

Dr hab. Radosław Rogoza, prof. AEH

Warszawa, 07.02.2024r.

Akademia Ekonomiczno-Humanistyczna w Warszawie

Recenzja pracy doktorskiej Magdy Tatary pt. „*Testowanie hipotezy leksykalnej na podstawie polskiego leksykonu osobowości.*”

Napisanej pod kierunkiem dr hab. Olega Gorbaniuka, prof. KUL

Recenzowana rozprawa doktorska odnosi się, jak na standardy psychologii, do bardzo wiekowego problemu naukowego jakim jest określona przez Lewisa Goldberga hipoteza leksykalna. Pomimo tej długiej brody, do tej pory (i to stosunkowo niedawno) poddano ją weryfikacji jedynie w kontekście języka angielskiego. Mając na uwadze, że hipoteza leksykalna wprost zakłada, że najważniejsze cechy powinny być zakodowane w wielu językach, trudno mówić o jej pełnej weryfikacji na bazie jednego tylko leksykonu. Ponadto, istniejące badania leksykalne odnoszące się do struktury osobowości jasno wskazują, że nie zawsze odtwarza się ta sama struktura, przykładowo, prace promotora niniejszej pracy sugerują, że w języku polskim, sześcioczynnikowe rozwiązanie wydaje się bardziej zasadne. Tym samym, mając na uwadze problem potwierdzenia hipotezy leksykalnej w jednym tylko języku w jednej ręce, w drugiej zaś mając wiedzę, że w różnych językach odtwarzana struktura może wyglądać inaczej, w mojej opinii, poruszany przez Doktorantkę problem badawczy w pełni wpisuje się w trwający dyskurs naukowy, a otrzymane wyniki mogą w sposób znaczący wpłynąć na rozwój wiedzy teoretycznej.

Oceniana przeze mnie praca składa się wyłącznie z przedłożonej rozprawy doktorskiej. Dość zaskakującym dla mnie faktem było to, że Doktorantka (zgodnie z ORCID), opublikowała tylko jedną pracę w czasopiśmie, niekoniecznie związanym z dyscypliną samej rozprawy. Nie udało mi się również odnaleźć informacji o dorobku naukowym Doktorantki w innych repozytoriach (PBN, Scholar, WoS), tak więc moja recenzja odnosi się wyłącznie do przedłożonej rozprawy. Stając niejako w obronie Doktorantki, badania leksykalne są badaniami, które są wyjątkowo angażujące, które trwają latami i wymagają sumienności, oddania sprawie i zaangażowania. Ilość pracy przy tego typu projekcie jest w istocie niewspółmierna do finalnych efektów parametryzacyjnych w porównaniu do prowadzenia badań ilościowych czy eksperymentalnych, jednak wartość tego typu badań jest oczywiście niezaprzeczalna.

W takim kontekście, warty jest również podkreślenia fakt, że Doktorantka kontynuuje bardzo ambitny projekt, który rozpoczęła wraz ze stworzeniem listy deskryptorów na potrzeby pracy magisterskiej. Lista ta wyróżnia się na tle innych tym, że jest szersza, gdyż nie ogranicza się jedynie do przymiotników, jak typowo postępuje się w literaturze. Ponadto, Autorka w swojej pracy eksperymentuje z różnymi instrukcjami, które nie odnoszą się do oceny jak dobrze prezentowane deskryptory odnoszą się do samego siebie, ale badana jest percepcja ważności tychże deskryptorów dla utrzymania relacji społecznej, badana jest ich pożądalność społeczna, obserwowalność, czy ważność znajomości. Takie wieloaspektowe i dość nowatorskie podejście jest miłym zaskoczeniem i wpisuje się w trwający dyskurs naukowy, który z jednej strony potwierdza stabilność różnorodnych cech, ale również podkreśla ich zmienność w zależności od interakcji sytuacyjnych czy kontekstowych. Poniżej, przedstawiam swoje uwagi w odniesieniu do pracy, mając nadzieję, że z jednej strony pomogą Autorce udoskonalić jej rozprawę, a z drugiej, wskażą nowe perspektywy na dalsze badania.

Uwagi Krytyczne

Z jednej strony, autorka skutecznie lawiruje pomiędzy wykorzystaniem mniej standardowych procedur statystycznych takich jak przykładowo analiza korelacji par antonimów jako wskaźnik trafności pomiaru, z drugiej strony – błędzi w przypadku bardziej standardowych procedur wybierając analizę głównych składowych. I to właśnie w odniesieniu do zastosowania tej konkretnej analizy – mam spore zastrzeżenia. Ilość wyborów, które trzeba wykonać w analizie czynnikowej potrafi przysporzyć bólu głowy, jednak mają one zasadnicze znaczenie dla otrzymanych wyników, stąd należy mieć świadomość ich konsekwencji.

Pierwszą decyzją jest wybór metody estymacji. Autorka zdecydowała się na metodą głównych składowych. Jest to metoda, której podstawowym celem jest redukcja obserwowalnych wskaźników do komponentów. Z technicznego punktu widzenia – nie jest to zatem analiza czynnikowa, która upoważnia do wnioskowania o zmiennych latentnych, które nie podlegają bezpośredniej obserwacji. Nieco uspokajając Autorkę, wyniki tych dwóch analiz będą w dużej mierze zbieżne, jednak dbając o poprawność wyciąganych wniosków, proszę o przeliczenie wyników stosując odpowiednią metodą estymacji. W przypadku SPSS, będzie to metoda osi głównych, w przypadku innych programów (Mplus, R), można pokusić się o estymatory największej wiarygodności, ale nie uważam tego za konieczność. Drugim krokiem jest wybór liczby czynników. Badania symulacyjne wykazują, że wykres osypiska jest dobrą metodą – znacznie lepsza niż kryterium wartości własnych, więc cieszę się, że autorka wybrała ten sposób. Nie mniej jednak, w przypadkach niejednoznacznych, wykres osypiska narażony jest na subiektywność interpretacji, zatem – za absolutnie konieczne widzę zaraportowanie takiego wykresu w rozprawie – bez tej informacji, nie jestem w stanie ocenić, czy wybór jest słuszny, czy wątpliwy. Podczas gdy nie uważam tego za konieczne, warto byłoby wesprzeć decyzję o wyborze

liczby czynników dodatkowymi analizami takimi jak np. analiza równoległa (ang. parallel analysis). Trzecim krokiem jest wybór rotacji. Autorka wybrała rotację ortogonalną zakładając, że komponenty nie będą ze sobą skorelowane. O tyle o ile we wczesnych pracach faktycznie panowało założenie ortogonalności cech podstawowych, sam Goldberg (np. w modelu AB5C) postulował ich współzmiennność. Wybór takiej rotacji w mojej opinii – jest co najmniej zaskakujący. Sugerowałbym zatem zamianę na rotację skośną (np. Oblimin).

Po czwarte, aczkolwiek nie jest to norma w literaturze, stąd też (w przeciwieństwie do poprzednich), proszę traktować ten komentarz jako nieobligatoryjny, eksploracyjna analiza czynnikowa jest modelem SEM, w związku z czym, możliwe jest zaraportowanie poziomu dopasowania tego modelu do danych. Uważam, że wzbogacenie raportowanej analizy o takie współczynniki podniosłoby wartość całego modelu. W SPSS niestety nie ma możliwości obliczenia współczynników dopasowania, stąd istniałaby konieczność wykorzystania innych programów (np. Mplus lub pakiet lavaan w R). Przejście do programu dedykowanego do SEM dostarczyłoby również innych ciekawych informacji, przykładowo – można byłoby posłużyć się kryteriami informacyjnymi (np. Bayesowskim kryterium informacyjnym; BIC), aby ocenić, czy model 4-czynnikowy wyjaśnia więcej niż model 3- czy 5-czynnikowy? Czy jest on lepiej dopasowany niż inne? Odpowiedź na te pytania znacząco wzmocniłaby przekaz.

Kolejną sprawą, która mnie zaskoczyła jest kolejność testowanych hipotez. Skoro Autorka w analizie czynnikowej postuluje np. istnienie czterech czynników ważności, jaki sens empiryczny (por. Tabela 4) ma porównywanie ogólnego wskaźnika? Biorąc pod uwagę, że w przywołanej ważności mamy czynniki bliskości emocjonalnej i wrogości – łączenie ich w ogólny wskaźnik wydaje się co najmniej zastanawiające i wymagające dodatkowej informacji. W moim odczuciu, zasadne byłoby przeorganizowanie sekcji empirycznej w taki sposób, aby najpierw zweryfikować strukturę, a następnie operować na wyróżnionych w toku analiz skalach.

Drobne Kwestie

Autorka, w jednym z najkrótszych akapitów niniejszej rozprawy (s. 86), zwraca uwagę, że należy kontrolować takie zmienne jak płeć i poziom wykształcenia, po czym w części empirycznej – nie za bardzo jestem w stanie zlokalizować analizy poświadczające o takiej kontroli. Z mojej perspektywy, nie jestem zwolennikiem „kontrolowania”, które nijak kontroluje zmienne, a wyłącza w modelu regresji wspólną wariancję między predyktorami. Moja sugestia jest zatem prosta – skoro i tak tego typu analiz nie ma – sugeruję usunięcie tego fragmentu.

Przyznam szczerze, że sposób prezentacji wyników w sekcji 7.2.2 jest dla mnie niejasny. Zaskakujące Czym są testowane struktury czynnikowe, które zawierają aż 100-czynników? Nie znalazłem informacji o procedurze tej analizy ani w sekcji metody, ani w sekcji wynikowej, stąd też nie jestem pewien w jaki sposób powinienem interpretować wyniki raportowane w tabelach. Będę wdzięczny za wyjaśnienia.

Dyskutując porównanie otrzymanej czteroczynnikowej struktury do nieistniejącej Wielkiej Czwórki można byłoby posłużyć się dodatkowym argumentem. Intelpekt/otwartość, jest zdecydowanie najbardziej poznawczą cechą B5. Na dodatek, w badaniach leksykalnych (co jest podkreślane zwłaszcza w kontekście HEXACO) – intelekt jest czynnikiem, który często ujawnia się jako ostatni. Być może, taka poznawcza interpretacja ma mniejsze znaczenie dla postrzegania ważności cech dla utrzymania relacji niż zaangażowanie, afekt i ew. wrogie działania? Jak można przełożyć wiedzę o najmniejszej roli poznania na utrzymanie relacji? Myślę, że jest tutaj potencjał na zarysowanie nowej, interesującej linii badawczej. Pytając obrazowo, czy utrzymywanie relacji jest ślepe na rozum?

Myślę, że warty dodatkowego rozbudowania teoretycznego jest akapit dotyczący optymalnej struktury pożądalności społecznej (s. 142). Wniosek ten jest stosunkowo prosty i

oczywisty (ważne jest nie tylko co dobre, ale i co złe), jednak spora część modeli nie jest w stanie sobie z tym poradzić. Przykładowo, w HEXACO wszystko co złe, łąduje w niskim biegunie uczciwości (co uwidacznia się np. w niemal idealnej kowariancji z latentną ciemną triadą osobowości). Skupienie wszystkiego co złe w jednym miejscu stanowi swoiste ograniczenie – trudno jest rozróżniać specyficzne wiązki tych złych cech, jeśli wpadają do jednego dużego worka. Wyniki tej pracy pokazują dość konkretnie, że ludzie cenią z jednej strony przedsiębiorczość i emocjonalność, a z drugiej dyskredytują buntowniczość, niegodziwość i marudność. Takie wyniki są szczególnie interesujące np. z perspektywy intensywnych badań longitudinalnych, które wskazują, że większa labilność (np. samooceny) jest przewidywana przez antagonistyczne cechy. W mojej ocenie, uzyskanie takich specyficznych wskaźników cech, które ludzie lubią (lub nie) w innych jest interesująca i mogłaby być bardziej wyeksponowana.

Sugestie dla Przyszłych Badań

Czytając rozprawę, wpadłem na kilka pomysłów, które być może okażą się przydatne w wyznaczaniu dalszych kierunków badań Doktorantki. Jasno zaznaczam, że ze względu na stopień zaawansowania tych analiz, nie są to rzeczy, których wymagam, aby Autorka wykonała na potrzeby niniejszej pracy doktorskiej, są to pewne sugestie, które mogą wyznaczyć pewne mniej oczywiste ścieżki działania. Autorka, jak zresztą sama zauważa na s. 103, posiada dane, analizuje dane, które są zagnieżdżone w konkretnych deskryptorach osobowości. Tym samym, mamy tutaj do czynienia z dwupoziomą strukturą danych, którą z technicznego punktu widzenia – można analizować z wykorzystaniem wielopoziomowej analizy czynnikowej. Dzięki takiemu zabiegowi, możliwe byłoby nie tylko wykazanie jaka jest „poszatkowana” struktura czynnikowa w zakresie poszczególnych wariantów, ale można byłoby wykorzystać całe bogactwo zebranych danych, aby zweryfikować, czy struktura ta odtwarza się zarówno na poziomie deskryptorów jak i na poziomie poszczególnych wariantów. Ponadto, wykorzystując analizę wielopoziomowej równoważności pomiaru – możliwe jest wykazanie, czy te dwie struktury są takie same na tych dwóch poziomach. Zaprezentowanie tego typu analizy, stanowiłoby kompletną ocenę zebranych danych, znacząco wspierając wyciągane z nich wnioski. Ponadto, taki syntetyczny sposób analizy wyników pozytywnie wpłynąłby na możliwości publikacji w formie artykułu naukowego, który z oczywistych względów dotyczących standardowej objętości – miałby ogromne trudności, aby pomieścić bogactwo analiz przedstawionych w rozprawie.

Drugi pomysł stanowi niejako rozwinięcie kwestii dotyczącej ważności relacji na różnych poziomach. Autorka wskazuje, że pogrupowały się w dwa skupiska – jedno dotyczące relacji z przyjacielem, partnerem i rodziną, drugie – z szefem i współpracownikiem. Wynik ten aż się prosi o interpretację przez pryzmat modelu Wielkiej Dwójki (por. np.

<https://doi.org/10.21827/ijpp.5.35594>). Z perspektywy analitycznej, warto byłoby pogłębić ten temat. Przykładowo, można byłoby wykorzystać analizę równoważności pomiaru, żeby z jednej strony potwierdzić, czy struktura czynnikowa odtwarza się na wszystkich poziomach (a być może – wygląda inaczej w zależności od skupiska) oraz, jeśli się odtworzy – czy są różnice w zakresie poziomym nasilenia poszczególnych czynników? Przykładowo, czy sumienność okazałaby się ważniejsza w relacjach wspólnotowych (team rodzina), czy jednak sprawczych (team praca)? Zaproponowane przeze mnie analizy wykraczają poza to, co spodziewałbym się zobaczyć w rozprawie doktorskiej, jednak na tym etapie, interesujące mogłoby okazać się zaraportowanie najwykleszych w świecie testów t. Zmiana kolejności analiz – najpierw analiza czynnikowa, później analizy na skalach, również wspierałaby taki sposób prezentacji wyników.

Trzecim pomysłem, jest próba konsumpcji otrzymanych wyników. Konkretniej mówiąc, widzę potencjał do stworzenia kwestionariusza do pomiaru wyróżnionych przez Autorkę czynników. W obecnej formie, jest ich ilość jest ogromna, więc konieczna byłaby redukcja. Wykazanie, że te czynniki są kluczowe dla oceny ważności relacji mogłoby mieć zastosowanie w badaniu tychże relacji, jednak wymagałoby to opracowania wersji w znośnej długości dla przeciętnego badacza. Ze względu na bardzo kompleksowe podejście Autorki, takie narzędzie miałoby dużą wartość badawczą.

Ostatnim pomysłem, który zrodził się w mojej głowie podczas lektury rozprawy doktorskiej było zerwanie z dominującą w literaturze tradycją psychometryczną. Wszystkie znane mi prace odnoszące się do szeroko rozumianych analiz leksykalnych, słusznie zresztą wykorzystują analizę czynnikową. Jest to odpowiednia metoda, która pozwala na udzielenie odpowiedzi na te najbardziej palące pytania. Ale co jeśli można zrobić to lepiej? W ostatnich latach, w literaturze istnieje prawdziwy empiryczny boom na psychometrię sieciową, być może mogłaby okazać się znacznie ciekawszą procedurą, która mogłaby okazać się wprost skrojona

pod testowanie hipotez ad. Ważności poszczególnych deskryptorów. Dla lepszego zobrazowania, posłużę się tutaj ilustracją własnej pracy (10.1037/pas0001149). Testując strukturę narcyzmu, podobnie jak i w przypadku badań autorki, najczęściej stosuje się analizę czynnikową. W omawianej pracy, posłużyłem się eksploracyjną analizą grafów (EGA; Golino, Epskamp, 2017), która jest przykładem zastosowania psychometrii sieciowej. Podczas gdy przedmiotem zainteresowania w zwykłych sieciach są ich wskaźniki (jak np. centralność, która może być interpretowana jako wskaźnik ważności danego elementu w sieci), EGA umożliwia również testowanie jak poszczególne elementy sieci łączą się w tzw. społeczności. Koncepcyjnie – jest to analiza bardzo tożsama z analizą czynnikową, która robi w zasadzie to samo. Już sam powstały rysunek jest bardzo informacyjny jeśli chodzi o relacje zachodzące między zmiennymi. W przypadku badania z ponad tysiącem zmiennych i szczególnym zainteresowaniem w synonimach, rysunek może okazać się nieczytelny. W takim przypadku, interesującą ścieżką byłoby wykorzystanie Unique Variable Analysis (UVA; Christensen i in., 2023), która umożliwia identyfikację zmiennych, które są „empirycznymi synonimami”. Takie ilościowe wykazanie dla której społeczności jest więcej empirycznych synonimów – mogłoby być ciekawą drogą do ilościowej weryfikacji jednej z hipotez, której nie udało się potwierdzić standardowymi metodami. Kolejną zaletą EGA jest fakt, że podobnie jak w analizie czynnikowej, możliwe jest wyciągnięcie informacji o wyniku dla pojedynczej społeczności (czyli odpowiadającemu ładunkom czynnikowym dla danej osoby). Następnie, takie zmienne można wykorzystać w celu estymacji kolejnej sieci (już znacznie zredukowanej), aby ocenić, która ze społeczności jest najbardziej centralnym elementem leksykonu? Czy są jakieś społeczności, które np. stają się mostem łączącym z innymi? Przyznam szczerze, że czuję dreszczyk emocji myśląc o odpowiedziach na te pytania i myślę, że jest to bardzo rozwojowy kierunek, którym można podążać w kolejnych pracach.

Podsumowanie

Po zapoznaniu się z rozprawą doktorską mgr Magdy Tatary, jestem w stanie tylko wyobrazić sobie jak wiele musiała wymagać ona od niej pracy i zaangażowania. W dorobku Doktorantki próżno szukać informacji o ilościowym dorobku, jednakże recenzowana rozprawa stanowi przykład jakościowego wkładu do literatury. Tym samym, stwierdzam, że recenzowana praca doktorska spełnia wymagania akademickie i ustawowe stawiane rozprawom doktorskim, określone w Ustawie Prawo o Szkolnictwie Wyższym i Nauce, wnosząc istotny wkład w rozwój wiedzy teoretycznej. Tym samym, wnoszę do Rady Dyscypliny Naukowej Psychologii Katolickiego Uniwersytetu Lubelskiego o dopuszczenie Doktorantki do dalszych etapów przewodu doktorskiego zgodnie z obowiązującą procedurą formalną.

A handwritten signature in black ink, reading "Radosław Rappier". The signature is written in a cursive style with a large, stylized initial 'R'.