



instytut biologii doświadczalnej
im. Marcielego Nenckiego PAN

dr hab. Michał Bola

Pracownia Obrazowania Mózgu

Instytut Biologii Doświadczalnej im. M. Nenckiego PAN

Recenzja pracy doktorskiej

Autorka pracy: Pani mgr Natalia Jakubowska

Tytuł: "Comparison of psychophysiological and behavioral predictors of training effects with a complex task in the form of a strategic computer game"

Promotorka: dr hab. Aneta Brzezicka

Promotor pomocniczy: dr Radosław Nielek

Gry komputerowe stały się ważnym elementem (pop)kultury, na równi z innymi, bardziej klasycznymi formami, jak książki, filmy i seriale. To co wyróżnia gry komputerowe to duża interaktywność, poczucie sprawczości i możliwość wpływania na bieg wydarzeń, co może przekładać się na duże zaangażowanie i motywację podtrzymujące aktywność grania w czasie. Jak podaje Autorka pracy, już ponad 3 miliardy osób na świecie identyfikuje się jako gracze. Nie dziwi więc, że badacze są zainteresowani wpływem jaki gry komputerowe mogą wywierać na umysł i mózg. Co ważne, o ile gry komputerowe są nadal traktowane głównie jako forma rozrywki, coraz więcej danych wskazuje, że mogą też pełnić rolę treningu funkcji poznawczych i potencjalnie przynosić bardziej korzystne efekty niż specjalnie zaprojektowane do tego celu programy treningowe. Praca Pani mgr Natalii Jakubowskiej wpisuje się w ten ważny nurt badań i dostarcza ciekawych danych dotyczących efektu grania w gry akcji na behawioralne i elektrofizjologiczne wskaźniki funkcjonowania poznawczego, oraz możliwości przewidywania postępów treningu na bazie danych elektrofizjologicznych uzyskanych przed jego rozpoczęciem.

Od strony formalnej, przedstawiona praca doktorska składa się z trzech manuskryptów. Dwa zostały opublikowane w czasopismach *Heliyon* i *Frontiers in Human Neuroscience*, a trzeci jest obecnie w recenzji w czasopiśmie *Frontiers in Psychology*. Wszystkie trzy czasopisma wymagają *peer review* i są obecne w obiegu międzynarodowym. Przedstawione publikacje są wieloautorskie (od 6 do 7 współautorów) i we wszystkich Pani Jakubowska jest pierwszą autorką. Wkład Pani Jakubowskiej w powstanie badań i publikacji jest znaczący - zadeklarowano odpowiednio 85%, 85% i 87% procent



wkładu dla poszczególnych publikacji. W przypadku wszystkich badań Pani Jakubowska wykonywała przegląd literatury, analizowała dane, opisywała i przedstawiała wyniki, przygotowywała pierwszy draft manuskryptu, a w przypadku już opublikowanych badań, przygotowywała także odpowiedzi na komentarze recenzentów.

Warto podkreślić, że zaprezentowane artykuły są ze sobą ściśle powiązane i przedstawiają wyniki jednego projektu badawczego. Projekt ten był podłużnym badaniem osób deklarujących brak (lub minimalne) doświadczenia z grami komputerowymi, które przeszły 30 godzin treningu gry StarCraft II. Osoby biorące udział w projekcie zostały zakwalifikowane do trzech grup - grupy trenującej w złożonym środowisku, grupy trenującej w stałym środowisku, lub grupy kontrolnej (brak treningu). Przed rozpoczęciem treningu wszystkie osoby badane wzięły udział w sesji eksperymentalnej w laboratorium, następnie przeszły 30 godzin treningu z wykorzystaniem gry StarCraft II (w przypadku grupy kontrolnej brak treningu), a po jego zakończeniu wzięły udział w kolejnej sesji eksperymentalnej z wykorzystaniem tych samych paradygmatów. Paradygmaty wykorzystane podczas sesji eksperymentalnych to dwa klasyczne zadania - *attentional blink* i *change detection* - mierzące, odpowiednio, czasową dynamikę uwagi i pojemność pamięci roboczej. Podczas wykonania tych zadań rejestrowano sygnał EEG w celu analizy potencjałów wywołanych (ERP). Ponadto, analizie poddano dane telemetryczne zebrane podczas treningu, które pozwoliły oszacować postępy i poziom osiągnięty przez każdą z osób badanych.

Jedną z głównych zalet pracy jest podłużny charakter przeprowadzonego badania, pozwalający na wnioskowanie przyczynowe o możliwym wpływie gry na funkcje poznawcze. Z tego względu badanie przyczynia się do lepszego zrozumienia czy różnice obserwowane między graczami i nie-graczami w badaniach przekrojowych są rzeczywiście spowodowane przez fakt grania w gry. Kolejną zaletą jest wykorzystanie dwóch grup - grających w złożonym lub stałym środowisku - i wykazanie, że obserwowane efekty mogą zależeć od środowiska gry, a także że wartość predykcyjna danych może się objawiać tylko w przypadku, kiedy trening jest wystarczająco wymagający. Kolejną zaletą jest włączenie do analiz danych telemetrycznych, pozwalających śledzić postępy w treningu każdej osoby biorąc pod uwagę wiele wskaźników dotyczących rozgrywanych gier.

Natomiast jednym z głównych ograniczeń pracy są małe grupy osób badanych (16 do 22 osoby). Dotyczy to szczególnie możliwości generalizacji wyników dotyczących predykcji postępów na podstawie danych zebranych przed treningiem. To ograniczenie jest wspomniane i dyskutowane zarówno we wstępie pracy, jak i w dyskusji wyników poszczególnych artykułów. O ile jest to słabość przeprowadzonych badań, to biorąc pod uwagę, że Pani Jakubowska zajmowała się głównie analizą danych i przygotowywaniem publikacji, nie jest to fakt świadczący na niekorzyść pracy wykonanej



przez nią w czasie doktoratu. Mam natomiast kilka krytycznych uwag odnośnie autoreferatu, przeprowadzonych analiz oraz ich opisu.

Efekty behawioralne

Moje główne zastrzeżenie dotyczy interpretacji i sposobu prezentacji uzyskanych efektów behawioralnych. Sprawdzenie czy trening doprowadzi do polepszenia mierzonych behawioralnie funkcji poznawczych było jednym z głównych celów pracy (pytanie badawcze nr 1, s. 7). Jednak z uzyskanych wyników nie można według mnie wyciągnąć wniosku, że tak się stało.

W badaniu 1 nie znaleziono efektów między-grupowych (interakcji *grupa x sesja*). Biorąc pod uwagę, że do analizy nie włączono pasywnej grupy kontrolnej, a wzrost poprawności zaobserwowano w obu analizowanych grupach (tzn. trenujących w złożonym i stałym środowisku), konkluzja *“results suggest that RTS games indeed enhance attentional skills... independently of the game variant trained”* (s. 12) nie jest uprawniona. Przy braku pasywnej grupy kontrolnej nie można wykluczyć, że wzrost poprawności w obu grupach był spowodowany powtórным wykonywaniem zadania. Jednocześnie, zastanawia mnie, dlaczego dane z pasywnej grupy kontrolnej nie zostały uwzględnione w badaniu 1 (oraz w badaniu 3), skoro grupa ta była analizowana w badaniu 2. W pracy nie znajduję uzasadnienia takiej decyzji, a uwzględnienie tych danych mogłoby pozwolić zweryfikować czy wyżej wymieniona konkluzja jest prawdziwa.

Na podstawie wyników badania 2 Autorka wnioskuję, że w grupach eksperymentalnych doszło do zwiększenia pojemności pamięci roboczej po treningu, ale że takiego efektu nie było w pasywnej grupie kontrolnej (np. s. 13, ale także streszczenie). Wniosek ten jest wyciągnięty z wyników testów post-hoc pomiędzy sesjami, wykonanych w ramach każdego poziomu obciążenia poznawczego i w ramach każdej grupy (zaprezentowanych na rycinie 9, s. 14). Jednak ponieważ w przeprowadzonej analizie znaleziono istotne interakcje *obciążenie x sesja* i *grupa x obciążenie*, ale nie interakcję *grupa x sesja x obciążenie* (ani *grupa x sesja*), przeprowadzenie takich testów post-hoc wydaje mi się nieuprawnione (tylko 3-czynnikowa interakcja dawałaby podstawę do ich przeprowadzenia w ramach każdej grupy i obciążenia). Uważam, że przy braku istotnej interakcji wyniki tych testów nie powinny stanowić podstawy do konkluzji, że wzrost pojemności pamięci roboczej jest większy w jednej grupie a mniejszy w drugiej. W tym zakresie polecam zapoznanie się z publikacją Nieuwenhuis i wsp. (2011).

Podsumowując, pomimo że pierwsze z trzech wymienionych pytań badawczych (s. 7) dotyczyło behawioralnych efektów treningu, to uzyskane wyniki nie są odpowiednio zaprezentowane i dyskutowane w zaprezentowanym autoreferacie. Szczególnie niepokojąca jest konkluzja z badania 1, dotycząca pozytywnego wpływu treningu przy braku odpowiedniej grupy kontrolnej, która podważa podstawowe według mnie zasady metodologii i wnioskowania. Otwarte opisanie braku tych efektów

nie byłoby uznane za zarzut wobec pracy ani nie podważało innych znalezionych efektów, a wręcz byłoby jej zaletą (i powinno zostać przedyskutowane w świetle wcześniejszych badań, co podnoszę w kolejnej sekcji recenzji).

Nieuwenhuis, S., Forstmann, B. U., & Wagenmakers, E. J. (2011). Erroneous analysis of interactions in neuroscience: a problem of significance. Nature neuroscience, 14(9), 1105-1107.

Wstęp teoretyczny i dyskusja wyników

Uważam, że zaprezentowany autoreferat dobrze opisuje, jak projekt został przeprowadzony od strony metodologicznej i technicznej. Autorka dobrze motywuje dlaczego warto badać wpływ gier komputerowych i dużo miejsca poświęca na opis gry SC II, różnych modeli rozgrywki, oraz możliwych do analizy parametrów. Brakuje mi natomiast rozwinięcie aspektów teoretycznych oraz zintegrowania wyników uzyskanych w trzech badaniach i wskazania jaką wartość dodaną stanowią. Przez to odnoszę wrażenie, że autoreferat jest napisany jakby prezentował badanie aplikacyjne/stosowane, a nie podstawowe.

Ta słabość pracy jest najbardziej widoczna w sekcji dyskusji (s. 16-18), która powinna przedstawić uzyskane wyniki w świetle poprzednich badań. Autorka wskazuje tu raczej na obszary, w których wyniki przeprowadzonych badań mogą mieć znaczenie i znaleźć zastosowanie, nie mówiąc wprost jakie to znaczenie jest. Nie dyskutuje też bardziej szczegółowo uzyskanych efektów (lub ich braku) w kontekście wyników obecnych w literaturze. Uważam, że moja opinia jest potwierdzona przez obserwację, że w całej sekcji dyskusji (s. 16-18) znajdują się cytowania jedynie 3 artykułów naukowych.

Przykładowo, Autorka stwierdza we sekcji "Introduction", że w poprzednich badaniach grupa graczy była mniej podatna na efekt *attentional blink* (AB) niż grupa nie-graczy (s. 4). Jednak brakuje odniesienia do tych prac w sekcji dyskusji i stwierdzenia jak uzyskane wyniki odnoszą się do tych już opublikowanych. Zakładając, że uzyskano brak efektu behawioralnego w zadaniu AB (jak wskazywałem w poprzedniej sekcji recenzji), może to świadczyć o tym, że różnice obserwowane w badaniach przekrojowych jednak nie wynikają z wpływu gier. Autorka stwierdza, że "*Amongst studies showing the positive influence of playing VGs, there are numerous publications showing lack of similar effects*" (s. 16). Jednocześnie badania te nie są omawiane, a uzyskane wyniki nie są dyskutowane w kontekście poprzednich badań. Być może zastosowany w przedstawionych w pracy badaniach trening był za krótki lub zbyt mało intensywny, żeby zaobserwować efekty behawioralne – warto by więc omówić jak długie były treningi w poprzednich badaniach. Ponadto, Autorka podkreśla wagę czynnika złożoności środowiska gry, ale jednocześnie ten aspekt też nie jest dyskutowany bardziej szczegółowo



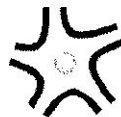
- w których poprzednich badaniach środowisko było bardziej lub mniej złożone i jak przełożyło się to na wyniki?

Z kolei zaobserwowany efekt P3 wskazuje, że jednak trening w złożonym środowisku miał jakiś wpływ na procesy percepcyjne i poznawcze biorące udział w zadaniu AB. Dysocjacja pomiędzy miarami behawioralnymi a neuronalnymi również nie jest bardziej dogłębnie dyskutowana. Idąc dalej, znalezienie efektu P3 w zadaniu AB, ale brak efektu grupy na komponent CDA w zadaniu pamięci roboczej może sugerować, że czasowe aspekty funkcjonowania uwagi (mierzone w zadaniu AB) mogą zmieniać się pod wpływem nawet krótkiego treningu (jaki zastosowano w projekcie), ale już zmiana pojemności pamięci roboczej może wymagać znacznie dłuższego (lub innego typu) treningu. Tego typu interpretacja wydaje się warta rozważenia, jednak to znowu wymagałoby odniesienia wyników do literatury i sprawdzenia czy mogą one w jakiś sposób potwierdzić taką interpretację.

Spójność analizowanych miar pomiędzy badaniami

Autorzy badań uzyskali bardzo bogaty zestaw danych, zawierający wiele wskaźników, które mogą być definiowane i analizowane względem siebie na różne sposoby. W niektórych miejscach brakuje mi jednak spójności analiz i wybieranych miar pomiędzy trzema przedstawionymi publikacjami. Głównie dotyczy to miar telemetrycznych, ponieważ w kolejnych publikacjach odmienne miary są analizowane w odniesieniu do wskaźników behawioralnych i EEG. Nie jest umotywowane, dlaczego te konkretne wskaźniki zostały wybrane. W publikacji 1 analizowana jest liczba meczy rozegranych przez osoby badane na danym poziomie i relacja tego wskaźnika z poprawnością w zadaniu *attentional blink* i amplitudą P3. Z kolei w publikacji 2 analizowano więcej wskaźników telemetrycznych (są zaprezentowane na rycinie 2) i wykazano, że CDA z sesji przed treningiem przewiduje liczbą rozegranych meczy, ale nie wiadomo czy przewiduje też inne wskaźniki dotyczące treningu. Uważam, że nawet w przypadku braku takiej relacji warto raportować modele i wyniki. Najpełniejsza analiza jest wykonana w publikacji 3, gdzie w zgodzie z moimi uwagami dane z zadania *change detection* są analizowane w odniesieniu do całego zestawu wskaźników i także modele pokazujące brak relacji są raportowane.

Dodatkowo, w drugim badaniu CDA z sesji przed-treningiem zostało użyte jako predyktor, ale aktywność została uśredniona tylko z warunków wysokiego obciążenia poznawczego (4 i 5, z wykluczeniem obciążenia 2 i 3), co nie zostało w żaden sposób umotywowane. Z kolei w badaniu trzecim moc oscylacji alfa i theta użyta jako predyktor miar telemetrycznych było utrudnione przez wszystkie warunki obciążenia poznawczego (od 2 do 5).



Obróbka i analiza sygnału EEG

Mam kilka technicznych uwag odnośnie przeprowadzania i raportowania analizy EEG. Jeśli chodzi o wstępną obróbkę (*preprocessing*) danych EEG to ten etap analizy mógł być wykonany w bardziej optymalny sposób. Przykładowo, re-referencja sygnału wykonana przed usunięciem zaszumionych kanałów może prowadzić do rozprzestrzenienia się szumu na wszystkie kanały. Ponadto wykonanie down-samplingu (tutaj do 250 Hz) przed odfiltrowaniem wysokich częstotliwości za pomocą filtra dolno-przepustowego (tzw. *anti-aliasing filter*) może doprowadzić do zjawiska aliasingu.

Brakuje informacji, ile prób (*epoch*) zostało usuniętych (lub ile ostatecznie weszło do analizy), ta informacja powinna być podana w rozbiściu na grupy i warunki.

W pracy nr 1 w metodzie nie ma też informacji z jakich elektrod analizowany był komponent P3. Informacja pojawia się dopiero na Rycinie 4, pokazującej że sygnał był uśredniony z 12 centralnych elektrod (panel C). Jednak nie jest dla mnie jasne, dlaczego na panelu A przedstawiono dane tylko z 1 elektrody Cz. Odnośnie tej ryciny, nie jest też jasne w jaki sposób w panelach A i B zostały oznaczone pojedyncze punkty w czasie wykazujące istotne różnice – w manuskrypcie nie wspomniana jest żadna analiza, która brała pod uwagę poszczególne punkty czasowe (opisana jest jedynie analiza polegająca na uśrednieniu potencjałów wywołanych w oknie czasowym 375-625 ms).

Jeśli chodzi o badanie nr 2, mam podobne zastrzeżenia odnośnie wstępnej obróbki sygnału EEG, ale w tym przypadku elektrody użyte w analizie statystycznej (P7/8) są podane.

W trzeciej pracy algorytm ICA został użyty do wyodrębnienia i usunięcia komponentów reprezentujących artefakty EOG, ale nie ma informacji w jaki sposób te komponenty zostały wybrane (subiektywny wybór, czy automatycznie?) i ile komponentów zostało usuniętych. W przypadku tego badania została natomiast zaraportowana liczba usuniętych epoch (w jednej z dodatkowych tabeli).

Ponadto, w przypadku pracy nr 3 brakuje mi wyjaśnienia w jaki sposób została usunięta aktywność 1/f – podane jest tylko odniesienie do pracy Donoghue i wsp. (2020) prezentującej toolbox *Fooof* (ale nie jest wprost napisane czy został użyty). Stosując ten toolbox w prowadzonych przeze mnie obecnie projektach wiem, że trzeba w nim zdefiniować wiele parametrów wpływających na sposób w jaki dane będą modelowane. W takim przypadku opisanie tego kroku tylko jednym zdaniem, bez opisanie jak dokładnie aktywność 1/f była usuwana jest niewystarczające, aby zreplikować zaprezentowane analizy.

Dodatkowo, moją uwagę zwróciło stwierdzenie “P3 amplitude power” (np. na stronie 12), które nie jest poprawne (analizowana była amplituda komponentu P3, nie moc amplitudy).



Podsumowanie

Praca doktorska Pani mgr Natalii Jakubowskiej dostarcza wartościowych danych dotyczących potencjału wykorzystywania gier komputerowych jako treningu poznawczego. Za główne zalety pracy uznaję podłużny charakter badania oraz wykorzystanie różnorodnych danych i metod analizy w celu odpowiedzenia na postawione pytania badawcze. Natomiast głównymi mankamentami pracy są według mnie brak klarownego przedstawienia i odpowiedniej interpretacji uzyskanych efektów behawioralnych, oraz brak integracji wyników uzyskanych w toku całej pracy doktorskiej w autoreferacie (brak interpretacji i dyskusji wyników w odniesieniu do literatury).

Uważam, że przedstawiona praca z stanowi "oryginalne rozwiązanie problemu naukowego," w zgodzie z art. 187. ust. 2. ustawy z dnia 20 lipca 2018 r. „Prawo o szkolnictwie wyższym i nauce”. W związku z tym wnoszę o dopuszczenie Pani mgr Natalii Jakubowskiej do dalszych etapów postępowania doktorskiego.

Warszawa, 17.02.2023

Signature valid

Digitally signed by
Michał Rola
Date: 2023.02.17
11:51:24 CET