

Prof. UG, dr hab. Anna Maria Zawadzka
Instytut Psychologii
Uniwersytet Gdański
ul. Bażyńskiego 4
80-592 Gdańsk

Gdańsk 27.05.2024

Recenzja pracy doktorskiej mgr Jakuba M. Krawca pt. „*Identyfikacja czynników decydujących o skuteczności interwencji behawioralnych w decyzjach ryzykownych*” napisanej pod kierunkiem dr hab., Wojciecha Białaszka, prof. Uniwersytetu SWPS.

Praca doktorska Pana Jakuba M. Krawca dotyczy dwóch kategorii interwencji behawioralnych/ podejść wdrożeniowych, zwanych - *nudge* i *boost*. Praca obejmuje: analizę porównawczą kategorii interwencji *nudge* i *boost* w zakresie ich skuteczności w podejmowaniu ryzykownych decyzji; sprawdzenie skuteczności tych interwencji w odniesieniu do perspektywy zysku i straty oraz wyboru dyskretnego w warunkach niepewności; i analizę wybranych moderatorów skuteczności każdej z tych interwencji behawioralnych. W ten sposób niniejsza rozprawa odnosi się do zagadnień z kilku obszarów psychologii – psychologii ekonomicznej, psychologii poznawczej i psychologii społecznej oraz ma charakter interdyscyplinarny – odwołuje się do dziedzin - ekonomii behawioralnej i finansów behawioralnych.

Struktura pracy jest prosta i przejrzysta. Rozprawa składa się z trzech rozdziałów: pierwszego poświęconego omówieniu podstaw teoretycznych oraz prezentacji koncepcji badań własnych; drugiego poświęconego prezentacji metod i wyników – pięciu badań własnych; oraz trzeciego, w którym doktorant podsumowuje wyniki przeprowadzonych badań.

Wprowadzenie teoretyczne jest zwięzłe (zawiera około 30 stron). Autor prezentuje w nim teoretyczne źródła dwóch podejść wdrożeniowych - *nudge* i *boost*, obejmujących zastosowanie heurystyk i zniekształceń poznawczych decydenta w polityce społecznej, finansowej i politycznej. Podejścia te związane są z podejmowaniem decyzji w otoczeniu ryzyka i niepewności i mają za zadanie ukierunkowywać decyzje decydenta ku opcjom, które uważane są jako pożądane społecznie lub ułatwiać wybór rozwiązania satysfakcjonującego dla decydenta w danym kontekście ekologicznym. Autor, w systematyczny sposób, prezentuje kolejne koncepcje jakie poprzedziły powstanie podejść *nudge* i *boost*. Mianowicie, omawia znaczenie w rozwoju podejść wdrożeniowych kolejno: Koncepcji dwóch światów w podejmowaniu decyzji (Savage), Teorii ograniczonej racjonalności (Simona), Teorii dwóch systemów

i heurystyk oraz zniekształceń poznawczych (Kahneman i Tverskiego) oraz Modelu racjonalności ekologicznej i prostych heurystyki (Gigerenzer i jego zespołu) aby następnie przejść do prezentacji koncepcji i badań w obszarze interwencji prowadzonych w nudge i boost. Tam też prezentuje przegląd heurystyk stosowanych w interwencjach – typu nudge i boost. We wprowadzeniu znajduje się także część poświęcona omówieniu wniosków z nielicznych badań dotyczących różnic indywidualnych (np. świadomości interwencji, cech osobowości, potrzeby poznania i potrzeby unikalności) jako moderatora efektywności interwencji behawioralnych. Niezwykle cennym podrozdziałem wprowadzenia teoretycznego jest rozdział „1.2.1 Wyzwania i ograniczenia badań nad interwencjami behawioralnymi”. Autor prezentuje w nim ważne zagadnienia etyczne i badawcze związane z stosowaniem tych interwencji, których jednostki i społeczeństwa powinny być świadome (np., aby podejścia te, umożliwiały procesy demokratyczne i ochronę autonomii jednostki, szczególnie gdy wykorzystywane są/ będą przez technologie AI). Część teoretyczna jest napisana w sposób jasny i przejrzysty, dobrze się ją czyta.

Pewnym ograniczeniem części teoretycznej jest sposób opracowania hipotez w podrozdziale 1.2.2 *Koncepcja badań własnych*. Hipotezy są postawione, ale żadna z nich nie jest uzasadniona, tj., nie wiadomo, na jakich badaniach, teoriach opiera się przypuszczenie o sile związku czy możliwej moderacji. W zamian autor omawia każde z badań własnych.

Część empiryczna rozpoczyna się badaniem 1. Jest nim metaanaliza na temat interwencji nudge i boost, przeprowadzona na podstawie badań poprzedników z lat 2008 -2021. Autor prezentuje kolejno kroki jakie podejmował w selekcji ostatecznej bazy danych i przedstawia wyniki metaanalizy uwzględniającej 32 próby (warunki). W ramach metaanalizy, dążąc do ustalenia zasadności prowadzenia analiz moderacji w ramach kategorii interwencji nudge i boost, posłużył się statystykami Q i I^2 Higginsa. Wykazały one na heterogeniczność analizowanych prób. Jednakże, na wykresie prezentującym szacowane wielkości efektu d-Cohena oraz odpowiadające mu interwały ufności można zauważyć, że warunki typu boost były (wyraźnie) mniej jednorodne niż interwencje typu nudge. Co więcej efekty miały zarówno znak pozytywny jak i negatywny. Zatem, pytanie czy, a jeśli tak to, na ile można wyciągnąć wiążące wnioski, co do różnic pomiędzy kategoriami wdrożeniowymi (tj. nudge i boost), skoro, są takie różnice wewnątrz kategorii? W związku z tym ograniczeniem, warto aby wnioski stawiane następnie w dyskusji metaanalizy, były prezentowane z należytą ostrożnością – nie

sugerowały, że jakaś kategoria jest lepsza (np.,...o potrzebie ponoszenia wyższych kosztów na technikę boost” str 52).

Wartym pochwalić jest to, że doktorant wykorzystuje metaanalizę do określenia wielkości efektów i tym samym prób, w badaniach własnych. Jest to wartościowy zabieg – podnosi poziom rzetelności kolejnych badań. Pewnym uchybieniem w opisie metaanalizy jest to, że zdarzają się tam potknięcia i niejasności językowe - np. na stronie 47 - „Analiza reszt ujawniła, że dwa trzy (?) siły efektów”, czy na stronie 46 - brak zakończenia zdania - „w analizie uwzględnionopochodzących z..”

Kolejne cztery badania własne zaprojektowano tak, aby weryfikacja postawionych hipotez odbywała się na kilka sposobów. W badaniach 2 (A i B) przetestowano techniki - dwie nudge i jedną boost w stosunku do ryzykownych decyzji finansowych. Badania te miały wykazać, jak skuteczność decyzji różni się gdy stosuje się strategię - nudge lub boost w stosunku do sytuacji kontrolnej (braku interwencji) w podejmowaniu decyzji finansowych z dwóch perspektyw – zysku i straty. W kolejnych badaniach 3 (A i B) przeanalizowano skuteczność technik boost i nudge w podejmowaniu decyzji dotyczących oszczędzania, w oparciu o różne informacje (tj., atrybuty sytuacji decyzyjnej). Badania 3 poszerzono także o nowe warunki interwencji behawioralnych. Mianowicie, uwzględniono dwie techniki typu nudge, dwie techniki typu boost i warunek mieszany (łączącym te dwie techniki wdrożeniowe) oraz przeanalizowano nie tylko różnice pomiędzy technikami boost czy nudge a grupą kontrolną ale także różnice pomiędzy technikami - nudge i boost i warunkiem mieszanym. Ponadto zarówno w badaniach 2 jak i badaniach 3 testowano moderacje tj., jak cechy indywidualne (np. cechy osobowości czy perspektywa czasowa), i środowiskowe (dane demograficzne) różnicują skuteczność decyzji, gdy wykorzystuje się wspomaganie decyzji technikami wdrożeniowymi. Doktorant w kolejnych badaniach (2 i 3) testuje inne charakterystyki indywidualne jako moderatory. Jednakże niewystarczająco uzasadnia, dlaczego rezygnuje z cech osobowości i temperamentu a wprowadza nowe zmienne (np. perspektywa czasowa) i dlaczego te a nie inne (np. tendencja do przetwarzania racjonalnego i doświadczeniowego - S. Epstein). Efekty moderacji w badaniach 2 były niespójne (por. Model B str. 67 i str. 73 – „żaden z istotnych predyktorów osobowościowych nie wchodził w interakcję z interwencjami”), dlatego byłaby wskazana, kolejna weryfikacja tych efektów (w kolejnym badaniu, np. 3). Ponadto, w badaniach 2 zastosowano wiele skal osobowości. Odpowiedzi na jedne skale mogły wpływać na odpowiedzi na następnych skalach. Nie zastosowano randomizacji skal osobowości (a przynajmniej nie ma o tym mowy w rozprawie).

Z kolei, warte pochwały i podkreślenia są starania doktoranta mające na celu zwiększenie rzetelności otrzymanych wyników. Pierwszym z nich jest zabieg polegający na przeprowadzenie weryfikacji hipotez badań 2 poprzez użycie dwóch (procedur) pomiarów zmiennej zależnej. Drugim, jest przeprowadzenie w ramach badań 3 dwukrotnego pomiaru zmiennej zależnej, sprawdzającego efekt interwencji w czasie. Wszystkie badania poprzedzono analizą GPower, która pozwoliła ustalić wielkości próby zwiększające rzetelność badań.

Ograniczeniem części empirycznej jest niewystarczająco staranny sposób opisu metod badań. W przypadku każdego opisu, nie wiadomo, dlaczego zastosowano inną strukturę w badaniach 2 i w badaniach 3, (np. w bad. 2 doktorat zaczyna od opisu celu a w bad. 3 od opisu warunków). W przypadku badań 2 - opis metody badania jest niedostatecznie uporządkowany i ma braki (np. nie opisano grupy kontrolnej). Z kolei, metoda badań 3 jest opisana w poprawny sposób - można dowiedzieć się dokładnie na czym polegały badania, z czego się składały i jak przebiegały. Jednakże brakuje tam dokładnego opisu prób (tj., ile K? ile M?, w jakim wieku M=?, SD=?). W badaniu 3 brakuje także omówienia efektów modelu 2, pomimo, że jest wykazany jako testowany (str. 133 i 135).

Wyniki badań na ogół zaprezentowane są w profesjonalny sposób. Doktorant testuje zakładane modele z zastosowaniem poprawnie dobranych testów statystycznych. Testowane modele są prezentowane w przejrzystych tabelach. Wyniki badania 3A wskazały, że nie zawsze interwencja działania zgodnie z założeniem – tj., wybiera się opcję wspieraną interwencją. Powstaje pytanie, na ile założenia interwencji behawioralnych są trafne co do wyboru decydenta i od czego to zależy, że decydent ulega interwencji?

Ograniczeniem wszystkich przeprowadzonych badań jest to, że badani są studentami psychologii i że przebadano zdecydowaną większość kobiet (proporcja zbadanych kobiet w stosunku do mężczyzn; 1-4/5). Pewnym niedociągnięciem, jest to, że doktorat czasami upraszcza wnioski ze swoich badań. Mianowicie, we wnioskach pisze o tym, że efekty moderacji cechami indywidualnymi w badaniach (2 i 3) były istotne. Jednakże w badaniu 2 efekty były istotne tylko w jednej z dwóch testowanych procedur (PDQ) zatem warto o pamiętać w interpretacji wyników. W dyskusji wyników, w części interpretacji, zabrakło odniesienia się do mechanizmu działania heurystyk, na których opierały się badania. Jak można zauważyć wyniki nie zawsze są spójne w ramach danej kategorii interwencji (np. nudge czy boost), a ta niespójność może być efektem specyfiki poszczególnych heurystyk czy technik wpływu społecznego. Na przykład, dlaczego dochody są moderatorem efektywności nudgingu

normatywnego a w przypadku innych technik ta zależność nie występuje (im wyższe dochody tym większa podatność na technikę nudgingu normatywnego, str. 120)?

Pomimo wspomnianych ograniczeń pracy uważam, że praca ma wiele silnych stron, które dowodzą jej wartości naukowej i aplikacyjnej. Mocnymi stronami rozprawy są:

- 1) praca poszerza dotychczasową wiedzę na temat efektywności interwencji nudge i boost przy podejmowaniu decyzji ryzykownych o decyzje w warunkach niepewności,
- 2) praca poszerza dotychczasową wiedzę o efektywności interwencji behawioralnych w sytuacji niepewności, przy decyzjach wieloatrybutowych/ dyskretnych (w dziedzinie oszczędzania),
- 3) praca poszerza dotychczasową wiedzę na temat moderatorów efektywności interwencji nudge i boost o charakterystyki indywidualne decydenta. Jest to novum w stosunku do istniejących badań,
- 4) novum pracy stanowi także badanie efektywności interwencji mieszanej – zawierającej jednocześnie cechy technik boost i nudge,
- 5) problem badawczy został zweryfikowany za pomocą różnych metod (metaanaliza na danych zastanych, quasi-eksperyment, powtarzane pomiary, metoda kwestionariuszowa), uwzględniając różne techniki interwencji z typu nudge (np. dyskluzyny, normatywny) i z typu boost (np. soczewkowy, Reguła Franklina) oraz z zastosowaniem różnych procedur (np. FFR Franklina i zespołu i PDQ Maddena i zespołu).
- 6) praca zawiera solidne analizy statystyczne oparte na testowaniu szeregu modeli i z zastosowaniem wielu dobrze dobranych statystyk,
- 7) praca dostarcza cennych informacji mogących mieć zastosowanie w aplikowaniu interwencji behawioralnych we wspieraniu polityki społecznej.

Tym samym stwierdzam, że rozprawa doktorska mgr Jakuba M. Krawca w pełni spełnia stawiane pracy doktorskiej wymogi i proszę o dopuszczenie do dalszych etapów przewodu doktorskiego.

Anne Marie Wredike

