

Poznań, 26 sierpnia 2024

dr hab. Łukasz D. Kaczmarek, prof. UAM
Uniwersytet im. Adama Mickiewicza w Poznaniu
Wydział Psychologii i Kognitywistyki
Kierownik Zakładu Psychologii Społecznej
Kierownik Laboratorium Psychofizjologii

**Recenzja rozprawy doktorskiej mgr Aleksandry Piejki
pt. „The association between loneliness, social cognitive processes, and parasympathetic
activity in young adults” przygotowanej w Instytucie Psychologii Polskiej Akademii
Nauk pod kierunkiem dr hab. Łukasza Okruszka, prof. IP PAN**

Na przedstawione dzieło doktorantki stanowiące przedmiot ubiegania się tytuł naukowy doktora składa się seria trzech opublikowanych artykułów. Artykuły ukazały się w latach 2021-2023 w czasopiśmie *International Journal of Psychophysiology*, *Applied Psychology: Health & Wellbeing*, oraz *Journal of Research in Personality*. Czasopisma są recenzowane i cechują się relatywnie wysokim prestiżem i wpływem. Jedna z prac jest już szeroko cytowana, co świadczy o jej użyteczności. Jednocześnie taka forma przygotowania pracy doktorskiej świadczy o tym, że przygotowane teksty przeszły już przez szereg rzetelnych recenzji *peer-review*, co z pewnością również pomogło wypracować rezultat w obecnej postaci. Również projekty badań – jeżeli zostały opisane we wnioskach grantowych - przeszły przez szereg recenzji, przyczyniając się do wysokiego poziomu.

Znaczącym mankamentem przedstawionej pracy doktorskiej jest z perspektywy recenzenta brak jasnego wydzielenia i opisu oryginalnego wkładu doktorantki w dzieło końcowe, co utrudnia ocenę samodzielności prowadzenia pracy naukowej. Mowa jest tylko o źródłach finansowania. Nie znajdujemy tego ani w autoreferacie ani w publikacjach. Jeżeli dobrze rozumiem (brak precyzyjnej informacji) wszystkie badania zostały zrealizowane w ramach projektów kierowanych przez promotora. Brakuje precyzyjnych informacji o tym, na ile kreatywny był udział doktorantki w zaprojektowanie badań, a na ile uczestniczyła w realizacji już przygotowanych projektów. Komplementując zatem dzieło doktorantki konieczne jest wprowadzenie pewnego założenia co do charakteru jej wkładu. Udział doktorantki w

poszczególnych opracowaniach można oszacować jedynie na podstawie kolejności na liście autorów. Niestety nie przedstawiono informacji o wkładzie a w samych tekstach brakuje również informacji o zadaniach poszczególnych autorów (np. CRediT). Można jednakże wnioskować, że wkład w przypadku przygotowania dwóch tekstów był wiodący (pierwsza autorka) a w trzecim znaczący (druga autorka), co jest zgodne z wymogami.

Prace stanowią tematyczną całość a motywem łączącym jest tematyka samotności rozpatrywanej w ramach Ewolucyjnej Teorii Samotności. Każde badanie stosuje inne podejście metodologiczne: laboratoryjnych eksperyment psychofizjologiczny, intensywne badanie podłużne oraz badanie przekrojowe z wykorzystaniem modelowania równań strukturalnych. Świadczy to o wszechstronności kompetencji doktorantki, która zdobyła różnorodne doświadczenie metodologiczne. Szczególnym walorem przedstawionych prac jest bardzo dobre osadzenie w teorii oraz w kontekście dotychczas prowadzonych badań. Doktorantka przekonująco przedstawia znaczenie podjętej problematyki, innowacyjność oraz wkład do już istniejącego stanu wiedzy.

Przechodząc do szczegółów, pierwszy tekst przedstawia wyniki próby manipulacji poziomem poczucia samotności na szereg zmiennych behawioralnych i fizjologicznych. Badanie posiada liczne mocne strony takie jak (odważna!) próba manipulacji zmienną niezależną oraz relatywnie duża liczba uczestników badania. Zawsze interesujące jest również samo powiązanie procesów psychologicznych z fizjologicznymi zwłaszcza w dynamicznym ujęciu. Za mocną stroną uważam również to, co pojawia się także w kolejnych pracach, a mianowicie pomiar dodatkowych zmiennych (emocjonalnych), których kontrola pozwala lepiej, pełniej zrozumieć specyfikę formułowanych hipotez i obserwowanych efektów. Praca poprzedzona jest dobrym wprowadzeniem teoretycznym. Dyskusja jest wyważona i bardzo szczegółowa. Wyniki są dość niejednoznaczne, co po części może wynikać z nie do końca optymalnych w mojej opinii analiz. Niemniej jednak praca stanowi bez wątpienia ważny wkład w obszar badań nad społecznymi uwarunkowaniami zmienności rytmu serca.

1. Badanie ma charakter psychofizjologiczny. Wprowadzenie jest dość jednoznacznie ukierunkowane w stronę aktywności parasympatycznej oraz aktywności sercowo-naczyniowej. Nieco zaskakujące jest zatem pojawienie się hipotez dotyczących przewodnictwa skórno (a tym samym aktywności sympatycznej) w podsumowującym akapicie wprowadzenia. Uważam, że uwzględnienie tego parametru poszerza rozumienie stanu uczestników badania, jednakże jego użycie wymagałoby również rozszerzenia części wprowadzającej w stronę pełniejszego opisu możliwych

reakcji autonomicznego układu nerwowego (części sympatycznej i parasympatycznej) w obliczu negatywnych interakcji społecznych. O ile reakcje parasympatyczne mogą pełnić funkcję „hamulca” chroniącego przed nadmiernymi reakcjami (*vagal break*), to aktywność układu sympatycznego ma znaczenie bardziej pierwotne w głównej mierze warunkując powstanie reakcji pobudzenia w odpowiedzi na zagrażające bodźce społeczne (*reakcja walki-ucieczki*), którą dopiero można hamować.

2. Z pewną przesadą stosowane jest pojęcie „blunted” przy opisie reaktywności (osłabiona/stępiona reaktywność). Na ogół jest ono zarezerwowane dla reaktywności skrajnie niskiej/odstającej od typowych, lub zupełnej obojętności na bodźce, co może świadczyć o poważnej dysregulacji (np. w stanach depresyjnych). Jej przeciwieństwem jest nadmierna reaktywność (*exaggerated reactivity*). Samo obniżenie reaktywności w jeden z dwóch grup nie musi świadczyć o jej stopieniu, a po prostu o jej obniżeniu/osłabieniu.
3. Ze względów metodologicznych, nie do końca można poprzeć odsianie niedowiarków i wydzielenie grupy osób, które „uwierzyły” w manipulację, a następnie powtarzanie na nich przeprowadzonych analiz. Po pierwsze, zwiększa to błąd I typu – należałoby zatem wprowadzić poprawkę na akceptowalny poziom prawdopodobieństwa, ponieważ każda hipoteza testowana jest dwukrotnie. Po drugie, po nieprzypadkowym wykluczeniu osób z próby (zwłaszcza tak wielu), tracimy wszystkie korzyści z randomizacji a badanie przestaje być eksperymentem i staje się quasi-eksperymentem, który nie pozwala już na wyciąganie wniosków o przyczynowości. Wykluczenie „niedowiarków” sprawia, że w obu grupach mogą znaleźć się osoby różniące się na wielu wymiarach i to te różnice (np. osobowościowe) a nie sama manipulacja mogą powodować różnice w poziomach zmiennych zależnych. Na przykład, po odsianiu grupy mogły zacząć się różnić na poziomie optymizmu-pesymizmu z większą liczbą optymistów pozostających w grupie pozytywnej i większą liczbą pesymistów w grupie negatywnej. Większą korzyść przynosi trzymanie się zasady tzw. *intention-to-treat*: „once randomized, always analyzed”.
4. Tak jak wspominałem, autorzy oprócz interesujących ich bezpośrednio zmiennych (samotność) mierzą również zmienne towarzyszące, w tym negatywne emocje. Pozwala to określić, na ile efekt manipulacji zachodzi za pośrednictwem zmiennych specyficznych dla samej manipulacji (poczucie samotności), a na ile za pośrednictwem innych, w tym bardziej uogólnionych stanów. Jak autorzy sami zauważają, w badaniu nie uzyskano wzrostu samotności a i sama manipulacja nigdy nie była dotąd stosowana

w tym celu. Można więc się zastanawiać, czy tekst powinien być tak konsekwentnie formatowany w stronę efektu samotności (co jest interpretacją) czy też efektu myślenia o samotnej przyszłości, co jest elementem faktycznym obecnym w badaniu, który wywiera – jak autorzy sami pokazują - szerokie spektrum wpływu nasilając m. in. a może przede wszystkim lęk i smutek.

5. W badaniu wystąpiły spore różnice w poziomach bazowych, jeszcze przed manipulacją. To się niestety zdarza. Szczególnie w takiej sytuacji warto się zastanowić, czy nie potraktować wartości bazowych jako współzmiennych. Taki model statystyczny pozwala porównać różnicę w średnich po zastosowanej manipulacji przy jednoczesnym uwzględnieniu wartości sprzed testu, co zapewnia większą moc statystyczną (Clifton, L., & Clifton, D. A. (2019). The correlation between baseline score and post-intervention score, and its implications for statistical analysis. *Trials*, 20(1), 1–6; Van Breukelen, G. J. (2006). ANCOVA versus change from baseline had more power in randomized studies and more bias in nonrandomized studies. *Journal of Clinical Epidemiology*, 59(9), 920–925.)

Drugi tekst przedstawia próbę przetestowania modelu ETL w dwóch badaniach przy użyciu metodologii intensywnych pomiarów podłużnych. Artykuł został bardzo czytelnie napisany. Część teoretyczna jest bardzo wyczerpująca i utrzymuje krytyczny i kreatywny ton. Szczególnie warto podkreślić, że autorzy nie dążą do prostego testu ETL, lecz podchodzą do testowanego modelu krytycznie. Poszerzają model badawczy o zmienne, które mogłyby pomóc w lepszym określeniu czy efekty opisane w ETL są specyficzne, czy też są tylko szczególnym przypadkiem bardziej ogólnych efektów związanych z samotnością. Autorzy pokazują, że było to uzasadnione przypuszczenie i rzeczywiście opisane efekty związane z oceną innych ludzi mają szersze zastosowanie. Za plus można uznać również włączenie zmiennych emocjonalnych, o których wiemy z innych obszarów badań, że stanowią często wyjaśnienie/mediator wielu efektów poznawczych. Kolejnym plusem jest również uwzględnienie samotności jako stanu i cechy. Ma to nie tylko wartość teoretyczną, ale również pozwala potwierdzić, że efekty cech (siłą rzeczy raczej nie manipulowanych, jak pokazało badanie 1) rzeczywiście wywierane są przez generowanie typowych dla tych cech stanów, co wyklucza wiele alternatywnych hipotez oraz zmniejsza ryzyko opisu zależności pozornych. Analizy są bardzo złożone metodologicznie. Zostały szczegółowo opisane i wyjaśnione.

Wydają się poprawne, lecz ocena pełnej poprawności wszystkich ich aspektów wykracza poza moją wiedzę w tym obszarze ze względu na zastosowanie wyjątkowo złożonych metod analizy.

W artykule 2 można dostrzec pewne kwestie dyskusyjne:

1. Nieco zastanawiające wydaje się, dlaczego w części teoretycznej (również w tekście 1 i 3) autorzy nie odwołują się do klasycznej, szeroko znanej pracy Baumeistera i Leary'ego (Baumeister, R. F., & Leary, M. R. (1995). The need to belong: Desire for interpersonal attachments as a fundamental human motivation. *Psychological Bulletin*, 117, 479-529). Co prawda dotyczy ona bardziej fundamentalnego i głębszego procesu przynależności, jednakże dość dobrze przedstawia bardziej ogólne niż w ETL procesy, w tym właśnie emocjonalne, które mogą odpowiadać za uogólnioną tendencję negatywną w osób samotnych wewnętrznie „karanych” za samotność, a tym samym motywowanych do zmiany aktualnego stanu rzeczy. Nie jest to praca niezbędną, ponieważ wprowadzenie jest dobrze uzasadnione teoretycznie i bez niej. Jednakże jej nie uwzględnienie (choćby w dyskusji) rodzi ryzyko „odkrycia” czegoś, co już wiadomo.
2. Nieco w kontekście uwagi przedstawionej powyżej. Szkoda, że autorzy nie uwzględnili efektów mediacyjnych sprawdzając, czy to nie emocje stanowią czynnik pośredniczący pomiędzy samotnością i następstwami poznawczymi. Przeprowadzone analizy zdają się nie wykluczać, że tak może być. Jednocześnie mogłoby to być najprostsze wyjaśnienie efektów uogólnionej tendencji negatywnej.
3. Tak jak wspomniałem wcześniej, wartościowe jest uwzględnienie stanów jako pochodnych cech. Jednakże autorzy wykorzystują cechę jako moderator a nie mediator. Można uznać, że jest w tym pewna wartość, ponieważ pozwala określić, jak dany stan (samotność) rzutuje na inne stany u osób, u których jest on w życiu częstszy (wysokie nasilenie cechy) lub nieczęsty. Choć wyniki analiz pokazują, że zmiany w wielkości efektów są raczej zaniedbywalne. Wydaje się, że co najmniej równie ważnym zastosowaniem perspektywy stanu/cechy byłoby przeprowadzenie analizy mediacji, pokazujące, że „efekty” cech można przypisać stanom dla nich specyficznym.
4. Autorzy zoperacjonalizowali samotność jako cechę mierzoną jednokrotnie przy użyciu kwestionariusza. Czy autorzy rozważali wyznaczenie cechy mniej „samoświadomościowo”, jako średnią lub zmienną latentną dla wszystkich

pomiarów stanów z mierzonego okresu? Mogłoby to lepiej pokazać, czy dany stan był wyjątkowy na tle wahań z analizowanego okresu.

5. Nie do końca jest dla mnie jasne, co autorzy mieli na myśli pisząc, że pierwsze badanie było podstawą dla prerejestracji badania drugiego. Jeżeli w tekście nie ma błędu edytorskiego, to oba badania zaczęły się praktycznie w tym samym momencie (26 czerwca 2020 i 21 lipca 2020) a badanie 1 zakończyło się dopiero 14 października 2020. Badania więc się mocno pokrywają.
6. Nie jestem pewien, jaki efekt miało wycięcie środka rozkładu zmiennej w badaniu 2. Rozumiem uzasadnienie teoretyczne (najwięcej dzieje się w grupach skrajnych), ale czy takie postępowanie spełnia założenia dla testów zakładających określony, ciągły rozkład zmiennych dla przypadków losowo dobieranych z populacji.
7. Można byłoby podać rzetelność pomiarów test-retest przy zmiennych mierzonych pojedynczą pozycją testową.
8. Plan analityczny jest bardzo złożony. Nie tylko obejmuje dwa podobne badania, ale analizowana jest bardzo duża liczba związków między zmiennymi dość zbliżonymi do siebie. To może zaspokajać ciekawość autorów i wielu czytelników, ale rodzi pytanie, na ile tak duża liczba pojedynczych testów nie rodzi ryzyka „korzystania” z podwyższonego ryzyka błędu I typu wynikającego z testowania rodziny hipotez (*family-wise error rate*). Autorzy w końcu oczekiwali, że wszystkie efekty będą istotne, i każdą istotność przyjęliby za dobrą monetę. Jeśli spojrzymy na przykład na tabelę 4, to zauważymy, że większość testowanych efektów nie przekroczyła progu istotności statystycznej. Po przyjęciu „kary” za testy nieskuteczne, np. poprzez wprowadzenie poprawki na okrycia pozorne (*false discovery rate*), można byłoby się spodziewać odsiania sporej części opisywanych efektów. Warto byłoby się przyjrzeć tej kwestii.
9. Pomiar był podłużny i intensywny, jednakże obejmował dość homogeniczny okres następujących po sobie 7 dni. Zapewne większego zróżnicowania stanów można byłoby oczekiwać przy większym odstępnie pomiarów.

Artykuł trzeci to opis badania empirycznego, w którym zagregowano dane z poprzednich badań. Miało ono charakter badań przekrojowych, w których przetestowano czy zmienne poznawcze stanowią predyktor subiektywnej i obiektywnej samotności wśród osób z populacji ogólnej. Autorzy wskazują, że *novum* ich podejścia jest przede wszystkim

przeprowadzenie badań na próbie obejmującej osoby z populacji ogólnej, ponieważ badania tego typu były już prowadzone na grupach klinicznych. Artykuł ma szereg mocnych stron. Tekst został bardzo jasno napisany i rozpoczyna się od bardzo rzeczowego przedstawienia aktualnego stanu badań. Mocną stroną samego badania jest przede wszystkim liczna próba badawcza. Autorzy również wykazali się dobrą strategią analityczną w obliczu napotkania na trudności wynikające z natrafienia na stan rzeczy odmienny od oczekiwanego. Istnieje jednakże kilka ważnych punktów, które budzą zastrzeżenia i wymagają doprecyzowania:

1. Nie przedstawiono rzetelności wykorzystanych narzędzi w przeprowadzonym badaniu. Autorzy przedstawiają rzetelność narzędzi z poprzednich badań. Jednakże brakuje udokumentowania, że również w przeprowadzonych badaniach pomiar był rzetelny. Rzetelność pomiaru nie jest tylko pochodną rzetelności samego narzędzia, ale i efektem rzetelności samych warunków pomiarowych oraz podejścia osób wypełniających. W dalszej części autorzy wspominają, że wskaźniki były trafne i rzetelne, jednakże siła ładunku nie jest wskaźnikiem rzetelności czy trafności pomiaru tylko siłą związku ze zmienną latentną, której odpowiednikiem przy testowaniu rzetelności jest tzw. korelacja pozycji z wynikiem ogólnym. Nie jest również pełnym wskaźnikiem trafności, ponieważ dotyczy tylko trafności teoretycznej (konstruktu) oraz nie uwzględnia trafności różnicowej. Jest to szczególnie ważne, w przypadku tworzenia zmiennej ad hoc „lower-level processing of social cues”.
2. W sekcji Statistical Analysis brakuje odwołania się do literatury przedmiotu opisując właściwości analiz SEM. Niektóre stwierdzenia są dość radykalne, np. że w SEM współliniowość nie stanowi problemu, co jest dość ryzykowną tezą wymagająca udokumentowania.
3. Dlaczego PSI oraz OSI nie zostały utworzone jako zmienne latentne (symbolizowane przez owale)? Rysunek 1 raczej przedstawia PSI i OSI jako wskaźniki pozostałych zmiennych latentnych, tzw. zmienne obserwowalne powstałe w wyniku zsumowania pozycji testowych ze skal (reprezentowane przez prostokąty). Raczej należało oczekiwać stworzenia zmiennych latentnych, być może po połączeniu pozycji testowych w trzy paczki (parcels). Umożliwiłoby to ich udział w modelu strukturalnym a nie w modelu pomiarowym. Być może jest to błąd w rysunku (prostokąty zamiast owali).

4. W badaniu uzyskano wysoką moc statystyczną dzięki uwzględnieniu relatywnie dużej liczby uczestników. Jednakże duża moc statystyczna w sposób szczególnie pociąga za sobą konieczność interpretowania siły efektu a nie tylko prawdopodobieństwa nieuzasadnionego odrzucenia hipotezy zerowej (p). Przy dużej mocy prawdopodobne jest uzyskanie istotności statystycznej w odniesieniu do efektów o małym znaczeniu w sensie istotności psychologicznej. Można odnieść wrażenie, że w przedstawionym tekście aspekt ten nie jest wystarczająco mocno wyeksponowany. Autorzy nie podają wielkości efektów i raczej eksponują (również w dyskusji) statystykę p .
5. Autorzy trafnie zauważają, że kierunek relacji (od zmian poznawczych do samotności oraz od samotności do zmian poznawczych) jest najprawdopodobniej dwustronny. Nie podejmują jednakże próby przeanalizowania obu kierunków i ich porównania, lecz wybierają tylko jeden z nich i poddają analizie. Pozostawienie jednego z kierunków (samotność jako predyktor sprawności poznania społecznego) w sferze domysłów jest frustrujące poznawczo. Korzystając z SEM można było porównać oba modele lub chociażby przedstawić drugi w suplemencie. Byłoby to dodatkowo uzasadnione tym, że autorzy już na początku przyjmują podejście eksploracyjne szukając lepszego modelu pomiarowego dla zebranych danych.
6. Autorzy w dyskusji nie opierają się pokusie formułowania (choć ostrożnie) zaleceń interwencyjnych („The results of the current study suggest that specific social cognitive mechanism may be a possible target for interventions which are aimed at reducing OSI/PSI”). Sami jednak piszą wcześniej (i trzeba się z tym zgodzić), że badanie jest przekrojowe a kierunek zależności trudny do ustalenia. W badaniu nie realizowano przecież żadnej interwencji, a jedynie mierzono spontaniczne przekonania oraz wyniki testów. Nie zawsze związki zmiennych przekładają się na skuteczne interwencje, choćby z tego powodu, że nie wszystkie czynniki są łatwo modyfikowalne. Dowodem na to jest choćby badanie 1.
7. W dyskusji autorzy porównują wielkość efektu bezpośredniego z pośrednim. Efekty pośrednie są zawsze dużo mniejsze, ponieważ są iloczynem efektów bezpośrednich. Gdyby autorzy chcieli w przybliżeniu porównać więc efekt pośredni z efektem pośrednim obejmującym jeden mediator, to powinni podnieść efekt bezpośredni do kwadratu.

8. Nie jest też jasne, co mają na myśli pisząc tu i w kilku innym miejscach, że efekt ten jest „more robust”. To dość mocne słowo jak na charakter przeprowadzonych analizy.

Podsumowując, stwierdzam, że pomimo pewnych ograniczeń przedstawiony cykl publikacji stanowi wartościowy wkład w obszar teorii oraz badań empirycznych w zakresie pełniejszego zrozumienia problematyki samotności z akcentem na procesy fizjologiczne i poznawcze. Doktorantka w przedstawionej serii publikacji udowodniła, że jest bardzo dobrze przygotowana warsztatowo do prowadzenia badań naukowych na każdym z etapów: od wyboru ważnej problematyki, przez prowadzenie analiz teoretycznych, przygotowanie badań (w tym wybór właściwych metod), prowadzenie analiz statystycznych, po opis i interpretację wyników. Jednocześnie wykazała się bardzo dużą różnorodnością metodologiczną stosując zaawansowane metody eksperymentu laboratoryjnego oraz intensywne badania podłużnych. Doktorantka potrafi czerpać z osiągnięć współautorów i dysponuje mocną siecią naukowego wsparcia, co świetnie rokuje dla jej dalszego rozwoju naukowego. Pozytywnie bilansując walory i ograniczenia pracy stwierdzam, że **praca doktorska mgr Aleksandry Piejki spełnia wymagania stawiane pracom doktorskim i wnoszę do Wysokiej Rady o dopuszczenie do dalszych etapów przewodu doktorskiego.**

/ Łukasz D. Kaczmarek /