

Opinia o pracy doktorskiej

Mgr Jakuba M. Krawca

„Identyfikacja czynników decydujących o skuteczności interwencji behawioralnych w decyzjach ryzykownych”

Zacznę od przytoczenia kilku próbek stylu w jakim została napisana ta praca.

S. 6 *„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.”*

s. 11 *„... Kahneman i Tverski 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 aksjomatów racjonalnego wyboru”*

s. 36 *„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 kontrolowanych w warunkach laboratoryjnych.”*

Przytoczone przykłady nie tylko pokazują skrajną niesprawność językową, ale w wielu miejscach czynią tekst kompletnie nieczytelnym.

Dodatkowym utrudnieniem czytania tej pracy jest nadmiar terminów anglojęzycznych. W szczególności dominuje nadmiar dwóch terminów: „nudgingy” i „boostingy”. Dwa te terminy powtarzają się w pracy mgr. Krawca przez całą pracę. Pierwszy termin odwołuje się do impulsów wywierających wpływ na zachowanie jednostki (jednak bez ograniczania dostępnych opcji). Drugi termin odwołuje się do interwencji zmierzających do usprawnienia predyspozycji jednostki do dokonania racjonalnego wyboru. Niestety oba te terminy trudno przetłumaczyć na język polski. W tej sytuacji autor recenzowanej pracy postanowił pozostać przy terminach angielskich, których używa w liczbie pojedynczej i mnogiej i odmienia przez wszystkie siedem przypadków. Rozwiązanie to stanowczo mi nie

odpowiada. Oczekiwałbym, że doktorant mógłby postarać się o jakieś tłumaczenia tych terminów. Nie mając żadnych sugestii, pierwszy termin będę tłumaczył jako „*impuls(y) wywierające presję na zachowanie jednostki*”, a krócej „*impuls(y) wywierające presję*”. Drugi termin będę tłumaczył jako „*usprawnienie predyspozycji jednostki do dokonania wyboru*”, a krócej „*usprawnienie predyspozycji jednostki*”.

Decydując się na przesłanie zawartości tego mało czytelnego tekstu poddam tę zawartość merytorycznej ocenie. Rozpoczyna ją wstęp teoretyczny, w którym w miarę jasno wyłożono koncepcję impulsów wywierających presję oraz usprawnianie predyspozycji jednostki. Mimo zawiłości językowych autor raczej poprawnie wyjaśnił podjęty temat. Także w miarę jasno przedstawił rozróżnienie między normatywnym i deskryptywnym podejściem do problemów podejmowania decyzji.

Jednakże w części wprowadzającej podstawowe pojęcia i koncepcje teorii podejmowania decyzji wywody autora są zdecydowanie nieudolne. Autor dyskutuje mało ważne błahostki, a z drugiej strony nie zauważa poważnych tematów – np. teorii użyteczności von Neumanna i Mongersterna czy teorii perspektywy Tversky’ego i Kahnemana itd.

W części empirycznej autor zaprojektował trzy badania. Pierwsze z nich dotyczyło metaanalizy publikacji w których używano terminów „nudge” i „boost”. W badaniu tym autor postawił hipotezę (H1), że *istnieją różnice w sile wpływu oddziaływania interwencji typu boost i nudge na decyzje w ryzyku i niepewności*. Pierwsze co się rzuca w oczy w tym sformułowaniu, to nieukierunkowany kształt postawionej hipotezy. Zakładam, że jest to jakieś niedopatrzenie. Rozwój nauki następuje zwykle w wyniku sprawdzania hipotez, które badacz wyprowadza z trafnie skonstruowanych teorii, a nie stawia ogólnikową hipotezę o istnieniu różnic.

Chociaż hipoteza H1 jest obarczona błędem, pomysł przeprowadzenia metaanalizy wydaje mi się słuszny. Oczywiście materiał którym dysponował badacz nie rozstrzyga postawionej kwestii a jedynie formułuje pewną sugestię. Trzeba także powiedzieć, że przeprowadzone badanie posiada zdecydowanie mocną stronę. Jest nią poziom statystycznych analiz otrzymanych wyników. Autor znacznie przekracza tu wysoki poziom kompetencji. Dodam, że przekracza w tym także kompetencję obecnego recenzenta. Nie jest

to jednak tylko kwestia kompetencji recenzenta. Tak jak w całej książce, autor mało dba o czytelność jego wywodów.

Przechodzę teraz do najważniejszego badania tej pracy – do badania Nr. 2. Jest ono w dużym stopniu replikacją badania Franklina i in. (2019). Dodam od razu, że jest replikacją interesującego badania. Niestety, w wielu sprawach autor zepsuł tę replikację. Najpierw znacznie utrudniając czytanie tej pracy. (Nawiasem mówiąc, żeby zrozumieć pracę mgr. Krawca musiałem w wielu miejscach wracać do pracy Franklina i in.)

Praca mgr. Krawca posiada mnóstwo spiętrzonych nieporozumień i błędów. Muszę je tu kolejno wymienić: W badaniu tym autor postawił pytania które, z jakichś powodów nazywa hipotezami. Odwołują się one do podejmowania ryzyka w dziedzinie zysków i w dziedzinie strat. H2.1: *Istnieją różnice w sile wpływu oddziaływania interwencji typu usprawnienia predyspozycji (boost) i typu impulsu (nudge) w domenie zysków oraz H2.2: Istnieją różnice w sile wpływu oddziaływania interwencji typu usprawnienia predyspozycji (boost) oraz typu impulsu (nudge) w domenie strat.*

W ten sposób autor nawiązał do znanego problem stosunku ludzi do ryzyka w dziedzinie zysków i w dziedzinie strat. Temat ten starannie przeanalizowali pięćdziesiąt lat temu Amos Tversky i Daniel Kahneman w swojej teorii perspektywy. Od tej pory powstało dziesiątki (o ile nie setki) badań na ten temat. Nie wiem jednak z jakich powodów autor dodał do swojego badania to rozróżnienie. A oto jak autor sam uzasadnia znaczenie swojego badania 2:

„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”

Autor słusznie napisał, że rozróżnienie decyzji w warunkach ryzyka i w warunkach niepewności pochodzi od Knighta i jest bardzo ważne. Istotnie rozróżnienie to ma już za sobą 100 lat i obie sytuacje wyboru - w warunkach ryzyka i w warunkach niepewności - zostały bardzo wszechstronnie przeanalizowane. Po tym stwierdzeniu czytelnik chciałby się ewentualnie dowiedzieć co z tego wynika dla podjętego przez autora tematu. Jednakże

dyskusji przeprowadzonej na str. 143 - 144 nie zrozumiałem. W szczególności nie rozumiem co znaczy, że rozróżnienie to „*pozwoliło trafnie odwzorować sytuacje decyzyjne w których mogą działać opisywane interwencje zgodnie ze swoimi założeniami*”?

Badanie 2 pokazało, że wybrane interwencje typu usprawnienia predyspozycji (boosting) prowadziły do większej liczby optymalnych wyborów (o większej oczekiwanej wartości) niż interwencje typu impulsu. Wynik ten okazał się zgodny z wynikami Franklina i in. Stąd mgr. Krawiec uogólnia swój wynik, że interwencje typu usprawnienia predyspozycji są bardziej skuteczne niż interwencje typu impulsu. Jednakże generalnie nie da się ustalić który rodzaj interwencji - zmierzające do wywarcia wpływu, czy zmierzające do usprawnienia predyspozycji - jest lepszy. Podsumowując swoje badanie na ten temat Franklin i in. (2019) napisali, że w badaniach tych nie chodziło o odkrycie uniwersalnie najlepszej procedury wpływania na zachowanie, a raczej o znalezienie najwłaściwszego kontekstu, populacji i metody badania. Doktorant nie poszedł tą drogą, a raczej poszukiwał, który rodzaj interwencji jest bardziej skuteczny. Jednakże jego postępowanie nie pozwala odpowiedzieć na to pytanie. Podobnie jak konstatacja, że jeżeli komuś bardziej smakuje zupa grochowa niż ciasto dyniowe, to nie wynika z tego, że preferuje on dowolną zupę nad dowolny deser.

W badaniu 2 autor sprawdzał trwałość efektów interwencji behawioralnych przy dwu rodzajach interwencji - zmierzające do wywarcia wpływu, oraz zmierzające do usprawnienia predyspozycji. Badanie nie potwierdziło tych hipotez. Autor pointuje to stwierdzeniem, że hipotezy te zostały odrzucone. Tu potrzebna jest drobna korekta: brak potwierdzenia hipotezy nie oznacza odrzucenia tej hipotezy. A nadto nie podoba mi się poszukiwanie hipotez „po omacku”. Hipotezy powinny być wyprowadzane z przemyślanych teorii, a nie z prostych domysłów. Toteż jeszcze gorzej wygląda cały zestaw cech, które mogłyby być moderatorami efektywności interwencji behawioralnych. Jak można było przewidzieć z góry, niektóre cechy odgrywały taką rolę inne – nie. Poszukiwanie po omacku znaczenia różnic indywidualnych nie przynosi na ogół ciekawych odkryć.

„Przebicie się” przez opis badania Nr. 3 kosztowało mnie wiele trudu. I nie mam wcale pewności, czy zrozumiałem sens tego badania. Autor zapowiada, że zajął się w nim zachowaniami oszczędnościowymi. Osoby badane dostawały wybór pomiędzy dwoma opisami programów oszczędnościowych. Programy te były opisane na dziewięciu atrybutach.

Takich np. jak „z programu korzysta 10% z Twoich znajomych, „program oszczędzania podjęto 6 mies. od rozpoczęcia pracy” itd. Badanie to miało testować hipotezę, że interwencje typu nudge i boost będą bardziej skuteczne niż interwencje uzyskiwane za pomocą osobnych interwencji. Jakoś nie przychodzi mi do głowy, żebym swoje wybory między programami oszczędnościowymi opierał na wymienionych dziewięciu atrybutów.

Na koniec chciałbym zwrócić autorowi uwagę na wielokrotne powtarzanie mniej więcej tego samego. Po każdym badaniu autor przeprowadza dyskusję uzyskanych wyników, a następnie dokonuje „uroczystego” podsumowania uzyskanych wyników. Zwykle te same wyniki są powtarzane w każdej takiej części i nic nowego z tego nie wynika.

Opisane słabości i błędy recenzowanej pracy nie pozwalają uznać jej za spełniającą wymagania ustawy o stopniach naukowych. Sama znajomość zaawansowanych metod analizy danych, za co chwaliłem doktoranta, nie wystarczy do osiągnięcia wartościowego celu naukowego. Żeby to osiągnąć trzeba znać i stosować podstawowe zasady metodologii. W szczególności trzeba starannie sformułować hipotezy badawcze. Powinny one być wyprowadzone z podstawowych koncepcji, teorii, modeli ... Oceniana praca jest pod tym względem wysoce zagmatwana. Nie wiadomo, co jest założeniem, a co wnioskiem itd. Wbrew temu, co twierdzi autor praca oparta na przedstawionych badaniach nie rozstrzyga, która kategoria interwencji behawioralnych jest bardziej skuteczna w usprawnianiu decyzji.

W sumie, nie mogę zatem uznać pracy mgr Jakuba M. Krawca za wykonaną zgodnie z regułami pracy naukowej. W efekcie muszę stwierdzić, że praca „Identyfikacja czynników decydujących o skuteczności interwencji behawioralnych w decyzjach ryzykownych” nie spełnia wymagań ustawy o stopniach naukowych.

31.08.24 Tadeusz Tyszka

