, R. (2021). How experimental methods shaped views on human competence and rationality. *Psychological Bulletin*, 147(6), 535.
- Lenth, R. V., Bolker, B., Buerkner, P., Giné-Vázquez, I., Herve, M., Jung, M., Love, J., Miguez, F., Riebl, H., & Singmann, H. (2024). *emmeans: Estimated Marginal Means, aka Least-Squares Means* (1.10.0) [Software]. <https://cran.r-project.org/web/packages/emmeans/index.html>
- Lorenz-Spreen, P., Geers, M., Pachur, T., Hertwig, R., Lewandowsky, S., & Herzog, S. M. (2021). Boosting people’s ability to detect microtargeted advertising. *Scientific Reports*, 11(1), 1–9.

- Madden, G. J., Petry, N. M., & Johnson, P. S. (2009). Pathological gamblers discount probabilistic rewards less steeply than matched controls. *Experimental and Clinical Psychopharmacology*, *17*(5), 283–290. <https://doi.org/10.1037/a0016806>
- Madrian, B. C., & Shea, D. F. (2001). The power of suggestion: Inertia in 401 (k) participation and savings behavior. *The Quarterly journal of economics*, *116*(4), 1149–1187.
- Magnusson, K. (2024). *Interpreting Cohen's d*. <https://rpsychologist.com/cohend/>
- 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>
- Maison, D., Marchlewska, M., Sekścińska, K., Rudzińska-Wojciechowska, J., & Łozowski, F. (2019). You don't have to be rich to save money: On the relationship between objective versus subjective financial situation and having savings. *PLOS ONE*, *14*(4), e0214396. <https://doi.org/10.1371/journal.pone.0214396>
- 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>
- Marcowski, P., Krawiec, J. M., Pyrkowski, M., & Białaszek, W. (w druku). *Rola atrybutów decyzji zależnych i niezależnych od polityki społecznej i interwencji behawioralnych w wyborach dotyczących szczepień*.
- Mata, R., von Helversen, B., & Rieskamp, J. (2010). Learning to choose: Cognitive aging and strategy selection learning in decision making. *Psychology and aging*, *25*(2), 299.
- Matusz, P., Traczyk, J., & Gasiorowska, A. (2011). Kwestionariusz Potrzeby Poznania—Konstrukcja i weryfikacja empiryczna narzędzia mierzącego motywację poznawczą.

- [Kwestionariusz Potrzeby Poznania—Construction and empirical verification of a scale measuring cognitive motivation.]. *Psychologia Społeczna*, 2, 113–128.
- Mcguire, W. J., & Papageorgis, D. (1961). The relative efficacy of various types of prior belief-defense in producing immunity against persuasion. *Journal of Abnormal and Social Psychology*, 62, 327–337. <https://doi.org/10.1037/h0042026>
- Mertens, S., Herberz, M., Hahnel, U. 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.
- Ministerstwo Zdrowia. (2020). *Zero tolerancji dla nieprzestrzegających zasad bezpieczeństwa [Internet]*. <https://www.gov.pl/web/zdrowie/zero-tolerancji-dla-nieprzestrzegajacych-zasad-bezpieczenstwa>
- Mol, J. M., Botzen, W. W., Blasch, J. E., Kranzler, E. C., & Kunreuther, H. C. (2020). All by myself? Testing descriptive social norm-nudges to increase flood preparedness among homeowners. *Behavioural Public Policy*, 1–33.
- Nisa, C. F., Bélanger, J. J., Schumpe, B. M., & Faller, D. G. (2019). Meta-analysis of randomised controlled trials testing behavioural interventions to promote household action on climate change. *Nature communications*, 10(1), 4545.
- Olejniczak, K., Borkowska-Waszak, S., Domaradzka-Widła, A., & Park, Y. (2020). Policy labs: The next frontier of policy design and evaluation? *Policy & Politics*, 48(1), 89–110. <https://doi.org/10.1332/030557319X15579230420108>
- Olejniczak, K., Śliwowski, P., & Roszczyńska-Kurasińska, M. (2019). Behaviour architects: A framework for employing behavioural insights in public policy practice. *Zarządzanie Publiczne/Public Governance*, 1 (47), 18–32.

- Ortloff, A.-M. (2021). *A Comparison of Nudging and Boosting for Privacy during Web Browsing* [PhD Thesis].
- Ortloff, A.-M., Zimmerman, S., Elswailer, D., & Henze, N. (2021). The Effect of Nudges and Boosts on Browsing Privacy in a Naturalistic Environment. *Proceedings of the 2021 Conference on Human Information Interaction and Retrieval*, 63–73.
<https://doi.org/10.1145/3406522.3446014>
- Ouvrard, B., Abildtrup, J., & Stenger, A. (2020). Nudging Acceptability for Wood Ash Recycling in Forests: A Choice Experiment. *Ecological Economics*, 177, 106748.
<https://doi.org/10.1016/j.ecolecon.2020.106748>
- Paunov, Y., & Grüne-Yanoff, T. (2023). Boosting vs. nudging sustainable energy consumption: A long-term comparative field test in a residential context. *Behavioural Public Policy*, 1–26.
- Pennycook, G., McPhetres, J., Zhang, Y., Lu, J. G., & Rand, D. G. (2020). Fighting COVID-19 misinformation on social media: Experimental evidence for a scalable accuracy-nudge intervention. *Psychological science*, 31(7), 770–780.
- Przepiorka, A., Sobol-Kwapinska, M., & Jankowski, T. (2016). A polish short version of the Zimbardo Time Perspective Inventory. *Personality and Individual Differences*, 101, 78–89.
- R Core Team. (2022). *R: A language and environment for statistical computing*. R Foundation for Statistical Computing [Software]. <https://www.R-project.org/>
- Rayner, K., Schotter, E. R., Masson, M. E. J., Potter, M. C., & Treiman, R. (2016). So Much to Read, So Little Time: How Do We Read, and Can Speed Reading Help? *Psychological Science in the Public Interest*, 17(1), 4–34. <https://doi.org/10.1177/1529100615623267>

- Rebonato, R. (2014). A critical assessment of libertarian paternalism. *Journal of Consumer Policy*, 37, 357–396.
- Reiff, J., Dai, H., Beshears, J., Milkman, K. L., & Benartzi, S. (2023). Save More Today or Tomorrow: The Role of Urgency in Precommitment Design. *Journal of Marketing Research*, 60(6), 1095–1113. <https://doi.org/10.1177/00222437231153396>
- Reijula, S., & Hertwig, R. (2022). Self-nudging and the citizen choice architect. *Behavioural Public Policy*, 6(1), 119–149.
- Roozenbeek, J., & van der Linden, S. (2019). Fake news game confers psychological resistance against online misinformation. *Palgrave Communications*, 5(1), Article 1. <https://doi.org/10.1057/s41599-019-0279-9>
- Ruggeri, K., & Cafarelli, V. (2023). 14. Addressing inequalities with behavioral science: A taxonomy of positive deviance. *Research Handbook on Nudges and Society*, 242.
- Ruggeri, K., Jarke, H., El-Zein, L., Verdelli, H., & Folke, T. (2021). Mental health and decisions under risk among refugees and the public in Lebanon. *Humanities and Social Sciences Communications*, 8(1), 1–11.
- Ruggeri, K., Ojinaga-Alfageme, O., Benzerga, A., Berkessel, J. B., Friedemann, M., Hlavová, R., Kunz, M., Plohl, N., Sundström, F., & Folke, T. (2021). Evidence-based policy. W *Psychology and Behavioral Economics* (s. 201–225). Routledge.
- Savage, L. J. (1954). *The foundations of statistics*. Courier Corporation. [https://books.google.com/books?hl=en&lr=&id=zSv6dBWneMEC&oi=fnd&pg=PA1&dq=Savage+L.+J.+\(1954\).+Foundations+of+Statistics.+New+York:+Wiley+&ots=XHYhRNuW_C&sig=mLeU7ysrx2NzEpo1qL5G1xW9ZzA](https://books.google.com/books?hl=en&lr=&id=zSv6dBWneMEC&oi=fnd&pg=PA1&dq=Savage+L.+J.+(1954).+Foundations+of+Statistics.+New+York:+Wiley+&ots=XHYhRNuW_C&sig=mLeU7ysrx2NzEpo1qL5G1xW9ZzA)

- Sekścińska, K., Goszczyńska, M., & Maison, D. (2017). Rola perspektywy czasowej w działaniach ekonomicznych—perspektywa psychologiczna. *Psychologia Wychowawcza*, *54*, 63–84.
- Sekścińska, K., Rudzińska-Wojciechowska, J., & Maison, D. (2018). Individual differences in time perspectives and risky financial choices. *Personality and Individual Differences*, *120*, 118–126.
- Shefrin, H. M., & Thaler, R. H. (1988). The Behavioral Life-Cycle Hypothesis. *Economic Inquiry*, *26*(4), 609–643. <https://doi.org/10.1111/j.1465-7295.1988.tb01520.x>
- Simon, H. (1990). *Reason in human affairs*. Stanford University Press.
- Simon, H. A. (1972). Theories of bounded rationality. *Decision and organization*, *1*(1), 161–176.
- Simon, H. A. (1997). *Models of bounded rationality: Empirically grounded economic reason* (T. 3). MIT press.
- Sorokowska, A., Słowińska, A., Zbieg, A., & Sorokowski, P. (2014). *Polska adaptacja testu Ten Item Personality Inventory (TIPI)—TIPI-PL—wersja standardowa i internetowa*.
- Steffel, M., Williams, E. F., & Pogacar, R. (2016). Ethically Deployed Defaults: Transparency and Consumer Protection through Disclosure and Preference Articulation. *Journal of Marketing Research*, *53*(5), 865–880. <https://doi.org/10.1509/jmr.14.0421>
- Sunstein, C. R. (2014). Nudging: A very short guide. *Journal of Consumer Policy*, *37*, 583–588.
- Sunstein, C. R. (2016). People prefer system 2 nudges (kind of). *Duke LJ*, *66*, 121.
- Sunstein, C. R., & Reisch, L. A. (2023). Introduction to the Research Handbook on Nudges and Society. W *Research Handbook on Nudges and Society* (s. 1–8). Edward Elgar Publishing.

- Sunstein, C. R., Reisch, L. A., & Kaiser, M. (2019). Trusting nudges? Lessons from an international survey. *Journal of European Public Policy*, 26(10), 1417–1443.
<https://doi.org/10.1080/13501763.2018.1531912>
- Szaszi, B., Higney, A., Charlton, A., Gelman, A., Ziano, I., Aczel, B., Goldstein, D. G., Yeager, D. S., & Tipton, E. (2022). No reason to expect large and consistent effects of nudge interventions. *Proceedings of the National Academy of Sciences*, 119(31), e2200732119.
<https://doi.org/10.1073/pnas.2200732119>
- Tagliabue, M., Squatrito, V., & Presti, G. (2019). Models of cognition and their applications in behavioral economics: A conceptual framework for nudging derived from behavior analysis and relational frame theory. *Frontiers in Psychology*, 10, 2418.
- Thaler, R. H. (2018). From Cashews to Nudges: The Evolution of Behavioral Economics. *American Economic Review*, 108(6), 1265–1287. <https://doi.org/10.1257/aer.108.6.1265>
- Thaler, R. H., & Benartzi, S. (2004). Save More TomorrowTM: 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*.
- Traets, F., Sanchez, D. G., & Vandebroek, M. (2020). Generating Optimal Designs for Discrete Choice Experiments in R: The idifix Package. *Journal of Statistical Software*, 96, 1–41.
<https://doi.org/10.18637/jss.v096.i03>
- Tversky, A., & Kahneman, D. (1974). Judgment under uncertainty: Heuristics and biases. *Science*, 185(4157), 1124–1131. <https://doi.org/10.1126/science.185.4157.1124>
- van den Broek-Altenburg, E., & Atherly, A. (2020). Using discrete choice experiments to measure preferences for hard to observe choice attributes to inform health policy

- decisions. *Health Economics Review*, 10(1), 18. <https://doi.org/10.1186/s13561-020-00276-x>
- Viechtbauer, W. (2010). Conducting meta-analyses in R with the metafor package. *Journal of statistical software*, 36, 1–48.
- Viechtbauer, W., & Cheung, M. W.-L. (2010). Outlier and influence diagnostics for meta-analysis. *Research Synthesis Methods*, 1(2), 112–125. <https://doi.org/10.1002/jrsm.11>
- Viechtbauer, W., & Viechtbauer, M. W. (2015). Package ‘metafor’. *The Comprehensive R Archive Network. Package ‘metafor’*. <http://cran.r-project.org/web/packages/metafor/metafor.pdf>.
- Warberg, L., Acquisti, A., & Sicker, D. (2019). Can Privacy Nudges be Tailored to Individuals’ Decision Making and Personality Traits? *Proceedings of the 18th ACM Workshop on Privacy in the Electronic Society*, 175–197. <https://doi.org/10.1145/3338498.3358656>
- Weber, E. U., & Johnson, E. J. (2009). Mindful Judgment and Decision Making. *Annual Review of Psychology*, 60(1), 53–85. <https://doi.org/10.1146/annurev.psych.60.110707.163633>
- WHO. (2020). *Infographics—English*. <https://www.who.int/brunei/news/infographics---english>
- World Health Organization. (2022). *WHO Coronavirus (COVID-19) Dashboard*. <https://covid19.who.int/table>
- Zaleśkiewicz, T. (2011). *Psychologia ekonomiczna*. Wydawnictwo Naukowe PWN Warszawa.
- Zaval, L., Markowitz, E. M., & Weber, E. U. (2015). How Will I Be Remembered? Conserving the Environment for the Sake of One’s Legacy. *Psychological Science*, 26(2), 231–236.
- Zhang, B., & Xu, H. (2016). Privacy Nudges for Mobile Applications: Effects on the Creepiness Emotion and Privacy Attitudes. *Proceedings of the 19th ACM Conference on Computer-Supported Cooperative Work & Social Computing*, 1676–1690. <https://doi.org/10.1145/2818048.2820073>

5. STRESZCZENIE W JĘZYKU POLSKIM

Interwencje behawioralne typu nudge i boost są coraz chętniej stosowane w politykach społecznych i biznesie. Mimo rosnącej popularności interwencji behawioralnych, większość dotychczasowych badań skupiało się głównie na kilku interwencjach typu nudge, nie biorąc pod uwagę alternatywnych narzędzi. Celem niniejszej pracy była identyfikacja czynników decydujących o skuteczności interwencji behawioralnych nudge i boost w decyzjach ryzykownych. Wśród uwzględnianych aspektów potencjalnie rzutujących na efektywność interwencji znalazły się transparentność interwencji, trwałość interwencji w czasie oraz sposób sformułowania decyzji. Ważnym aspektem projektu było zbadanie efektu synergii pomiędzy interwencjami nudge i boost. W charakterze eksploracyjnym, uwzględniono również szereg różnic indywidualnych. Okazało się, że interwencje typu boost wykazują prawie dwukrotnie większą siłę efektu w porównaniu z nudgingiem, podczas gdy interwencje behawioralne ogólnie charakteryzują się raczej małą siłą efektu. Stwierdzono, że boosty skuteczniej wpływały na podejmowanie decyzji w domenie zysków, podczas gdy nudge były bardziej efektywne w domenie strat. Nie wykazano znaczącego wpływu transparentności interwencji ani upływu czasu na ich skuteczność. W badaniach koncentrujących się na zachowaniach oszczędnościowych, nudging okazał się skuteczniejszy w promowaniu wyboru programów oszczędnościowych, zwłaszcza gdy wykorzystano nudge oparte na normach społecznych. Połączenie nudgingu i boostingu wykazało większą skuteczność niż pojedyncze interwencje, co potwierdza hipotezę o synergii interwencji i wskazuje na komplementarność obu metod. Niniejsza dysertacja wnosi istotny wkład nie tylko w dziedzinę psychologii i ekonomii behawioralnej, ale także w obszary zarządzania publicznego i polityki publicznej, dostarczając informacji nt. strategii wspierania zachowań decydentów w sytuacjach ryzykownych.

6. STRESZCZENIE W JĘZYKU ANGIELSKIM

Behavioral interventions such as nudges and boosts are increasingly being used in social policies and business. Despite the growing popularity of behavioral interventions, most existing research has primarily focused on a few nudge-type interventions, without considering alternative tools. The aim of this work was to identify factors determining the effectiveness of behavioral interventions in risky decisions. Among the aspects considered that could potentially affect intervention efficacy were the transparency of the intervention, its durability over time, and the way decisions were phrased. An important aspect of the project was to examine the synergy effect between nudge and boost interventions. In an exploratory manner, a range of individual differences was also considered. It was found that boost-type interventions demonstrated nearly twice the effect size compared to nudging, generally characterized by a rather small effect size. It was observed that boosts more effectively influenced decision-making in the gain domain, while nudges were more effective in the loss domain. No significant impact of intervention transparency or the passage of time on their effectiveness was shown. In studies focusing on saving behaviors, nudging proved more effective in promoting the choice of saving programs, especially when social norm-based nudges were used. The combination of nudging and boosting showed greater effectiveness than individual interventions, confirming the hypothesis of intervention synergy and indicating the complementarity of both methods. This dissertation contributes significantly not only to the field of psychology and behavioral economics but also to public management and public policy areas, providing insights into strategies to support decision-makers' behaviors in risky situations.

7. ZAŁĄCZNIKI

Załączniki obejmujące m.in. szczegółowe oszacowania modeli, dane oraz skrypty analityczne zostały zamieszczone w wersji cyfrowej na stronie OSF.io: <https://osf.io/n2hcj/>

7.1 Procedura FFR

Opis procedury: Procedura FFR (Franklin i in. 2019) została zastosowana w Badaniach 2A i 2B. Loterie występowała w tej samej formie w grupie kontrolnej oraz w grupie boost (w grupie boost była prezentowana instrukcja wprowadzająca o zaletach obliczania wartości oczekiwanej i trzy pozycje treningowe przed przystąpieniem do faktycznej procedury). W grupach nudge dyskluzynywny i nudge normatywny, przy każdej pozycji z procedury występował dodatkowy bodziec/informacja wprowadzający mechanizm danej interwencji.

Instrukcja dla osób badanych: Za chwile zaprezentowana zostanie seria wyborów finansowych. Prezentowane kwoty będą hipotetyczne co oznacza, że w rzeczywistości nie otrzymasz pieniędzy ani ich nie stracisz, ale prosimy abyś podejmował/a decyzje tak jakby dotyczyły prawdziwych pieniędzy. Prosimy również abyś dokładnie zwracał/a uwagę na wielkość kwoty oraz szanse jej otrzymania. Nie ma tu dobrych ani złych odpowiedzi.

Wszystkie pozycje (bez uwzględniania bodźców interwencji nudge)

1. Masz 690 zł. Który z poniższych wariantów wolisz?

- A) 25% szansy na zatrzymanie wszystkiego lub 75% szansy na utratę wszystkiego
- B) 100% szansy na utratę 490 zł

2. Który z poniższych wariantów wolisz?

- A) 25% szansy na wygraną 4550 zł lub 75% szansy na wygraną 0 zł
- B) 100% szansy na wygraną 1000 zł

3. Który z poniższych wariantów wolisz?

- A) 50% szansy na wygraną 2000 zł lub 50% szansy na wygraną 0 zł
- B) 100% szansy na wygraną 1000 zł

4. Masz 230 zł. Który z poniższych wariantów wolisz?

- A) 75% szansy na zatrzymanie wszystkiego lub 25% szansy na utratę wszystkiego
- B) 100% szansy na utratę 30 zł

5. Który z poniższych wariantów wolisz?

- A) 75% szansy na wygraną 303,3 zł lub 25% szansy na wygraną 0 zł
- B) 100% szansy na wygraną 200 zł

6. Masz 690 zł. Który z poniższych wariantów wolisz?

- A) 25% szansy na zatrzymanie wszystkiego lub 75% szansy na utratę wszystkiego

B) 100% szansy na utratę 490 zł

7. Który z poniższych wariantów wolisz?

A) 25% szansy na wygraną 910 zł lub 75% szansy na wygraną 0 zł

B) 100% szansy na wygraną 200 zł

8. Masz 1150 zł. Który z poniższych wariantów wolisz?

A) 75% szansy na zatrzymanie wszystkiego lub 25% szansy na utratę wszystkiego

B) 100% szansy na utratę 150 zł

9. Który z poniższych wariantów wolisz?

A) 75% szansy na wygraną 1516,6 zł lub 25% szansy na wygraną 0 zł

B) 100% szansy na wygraną 1000 zł

10. Masz 3450 zł. Który z poniższych wariantów wolisz?

A) 25% szansy na zatrzymanie wszystkiego lub 75% szansy na utratę wszystkiego

B) 100% szansy na utratę 2450 zł

11. Który z poniższych wariantów wolisz?

A) 25% szansy na wygraną 4550 zł lub 75% szansy na wygraną 0 zł

B) 100% szansy na wygraną 1000 zł

Przykładowa pozycja 4 w warunku nudge dyskluzyny:

4. Masz 230 zł. Który z poniższych wariantów wolisz?

A) 75% szansy na zatrzymanie wszystkiego lub 25% szansy na utratę wszystkiego
(Przewidywana wartość oczekiwana tej opcji to 172,5 zł)

B) 100% szansy na utratę 30 zł (Przewidywana wartość oczekiwana tej opcji to 200 zł)

Przykładowa pozycja 4 w warunku nudge dyskluzyny:

4. Masz 230 zł. Który z poniższych wariantów wolisz?

A) 75% szansy na zatrzymanie wszystkiego lub 25% szansy na utratę wszystkiego

B) 100% szansy na utratę 30 zł (Większość uczestników badania wybierało tę opcję)

7.2 Procedura PDQ

Opis procedury: Oryginalny cel procedury PDQ to pomiar tempa dyskontowania. W niniejszej pracy, procedura PDQ posłużyła jako usystematyzowane loterie pieniężne w celu zreplikowania procedury FFR (Franklin i in., 2019). Procedura PDQ składała się w sumie z 60 pozycji, po 10 na każdą kategorię: małe zyski, średnie zyski, duże zyski, małe straty, średnie straty i duże straty. Pełna lista pozycji została opublikowana w wersji cyfrowej na stronie OSF.io: <https://osf.io/n2hcj/>. Poniżej jako przykład, zostały zamieszczone pozycje z kategorii małych zysków. Podobnie jak w procedurze FFR, wybór optymalny oznaczał wybór opcji z większą wartością oczekiwaną.

Instrukcja dla osób badanych: Przed Tobą kolejna seria wyborów finansowych. Prezentowane kwoty będą hipotetyczne co oznacza, że w rzeczywistości nie otrzymasz pieniędzy ani ich nie stracisz, ale prosimy abyś podejmował decyzje tak jakby dotyczyły prawdziwych pieniędzy. Prosimy również abyś dokładnie zwracał uwagę na wielkość kwoty oraz szanse jej otrzymania. Nie ma tu dobrych ani złych odpowiedzi.

Małe zyski:

1. Co wolisz?

- A) 100% szansy na wygranie 20 zł
- B) 10% szansy na wygranie 80 zł lub 90% szansy na wygranie niczego

2. Co wolisz?

- A) 100% szansy na wygranie 20 zł
- B) 13% szansy na wygranie 80 zł lub 87% szansy na wygranie niczego

3. Co wolisz?

- A) 100% szansy na wygranie 20 zł
- B) 17% szansy na wygranie 80 zł lub 83% szansy na wygranie niczego

4. Co wolisz?

- A) 100% szansy na wygranie 20 zł
- B) 20% szansy na wygranie 80 zł lub 80% szansy na wygranie niczego

5. Co wolisz?

- A) 100% szansy na wygranie 20 zł
- B) 25% szansy na wygranie 80 zł lub 75% szansy na wygranie niczego

6. Co wolisz?

- A) 100% szansy na wygranie 20 zł
- B) 33% szansy na wygranie 80 zł lub 67% szansy na wygranie niczego

7. Co wolisz?

- A) 100% szansy na wygranie 20 zł
- B) 50% szansy na wygranie 80 zł lub 50% szansy na wygranie niczego

8. Co wolisz?

- A) 100% szansy na wygranie 20 zł
- B) 67% szansy na wygranie 80 zł lub 33% szansy na wygranie niczego

9. Co wolisz?

- A) 100% szansy na wygranie 20 zł
- B) 75% szansy na wygranie 80 zł lub 25% szansy na wygranie niczego

10. Co wolisz?

- A) 100% szansy na wygranie 20 zł
- B) 83% szansy na wygranie 80 zł lub 17% szansy na wygranie niczego

Przykładowa pozycja 1 w warunku nudge dyskluzyjny

1. Co wolisz?

- A) 100% szansy na wygranie 20 zł (Przewidywana wartość oczekiwana tej opcji to 20 zł)
- B) 10% szansy na wygranie 80 zł lub 90% szansy na wygranie niczego (Przewidywana wartość oczekiwana tej opcji to 8 zł)

Przykładowa pozycja 1 w warunku nudge normatywny

1. Co wolisz?

- A) 100% szansy na wygranie 20 zł (Większość uczestników badania wybierało tę opcję)
- B) 10% szansy na wygranie 80 zł lub 90% szansy na wygranie niczego

7.3 Procedura paradygmatu wyborów dyskretnych (ang. *Discrete Choice Experiment; DCE*)

Opis procedury: W Badaniach 3A i 3B, zadaniem osób badanych był wybór pomiędzy dwoma programami oszczędnościowymi, różniącymi się poziomami 9 atrybutów decyzji (szczegółowy opis atrybutów w Tabeli 10). Osoby badane były proszone by dokonać 35 wyborów. W każdym wyborze wyświetlał się inny zestaw poziomów 9 atrybutów. Poniżej została zaprezentowana sytuacja wyboru w grupie nudge dyskluzyjny. Pliki w większej rozdzielczości przedstawiające pozostałe warunki, zostały zamieszczone w wersji cyfrowej na stronie OSF.io:

<https://osf.io/n2hcj/>

Przykładowa sytuacja wyboru w grupie nudge dyskluzyjny

Wybór: 1 / 35

	Program A	Program B
Liczba Twoich znajomych którzy uczestniczą w programie oszczędnościowym	30% z Twoich znajomych obecnie korzysta z programu oszczędnościowego	10% z Twoich znajomych obecnie korzysta z programu oszczędnościowego
Jaką kwotę dokładałoby Państwo?	2% miesięcznej pensji	1% miesięcznej pensji
Moment rozpoczęcia oszczędzania	Natychmiast w momencie rozpoczęcia pracy	3 miesiące od rozpoczęcia pracy w danej firmie
Organizacja rekomendująca	WBW	DEOC
Czy w umowie z pracodawcą jest przewidziana podwyżka?	Zaplanowana podwyżka w wysokości 10% pensji po 6 miesięcznym okresie próbnym	Zaplanowana podwyżka w wysokości 10% pensji po 6 miesięcznym okresie próbnym
Aktualne oszczędności na Twoim koncie	Ponad 10 tys. złotych	Mniej niż 1000 zł
Jaką kwotę dokładałby pracodawca?	1% miesięcznej pensji	2% miesięcznej pensji
Ile odkładał(a)byś miesięcznie od swojej pensji?	5%	1%
Otrzymałeś/aś rekomendację do uczestniczenia w tym programie od	Bliski znajomy/Ktoś z rodziny	Doradca finansowy z banku
	Wartość tego programu to: 200	Wartość tego programu to: 250

W której sytuacji jest bardziej prawdopodobne, że zapisał(a)byś się do programu oszczędnościowego? Wybór należy do Ciebie:

Program A Program B