

Kraków, dnia 25. sierpnia 2022r.

Prof. UJ dr hab. inż. Wiktor Adamus

Instytut Ekonomii, Finansów i Zarządzania

Uniwersytet Jagielloński

RECENZJA

rozprawy doktorskiej Mgr Marty Kabut

pt.: „False Respondents in Web Human Resource Surveys” („Falszywi respondenci w ankietach internetowych w badaniach z zakresu zarządzania zasobami ludzkimi”)

napisanej pod kierunkiem prof. dr hab. Grażyny Wieczorkowskiej-Wierzbińskiej oraz promotora pomocniczego dr Anny Kuźmińskiej

w Zakładzie Psychologii i Socjologii Zarządzania na Wydziale Zarządzania Uniwersytetu Warszawskiego

1. Przesłanki formalne napisania recenzji

Podstawą formalną przygotowania recenzji rozprawy doktorskiej pod w/w tytułem stanowi:

- uchwała Rady Naukowej Dyscypliny Nauki o Zarządzaniu i Jakości Uniwersytetu Warszawskiego z dnia 22.12.2021 oraz pismo prof. dr hab. Grzegorza Karasiewicza Przewodniczącego Rady Naukowej Dyscypliny Nauki o Zarządzaniu i Jakości z dnia 25.05.2022
- ustawa o stopniach i tytułach naukowych z dnia 14 marca 2003 r. (Dz. U. 2003, Nr 650, poz. 595 z późn. zm. zawartymi w Dz. U. Z 2017 r., poz. 1789)

2. Cel rozprawy, problematyka badawcza, metodyka badawcza oraz wykorzystana w rozprawie literatura przedmiotu

Należy z uznaniem powitać rozprawę mgr Marty Kabut na temat „Falszywi respondenci w ankietach internetowych w badaniach z zakresu zarządzania zasobami ludzkimi”. Problem identyfikacji respondentów udzielających odpowiedzi niezgodnych z ich wewnętrznymi przekonaniem lub też odpowiedzi nieuważnych, czy wręcz niepoważnych, zyskał ogromnie na znaczeniu w dobie popularności badań realizowanych za pośrednictwem Internetu. Przed upowszechnieniem się badań internetowych mieliśmy do czynienia z inną sytuacją, w której respondent/ka odpowiadał/a na pytania ankietera lub ankieterki w toku bezpośredniego spotkania – Autorka wspomina o tym. Choć z pewnością taki bezpośredni kontakt nie eliminował kłopotów związanych z tzw. fałszywymi respondentami, to jednak znacznie redukuje przenikanie danych uzyskanych od nierzetelnych respondentów do raportów i opracowań empirycznych. Złudzenie anonimowości, a także łatwość towarzysząca badaniom via Internet i w ogóle posługiwania się Internetem, sprawiły, że badania przekształciły się w wielu przypadkach w rodzaj zabawy, w odmianę popularnych quizów. Nie jest wykluczone, że są respondenci tak właśnie traktujący swój udział w badaniach, jako rodzaj psychozabawy. Można by więc zaryzykować twierdzenie, że przynajmniej część badanych (trudno jednak określić, jak duża) podlegała swoistej socjalizacji do ludycznego rozumienia badań realizowanych online.

Autorka podejmuje problematykę ściśle metodologiczną – zmierza do wykorzystania dostępnych metod badawczych w identyfikacji tzw. fałszywych respondentów po to, by wykluczyć ich wpływ na wyniki badań. Zmierza w szczególności do opracowania i przetestowania procedury wykrywania takich respondentów. Przyjmuje – na podstawie analizy obszernej literatury przedmiotu - że, istnieją cztery wskaźniki sugerujące, iż mamy do czynienia z nierzetelnym (fałszywym) respondentem. Pierwszy z nich to czas odpowiedzi – jeśli odpowiedź sformułowana jest w bardzo krótkim czasie, krótszym, niż w przypadku zdecydowanej większości respondentów, zachodzi uzasadnione przypuszczenie, że udzielił jej „fałszywy respondent”. Drugi to reakcje respondenta/respondentki na pytania kontrolne, trzeci – styl odpowiedzi i liczba odpowiedzi „nie wiem” oraz wskaźnik czwarty – miary deklarowanej i rzeczywistej współpracy z badaczem. Cel badań został precyzyjnie sformułowany, co z pewnością przełożyło się na trafny dobór instrumentów analitycznych rozprawy.

Niejako z obowiązku recenzenta zaznaczymy, że można, rzecz jasna, poprzestać na wyróżnieniu wskaźników „bycia fałszywym respondentem” i dla eliminacji wpływu takich osób na wyniki badań może to okazać się wystarczające. W badaniach społecznych, także w badaniach realizowanych dla potrzeb zarządzania zasobami ludzkimi, można by spróbować

pójść w nieco inną stronę i podjąć próbę budowy, ewentualnie wykorzystania typologii w opisie wyników badań.

W literaturze przedmiotu istnieją typologie fałszywych respondentów, wyszczególniające na przykład „zawodowych poławiaczy nagród” („professional competition entrants”) zapowiadanych w zachęcie do udziału w sondażu (o tych respondentach Autorka wspomina), „sondażowych sprinterów” („survey speeders”), którzy mogą nawet z zapałem przystępować do udzielania odpowiedzi, lecz po chwili tracą chęć i zainteresowanie treścią kwestionariusza, zwłaszcza, gdy jest to kwestionariusz dłuższy, liczący kilkadziesiąt pytań. Są też respondenci, którzy odpowiadają na pytania automatycznie („straight – liners”), zaznaczając te same odpowiedzi w pytaniach przybierających postać tabelaryczną. Wreszcie, niektóre badania sondażowe wymagają reakcji zbliżonych do głosowania (np. rangowanie atrakcyjności przedstawionych w pytaniu opcji). W takim przypadku mogą się pojawić respondenci „głosujący blokiem” („block voters”), co ma miejsce zwłaszcza wtedy, gdy są oni mocno zainteresowani rozwiązaniem proponowanym w pytaniu, wprowadzeniem jakiejś innowacji, zmiany, którą postrzegają jako bardzo korzystną dla siebie. W takiej sytuacji będą oni nie tylko wskazywać tę opcję ale i inne związane z nią możliwości, będą też skłonni namawiać współpracowników do wskazywania tych samych odpowiedzi. W typologiach nie zawsze wskazuje się na jeszcze jeden rodzaj fałszywych respondentów, mianowicie takich, którzy kierują się zasadami politycznej poprawności i ukrywają swoje rzeczywiste opinie i poglądy, starając się udzielać tylko takich odpowiedzi, o których wiedzą, że są społecznie akceptowane. Ta grupa respondentów być może ma mniejszy udział w badaniach z zakresu zarządzania zasobami ludzkimi, choć i na tym obszarze istnieją sprawy niezmiernie wrażliwe na wspomnianą poprawność (kwestia zarządzania różnorodnością). Podejmowanie wysiłków na rzecz budowania typologii fałszywych respondentów i uwzględnienie typologii fałszywych respondentów w interpretacji wyników badań nad ich fenomenem jawi się jako istotne wyzwanie badawcze, choć także jako zadanie dość złożone.

Wspomniany wyżej cel pracy uległ operacjonalizacji na dwa zadania badawcze: (a) oszacowanie poziomu „nieuwagi” respondenta (*inattention*) i (b) oszacowanie konsekwencji zignorowania problemu i przetestowanie użyteczności procedury FLEXMIX. Procedura FLEXMIX jest znana od stulecia, lecz od 20 lat przeżywa renesans, znajdując zastosowania w różnych dyscyplinach naukowych (biologii, medycynie, fizyce, ekonomii i marketingu). Oczywiście, procedura ta znalazła też liczne zastosowania w badaniach socjologicznych i antropologicznych, w analizach różnic kulturowych, kulturowego kontekstu komunikacji czy zarządzania. Autorka nie cytuje opracowań z tej akurat dziedziny zastosowań procedury,

jednak nie należy czynić z tego zarzutu, gdyż odwołania do bardzo kompetentnie skonstruowanej literatury przedmiotu są wystarczające i dokumentują profesjonalne znanstwo Autorki. Bibliografia pracy liczy z górą 150 pozycji, w większości artykułów ze źródeł internetowych, obejmuje prace w języku angielskim (jednak nie tylko autorów angielskich czy amerykańskich) oraz kilkanaście prac w języku polskim. 30% wymienionych w pracy pozycji bibliograficznych nie przekracza 6 lat, zatem wykorzystanie aktualnego piśmiennictwa naukowego należy zaliczyć do walorów rozprawy.

Metodyka badawcza przyjęta w pracy opierała się w części na wtórnej analizie zastanych wyników badań (trzy zbiory danych z badań międzynarodowych, z których uwzględniono respondentów polskich w liczbie 5604 osób), w części zaś na analizie danych zgromadzonych dla potrzeb recenzowanej rozprawy (dziewięć zbiorów, obejmujących ogółem 3169 respondentów). Podkreślić należy, iż obie grupy zbiorów zostały bardzo starannie dobrane, na podstawie oceny zastosowanej metodyki badań (źródła zastane) oraz na starannym zaprojektowaniu tej metodyki (badania własne). Analiza przeprowadzona na tych dwóch głównych zbiorach danych pozwoliła na wyciągnięcie szeregu interesujących i dobrze uzasadnionych wniosków.

3. Układ i zawartość merytoryczna rozprawy

Rozprawa składa się z „Części teoretycznej”, podzielonej na trzy rozdziały i jedenaście podrozdziałów oraz z „Części empirycznej”, podzielonej na piętnaście rozdziałów i trzynaście podrozdziałów. Całość uzupełnia Aneks, zawierający jedenaście załączników prezentujących dane z badań, bibliografia oraz spisy wykresów i tabel.

Merytoryczna zawartość rozprawy obejmuje w części teoretycznej prezentację badań internetowych jako sposobu gromadzenia danych, zalety i ograniczenia tych badań, głównie sondażowych. Zawiera też model odpowiedzi respondentów na pytania sondażu, opis „fałszywych respondentów” oraz wpływ tych respondentów na wyniki badań, klasyfikację wskaźników bycia „fałszywym respondentem”, następnie przegląd prac w których przedstawione są rezultaty badań nad kształtem części tych wskaźników. Finalny akord tej części rozprawy to model fałszywego respondenta, stanowiący główny obiekt badawczy. Jednakże użycie tego sformułowania („false respondent model”) wydaje się nieco ponad miarę. W odpowiedniej części pracy znajdujemy na kilku stronach prezentację tego zagadnienia, jednak w większym stopniu jest ono enumeracją cech identyfikowanych u różnych respondentów, udzielających odpowiedzi przypadkowych, nielogicznych, nieuważnych itp. Cechy te mogą znamionować zarówno intencjonalne zachowanie badanych

(celowe okłamywanie), jak i zachowanie niezamierzone, wynikające z niezrozumienia treści pytania, niestarannej metodyki badań, braku zainteresowania, zdarzeń losowych towarzyszących odpowiadaniu na pytania i innych podobnych czynników, niezawinionych przez respondenta. Zestaw cech, który cechuje respondentów „niewspółpracujących”, a więc fałszywych, obejmuje ich następujące zachowania: wskazywanie pierwszej odpowiedzi, która wydaje się rozsądna, sposób wyświetlania odpowiedzi, pośpiech, skłonność do akceptowania dowolnego twierdzenia, niezależnie od jego treści, wskazywanie na wartość wyjściową w odpowiedziach na pytania o natężenie jakiegoś zjawiska, brak różnicowania obiektów na skalach ocen, udzielanie odpowiedzi „nie wiem”, mentalne „rzucanie monetą”, pomijanie całych sekwencji pytań. Uwaga, iż posłużenie się terminem „model fałszywego respondenta” jest nieco na wyrost, nie jest zarzutem, lecz raczej sugestią, i zachętą, by w przyszłych swoich analizach tej ogromnie ważnej problematyki Autorka śmieiej skorzystała z typologii jako narzędzia syntezy (a nie analizy) wyników badań. Warto by wówczas poświęcić więcej uwagi związkom korelacyjnym między wykazywanymi w badaniach zachowaniami i cechami demo-społecznymi respondentów.

Część empiryczna rozprawy jest dość obszerna i zawiera sprawozdanie z rezultatów zastosowania wybranej procedury badań. Zawiera ona operacjonalizację czterech „sygnałów ostrzegawczych” oraz testuje ich kształt na dwunastu zbiorach danych – trzech zbiorach offline i dziewięciu próbach badanych przy pomocy sondażu internetowego. Autorka wyczerpująco opisuje poszczególne fazy badania, składające się na realizację dwóch zadań: szacowania zakresu występowania fałszywych respondentów w badaniach oraz szacowania konsekwencji niewykluczenia fałszywych respondentów ze zbiorów danych. Autorka wykazała, że odsetek respondentów kwalifikowanych jako „fałszywi” zależał od badania (rodzaju próby) i rodzaju sygnału ostrzegawczego. Dla sześciu sondaży ankietowych najsilniej wykluczającym kryterium okazała się ocena zgodności logicznej i poziomu zaangażowania respondenta (od 14,3% do 25,4%). Z kolei w przypadku dwóch zbiorów panelowych najsilniej wykluczającym czynnikiem okazał się czas (56,4% vs 26,1%) i również ocena zgodności logicznej i zaangażowania respondenta (48,1% vs 30,3%). Różnice w rozkładach danych z prób panelowych są, jak pisze Autorka, dość zaskakujące (badania zostały przygotowane przez zespół badawczy z udziałem Autorki i były realizowane przez tę samą firmę komercyjną). Autorka przypisuje tę różnicę odmiennym typom pytań sprawdzających uwagę, użytych w obu kwestionariuszach. Trudno oprzeć się wrażeniu, że nie ustrzeżono się błędu tzw. instrumentacji, czyli zmiany w narzędziu pomiarowym między pierwszym i drugim pomiarem. Jednak wskazanie na to możliwe źródło rozbieżności (czyli określenie rozbieżności jako artefaktu uzyskanego w wyniku zastosowania dwóch różnych

miar tego samego fenomenu) stanowi wystarczający dowód ostrożności Badaczki i ostrzeżenie przed pochopną interpretacją.

Jeśli zaś chodzi o drugie zadanie badawcze, czyli wykazanie skutków ignorowania fałszywych respondentów w badaniach, zastosowano porównanie rzetelności pomiaru (*alfa Cronbacha*) w grupach respondentów fałszywych i uważnych (niewykluczonych). Przetestowano też przydatność dwóch procedur podziału grupy badanej na respondentów fałszywych i uważnych (procedura czterech wskaźników oraz procedura FLEXMIX). FLEXMIX pozwolił podzielić respondentów na dwie grupy na podstawie korelacji między ich odpowiedziami na dwa pytania; pytania miały wykluczającą się treść, więc uzyskanie dodatniej korelacji świadczyło o tym, że respondenci nieuważnie czytali pytania. I znów, jak poprzednio, pojawiły się znaczne różnice w szacowaniu „fałszywości”. Procedura 4WS szacowała liczbę fałszywych respondentów na 46% w zbiorze A1 i na 17% w zbiorze A2. Procedura FLEXMIX – odpowiednio na 32% i 34% w obu tych zbiorach. Autorka tłumaczy te wyniki w następujący sposób: procedura FLEXMIX ma charakter „lokalny”, to znaczy wykorzystuje korelację między odpowiedziami na jedynie dwa pytania. Z kolei procedura 4WS ma z kolei charakter „globalny”, gdyż odnosi się do zachowania respondenta w całym badaniu. W konkluzji Autorka stwierdza, że procedura FLEXMIX może być zalecona tylko do sygnału ostrzegawczego 4, jednak z wykorzystaniem większej liczby pytań niż tylko dwa. Badania Autorki odłoniły istotny problem w szacowaniu rzetelności pomiarów rozmaitych cech i dyspozycji osób poddawanych badaniom społecznym. Należałoby przyjąć, że im mniej różnią się pomiary tej samej cechy, dokonane przy użyciu różnych instrumentów pomiarowych, tym wyższa jest rzetelność pomiaru. To jednak, niestety, zdarza się dość rzadko. Wiedzą o tym wszyscy ci, którzy choć raz dokonywali szacunku rzetelności zbudowanych przez siebie skal do pomiaru postaw. Należałoby też zastanowić się, czy w przyszłych badaniach nad podejmowaną przez Autorkę problematyką nie należałoby skorzystać z dodatkowej metody szacowania także trafności wykorzystywanych instrumentów pomiaru, metody zwanej „konfirmacyjną analizą czynnikową” (K.G., Jöreskog. 1969. A general approach to confirmatory maximum likelihood factor analysis. *Psychometrica* 34, 183-202).

4. Uzyskane wyniki oraz wnioski z badań

Badania dostarczyły szeregu interesujących, niekiedy zaskakujących wniosków. Za cenny wkład poznawczy należy uznać pomiar częstości występowania fałszywych respondentów w dziesięciu dobrze przygotowanych badaniach internetowych. Autorka wykazała, że brak oczyszczenia próby w danych uzyskiwanych od tych respondentów obniża rzetelność

pomiarów. Ma to szczególne znaczenie w tych dziedzinach społecznej praktyki, gdzie od badań do wdrożenia wynikających z nich sugestii wiedzy niedługa droga; jest tak z pewnością w przypadku zarządzania zasobami ludzkimi. Wykorzystanie danych obarczonych tendencyjnością może bardzo utrudnić albo nawet zniszczyć diagnostyczną siłę interwencji w ludzki potencjał organizacji.

Rozprawa wnosi też oryginalny wkład w metodykę wykrywania fałszywych respondentów – sprowadza się on do przetestowania procedury „czterech wskaźników” oraz wykorzystania procedury FLEXMIX do identyfikowania logicznych niespójności w odpowiedziach respondentów.

Praktyczna użyteczność przedstawionych badań i ich wyników polega zaś głównie na tym, że zarządzający zasobami ludzkimi będą mogli zwiększyć diagnostyczną moc swoich narzędzi diagnostycznych, przyczyniając się do pełniejszej realizacji swoich zamierzeń.

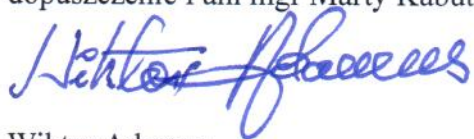
5. Strona edytorska rozprawy

Strona edytorska nie budzi zasadniczych zastrzeżeń. Spis treści jest wyraźnie rozdrobniony, lecz wynikało to z formuły narracji przyjętej przez Autorkę, akcentującej kolejne fazy realizacji metodyki badań. Spis treści jest odzwierciedlaniem raczej logiki realizacji badań, a nie projekcją jakiejś apriorycznej struktury porządkującej. Praca napisana jest prostym językiem, posługuje się dobrze zdefiniowanymi pojęciami. Zawiera starannie opracowaną metodykę badań i w jasny sposób prezentuje wnioski pobadawcze.

6. Konkluzja

Pomimo kilku krytycznych uwag sformułowanych wyżej, ogólna ocena rozprawy doktorskiej Pani mgr Marty Kabut jest pozytywna. Doktorantka zrealizowała sformułowane przez siebie cele badawcze, w bardzo staranny sposób dobrała metody badań i poprawnie zastosowała je do zebranych danych. Za istotny wkład Autorki do dorobku naukowego uznaję podjęcie problematyki niezbyt często sygnalizowanej w opracowaniach polskich ale i zagranicznych (Autorka skupiła wszak uwagę na badaniu fenomenu fałszywych respondentów w badaniach dla potrzeb zarządzania zasobami ludzkimi). Rozprawa Pani Mgr Marty Kabut z pewnością przyczyni się – po opublikowaniu – do wzrostu metodologicznej samoświadomości praktyków z zakresu zarządzania zasobami ludzkimi, może też posłużyć jako istotna pomoc dydaktyczna w realizacji przedmiotu metody badań społecznych, nie tylko na wydziałach zarządzania. Podsumowując, warsztat badawczy oraz waga (nie tylko ze względów naukowych ale również praktycznych) wyników badań i wyciągniętych wniosków, pozwala na stwierdzenie, że przedstawiona do recenzji rozprawa doktorska Pani mgr Marty Kabut

spełnia wymogi stawiane pracy doktorskiej. W związku z powyższym wnioskuję do Rady Naukowej Dyscypliny Nauki o Zarządzaniu i Jakości Uniwersytetu Warszawskiego o dopuszczenie Pani mgr Marty Kabut do publicznej obrony rozprawy.



Wiktor Adamus