

Kraków, 13 października 2024

dr hab. Magdalena Senderecka, prof. UJ
Katedra Kognitywistyki
Uniwersytet Jagielloński w Krakowie



UNIWERSYTET
JAGIELLOŃSKI
W KRAKOWIE

Recenzja poprawionej wersji pracy doktorskiej mgr Katarzyny Zemły,
pt. *Examining the efficacy of Guided Imagery relaxation technique in reducing
stress, modulating brain wave activity, and enhancing attention control*

Promotor: dr hab. Grzegorz Marcin Wójcik, prof. PJATK

Kopromotor: prof. dr hab. Grzegorz Sędek

Wydział Filozoficzny

Katedra Kognitywistyki

Wskaźniki ilościowe

Poprawiona wersja pracy doktorskiej mgr Katarzyny Zemły składa się z dwóch artykułów opublikowanych w 2023 roku w multidyscyplinarnych czasopismach indeksowanych w bazie *Web of Science*, jednego rozdziału opublikowanego w monografii wydanej również w 2023 roku oraz jednego preprintu (którego nie było w pierwotnej wersji pracy). Pierwszy artykuł ukazał się w piśmie *Applied Sciences*, dla którego aktualne współczynniki wpływu, takie jak *Impact Factor* (IF) i *SCImago Journal Rank* (SJR), oraz punktacja MNiSW prezentują się następująco: IF = 2.5, SJR = 0.508, 100 punktów. Drugi artykuł został opublikowany w piśmie *Sensors*, którego wskaźniki oddziaływania i punktacja wynoszą: IF = 3.4, SJR = 0.786, 100 punktów. Warto zwrócić uwagę, że o ile IF wyliczany jest wyłącznie na podstawie średniej liczby cytowań artykułów opublikowanych w danym piśmie, o tyle SJR uwzględnia rangę czasopism, z których te cytowania pochodzą. Znaczna rozbieżność między wartością IF a SJR w przypadku obu czasopism wskazuje, że choć artykuły w nich publikowane są cytowane relatywnie często, to jednak cytowania te nie pochodzą z renomowanych czasopism. Zarówno *Applied Sciences*, jak i *Sensors* firmowane są przez MDPI, tj. wydawcę kojarzonego z praktykami „drapieżnymi”. Oddznaczą się bardzo krótkim procesem recenzyjnym, co potwierdzają informacje zamieszczone w pracach Doktorantki i współpracowników. Pierwszy artykuł, wchodzący w skład pracy doktorskiej, został przyjęty do druku po dwóch i pół miesiąca od wysłania do redakcji (*Applied Sciences*), natomiast drugi – po miesiącu (*Sensors*). Z kolei rozdział ukazał się w liczącej 77 stron monografii *Selected Topics in Applied Computer Science*, która została wydana nakładem wydawnictwa Maria Curie-Skłodowska University Press w Lublinie.

ul. Ingardena 3

30-060 Kraków

tel. +48 (12) 663 17 99

kognitywistyka@uj.edu.pl

<https://kognitywistyka.uj.edu.pl/pl>

Oba artykuły oraz preprint mają formę doniesień z badań empirycznych i powstały w wyniku realizacji tego samego eksperymentu (a przynajmniej nie znalazłam w pracy danych, które pozwalałyby wnioskować, że było inaczej). Rozdział opublikowany w monografii ma formę przeglądową, przy czym zawiera też opis wyników pomiarów przeprowadzonych na jednej osobie. Wszystkie te prace są efektem współautorskiej pracy zespołu, składającego się z 5-6 osób. W przypadku obu artykułów oraz rozdziału mgr Katarzyna Zemła pełniła wśród nich rolę pierwszego autora, a jej wkład obejmował takie aktywności, jak: konceptualizacja projektu badawczego, przeprowadzenie przeglądu literatury, przygotowanie procedury relaksacji bazującej na prowadzonej wyobraźni, analiza danych i weryfikacja wyników, przygotowanie manuskryptów. W preprincie, który został dodany do poprawionej wersji pracy, Doktorantka została wymieniona na piątym (przedostatnim) miejscu. Uczestniczyła w konceptualizacji projektu od strony psychologicznej oraz była odpowiedzialna za przygotowanie procedury relaksacji. Poza dwoma artykułami i preprintem, w skład rozprawy wchodzi autoreferat, napisany w języku angielskim, który w poprawionej wersji pracy uległ rozszerzeniu. Pod względem ilościowym rozprawa spełnia wymagania ustawowe stawiane pracom doktorskim. Tej pozytywnej oceny niestety nie można rozciągnąć na cechy formalne i zawartość treściową pracy.

Cechy formalne

Praca doktorska mgr Katarzyny Zemły, na czele z opublikowanymi już artykułami, została przygotowana, w mojej ocenie, w sposób niestaranny, bez zachowania obowiązujących standardów edytorskich dla tekstu naukowego z zakresu psychologii. Recenzja nie jest miejscem odpowiednim do wskazania pełnej listy uchybień (byłoby to możliwe jedynie poprzez nanoszenie korekt w pliku z rozprawą), poniżej wskazuję jednak na trzy główne typy popełnionych błędów.

(I) Zaburzenia kompozycji tekstu i istotne braki w jego zawartości.

W artykule pierwszym (*Applied Sciences*) sekcja *Discussion* zawiera elementy opisu wyników, z przedstawieniem wartości liczbowych różnych statystyk (podrozdział 4.1. *Signal Classification*), które powinny znaleźć się we wcześniejszej sekcji *Results*. Opis procedury jest niepełny. Nie wiadomo choćby, na jakiej puli osób przeprowadzono finalne analizy. W tekście pojawiają się mylne informacje. Na przykład zdanie „The event-related potential (ERP) experiments were designed in PST e-Prime 2.0.8.90.” sugeruje, że badanie obejmowało pomiary, które miały stanowić bazę dla analizy potencjałów związanych ze zdarzeniem. Ostatecznie jednak artykuł nie zawiera dalszych odniesień do wspomnianego typu analizy sygnału elektroencefalograficznego (EEG). Bardzo podobne zdanie pojawia się także w drugim artykule oraz w poprawionej wersji autoreferatu – w tych przypadkach również nie wiadomo dlaczego. Ponadto w artykule drugim (*Sensors*) cała sekcja *Materials and Methods* napisana

jest w sposób niezwykle chaotyczny. Opis procedury, osób badanych, użytych narzędzi badawczych i aparatury, wykonanych analiz przenika się wzajemnie, bez zachowania struktury, którą pozornie sugerują tytuły podrozdziałów. Na przykład podrozdział 3.3. *The Cohort* zawiera wymieszane ze sobą informacje o zrekrutowanej próbie, użytym systemie EEG, uzyskanej zgodzie od komisji etycznej, sposobie przetworzenia danych EEG, czasie trwania zadań, typie przeprowadzonych analiz, instrukcjach dla osób badanych i jeszcze paru innych wątkach. Drugi artykuł nie zawiera też sekcji *Results* oraz *Discussion*. Po opisie wyników, który znajduje się w mylnie zatytułowanym podrozdziale 4. *Statistical Analysis of the Data* (zamiast *Results*), pojawia się od razu podrozdział 5. *Limitations of the Study*, a następnie 6. *Conclusions*. Mniej zastrzeżeń budzi kompozycja preprintu, jednak w jego przypadku Doktorantka nie uczestniczyła w przygotowaniu manuskryptu.

W poprawionej wersji autoreferatu również zdarza się problematyczne przyporządkowanie treści do podrozdziału. Na przykład w podrozdziale 1.10 *Potential benefits of Guided Imagery in Virtual Environments for Relaxation and Well-Being* Doktorantka w bardzo ograniczonym zakresie odnosi się do zapowiedzianych w tytule korzyści, płynących ze stosowania metody prowadzonej wyobraźni w specyficznym środowisku, jakim jest wirtualna rzeczywistość. Większość podrozdziału nieoczekiwanie poświęca na omówienie różnych metod klasyfikacji sygnału EEG, opierających się na uczeniu maszynowym. Takie nieścisłości pojawiają się też w innych miejscach autoreferatu. Na przykład z jego treści wynika, że Rycina 3.1. ilustruje wzrost mocy w zakresie rytmu alfa po sesji relaksacyjnej, podczas gdy wykres obrazuje skuteczność działania klasyfikacji badanych na osoby przynależące do grupy eksperymentalnej i kontrolnej na podstawie sygnału EEG. Wykres ten nie dostarcza jednak informacji na temat wzrostu mocy w zakresie rytmu alfa w grupie eksperymentalnej.

(II) Błędy oraz brak spójności i/lub dostosowania się do standardów w sposobie cytowania literatury.

W opublikowanych artykułach odniesienia do literatury częstokroć przyjmują formę przywołania pełnego tytułu cytowanego artykułu, na przykład „the study conducted by Yaxin Fan “Short Term Integrative Meditation Improves Resting Alpha Activity and Stroop Performance” [62] provides...”. W literaturze psychologicznej taki sposób cytowania zarezerwowany jest z reguły dla klasycznych tekstów, których znaczenie autor pracy chce podkreślić w sposób szczególny. Regularne stosowanie takiej formy nie jest samo w sobie błędem, ale stanowi odejście od przyjętych standardów (tj. zasady wskazywania wyłącznie nazwisk autorów i roku opublikowania artykułu lub alternatywnie podawania liczby, odwołującej czytelnika do odpowiedniej pozycji w sekcji *References*).

Odwołania do literatury – we wszystkich częściach pracy doktorskiej za wyjątkiem dodanego preprintu – pomijają często informację o współautorach cytowanego artykułu. Wspominałam o tym mankamencie już w poprzedniej recenzji, jednak problem ten nadal jest obecny, również w poprawionej wersji autoreferatu. Na przykład w zdaniach:

- *Specifically, Lin's analysis (Lin et al., 2022) highlighted...;*
- *Miyake (Miyake et al., 2000) proposed...;*
- *For instance, Davidson (Davidson et al., 2003) demonstrated...;*

nie uwzględniono, że podmiotem tych zdań powinien być pierwszy autor i współpracownicy, czyli:

- *For instance, Davidson et al. (2003) demonstrated...;*
- *Specifically, Lin et al.'s (2022) analysis highlighted...;*
- *Miyake et al. (2000) proposed....*

Są też miejsca, w których Doktorantka prawidłowo wymienia wszystkich współautorów w treści zdania, ale jednocześnie nadmiarowo powtarza ich nazwiska w nawiasie. Na przykład zdania:

- *Dahl, Lutz and Davidson (Dahl, Lutz, & Davidson, 2015) proposed...*
- *The findings also align with the work of Kane and Engle (Kane & Engle, 2003)...*
- *Diamond describes executive functions (...) environmental demands (Diamond, 2013);*

należałoby skrócić do następującej formy:

- *Dahl, Lutz and Davidson (2015) proposed...*
- *The findings also align with the work of Kane and Engle (2003)...*
- *Diamond (2013) describes executive functions (...) environmental demands.*

O tym mankamencie również wspomniałam w poprzedniej recenzji, ale ponieważ problem ten nie został wyeliminowany, tym razem podaję więcej przykładów zaczerpniętych z autoreferatu.

Kolejną wadą, na którą zwróciłam już poprzednio uwagę, jest brak spójności w kolejności cytowania kilku prac w jednym nawiasie. Dla przykładu, w nowej wersji autoreferatu Doktorantka nadal kieruje się:

- czasami alfabetycznie, np. (Diamond, 2013; Miyake et al., 2000), (Mellenthin, 2021; Shafer & Greenfield, 2000);
- czasami chronologią, np. (Miyake et al., 2000; Friedman, 2015), (Miyake et al., 2000; Hellmuth et al., 2012);
- a czasami pomija obie te reguły, np. (Lin et al., 2022; Fox et al., 2016; Manna et al., 2010).

Rozdział opublikowany w monografii cechuje się dodatkowo niespójnym formatowaniem sekcji *References*, w której część wpisów zawiera informacje o pełnych imionach autorów,

a część jedynie o ich inicjałach. Wszystkie wspomniane wyżej błędy świadczą o nieznajomości obowiązujących reguł lub o znacznym braku konsekwencji w ich stosowaniu.

Chciałabym w tym miejscu mocno podkreślić, że błędy te rozważane z osobna mogłyby stanowić wybaczone pomyłki, które przytrafiają się każdemu i o których nie warto nawet wspominać, o ile pojawiają się sporadycznie. W przypadku przedłożonej pracy doktorskiej stanowią one jednak regułę, a nie wyjątek, dlatego urastają do rangi problemu, ważącego na ocenie całego dzieła.

(III) Błędy językowe i niespójności terminologiczne.

W tekstach stanowiących pracę dokorską można znaleźć zdania, których przekaz jest bardzo trudny do odczytania. Dla przykładu przytaczam poniżej dwa zdania wyjęte z sekcji *Introduction* artykułu drugiego (z *Sensors*):

- *Hence, improving attention and executive functions is vital in our current world, given the cognitive demands and challenges we face, ultimately benefiting individuals and society as a whole in wide range of EEG experiments designed to quantitatively measure cognitive functions like those in [8].*
- *Given the evidence of both anxiety reductions and immune system enhancements, GI has not been studied during brain behavior and brainwave changes while patients are conducting GI sessions.*

Tego typu zdania wskazują, iż manuskrypt (opublikowanego już artykułu) został zredagowany niestarannie, bez upewnienia się, czy wywód jest logiczny i zrozumiały dla czytelnika.

W artykułach zwraca też uwagę niekonsekwentne stosowanie terminologii. Dobrą ilustrację tego problemu może choćby stanowić wymienne używanie nazw do określenia typu analiz statystycznych, przeprowadzonych w ramach badań opisanych w artykule z *Applied Sciences*: (1) *general linearized model*, (2) *general linear model*, (3) *generalized linear models*. Jednocześnie w artykule z *Sensors* pojawia się zastrzeżenie:

- *The **generalized linear model** enhances the **general linear model** by introducing a specified link function to establish a linear association between the dependent variable and the factors and covariates.*

Oznacza to, że sami Autorzy tych dwóch artykułów uważają, że nazwy te oznaczają coś innego, co w konsekwencji prowadzi do wniosku, że nie powinny być używane wymiennie, a jednak są.

Do językowej kategorii błędów formalnych pozwolę sobie także zaliczyć niekonsekwentne stosowanie akronimów. Dany akronim powinien zostać wprowadzony w głównej części artykułu tylko jeden raz, przy pierwszym wystąpieniu ciągu wyrazów, który ma zastąpić, np.

guided imagery (GI). W dalszej części tekstu zamiast ciągu wyrazów powinien pojawiać już tylko wprowadzony wcześniej akronim, czyli w naszym przykładzie GI. Wbrew tej regule Doktorantka wielokrotnie wprowadza właściwe wszystkie akronimy, których używa (typu CNN, EEG, FA, GI, GLM, MBIs, MBSR, OM itd.). W pierwotnej wersji pracy akronim GI został wprowadzony w autoreferacie siedem razy (i sześć razy w artykule z *Sensors*). W poprawionej wersji pracy akronim GI został wprowadzony w autoreferacie ponad 50 razy, czasami dwa razy w jednym zdaniu lub kilka razy w jednym akapicie. Nie wiem, dlaczego Doktorantka zdecydowała się na takie rozwiązanie. Z perspektywy czytelnika mogę jedynie podzielić się refleksją, że stosowanie akronimów z reguły utrudnia, a nie ułatwia lekturę.

Znaczne nagromadzenie tego typu problemów (podobnych do tych wybiórczo opisanych powyżej) prowadzi do konkluzji, że rozprawa doktorska nie została właściwie opracowana od strony formalnej. Wskazuje to również na brak odpowiedniej znajomości standardów edytorskich tekstów naukowych.

Cechy treściowe – przedmiot badań, uzyskane wyniki i ich ocena

Rozprawa podejmuje tematykę wpływu technik relaksacyjnych, w szczególności metody prowadzonej wyobraźni, na funkcjonowanie afektywne, poznawcze oraz na aktywność mózgu. Zgodnie z informacjami zamieszczonymi w zmodyfikowanej wersji polskojęzycznego abstraktu celem pracy było zbadanie *wpływu techniki relaksacyjnej - wyobraźni prowadzonej (GI) na funkcje poznawcze i emocjonalne. Dodatkowym aspektem badania była eksploracja potencjalnych zastosowań klasyfikatorów uczenia maszynowego (GLM) w wsparciu terapeutycznym i tworzeniu na tej podstawie interfejsów mózg-komputer, które byłyby indywidualnie dopasowane do potrzeb użytkownika.* Wedle przyjętych hipotez (przedstawionych mniej lub bardziej bezpośrednio w różnych miejscach pracy), wprowadzenie osób badanych w stan relaksacji przy użyciu metody prowadzonej wyobraźni powinno przełożyć się na obniżenie poziomu lęku i stresu, zredukowanie poczucia bezsilności, wzrost poczucia relaksu, wzmocnienie funkcji uwagowych, zwiększenie mocy sygnału EEG w zakresie rytmu alfa oraz zmniejszenie mocy sygnału EEG w zakresie rytmu beta. Ponadto założono, iż w oparciu o charakterystyki sygnału EEG możliwa będzie klasyfikacja badanych na osoby przynależące do grupy eksperymentalnej, tj. wprowadzone w stan relaksacji, oraz stanowiące grupę kontrolną, tj. wykonujące zadanie poznawcze.

W badaniu udział wzięło 60 mężczyzn, w wieku od 17 do 24 lat, losowo przydzielonych do wspomnianych wyżej dwóch, równolicznych grup. W analizach opisanych w artykule z *Sensors* finalnie uwzględniono 20 osób z grupy eksperymentalnej i 28 osób z grupy kontrolnej. Z kolei w artykule z *Applied Sciences* liczba osób włączonych do analiz nie została wprost określona, ale kierując się stopniami swobody można przypuszczać, że wynosiła 55. Sygnał EEG w grupie eksperymentalnej zarejestrowano podczas sesji relaksacyjnej trwającej około 20 minut,

natomiast w grupie kontrolnej – podczas trwającego tyle samo zadania poznawczego polegającego na przywoływaniu z pamięci nazw europejskich krajów, ich stolic, nazw stanów USA oraz znaków zodiaku. Polecenia głosowe, składające się na sesję relaksacyjną, zostały uprzednio nagrane, aby zapewnić jednorodność manipulacji eksperymentalnej między osobami. Po zakończeniu rejestracji sygnału EEG osoby badane wykonywały trzy zadania poznawcze: zadanie Stroopa, zadanie antysakadowe i zadanie go/no-go. Ponadto przed i po rejestracji EEG wypełniały kwestionariusze, służące do samoopisowej oceny stanu afektywnego. Przeprowadzona klasyfikacja obejmowała podział osób badanych na zbiór treningowy (80% osób) i testowy (20% osób).

Zgodnie z interpretacją Doktorantki oraz współpracowników uzyskane wyniki potwierdziły hipotezy, dotyczące obniżenia poziomu stresu, wzrostu poczucia relaksu, wzmocnienia funkcji uwagowych w zadaniu Stroopa i zadaniu antysakadowym oraz zwiększenia mocy sygnału EEG w zakresie rytmu alfa w wyniku przeprowadzenia sesji relaksacyjnej. Nie potwierdziły natomiast hipotez, dotyczących obniżenia poziomu lęku, zredukowania poczucia bezsilności, wzmocnienia funkcji uwagowych w zadaniu go/no-go oraz zmniejszenia mocy sygnału EEG w zakresie rytmu beta. Ponadto możliwe okazało się przyporządkowanie osób do grupy eksperymentalnej i kontrolnej w oparciu o przeprowadzone klasyfikacje (opisane w artykule i preprincie). Dodatkowe analizy korelacyjne wskazały, zdaniem Doktorantki i współpracowników, na szereg istotnych związków między uwzględnionymi zmiennymi. W ramach interpretacji uzyskanych wyników Doktorantka i współpracownicy stwierdzili m.in., że relaksacja poprawia funkcje poznawcze oraz podnosi samopoczucie na poziomie emocjonalnym, a za efektami tymi stoi zwiększenie mocy sygnału EEG w zakresie rytmu alfa. Autorzy zasugerowali również, że wyniki te wskazują na przewagę relaksacji opierającej się na metodzie prowadzonej wyobraźni nad relaksacją bazującą na treningu uważności.

Przechodząc do oceny merytorycznej opisanego wyżej badania z przykrością muszę stwierdzić, że jest ono obciążone poważnymi błędami metodologicznymi, które niestety nie pozwalają na poprawne przetestowanie większości postawionych hipotez i przekreślają możliwość sensownej interpretacji uzyskanych wyników. Na kluczowe problemy wskazuję w zamieszczonych poniżej punktach.

(I) Pierwszy poważny błąd metodologiczny polega na przeprowadzeniu badań w schemacie poprzecznym, a nie podłużnym. Błąd ten wiąże się z **brakiem uwzględnienia dwukrotnego pomiaru wszystkich zmiennych, na które zgodnie z przyjętymi hipotezami mogła wpłynąć manipulacja eksperymentalna**. Pomiar powinien zostać przeprowadzony przed zastosowaniem manipulacji (w preteście) oraz po niej (w postteście). Dopiero w oparciu o porównanie wyników pretestu i posttestu (oraz przy uwzględnieniu wyników grupy kontrolnej, do której szerzej odniosę się w kolejnym punkcie) możliwe jest wnioskowanie

o wpływie manipulacji eksperymentalnej na zmienne zależne. W przeprowadzonym badaniu dwukrotnie zostały zmierzone wyłącznie zmienne afektywne (samoopisowe). W pracy nie podano jednak statystyk opisowych z pretestu i posttestu, ani nie przeprowadzono analizy z powtarzanym pomiarem, a jedynie wskazano na różnicę występującą między pretestem a posttestem. Różnica ta okazała się istotnie większa w grupie eksperymentalnej niż kontrolnej wyłącznie w jednym przypadku – poziomemu stresu. Wynik ten jednak trudno zinterpretować bez informacji o tym, czy grupy były wyrównane pod względem poziomu stresu w preteście, a takiej informacji praca nie zawiera (Doktorantka nie podała jej też w zmienionej wersji autoreferatu). W przypadku funkcji poznawczych, brak zastosowania dwukrotnego pomiaru wyklucza możliwość interpretowania uzyskanych wyników w kategoriach wpływu przeprowadzonego treningu. Z kolei w przypadku zmiennych neurofizjologicznych brak pretestu i posttestu (czyli pomiaru aktywności elektrycznej mózgu w stanie spoczynku przed i po treningu) można byłoby częściowo zrekompensować poprzez analizę sygnału zarejestrowanego w pierwszej i ostatniej minucie treningu. Aby jednak ewentualną różnicę między tymi dwoma pomiarami przypisać wpływowi relaksacji, potrzebne byłoby odniesienie do wyników uzyskanych w odpowiednio dobranej grupie kontrolnej. W tym miejscu dochodzimy jednak do drugiego problemu.

(II) Drugi poważny błąd metodologiczny polega na **zastosowaniu w grupie kontrolnej zadania, które różni się na bardzo wielu wymiarach od procedury relaksacji użytej w grupie eksperymentalnej**. Manipulacja kontrolna co do zasady powinna angażować osoby badane w podobny sposób do manipulacji eksperymentalnej (czas trwania, warunki jej przeprowadzania, typ aktywności itp.), choć jednocześnie musi być pozbawiona tego elementu, który uznaje się za kluczowy do wywołania efektu eksperymentalnego. Jeśli analiza wyników uzyskanych w preteście i postteście ujawni, że wpływ manipulacji eksperymentalnej na zmienną zależną jest istotnie statystycznie większy od wpływu manipulacji kontrolnej, wówczas w uprawniony sposób można wnioskować o specyficznym wpływie treningu eksperymentalnego na zmienną zależną. W opisanym w pracy badaniu manipulacja eksperymentalna polegała na podążaniu za głosowymi instrukcjami oddziałującymi na wyobraźnię wzrokową, mającymi na celu wywołanie efektu relaksacji. Manipulacja kontrolna mogłaby z kolei polegać np. na słuchaniu neutralnego tekstu, odczytywanego przez tę samą osobę, co w przypadku manipulacji eksperymentalnej, aby maksymalnie zrównać fizyczne parametry sygnału dźwiękowego prezentowanego w obu grupach. W opisanym w pracy badaniu w ramach manipulacji kontrolnej zastosowano jednak zadanie poznawcze o zupełnie innym charakterze, polegające (jak już wspomniałam powyżej) na przywoływaniu z pamięci nazw europejskich krajów, ich stolic, nazw stanów USA oraz znaków zodiaku. Było to zatem zadanie, które w żaden sposób nie angażowało modalności słuchowej, za to stanowiło dla badanych intelektualne wyzwanie, które mogło działać w kierunku przeciwnym do relaksacji. Potwierdzenie tej hipotezy można znaleźć w wynikach klasyfikacji, w której dokładność

działania klasyfikatora osiągnęła poziom powyżej losowego (ok. 76-77%, bazując na wykresie z artykułu z *Sensors*) dla sygnału EEG zebranego już w pierwszej minucie treningu. Oznacza to, że już w punkcie wyjścia aktywność mózgu osób z grupy eksperymentalnej i kontrolnej różniła się znacząco.

(III) Powyższe rozważania wprost prowadzą do wniosku, że wyniki klasyfikacji opisanej w artykule z *Sensors* mają dość trywialny charakter, gdyż pokazują, że aktywność elektryczna mózgu zarejestrowana w dwóch zupełnie różnych sytuacjach różni się od siebie. Dla sytuacji mniej angażującej poznawczo, aktywność ta charakteryzuje się większą mocą sygnału w zakresie rytmu alfa, a dla sytuacji bardziej angażującej poznawczo – mniejszą. Bardzo doceniam rozbudowanie tego wątku w dodanym do rozprawy preprincie, w którym zastosowano szeroki zestaw klasyfikatorów i porównano skuteczność ich działania na mniej lub bardziej rozbudowanym zestawie danych (tj. sygnały, pochodzącym z mniejszej lub większej liczby elektrod), z uwzględnieniem wielu kluczowych metryk. Uważam, że wyniki tych analiz są z pewnością wartościowe i istotne z metodologicznego punktu widzenia. Pozytywna ocena preprintu niestety nie zmienia całościowego odbioru rozprawy doktorskiej, tym bardziej że wkład mgr Katarzyny Zemły w powstanie tej ostatniej części pracy był niewielki.

(IV) Do błędów powiązanych z warstwą metodologiczną pracy zaliczyć też można całkowite pominięcie – w opublikowanych już artykułach – opisu struktury trzech wykorzystanych w badaniu zadań poznawczych (zadania Stroopa, zadania antysakadowego i zadania go/no-go). Czytelnik artykułów nie odnajdzie w nich informacji o pojawiających się w zadaniach bodźcach, czasie ich prezentacji, liczbie prób, sposobie udzielania odpowiedzi itd. Brak takiego opisu czyni niemożliwym zreplikowanie badania. Za dość zdawkowy można też uznać opis przetworzenia surowego sygnału EEG. Nie obejmuje on informacji, dotyczących np. kryteriów odrzucania artefaktów, liczby odrzuconych segmentów oraz liczby zinterpolowanych kanałów.

W nowej wersji autoreferatu Doktorantka opisała niektóre z tych brakujących elementów, np. podała liczbę prób występujących w zadaniach, liczbę bloków oraz przypisanie prób do bloków. Te dodatkowe informacje niestety jedynie wzmocniły dotychczasowe wątpliwości odnośnie metodologicznej poprawności przeprowadzonych badań. W szczególności, niepokoić może mała liczba prób, uwzględnionych w zadaniu antysakadowym oraz w zadaniu Stroopa, oraz nielosowe przypisanie prób zgodnych i niezgodnych do bloków. Na przykład zadanie Stroopa obejmowało w sumie 40 prób, podzielonych na cztery bloki w następujących proporcjach: blok 1. – 5 prób zgodnych, blok 2. – 15 prób niezgodnych, blok 3. – 5 prób niezgodnych i 10 zgodnych, blok 4. – 5 prób zgodnych. Oznacza to, że zadanie zawierało ogółem bardzo mało prób, a próby z dwóch kategorii były w większości bloków prezentowane w seriach. Taka konstrukcja zadania może pociągać za sobą problemy, związane z niską

rzetelnością pomiaru oraz niską skutecznością wzbudzania interferencji. Konstrukcja zadania antysakadowego obciążona była bardzo podobnymi mankamentami, co omówione właśnie zadanie Stroopa. Jedynie w przypadku zadania go/no-go liczba prób była nieco większa (150), a proporcje i sposób prezentacji prób go i no-go literaturowe (80:20%, pseudolosowa kolejność, zapobiegająca seryjnej prezentacji prób no-go). Warto zwrócić uwagę, że istotne różnice międzygrupowe w wykonaniu zadań udało się zaobserwować wyłącznie w zadaniu Stroopa i zadaniu antysakadowym. Biorąc pod uwagę niską siłę efektów i mankamenty konstrukcji obu zadań, nie można wykluczyć, że zaobserwowane różnice mogą odzwierciedlać wyniki fałszywie dodatnie. W rozstrzygnięciu tej kwestii na pewno pomogłoby przeprowadzenie testów rzetelności pomiaru we wszystkich zadaniach.

(V) Oprócz wspomnianych wyżej błędów metodologicznych, problemy pojawiają się też na poziomie analiz. W pierwszej kolejności zaliczyć do nich można znaczne rozbieżności, występujące między zaproponowaną przez Doktorantkę i współpracowników interpretacją wyników a wnioskami, które narzucają się po przestudiowaniu wartości statystyk podanych w tabelach. Dla przykładu rozważmy następujące zdanie z artykułu z *Sensors*:

- *The robust findings from this research provide compelling evidence supporting the efficacy of guided imagery (GI) as an intervention for stress reduction and **relaxation**, surpassing the effects observed in the mental task group.*

Zdanie to jednoznacznie wskazuje na istotną różnicę między grupą eksperymentalną a kontrolną dla efektu poczucia relaksu, podczas gdy statystyki podane w Tabeli 1, stojące za tym efektem, temu przeczą: $F = 2.28$, $p = 0.14$, $\eta^2 = 0.048$.

Podobny problem pojawia się w opisie wyników analiz korelacyjnych. W artykule z *Sensors* można przeczytać:

- *Furthermore, the anxiety measures (STAI Trait and STAI State) were positively correlated (...) with the anti-saccade task and the numerical Stroop task. (...) Notably, the correlation between the STAI State anxiety measure and the alpha power at the 14th min was also significant, indicating that a higher anxiety was associated with a lower alpha power.*

Jednak po spojrzeniu na statystyki podane w Tabeli 3 okazuje się, że żadna z opisanych wyżej korelacji nie przekroczyła progu istotności statystycznej. Oznacza to, że albo w tabelach podano błędne wartości statystyk, albo zostały one źle zinterpretowane (jako istotne, choć de facto istotnymi nie są).

Pewne wątpliwości wzbudza też opis niektórych narzędzi statystycznych lub uzasadnienie ich użycia. Na przykład w artykule z *Sensors* Doktorantka i współpracownicy twierdzą, że MANOVA jest testem nieparametrycznym, podczas gdy w istocie jest testem parametrycznym.

Z kolei uzasadnienie dla użycia testu ANOVA brzmi następująco:

- *We decided to use an ANOVA for our data analysis due to its balance and neuroscientific character.*

W tym przypadku czytelnik może się zastanawiać, o jaką równowagę właściwie chodzi i w czym tkwi neuronaukowy charakter tego narzędzia statystycznego.

Dość niezręcznym uzasadnieniem przeprowadzenia analiz statystycznych jest też pojawiający się w artykule z *Applied Sciences* argument z niemożliwości przeprowadzenia oceny różnic gołym okiem:

- *These maps, however, are too similar and cannot be easily interpreted using the naked eye. (...) Nevertheless, differences in activity are visible, even though they are not easily interpretable. An appropriately trained machine learning classifier can be used for this task.*

W przypadku klasyfikacji można też zgłosić zastrzeżenia, dotyczące liczebności zbioru treningowego i testowego (przy podziale 80% do 20% w zbiorze testowym w grupie eksperymentalnej znalazły się zapewne tylko 4 osoby) oraz nierównoliczności grupy eksperymentalnej i kontrolnej, która mogła zaważyć na wynikach klasyfikacji, zawyżając wartości metryk.

(VI) W nowej wersji autoreferatu Doktorantka rozszerzyła nieco krytyczną ocenę dotychczasowych badań, poświęconych wpływowi technik relaksacyjnych na funkcjonowanie poznawcze i aktywność mózgu. Jest to obszar badań, w którym dominują prace wykonane na małych próbach, obciążone niedociągnięciami metodologicznymi, opublikowane w niszowych czasopismach. Warto w tym miejscu zwrócić jednak uwagę na pracę autorstwa Hudetz, Hudetz i Reddy, opublikowaną w 2004 roku w *Psychological Reports*. Doktorantka często odwołuje się do tej pracy, podkreślając, iż było to pierwsze badanie dotyczące wpływu relaksacji bazującej na wyobraźni na funkcjonowanie poznawcze, które zostało zrealizowane z uwzględnieniem pomiaru sygnału EEG. Z przeanalizowania zawartości tej pracy wynika, że Hudetz et al. (2004) przeprowadzili badanie podłużne na 42 osobach podzielonych na dwie grupy – eksperymentalną, tj. poddaną treningowi relaksacyjnemu, oraz kontrolną, która między pretestem i posttestem nie została poddana żadnym zabiegom. Pomiar w preteście i postteście obejmował zmienne samoopisowe oraz zmienne behawioralne z zadania na pamięć roboczą. Przez cały czas trwania badania rejestrowano zmiany w aktywności elektrycznej mózgu oraz w pracy serca. Wyniki analiz wskazały, że w preteście osoby badane z obu grup nie różniły się na żadnym wymiarze, natomiast w postteście osoby z grupy eksperymentalnej w porównaniu do osób z grupy kontrolnej charakteryzowały się niższym poziomem lęku, wyższą skutecznością działania pamięci roboczej oraz zmniejszeniem mocy sygnału EEG w zakresie wysokich częstotliwości czy też zwiększeniem mocy sygnału EEG w zakresie niskich

częstotliwości. Mimo iż to pionierskie badanie nie jest pozbawione wad (przydałaby się chociażby aktywna grupa kontrolna i większa liczba osób badanych), to jednak wskazuje na pewien metodologiczny wzorzec badań w tym nurcie, którego warto byłoby się co do zasady trzymać, oczywiście koncentrując się na jego udoskonaleniu i rozwinięciu.

W ocenie merytorycznej pracy doktorskiej celowo pominęłam zawartość opublikowanego w monografii rozdziału, gdyż zawarte w nim informacje mają w dużej mierze charakter podręcznikowy, a wyniki pomiarów przeprowadzone na jednej osobie nie stanowią istotnego wkładu do badań empirycznych.

Na marginesie warto jeszcze dodać, że inaczej niż sugeruje Doktorantka i współpracownicy w artykule z *Applied Sciences* (*Modern brain imaging techniques, such as EEG, are crucial in verifying computational models...*) metoda EEG nie jest narzędziem nowym, gdyż jego historia sięga XIX wieku. Zdecydowanie zgadzam się jednak, że jego właściwe wykorzystanie z pewnością może przyczynić się do zweryfikowania wielu modeli, formułowanych w ramach neuronauki obliczeniowej, a przy okazji przełożyć się także na wzrost rzetelności i replikowalności badań poświęconych wpływowi technik relaksacyjnych na samopoczucie i zachowanie.

Uwagi podsumowujące

W oparciu o przeprowadzoną ocenę z przykrością muszę stwierdzić, że poprawiona wersja pracy doktorskiej mgr Katarzyny Zemły nie spełnia pod względem formalnym i merytorycznym wymagań ustawowych stawianych pracom doktorskim. Pozytywna ocena rozprawy pod względem ilościowym wynika wyłącznie z faktu, iż artykuły wchodzące w skład rozprawy zostały opublikowane w czasopiśmie znajdującym się w bazie *Web of Science* i na liście MNiSW, mimo jawnie „drapieżnego” charakteru ich wydawcy. W poprzedniej recenzji zgłosiłam gotowość zrewidowania mojej oceny, pod warunkiem przedłożenia pracy doktorskiej, która opierałaby się na zupełnie nowych badaniach. Ten warunek nie został spełniony. Obecne w pracy błędy formalne i merytoryczne sprawiają, iż nie mam podstaw do przyjęcia założenia, że mgr Katarzyna Zemła jest już teraz w pełni gotowa, aby wraz z uzyskaniem stopnia doktora móc prowadzić samodzielne badania i sprawować opiekę nad pracami dyplomowymi studentów. Mam jednak wielką nadzieję, że ta gotowość pojawi się i zostanie doceniona w przyszłości. Biorąc pod uwagę poprawioną wersję rozprawy, nie widzę podstaw do dopuszczenia mgr Katarzyny Zemły do dalszych etapów postępowania o nadanie stopnia doktora.