

Prof. ALK dr hab. Grażyna Wąsowicz
Akademia im. Leona Koźmińskiego
Ul. Jagiellońska 59
03-301 Warszawa

Warszawa, 30 września 2022 r.

RECENZJA ROZPRAWY DOKTORSKIEJ

mgr Marty Kabut

pt. „False respondents in web human resource surveys”

napisanej pod kierunkiem

prof. dr. hab. Grażyny Wiczorkowskiej-Wierzińskiej i dr Anny Kuzmińskiej

Rozprawa doktorska mgr Marty Kabut, zatytułowana „False respondents in web human resource surveys” przedstawia bardzo interesujący, ważny z teoretycznego i praktycznego punktu widzenia, a także rzadko eksplorowany, temat udziału tzw. fałszywych respondentów w badaniach ankietowych. Rozprawa ma charakter metodologiczny, jest zakotwiczona teoretycznie i empirycznie w metodologii badań sondażowych (prowadzonych w zakresie zarządzania zasobami ludzkimi). Została napisana w języku angielskim; zgodnie z wymaganiami ustawowymi zawiera także streszczenie w języku polskim.

Dysertacja składa się z dwóch części, określonych mianem teoretycznej i empirycznej. Pierwszą część Autorka rozpoczyna od przedstawienia struktury rozprawy, wskazując w niej m.in., że głównym celem jest skonstruowanie modelu fałszywego respondenta (s. 9: „*At the end of this part the author will construct a FALSE respondent model, the main aim of the research*”). Następnie, wyczerpująco i odwołując się do właściwych danych, Autorka prezentuje argumenty uzasadniające dobór tematu. Wskazuje także cele szczegółowe swoich badań, obejmujące określenie zakresu zjawiska występowania fałszywych respondentów, konsekwencje nieusuwania fałszywych respondentów ze zbiorów danych przed podjęciem analizy (zapewne chodziło o dane tych respondentów, a nie o nich samych) oraz propozycję procedury wykrywania (odpowiedzi) fałszywych respondentów w zbiorach danych powstających w badaniach dotyczących zasobów ludzkich, prowadzonych w Internecie (s. 12).

Trzon części pierwszej stanowi przegląd literatury, oparty na niemal 150 pozycjach (polsko- i angielsko-języcznych). Adekwatnie do przyjętego – metodologicznego - charakteru rozprawy, omawiane pozycje bibliograficzne odnoszą się przede wszystkim do różnych aspektów prowadzenia badań ankietowych. W mniejszym zakresie przywoływane są pozycje dotyczące zarządzania zasobami ludzkimi. Autorka omawia zalety i wady badań realizowanych w sieci, przedstawia dane dotyczące zakresu zjawiska fałszywych odpowiedzi i

respondentów, opisuje także stosowane w badaniach na świecie metody wykrywania tzw. fałszywych respondentów. W mojej ocenie nadanie tej części pracy tytułu „Theoretical part” nie jest uzasadnione, bowiem niewiele jest w niej odniesień do teorii i rozważań teoretycznych, natomiast licznie są przywoływane interesujące fakty i wyniki badań. Autorka przedstawia dwa podrozdziały, których tytuły zapowiadają modele teoretyczne (1.2.3. i 1.2.4.), co daje podstawy do przypuszczenia, że modele te (lub ich fragmenty) będą testowane w badaniach własnych. Zgodnie z dobrym standardem badawczym, właśnie część teoretyczna – poza określeniem luki w wiedzy i uzasadnieniem podjętego problemu – powinna zawierać podstawy teoretyczne własnej koncepcji teoretycznej, powiązanej z pytaniami (i ewentualnie hipotezami) badawczymi, której testowanie nastąpi w adekwatnie opracowanym programie badawczym. W podrozdziale 1.2.3, zatytułowanym „Psychological model of answering questions” Autorka przedstawia m.in. różne strategie udzielania odpowiedzi, przyjmowane przez badanych, omawia znaczenie długości pytań. Wskazuje także na kilka – jak się zdaje dość przypadkowo wybranych - elementów skomplikowanego pod różnymi względami procesu badania, które mogą determinować gotowość respondenta do udzielania (nie)prawdziwych odpowiedzi. W tym fragmencie mgr Marta Kabut bardzo krótko (w kilku zdaniach) omawia dwie charakterystyki respondenta („respondent’s ability”, „motivation”) i dwie związane z kwestionariuszem („question difficulty”, „rating scales”), co w żadnej mierze nie może być uznane za „psychologiczny model udzielania odpowiedzi na pytania sondażowe”. Nie ma tu odwołań do teorii testów psychologicznych (stosowanych przecież w zarządzaniu zasobami ludzkimi), do psychometrii. Pozyskiwanie informacji od badanych to bardzo subtelny proces, którego wyniki zależą od wielu innych czynników, poza wspomnianymi przez Autorkę. Ważne jest to, na ile badany czuje się bezpiecznie, na ile chce wypaść dobrze w efekcie udziału w badaniu, jak bardzo silna jest jego potrzeba aprobaty społecznej, która wpłynie na udzielanie odpowiedzi akceptowalnych społecznie, czy zastosowano pytania filtrujące, zwalniające z odpowiedzi na pytania zakładające wiedzę, czy pytania są jasne (do testowania służy współczynnik mglistości Roberta Gunninga) i jaki jest poziom ich intymności.

Zatem nie jest dla mnie jasne kryterium doboru tych kilku aspektów w rozdziale mającym dotyczyć psychologicznego modelu udzielania odpowiedzi, ale skoro zostały wymienione, to nawet szanując ten wybór mogę wyrazić jedynie żal, że żaden z tych wątków nie był przedmiotem prowadzonych badań (czyli w programie badawczym nie podjęto tematu wpływu np. motywacji i/ lub długości pytań na wielkość zjawiska występowania fałszywych respondentów. Brak ten jest jeszcze bardziej widoczny po lekturze podrozdziału 1.2.4, o tytule zapowiadającym model fałszywego respondenta. Nie przedstawiono w nim modelu, lecz opisano strategie udzielania odpowiedzi, ponownie nawiązując do motywacji respondenta, a

także przywołując wyniki badań, wskazujące na znaczenie cech osobowości. Rozumiem, że celem opisywanego w dysertacji projektu badawczego nie była analiza wpływu cech psychologicznych na przejawianie zachowań mieszczących się w nazwie „fałszywego respondenta”, jednakże posługiwanie się nazwą „model” w części określonej mianem „teoretycznej” sugeruje, że będzie on testowany. Oczekiwanie to wzmacnia przywołana wyżej (jak sądzę niefortunna pod względem redakcyjnym, bo niezrealizowana) zapowiedź głównego celu dysertacji, jakim miałyby być opracowanie modelu fałszywego respondenta.

W pierwszej części pracy znajduje się także bardzo ważny, obszerny i przydatny dla każdej osoby prowadzącej badania sondażowe, przegląd metod stosowanych do wykrywania przypadkowych odpowiedzi (fałszywych respondentów): obiektywnych i subiektywnych, statystycznych i nieopartych na narzędziach statystycznych, bezpośrednich i pośrednich. Zbiór tych działań jest całkiem pokaźny. Jak słusznie zauważa Autorka, są one jednak stosowane zwykle pojedynczo i brakuje procedury obejmującej różne metody, która w większym stopniu zapewni możliwość identyfikacji jak największej liczby osób, udzielających odpowiedzi przypadkowych. Jednocześnie, co świadczy o dużej wrażliwości metodologicznej, mgr Marta Kabut słusznie wskazuje, że omawiane cztery kategorie znaków mogących ostrzegać badacza o występowaniu fałszywych respondentów nie muszą być stosowane zawsze w komplecie i w każdej sytuacji zbierania danych sondażowych. Podkreśla także, że ich dobór powinien być efektem decyzji świadomie podejmowanych przez badacza tworzącego procedurę. Uważam, że wyniki prezentowane w drugiej części pracy mogą być w tym zakresie pomocne.

Przeгляд metod nie służy jedynie prezentacji informacji. Zgromadzona przez Autorkę bogata wiedza została bowiem wykorzystana do sformułowania propozycji własnej procedury, obejmującej różne działania, mające zapewnić większą wiarygodność danych poddawanych analizie statystycznej, z której wyprowadzane są wnioski teoretyczne. Procedura ta uwzględnia, opisane precyzyjnie w podrozdziale 1.3 (a także 2.3 i 2.15), szczegółowe (wieloetapowe) działania zapewniające ocenę czterech sygnałów ostrzegawczych (warning signs) obejmujących analizę czasu udzielania odpowiedzi, spójności odpowiedzi na pytania testujące uważność respondenta, występowanie odpowiedzi „nie wiem”, „nie mam zdania” i odnoszących się do zaangażowania respondenta w badanie.

Część empiryczna rozprawy zawiera raport z prac podjętych by zrealizować dwa cele badawcze, zdefiniowane w podrozdziale 2.1. Pierwszym z nich jest określenie skali zjawiska fałszywych respondentów (nieuważnych odpowiedzi). Cel drugi obejmował opisanie konsekwencji ignorowania (problemu) występowania fałszywych respondentów. W części empirycznej Autorka opisuje także zastosowanie procedury FLEXMIX (Finite Mixtures of

Generalised Regression Models) – jeśli dobrze rozumiem – do oceny proponowanej przez siebie procedury wykrywania fałszywych respondentów.

Opis działań badawczych poprzedza prezentacja informacji o imponującej liczbie 12 zbiorów danych (pochodzących od ponad czterech tysięcy osób), które wykorzystano do realizacji przyjętych celów części empirycznej. W dziewięciu z tych badań dane zostały zebrane za pośrednictwem Internetu, a w trzech - głównie metodami innymi niż internetowe (ankieta bezpośrednia z odpowiedziami zaznaczanymi przez ankietera, ankieta pocztowa, ankieta samodzielnie wypełniana przez respondenta). W tym drugim przypadku są to bazy danych z międzykulturowych badań prowadzonych w wielu krajach (w niektórych także metodą on-line): European Working Conditions Survey oraz World Values Survey. Z przedstawionego opisu baz danych pochodzących z badań online wynika, że w sześciu z nich brali udział studenci Wydziału Zarządzania Uniwersytetu Warszawskiego (zbiory oznaczone B1, B2, B3, B4, B5, B6). Z całego zbioru 12 baz tylko w trzech przypadkach widnieje informacja o danych pochodzących od osób zatrudnionych.

Prezentacja baz pozostawia poczucie niedostatku informacji. Nie wynika z niej, czego dotyczyły pytania zawarte w zastosowanych kwestionariuszach, w jakim zakresie wiązały się z tematem zarządzania zasobami ludzkimi, jak długie były kwestionariusze i zamieszczone w nich pytania, jakie skale odpowiedzi zastosowano. Jedynie w przypadku baz A1 i C udostępniono w załączniku 3 treść pytań dotyczących jednego ze znaków ostrzegawczych (zaangażowanie) oraz w przypadku bazy oznaczonej D - kwestionariusz, zawierający pytania, na które odpowiedzi stanowiły przedmiot analiz (w załączniku 7). Jak sama Autorka wskazuje w pierwszej części dysertacji, charakterystyki pytań i skal odpowiedzi są powiązane ze zjawiskiem „fałszywych respondentów”. Dlatego można by było oczekiwać, że charakterystyki te zostaną uwzględnione we własnym programie badawczym (skoro zwrócono na to uwagę w części teoretycznej). Ponieważ Autorka nie informuje w rozprawie (ja nie znalazłam tej informacji), jaki zastosowała klucz doboru baz danych do swoich analiz, to trudno ocenić czy obejmował on charakterystyki pytań w poszczególnych badaniach.

Prezentację baz kończy zestawienie informacji o dostępności danych dotyczących możliwości wykorzystania znaków ostrzegawczych w poszczególnych bazach. To z jednej strony cenna informacja, a z drugiej – prowadzi do pytania, jaki wpływ miała Autorka na to, jakie dane niezbędne do identyfikacji fałszywych respondentów zostały zgromadzone w badaniach (oczywiście poza badaniami prowadzonymi off-line); czy informacje te były elementem świadomie zaplanowanej procedury, czy też były zastane. W kontekście wymagań stawianych rozprawom doktorskim jest to ważne pytanie, bowiem zadanie recenzenta polega m.in. na dokonaniu oceny umiejętności samodzielnego prowadzenia pracy naukowej. Nie

znalazłam w rozprawie podanej wprost informacji, które z tych badań Autorka zaplanowała, by zrealizować wytyczone przez siebie cele badawcze, a które z nich prowadzono w zespole badawczym, który udostępnił dane mgr Marcie Kabut (dopiero na s. 103 pojawia się informacja, że badania zaplanowano w zespole, jednak bez doprecyzowania roli Doktorantki). Można jedynie przypuszczać na podstawie przesłanek obecnych w różnych częściach pracy, że – przynajmniej częściowo – badania online były planowane pod kątem zamierzeń Autorki. Na przykład w opisach baz można przeczytać, że uczestnicy badań byli informowani o stosowaniu procedury sprawdzania danych pod kątem fałszywych odpowiedzi i nagradzaniu za udział w badaniu uzależnionym od jakości udzielanych odpowiedzi (s. 64 i 65). Autorka podaje także informacje o stosowaniu pytań służących testowaniu uwagi i zmianie ich treści w kolejnych badaniach (s. 69) oraz par pytań służących analizie spójności logicznej odpowiedzi (s. 75), co także wskazuje na własny program badawczy. Ponadto, bazy pochodzą z badań prowadzonych od roku 2017 do roku 2021, co również może sugerować zamierzony i zrealizowany przez mgr Martę Kabut cykl badań. W kontekście wagi tych informacji dla oceny rozprawy istotne jest rozstrzygnięcie, czy Doktorantka planowała i realizowała zaplanowany przez siebie cykl badań on-line, mający służyć testowaniu sygnałów ostrzegających o występowaniu fałszywych respondentów, czy może, i w jakim zakresie skorzystała (też) z baz zastanych, tworzonych na podstawie projektów badań online opracowanych przez innych naukowców. O doprecyzowanie tej informacji poproszę w czasie obrony.

Główne wyniki analiz dotyczące pierwszego celu (określenie skali zjawiska fałszywych respondentów) przedstawione są w rozdziałach od 2.5 do 2.8, poświęconych poszczególnym znakom ostrzegawczym. Z dużą skrupulatnością i precyzją Autorka wyjaśnia zastosowane procedury i przedstawia uzyskane wyniki, pokazujące różne odsetki dla wskaźnika fałszywych respondentów w poszczególnych bazach. Lektura tych części prowadzi do pozytywnej oceny wysokiego poziomu wrażliwości analitycznej mgr Marty Kabut. W niektórych tabelach prezentowane są liczby pytań, na podstawie których oznaczano badanych jako należących do grupy fałszywych (nieuważnych) respondentów. Autorka tego nie komentuje, nie stosuje też żadnych testów statystycznych do oceny, czy liczba pytań różnicuje odsetek fałszywych respondentów. Jest to decyzja spójna z podanym celem analiz, więc nie czynię z tego zarzutu. Lecz poznanie odpowiedzi na to pytanie byłoby bardzo interesujące i ważne z praktycznego punktu widzenia.

Z dużym zainteresowaniem zwróciłam uwagę na wyniki uzyskane w badaniach z udziałem panelistów i studentów. Jak się okazuje (np. s. 78, 94), w badaniach panelowych odsetek fałszywych respondentów jest znacznie wyższy niż w próbach studenckich. Temat ten wraca w podsumowaniu wyników s. 105, gdzie Autorka przedstawia w tabeli 30 zestawienie

procentowe osób, dane od których powinny być usunięte z każdej z analizowanych baz. Okazuje się, że jest on najniższy dla badań prowadzonych metodą wywiadów bezpośrednich. Wyniki te, choć nie są zaskakujące, są bardzo ważne i zasługują na propagowanie. Dość powszechnie uważa się, że dane gromadzone z udziałem studentek i studentów są gorszej jakości, a wiele wysokopunktowanych pism odmawia publikacji artykułów powołujących się na takie dane. Wydaje się, że mniej krytycznie świat nauki podchodzi do badań panelowych. Jak jednak pokazują wyniki uzyskane przez mgr Martę Kabut, właśnie w ich przypadku trzeba zachować szczególną ostrożność i dokładać starań do sprawdzania baz pod kątem fałszywych respondentów.

W podrozdziale 2.10 opisane są wyniki analiz z zastosowaniem procedury FLEXMIX (model w R). Na s. 96 Autorka podaje informację, że celem tego zadania badawczego było sprawdzenie przydatności tej metody analizy danych do wyszukiwania fałszywych respondentów. Taki opis celu w tym miejscu jest zaskakujący, bowiem nie był wcześniej zapowiadany. Sensowność zastosowania tej procedury dostrzegam raczej w kontekście oceny spójności (porównania) między wynikami procesu oznaczania fałszywych respondentów z jej wykorzystaniem w porównaniu do efektów autorskiej procedury opartej na czterech znakach ostrzegawczych. Uzyskanie spójności w wynikach świadczyłoby bowiem o trafności kryterialnej procedury proponowanej przez mgr Martę Kabut. Swoje przypuszczenie wywodzę z lektury wyników. Na s. 99 Autorka przedstawia w Tabeli 26 zestawienie odsetków badanych oznaczonych jako respondenci uważni i fałszywi w efekcie zastosowania obu procedur wobec dwóch zbiorów danych (A1 i A2). Wynika z niej, że w przypadku pierwszej bazy procedura FLEXMIX pozwoliła oznaczyć mniejszy odsetek respondentów jako fałszywych niż procedura znaków ostrzegawczych: 32% do 54% (nie jest jasne, co oznaczają procenty w tym drugim wypadku, czy to są uśrednione wartości dla czterech wskaźników?). Natomiast analizy wykonane na danych w bazie A2 pokazały wynik przeciwny (34% do 18%). Zatem na tej podstawie nie da się rozstrzygnąć, która z procedur jest lepsza, choć Autorka podaje, że dają spójne wyniki na poziomie 55% i 60% (nie rozumiem, skąd wzięły się te wartości). Niestety, nie jest wyprowadzony wniosek z tych analiz, stąd brak jasności czemu miały służyć.

W tej części pracy także nie jest dla mnie jasne odwołanie do współczynnika alfa Cronbacha i chętnie posłucham wyjaśnienia w trakcie obrony. Alfa Cronbacha jest to miara rzetelności skali zbudowanej z pytań lub stwierdzeń, jej spójności wewnętrznej, stopnia, w jakim pytania mierzą ten sam konstrukt; wartość tego współczynnika jest zależna od liczby i jakości pytań. Zatem jest to miara odnosząca się do kwestionariusza, a nie do jakości funkcjonowania osób uczestniczących w badaniu. Jeśli dobrze odczytuję zamysł Autorki (szkoda, że nie jest to wyjaśnione wprost), przyjęła ona założenie, że fałszywi respondenci

udzielają odpowiedzi przypadkowych, co przekłada się na obniżone wartości miary rzetelności. I takie wyniki są przedstawione w tabeli 25. Istotne jest to, że alfa Cronbacha przyjmuje wartości z przedziału $<0;1>$, czyli dodatnie i nie większe niż jeden. Zatem uzyskanie wartości alfa Cronbacha wynoszącej -2.841 (na podstawie analiz dla bazy A1) raczej wskazuje na jakiś błąd obliczeniowy, niż na coś innego.

Podobne wrażenie odnoszę czytając wyniki przedstawione w tabeli 27 w podrozdziale 2.10. Jest on poświęconym drugiemu celowi badawczemu, jakim jest określenie konsekwencji ignorowania problemu występowania fałszywych respondentów. Rozdział ten ma objętość tylko niecałej strony, a wyniki w nim przytoczone są efektem analiz wykonanych na zbiorach A1 i A2 (podobnie jak w przypadku procedury FLEXMIX). Autorka nie podaje uzasadnienia, dlaczego tylko te dwie bazy z 12 wzięto pod uwagę (pojawia się ono dopiero na końcu pracy, na s. 115), a także jakie było kryterium doboru skal (pytań kwestionariuszowych), wobec których te analizy zostały wykonane (bardzo krótkie, 3- i 4-itemowe), skoro zapewne kwestionariusze były dłuższe. W opisie wyników ponownie pojawia się ujemna wartość alfa Cronbacha, co nie powinno mieć miejsca i wymaga szerszego wyjaśnienia (a być może także ponownego sprawdzenia analiz i uzyskanych wyników).

Wątek konsekwencji nieusuwania z baz danych od fałszywych respondentów wraca w podrozdziale 2.11.2, poświęconym podsumowaniu wyników, w którym Autorka ponownie opisuje zastosowaną procedurę i przytacza te same wyniki, tylko że tym razem w formie graficznej. W obu tych częściach pracy brakuje jednak konkluzji. Poza podaniem informacji, że alfa Cronbacha przyjmuje wartości niższe w przypadku danych pochodzących od fałszywych respondentów niż od respondentów uważnych, Autorka nie formułuje żadnego wniosku na podstawie tego wyniku dotyczącego konsekwencji. Jak rozumiem, wniosek z analiz może być taki, że konsekwencją pozostawienia w bazie danych pochodzących od fałszywych respondentów jest obniżenie rzetelności stosowanych narzędzi pomiarowych. Skoro jednak tytuł podrozdziału mówi o konsekwencjach (liczba mnoga), a nie o jednej konsekwencji, to jakie jeszcze inne konsekwencje (i jakich zjawisk dotyczące) można by było wskazać? Autorka zapewne posiada wiedzę w tym zakresie, o czym świadczy lista przedstawiona na s. 117, w podrozdziale kończącym rozprawę. Jednakże, poza porównaniem wartości alfa Cronbacha nie przedstawia żadnych analiz dotyczących zagadnień, znajdujących się na tej liście. Nie czynię z tego zarzutu, rozumiem, że praca musi mieć pewne ramy, lecz – ze względu na duże znaczenie praktyczne i teoretyczne – poproszę o szerszą wypowiedź na temat konsekwencji w trakcie obrony. W kontekście tytułu tego podrozdziału (konsekwencje uwzględniania danych od fałszywych respondentów) spójne z jego brzmieniem i interesujące

by było także zestawienie wartości alfa Cronbacha dla trzech zbiorów danych: respondentów fałszywych, uważnych i wszystkich uczestniczących w badaniu.

Tytuł podrozdziału 2.11 rozprawy zapowiada podsumowanie wyników i ich dyskusję. Wcześniej Autorka przedstawiła operacjonalizację czterech znaków ostrzegawczych, w tabeli 29 przedstawia podsumowanie wyników dla poszczególnych znaków ostrzegawczych w 12 bazach. Muszę przyznać, że ta część prowokuje pytania i potrzebę doprecyzowania podawanych informacji. Wcześniej, na s. 66, w tabeli 3 Autorka podaje informacje o dostępności danych w poszczególnych bazach (pod kątem znaków ostrzegawczych). Następnie w rozdziale 2.3. przedstawia operacjonalizację znaków ostrzegawczych, a w rozdziałach od 2.5 do 2.8 - szczegółowe wyniki. Na przykład w rozdziale 2.5 omawiane są wyniki analiz prowadzonych na sześciu bazach (dlaczego tylko sześciu, skoro w tabeli 3 przedstawionej na s. 66 wskazana jest dostępność danych dla wszystkich dziewięciu baz z badań on-line?). Dotyczą one m. in. całkowitego czasu udzielania odpowiedzi (tabela 5), całkowitego czasu z pominięciem przerw i udzielania odpowiedzi na pytania otwarte (tabela 6), czasu udzielania odpowiedzi na pytania w blokach (tabela 7), itd. Dla wymienionych wskaźników, na przykład w bazie A1, odsetki fałszywych respondentów wynoszą kolejno 43,2%, 49,1% i 54,5%. Natomiast w tabeli 29, zawierającej podsumowanie analiz wykonanych w celu ustalenia zakresu zjawiska fałszywych respondentów, przy czasie udzielania odpowiedzi dla bazy A1, podana jest wartość 56,4%. Niestety, nie jest wyjaśnione, skąd się ta wartość wzięła, z jakiej procedury, i jak się ma do wcześniej wskazywanych liczb. Co więcej, w tabeli 29 dla pierwszego znaku ostrzegawczego czasy udzielania odpowiedzi są podane również dla trzech baz, których nie uwzględniono w analizach szczegółowych, co także powoduje pytanie dlaczego zawartych w nich danych nie wykorzystano wcześniej. Podobne uwagi dotyczą podsumowania wyników dla drugiego znaku ostrzegawczego. W tym wypadku analizy szczegółowe przeprowadzono na danych zawartych w siedmiu bazach (choć – jak wynika z tabeli 3 są one dostępne w ośmiu), wskazując odsetki fałszywych respondentów określone na podstawie łagodnego i ostrego kryterium. Z tabeli 29 jednak nie wynika, które z nich zostało wzięte pod uwagę jako podstawa podsumowania.

Bardziej jasna jest informacja dotycząca trzeciego znaku ostrzegawczego. Jeśli dobrze rozumiem, podsumowanie odnosi się do odsetka fałszywych respondentów oznakowanych na podstawie wariacji odpowiedzi (nie wiem, nieinformatywnych i brakujących). Nie rozumiem jednak, dlaczego te odsetki są inne w tabelach 16 i 29 (np. dla bazy A1 w pierwszej z nich wskazane jest 8,9%, w tabeli 29 dla tego samego zbioru danych jest to wartość 11%). Pytanie o to, na jakiej podstawie wyliczono odsetki fałszywych respondentów pojawia się także w przypadku podsumowania dla czwartego znaku ostrzegawczego. W szczegółowej prezentacji

wyników analiz, w podrozdziale 2.8. Autorka wskazuje odsetki dla zaangażowania we współpracę w zakresie badania (tabele 17 i 18), spójności logicznej (tabela 20) i odpowiedzi „dziwnych”, nieinformatywnych na pytania otwarte (tabela 21). Co zatem zawiera podsumowanie dla tego wskaźnika? Trudność w analizie tabeli 29 wynika także z tego, że nie ma do niej legendy (która pojawia się dalej, pod tabelą 30, która pewnie zawiera te same oznaczenia). Zakładam, że mgr Marta Kabut skrupulatnie sprawdzała prezentowane wyniki, natomiast pozostawienie czytelnika bez precyzyjnych informacji o tym, co zawiera podsumowanie, prowadzi do konfuzji. Poproszę zatem o doprecyzowanie w trakcie obrony.

W rozdziale obejmującym podsumowanie wyników, w tabeli 32 znajdują się także wyniki dotyczące odsetka osób, których dane należałoby usunąć stosując jeden ze wskaźników i różne ich kombinacje (wyliczone dla dwóch baz A1 i A2). Jak słusznie zauważa mgr Marta Kabut, powinno stosować się więcej niż jedno kryterium. Wyniki przedstawione w tabeli 32 oceniam jako bardzo cenne w przypadku prowadzenia badań online. Ich publikacja może pomóc w podejmowaniu decyzji, czy poza czasem udzielania odpowiedzi wprowadzać dodatkowe opcje weryfikacji jakości danych i które.

Lektura rozprawy mgr Marty Kabut wielokrotnie prowokowała myśl, że bardzo potrzebne byłyby badania, w wyniku których można by było wnioskować o zależnościach przyczynowo-skutkowych, np. między zmiennymi wskazanymi w części pierwszej jako znaczące dla udzielania (nie)prawdziwych odpowiedzi (np. motywacja, długość kwestionariusza) a skalą zjawiska fałszywych respondentów. Dlatego z zadowoleniem przeczytałam, że taki program badawczy został już rozpoczęty, a jeden z eksperymentów opisano w załączniku 11 do rozprawy.

Zmierzając ku konkluzji wskażę słabsze i mocne strony przedstawionej do oceny rozprawy. Wśród tych pierwszych przede wszystkim zwraca uwagę brak interpretacji (wyjaśniania) uzyskanych wyników przez odniesienie do wiedzy teoretycznej i treści przedstawionych w pierwszej części rozprawy, a także niewystarczające - w mojej ocenie - udokumentowanie, że badania przeprowadzone online dotyczyły obszaru zarządzania zasobami ludzkimi (wystarczyłoby załączyć zastosowane kwestionariusze). Zabrakło mi także refleksji, nad tym, czym (jeśli w ogóle) różnią się badania sondażowe w zarządzaniu zasobami ludzkimi od badań prowadzonych w innych dziedzinach i na ile poczynione obserwacje można generalizować na inne typy badań (np. badania jakościowe on-line, czy badania dotyczące innej dziedziny). Ponadto, praca powinna być bardziej dopracowana redakcyjnie. Duże jej fragmenty są przygotowane bardzo dobrze, z dużą dbałością o jasność wyводу, zawierają cenne definicje i precyzyjne wyjaśnienia. Lecz w innych częściach np. powtarzają się te same

treści, brakuje wyjaśnień (np. w podsumowaniu wyników), a pewne informacje podawane są na końcu pracy, choć bardziej logiczne i przydatne byłoby zaprezentowanie ich wcześniej.

Wśród zalet rozprawy należy wymienić interesujący przegląd literatury na temat tego, jakie cechy respondentów zwiększają prawdopodobieństwo udzielania nieważnych, przypadkowych odpowiedzi w badaniach sondażowych. Zwraca w nim uwagę bardzo cenne w pracy naukowej skupienie na najdrobniejszych szczegółach metodologicznych i duża wrażliwość Autorki na subtelności badań sondażowych. Przegląd ten zawiera także wiele praktycznych informacji, przydatnych dla osób uczących się w jaki sposób planować badania sondażowe, by uzyskiwać dane wysokiej jakości, ale także pokazuje (i ostrzega), jak groźna dla wnioskowania może być niewystarczająca dbałość o jakość danych. Należy jednak przede wszystkim docenić wybór tematu, opracowanie i testowanie procedury identyfikowania odpowiedzi pochodzących od fałszywych respondentów oraz udokumentowanie występowania i skali zjawiska fałszywych respondentów.

Dysertacja mgr Marty Kabut dotyczy bardzo ważnego z teoretycznego i praktycznego punktu widzenia tematu jakości danych, na podstawie których formułowane są wnioski teoretyczne w badaniach naukowych. Jest ona także ważna z punktu widzenia rozwoju metodologii badań prowadzonych w naukach społecznych. Opracowanie autorskiej procedury wykrywania danych pochodzących od tzw. fałszywych respondentów uważam za oryginalne rozwiązanie problemu naukowego. Mgr Marta Kabut wykazuje także ogólną wiedzę teoretyczną w zakresie metodologii badań sondażowych. Sądzę, że także pozytywnie można ocenić umiejętność Doktorantki samodzielnego prowadzenia pracy naukowej. Uważam zatem, że przedłożona do oceny dysertacja przygotowana przez mgr Martę Kabut spełnia warunki określone w Artykule 13 Ustawy z dnia 14 marca 2003 r. o stopniach naukowych i tytule naukowym, stawiane rozprawom doktorskim i rekomenduję dopuszczenie mgr Marty Kabut do dalszych etapów przewodu doktorskiego.

prof. ALK dr hab. Grażyna Wąsowicz