

Kraków, 7.07.2021

Dr hab. Szymon Wichary, prof. UJ  
Instytut Psychologii, Wydział Filozoficzny  
Uniwersytet Jagielloński w Krakowie

Recenzja rozprawy doktorskiej pt.

“Wysiłek jako koszt i inwestycja w hipotetycznych i rzeczywistych sytuacjach wyboru”

Autor: Mgr Przemysław Marcowski

Promotor: Dr hab. Wojciech Białaszek, prof. Uniwersytetu SWPS

Tematyka wysiłku w kontekście podejmowania decyzji jest bardzo ciekawym i słabo zbadanym obszarem badawczym. Przedstawiona mi do oceny rozprawa doktorska Pana magistra Przemysława Marcowskiego porusza bardzo ciekawy temat roli wysiłku, rozumianego jako koszt lub inwestycja, w dokonywaniu wyborów. Do zbadania interesujących go zależności Doktorant użył w pracy całej palety zaawansowanych metod. Przede wszystkim, głównym narzędziem do zbadania wpływu wysiłku na podejmowanie decyzji była metoda eksperymentalna, w szczególności komputerowe zadania decyzyjne, które pozwalało na precyzyjną manipulację interesującymi Doktoranta zmiennymi, prezentację bodźców i pomiar reakcji behawioralnych. Do tego, w Eksperymentie 2., doszedł pomiar sygnału EEG i analiza potencjałów związanych ze zdarzeniem (*event related potentials*, ERP). Dane behawioralne zostały przeanalizowane w bardzo zaawansowany sposób przy pomocy złożonych statystyk oraz modelowania matematycznego, zwieńczonego selekcją modeli najlepiej opisujących zebrane dane. Na tej podstawie można więc stwierdzić, że mgr Przemysław Marcowski wykazał się ambicją i odwagą oraz niezwykle zaawansowanym warsztatem badawczym, który wykracza poza granice jednej dyscypliny i umiejętnie łączy różne podejścia badawcze. Oprócz ogólnie bardzo mocnych stron, praca zawiera również błędy i słabsze fragmenty, które w poniższej recenzji wymienię, i poproszę o ustosunkowanie się do nich Doktoranta.

## **Formalny opis rozprawy**

Przedstawiona mi do recenzji rozprawa doktorska zawiera się w stu sześćdziesięciu czterech stronach i składa się z dwóch głównych części, Części teoretycznej i Części empirycznej. Część teoretyczna poprzedzona jest krótkim Wprowadzeniem i dzieli się następnie na sześć rozdziałów omawiających poszczególne zagadnienia, Część empiryczna również dzieli się na sześć rozdziałów (Metoda, Eksperyment 1, Eksperyment 2, Dyskusja, Ograniczenia i dalsze kierunki badań, Podsumowanie). Praca zawiera również streszczenie, spis literatury cytowanej, spis tabel, spis rycin, spis równań i załącznik. Układ pracy jest zatem logiczny i przejrzysty. Na pochwałę w tym względzie zasługuje umieszczenie w rozdziałach Eksperyment 1 i Eksperyment 2 krótkich podsumowań opisanych eksperymentów i dotyczących ich analiz. Praca napisana jest dobrym językiem polskim, chociaż Doktorant nie ustrzegł się błędów stylistycznych i typograficznych. Praca spełnia wymogi formalne rozprawy doktorskiej.

## **Ocena merytoryczna**

### **Część teoretyczna**

Część teoretyczna, łącznie z poprzedzającym ją krótkim wprowadzeniem, jest bardzo dobrze napisana, zawiera wyczerpujący przegląd badań i stanowi dobrą podstawę dla hipotez weryfikowanych w części empirycznej. We Wprowadzeniu Autor dobrze przedstawił problem badawczy i dobrze go uzasadnił. Ogólnie, bardzo mi się podoba wybór problemu badawczego, po lekturze całości pracy można odnieść wrażenie, że Autorowi udało się znaleźć ciekawą niszę badawczą i ją „zasiedlić”. Podobają mi się również rozdziały dotyczące decyzji opartych na wartości oraz czynników kształtujących wartość działania i roli wysiłku w podejmowaniu decyzji – w tych rozdziałach Autor dokonał wyczerpującego przeglądu literatury, cytując prace klasyczne, ale uwzględniając również najnowsze pozycje z interesujących go obszarów. Bardzo podobał mi się sposób omówienia podstawowych konstruktów, którymi się Autor posługuje w całej pracy – wysiłku jako kosztu lub jako inwestycji. Szczególnie użycie przykładów (przewijających się przez całą pracę) bardzo ułatwia zrozumienie używanych przez Autora pojęć. Stosunkowo słaby fragment tych rozdziałów to opis adaptacyjnej procedury schodkowej (str. 20-21), który wydał mi się nieco zagmatwany.

Te rozdziały są ogólnie dobre, jednak na szczególną pochwałę zasługuje rozdział dotyczący wybranych modeli ewaluacji konsekwencji. Ogólnie, uważam takie przeglądy literatury nt. precyzyjnych modeli jakiegoś zjawiska za bardzo wartościowe, ponieważ pozwalają one usystematyzować wiedzę na temat badanego zjawiska oraz w skondensowanej formie przedstawić główne hipotezy wyjaśniające dane zjawisko. Tak więc ten rozdział uważam za bardzo wartościowy, szczególnie dlatego, że modele zostały przedstawione w formie równań matematycznych, łatwo porównywalnych ze sobą. Chociaż taki opis wielu

osobom może wystarczyć, myślę, że dobrym pomysłem byłoby uzupełnienie go o graficzne przedstawienie omawianych modeli i postulowanych przez nie funkcji. Na tle tych modeli Autor zaproponował również swój autorski, matematyczny model wpływu wysiłku na dokonywanie wyborów, w rozdziale 6. Koncepcja badań własnych. Takie podejście do problemu badawczego zasługuje na szczególne uznanie, ponieważ Autor pokazał, że nie boi się tworzenia swojego precyzyjnego modelu badanego zjawiska – a jak wiadomo, czym model bardziej precyzyjny, tym łatwiej go sfalsyfikować. Autor pokazał tym samym, że ma coś istotnego do powiedzenia na temat zjawiska, którym się zajmuje, i jest gotowy konkurować z innymi badaczami na tym polu.

W części teoretycznej są też miejsca słabsze. Stosunkowo naj słabszy jest rozdział dotyczący psychofizjologicznych korelatów wyboru. Na początku rozdziału Autor źle podaje nazwę metody, której używa w pracy – pisze „analiza potencjałów wywołanych związanych ze zdarzeniem”, a prawidłowa, obecnie używana nazwa tej metody to „analiza potencjałów związanych ze zdarzeniem” (kiedyś używano nazwy „potencjały wywołane” zamiast „potencjały związane ze zdarzeniem”, Autor połączył te dwie nazwy).

Autor również używa nazwy skrzyżowanie skroniowo-ciemieniowe (str. 43), zapewne jako odpowiednik angielskiej nazwy *temporoparietal junction*. Poprawna polska nazwa tej struktury to styk skroniowo-ciemieniowy. Ten błąd pojawia się we fragmencie, w którym Autor omawia mózgowe podstawy komponentu ERP P3. Z pozycji, które Autor cytuje w tym fragmencie (dość starych, bo z lat 80. i 90. XX w.), można się domyślić, że Autor nie zna podstawowych prac nt. mózgowych lokalizacji źródeł tego komponentu, np. prac Bledowskiego i wsp. (2004) i Lindena (2005).

Bledowski, C., Prvulovic, D., Hoechstetter, K., Scherg, M., Wibral, M., Goebel, R., & Linden, D. E. (2004). Localizing P300 generators in visual target and distractor processing: a combined event-related potential and functional magnetic resonance imaging study. *Journal of Neuroscience*, 24(42), 9353-9360.

Linden, D. E. (2005). The P300: where in the brain is it produced and what does it tell us? *The Neuroscientist*, 11(6), 563-576.

W tym rozdziale (ale również tak jest w innych miejscach pracy) brak jest cytowań w kilku miejscach, gdzie powinny się one znaleźć. Na przykład na str. 38. Autor stawia, według mnie, dość kontrowersyjną tezę, że zastosowanie ERP jest dobrze ugruntowane w badaniach nad podejmowaniem decyzji, nie cytując choćby jednej pracy na poparcie tej tezy.

W części teoretycznej (ale również w pozostałych częściach pracy) są błędy stylistyczne i typograficzne. Na przykład na stronie 41. Autor pisze „...wzrosty lub spadki w śródmózgowiowych sygnalizacjach dopaminowych kodują błędy predykcji...”, a można by napisać np. „...wzrosty lub spadki poziomu dopaminy w śródmózgowiu kodują błędy predykcji...”. A na stronie 42. czytamy „...niezgodność między oczekiwaną o otrzymaną nagrodą...” – takich błędów jest dość dużo.

## Część empiryczna

Jak wspomniałem wyżej, Część empiryczna pracy jest podzielona na sześć rozdziałów. W rozdziale 7. Metoda Autor w bardzo udany sposób przedstawił i uzasadnił metody używane w swoich badaniach i analizach. Bardzo doceniam podejście empiryczne Autora polegające na połączeniu modelowania behawioralnego i analizy aktywności mózgu, takie podejście jest mi bliskie i uważam, że Autor bardzo dobrze przedstawił korzyści płynące z takiego podejścia. W tym rozdziale Autor również elegancko podsumował swoje pytania badawcze i hipotezy oraz metody ich weryfikacji. W związku z tym, że hipotez było wiele, plan eksperymentalny jest dość skomplikowany - jak dla mnie nieco zbyt skomplikowany. Na przykład, nie rozumiem do końca potrzeby rozróżniania na warunek retrospektywny i prospektywny, mam wrażenie, że Autor dobrze tego nie uzasadnił, i mam prośbę, żeby mnie Autor przekonał podczas obrony, dlaczego to rozróżnienie jest zasadne. Co do głównej manipulacji eksperymentalnej w warunku wysiłku realnego, to zastosowanie steppera uważam za bardzo dobry wybór, który pozwala dobrze kontrolowane manipulacje wysiłkiem i dostosowywanie jego parametrów do osób badanych (to raczej uwaga do przyszłych, możliwych badań Autora).

W rozdziale 8. Eksperyment 1 Autor przedstawił pierwsze przeprowadzone przez siebie badanie, opisując na początku uczestników i procedurę badania. W odniesieniu do tego, co napisałem powyżej, wartościowym dodatkiem w tym rozdziale jest Rycina 3., przedstawiająca plan eksperymentalny, co nieco ułatwia jego zrozumienie. Co do procedury testowania decyzji osób badanych, niezrozumiałe są dla mnie reguły rządzące dostosowywaniem tej procedury do zachowania osób badanych (na podstawie Taylor i Creelman, 1967) – te reguły nie są nigdzie w pracy przedyskutowane ani uzasadnione, i czytelnik nie wie, dlaczego akurat takie reguły zostały zastosowane. Domyślam się też, że od roku 1967 pojawiło się przynajmniej kilka badań, gdzie te reguły zmieniano i testowano – według mnie, Autor powinien poświęcić więcej miejsca refleksji na temat tych reguł i ich ewentualnych wariantów obecnych w literaturze. Proszę Autora o ustosunkowanie się do tego punktu.

W podrozdziale 8.3. Modelowanie behawioralne Autor opisuje zastosowane przez siebie zabiegi analityczne na danych prowadzące do oceny dopasowania poszczególnych modeli do swoich danych, które potem przedstawia w Wynikach. Podobnie jak w przypadku części teoretycznej, w części empirycznej fragmenty dotyczące modelowania behawioralnego uważam za najbardziej wartościowe i zasługujące na szacunek. Myślę, że niewielu Doktorantów w Polsce (przynajmniej w naukach społecznych) potrafi wykonać podobne analizy.

Co do samych procedur analitycznych, to w tym miejscu zastanowiło mnie dlaczego Autor nie używa do oceny modeli kryteriów informacyjnych (np. Akaike Information Critetion, AIC; co ma zresztą miejsce później w pracy, w kolejnym rozdziale), stosuje natomiast logarytmiczną funkcję straty i proporcję poprawnie przewidzianych wyborów. Żadne z tych

kryteriów nie uwzględnia (i nie penalizuje) złożoności modeli, co sprawia, że modele bardziej złożone mają przewagę nad prostszymi, ponieważ, jak wiadomo, łatwiej je dopasować do danych. Autor stosuje kryterium informacyjne, AIC, które uwzględnia złożoność modeli, do selekcji modeli w Eksperymentach 1. i 2., proszę więc Autora o omówienie różnic w podejściach do oceny modeli między Eksperymentem 1 i 2 i uzasadnienie tych różnic.

W podrozdziale 8.4. Wyniki Autor najpierw omawia wyniki analizy regresji na danych behawioralnych, która pozwala ocenić wpływ warunków eksperymentalnych na dokonywane przez badanych wybory, a następnie przechodzi do opisu wyników procedury selekcji modeli oraz oszacowania parametrów najlepszego modelu dla każdej osoby badanej, co pozwala mu na sklasyfikowanie osób badanych pod kątem preferencji co do wysiłku związanego z dokonaniem wyboru. Znow, jak pisałem, już wcześniej, te analizy uważam za niezwykle wartościowe, pozwalają one bowiem na precyzyjny wgląd w latentne procesy leżące u podstaw obserwowanych wyborów. W podrozdziale 8.5. Podsumowanie Eksperymentu 1. Autor podsumowuje wszystkie zabiegi analityczne w Eksperymentach 1., co jest bardzo dobrym pomysłem, biorąc pod uwagę złożoność tych analiz, wielość hipotez i złożoność planu badawczego.

W rozdziale 9. Eksperyment 2 Autor opisuje Eksperyment 2., w podobnym schemacie jak Eksperyment 1., zaczynając od opisu badanej próby i procedury. Następnie przechodzi do opisu metod rejestracji i wstępnego przetwarzania sygnału EEG. Ten podrozdział w większości przedstawia standardowe podejście i nie mam tu zastrzeżeń, oprócz jednego, moim zdaniem dość ważnego.

Problematyczny jest dla mnie wybór elektrod do określenia komponentu Reward Positivity (RewP). W literaturze przedmiotu, w tym również w pracach autorów, których cytuję Doktorant (Proudfit (Hajcak), 2012, Yi et al., 2020)

Proudfit, G. H. (2015). The reward positivity: From basic research on reward to a biomarker for depression. *Psychophysiology*, 52(4), 449-459.

Yi, W., Mei, S., Zhang, M., & Zheng, Y. (2020). Decomposing the effort paradox in reward processing: Time matters. *Neuropsychologia*, 137, 107311.

komponent RewP jest obliczany z elektrody FCz, czyli z okolic czołowych, natomiast w recenzowanej pracy do przedstawienia RewP i obliczenia statystyk Autor wykorzystał elektrody ciemieniowe, z przewagą tych znad lewej półkuli, co podaje w tym podrozdziale i w podrozdziale 9.5. Wyniki. Nigdzie w tych podrozdziałach Autor nie podał uzasadnienia dla takiego wyboru i odstępstwa od istniejącego konsensusu. W związku z tym, taki dobór elektrod wydaje się arbitralny i niewłaściwy. To jest kwestia, którą Doktorant powinien wyjaśnić w trakcie obrony – dlaczego wybrał te właśnie elektrody do obliczenia i raportowania wyników dotyczących komponentu RewP. Ma to o tyle znaczenie, że komponenty ERP (ich amplitudy, latencje i inne cechy) różnią się znacznie między elektrodami, a określona lokalizacja elektrod jest wręcz cechą definicyjną danego komponentu, w wielu przypadkach.

Tak więc Autor, wybierając do obliczenia RewP inne elektrody niż te raportowane tradycyjne w literaturze, mógł w istocie obliczyć jakiś inny komponent, którego nie można by nazwać RewP.

Ta konkretna kwestia jest przykładem szerszego problemu związanego z pracami tak złożonymi pod względem metodologicznym jak niniejsza praca. Przy tak złożonych analizach, jak w niniejszej pracy, występuje wiele „stopni swobody badacza”, czyli miejsc, gdzie badacz może podjąć decyzję o tym, jak poprowadzić daną analizę, a wybór różnych ścieżek analizy może doprowadzić do skrajnie odmiennych wyników. W związku z tym, w takich złożonych pracach jest szczególnie ważne, żeby każda decyzja co do metod analizy była dobrze uzasadniona, z wyraźnymi odwołaniami do literatury i cytowaniami źródeł, żeby uniknąć poczucia, że badacz tak dobiera metody analizy, żeby dowieść swoich tez. W kilku miejscach części empirycznej miałem wrażenie, że Autor nie poświęca takim uzasadnieniom dość uwagi. Tak jest na przykład na str. 91, w podrozdziale 9.4. Modelowanie behawioralne, gdzie Autor pisze o tym, że w analizach wykorzystał regułę uczenia delta i regułę wzmocnienia-rozpadu, bez podania szczegółowych odniesień do literatury (obydwie te reguły mają bogatą literaturę). Mimo tych niedociągnięć, ogólnie ten podrozdział, 9.4. Modelowanie behawioralne, znów pokazuje metodologiczne zaawansowanie Autora i oceniam go bardzo pozytywnie, szczególnie dlatego, że Autor w tym podrozdziale stosuje wspomniane już przeze mnie kryterium informacyjne AIC do oceny modeli, co wydaje się właściwsze niż stosowanie metod opisanych w rozdziale 8.

W podrozdziale 9.5. Wyniki Autor na początku porównuje dane behawioralne między warunkami eksperymentalnymi przy pomocy analizy regresji, po czym przechodzi do opisu procedury selekcji modeli. W tej części pracy są miejsca, gdzie w trakcie czytania miałem wrażenie, że Autor, z powodu dużej złożoności analiz, się zgubił i prowadzi mnie na manowce. Tak jest na str. 99, gdzie Autor opisuje wyniki procedury selekcji modeli, używając skrótowych nazw modeli, DEL2DP i DRL2DP, bez podawania rozwinięcia tych skrótów – nie wiem, co tutaj oznacza owo DP i myślę, że w tym fragmencie Autor się pomylił w nazwach modeli, co spowodowało duże zamieszanie w opisie wyników. Proszę Autora o uważne przeczytanie tekstu na stronie 99., ustosunkowanie się do tego komentarza i szczegółowe wyjaśnienie tego fragmentu pracy w trakcie obrony. Inne fragmenty tego podrozdziału na szczęście bronią go jako całość, szczególnie połączenie modelowania behawioralnego z analizami ERP, na stronach 103-110, są bardzo wartościowe. Wyjątek stanowi tu podrozdział dotyczący komponentu Reward Positivity, który jest obciążony błędem wyboru elektrod, o którym pisałem powyżej. Podsumowanie Eksperymentu 2., podobnie jak w przypadku Eksperymentu 1., spełnia bardzo ważną funkcję systematyzowania przeprowadzonych analiz i otrzymanych wyników, i jest wartościowym fragmentem pracy.

W kolejnych rozdziałach, 10. Dyskusja, 11. Ograniczenia i dalsze kierunki badań i 12. Podsumowanie, Autor interpretuje uzyskane wyniki i odnosi je do szerszego kontekstu. W oparciu o dobrze przeprowadzony w części teoretycznej przegląd literatury, Autor potrafił

wyciągnąć ze swoich badań cenne wnioski na temat mechanizmów leżących u podstaw obserwowanych przez siebie zależności. Lektura tych rozdziałów sprawiła, że byłem zaintrygowany przeprowadzonymi przez Autora analizami i razem z Autorem próbowałem szukać odpowiedzi na postawione przez niego pytania, co według mnie świadczy o wysokiej wartości tej części oraz całości pracy. Części dodatkowe: Literatura cytowana, Spis tabel, Spis rycin, Spis równań i Załącznik są wyedytowane starannie, według reguł obowiązujących w odniesieniu do manuskryptów i prac empirycznych.

### **Konkluzja**

Na podstawie powyższych uwag konkluduję więc, że omawiana praca doktorska jest bardzo ambitna, zaawansowana metodologicznie oraz interdyscyplinarna – łączy ze sobą zaawansowane metody statystycznej analizy danych, modelowania matematycznego danych behawioralnych i analizy danych psychofizjologicznych. Po przeczytaniu jej zdecydowanie uważam, że Doktorant posiada kompetencje badawcze, jakich oczekuje się od kandydatów na stopień naukowy doktora oraz od naukowców na dalszych etapach kariery.

**Stwierdzam zatem, że przedstawiona mi do oceny rozprawa doktorska spełnia warunki określone w art. 13 ustawy z dnia 14 marca 2003 roku o tytule naukowym i stopniach naukowych. Dlatego też wnioskuję o dopuszczenie pana mgra Przemysława Marcowskiego do dalszych etapów przewodu doktorskiego.**

