

Poznań, 14 lutego 2023

dr hab. Marek Kowalczyk, prof. UAM
Uniwersytet im. Adama Mickiewicza w Poznaniu
Wydział Psychologii i Kognitywistyki

Recenzja rozprawy doktorskiej mgr. Tomasza Janusa
pt. *Funkcjonowanie poznawcze w warunkach narkozy azotowej. Ocena różnic procesów poznawczych przy ekspozycji na hiperbarię powietrzną w komorze hiperbarycznej i w warunkach podwodnych*

Opiniowana rozprawa doktorska ma formę klasycznej monografii naukowej. Główny tom liczy 271 stron, do tego dochodzi tom z załącznikami liczący 200 stron. Rozprawa zawiera streszczenie (wyłącznie w języku polskim), tekst zasadniczy podzielony na 22 rozdziały oraz spis literatury cytowanej. W pierwszej części rozprawy Autor szeroko przedstawia teoretyczne i empiryczne tło podjętych problemów badawczych. W następnych opisuje kolejne badania własne: przedstawia wysuwane hipotezy, charakteryzuje metodę (uczestników, narzędzia, procedury), opisuje wyniki oraz wysuwa propozycje ich interpretacji.

Główny problem rozprawy i ogólna charakterystyka ujęcia

Dysertacja dotyczy funkcjonowania pamięci trwałej, uwagi i procesów wykonawczych w warunkach narkozy azotowej. Główna idea to porównanie wykonania zadań poznawczych w warunkach normalnego ciśnienia atmosferycznego oraz w dwóch warunkach nadciśnienia w komorze hiperbarycznej: kiedy używane jest w niej sprężone powietrze (czyli może wystąpić narkoza azotowa) oraz kiedy używana jest w niej nienarkotyczna mieszanka helowo-tlenowa (helioks). Użyte zadania to zapamiętywanie i odroczone swobodne odtwarzanie list wyrazów oraz trzy testy wymagające wykrywania bodźców o określonych właściwościach wśród innych w polu wzrokowym i zaznaczania ich bądź łączenia na kartce według podanych reguł. Test kolorowych połączeń (wersje CTT-1 i CTT-2) miał dostarczyć wskaźniki sprawności przeszukiwania pola percepcyjnego oraz procesów zarządczych, a dwa inne zadania według zamierzeń Autora miały pozwolić diagnozować uwagę „intensywną”

i „ekstensywną”. W każdym z tych zadań rejestrowano kilka wskaźników charakteryzujących wykonanie. Analizy statystyczne wykonywano na pojedynczych wskaźnikach: nie podejmowano prób ich grupowania czy integrowania, ani też nie nadawano im specyficznych interpretacji.

Nie testowano hipotez identyfikujących precyzyjnie charakter zmiany wywołanej przez narkozę azotową, osadzonych w jakiejś rozwiniętej teorii wyjaśniającej badane procesy umysłowe. Na przykład, swobodne odtwarzanie można wyjaśniać w modelach rozwijanych w architekturze ACT-R, w modelach z rodziny SAM, czy też w modelach zakładających wydobywanie kontekstu. Tego rodzaju refleksja teoretyczna nie jest jednak w pracy podejmowana. Wysuwane przez Autora (bardzo liczne) hipotezy dotyczą oczekiwanych różnic (albo braku różnic) pomiędzy analizowanymi wskaźnikami w porównywanych warunkach. Niewątpliwie nawiązują one do ustaleń poczynionych we wcześniejszych badaniach (relacjonowanych we wprowadzającej, przeglądowej części pracy), ale nie są bezpośrednio w rozwinięty sposób wyjaśnione. Sformułowane są w czasie przyszłym i dotyczą tego, co jest oczekiwane w badaniu, a nie ogólnych zależności.

Jak zapowiada podtytuł pracy, podjęto też próbę porównania procesów poznawczych w warunkach hiperbarii powietrznej wywołanej w komorze hiperbarycznej i w warunkach podwodnych. Pod wodą wykonywane było tylko jedno zadanie (test kolorowych połączeń).

Dokonując porównań, Autor posługuje się terminami „grupa eksperymentalna” i „grupa kontrolna”, ale niestety nie jest tak, że te porównania obejmują różne warunki w dobrze kontrolowanych eksperymentach. Te porównania mają zawsze charakter międzygrupowy, a przy tym: 1) przydział osób do grup nie miał charakteru losowego; 2) grupy systematycznie różniły się tym, z jakiego środowiska pochodzili uczestnicy; 3) grupy różniły się niekiedy istotnie pod względem wieku uczestników i proporcji kobiet i mężczyzn, 4) pomiędzy grupami występowały różnice w procedurach i cechach sytuacji innych niż te, ze względu na które grupy porównywano. Nie mamy tu więc do czynienia z klasycznym eksperymentem czy eksperymentami, ale z porównaniami obejmującymi wyniki różnych badań. W takich porównaniach jednoznaczne stwierdzenie, że jakaś zmienna wpływa na zmienną zależną albo że tego wpływu nie ma, nie jest możliwe. Autor niekiedy określa te poszczególne badania mianem eksperymentów, ale eksperymentami one nie są. Warunki kontrolne dla „eksperymentu” zawsze znajdują się tu poza nim lub też „eksperyment” kreuje warunki kontrolne dla innego badania.

W dalszej części opinii skomentuję kolejno wprowadzającą, przeglądową część rozprawy, wątek dotyczący pamięci w badaniach Autora oraz analizy dotyczące

funkcjonowania uwagi, przeszukiwania wzrokowego oraz procesów zarządczych (również w warunkach podwodnych). Ocenę także formalną stronę pracy.

Wprowadzająca i przeglądowa część pracy

Autor bardzo szeroko wprowadza czytelnika w problematykę nurkowania i narkozy azotowej, a także problemów, jakie wiążą się z badaniem wpływu tego zjawiska na funkcjonowanie człowieka. To wprowadzenie jest jeszcze rozszerzone przez materiał zawarty w Załącznikach. Widać w tych częściach rozprawy znanstwo i niekłamane zainteresowanie tematem. Autor dokonuje też przeglądu literatury dotyczącej procesów poznawczych w warunkach narkozy azotowej, a przedtem w perspektywie psychologii ogólnej omawia kategorie procesów poznawczych, których dotyczyły jego badania, oraz przedstawia czynniki zakłócające te procesy.

Przegląd pojęć, teorii i badań w psychologii ogólnej, które dotyczą procesów umysłowych uwzględnionych w badaniach Autora, w mojej ocenie świadczy o dość powierzchownej orientacji – raczej na poziomie podręcznika niż specjalistycznej pracy poświęconej wybranym procesom poznawczym. W niektórych miejscach razi w tym przeglądzie brak precyzji; zdarza się także powielanie poważnych nieporozumień, które niestety pokutują w literaturze. Jaskrawym przykładem jednego i drugiego jest to, co Autor pisze o efektach pierwszeństwa i świeżości (w rozprawie stają się one jednym „efektem pierwszeństwa i świeżości”, s. 38).

Efekty pierwszeństwa i świeżości to są dwa różne efekty, na których wielkość wpływają odmienne czynniki. Nieprecyzyjne jest stwierdzenie, że badani „zapamiętują dużą liczbę słów z początku i końca listy, przy jednocześnie niewielkiej liczbie słów ze środka listy” (s. 38) – obydwie efekty dotyczą stosunkowo niewielkiej liczby wyrazów z początku i z końca listy. Kiedy lista się wydłuża, to rośnie bezwzględna liczba (nie odsetek) zapamiętanych wyrazów ze środka listy. Odtworzenie bezpośrednie nie jest warunkiem wystąpienia efektu pierwszeństwa (w pewnych szczególnych okolicznościach występuje również długotrwały efekt świeżości). Wyjaśnienie przywołane przez Autora nie tłumaczy tego, że większe szanse odtworzenia i rozpoznania mają wyrazy z początku niż ze środka listy – w modelu wielomagazynowym, do którego Autor się najwyraźniej odwołuje, jedno i drugie wydobywane są z pamięci trwałej. Nie jest tak, że sekwencyjna prezentacja wyrazów w tempie jeden na dwie sekundy uniemożliwia powtórki wewnętrzne (umożliwia – poziom odtworzenia jest skorelowany z ich liczbą). Nie jest też tak, że znajomość długości listy znosi

efekty pierwszeństwa i świeżości (choć może wywierać pewien wpływ na kształt krzywej kolejnych pozycji – zob. Bernbach, 1975)¹.

W przeglądzie nie ma mowy o innych niż efekty pierwszeństwa i świeżości ważnych zjawiskach czy prawidłowościach występujących w swobodnym odtwarzaniu, np. efektach sąsiedztwa i bliskości czy warunkach, w jakich zwykle pojawiają się intruzje. Można by omówienia tych zagadnień oczekiwać, gdyż zapamiętywanie i swobodne odtwarzanie to ważne zadanie wykorzystywane w badaniach Autora, a niektóre testowane hipotezy dotyczą właśnie przywołań wyrazów, których nie było na liście (intruzji). Brakuje też wyjaśnienia, co dzieje się w swobodnym odtwarzaniu w świetle choćby jednego rozwiniętego modelu pamięci (np. ACT-R czy SAM) i jakie są mechanizmy czy uwarunkowania niepowodzeń w przypominaniu sobie elementów zapisanych w pamięci.

Podobnie przegląd idei teoretycznych i ustaleń dotyczących przeszukiwania pola percepcyjnego jest niesystematyczny i poważnie niekompletny. Nie uwzględnia ważnych generalizacji behawioralnych w omawianym obszarze, nie odwołuje się do żadnej rozwiniętej współczesnej teorii przeszukiwania (jak np. teoria kierowanego przeszukiwania Wolfe'a). Spojrzenie przez pryzmat takiej teorii na zadania dotyczące uwagi wykorzystywane przez Autora mogłoby pomóc uchwycić ważne wymiary podobieństw i różnic między nimi. Autor uwzględnia w swoich badaniach opozycję uwaga intensywna – uwaga ekstensywna, sięgając po koncepcję Kolańczyk (1991), ale ignoruje współczesną wiedzę o możliwościach i ograniczeniach równoległego przetwarzania i identyfikowania bodźców oraz regulowania zakresu ogniska uwagi.

Pisząc o funkcjach zarządczych, Autor przywołuje ważny artykuł Miyakego i współpracowników (2000), ale pomija późniejszy dorobek tego zespołu. Nowsze prace Miyakego i Friedman dotyczące problemu homogeniczności vs. heterogeniczności funkcji zarządczych kreują nową perspektywę w myśleniu o relacjach pomiędzy zadaniami mierzącymi te funkcje i o tym, co wyniki w tych zadaniach wyrażają.

Odnotowuję tu wiele luk i luki w tych tylko fragmentach przeglądu, które dotyczą procesów umysłowych będących przedmiotem badawczego zainteresowania Autora. Poglębiając namysł nad tymi procesami, Autor mógłby zapewne wypracować bogatsze i mocniej teoretycznie ugruntowane hipotezy. Być może bardziej świadomie podejmowałby decyzje odnośnie do tego, jak zaaranżować sytuacje w badaniach czy jakich użyć zadań, żeby

¹ Polecam Autorowi świetną książkę Jacoba Kahany *Foundations of human memory*, Oxford University Press (2012).

zrealizować swoje poznawcze cele. Mógłby wydobyć więcej informacji z analiz materiału zdobywanego w tak trudnych organizacyjnie badaniach. Mógłby wreszcie uzyskać lepszy wgląd w znaczenie otrzymanych wyników. Zawarty w rozprawie przegląd literatury dotyczącej procesów poznawczych eksplorowanych w badaniach Autora zapowiada (niestety trafnie) dość płytki poziom analizy.

Zadanie pamięciowe

Uczestnicy zapamiętywali wyrazy z listy, wykonywali zadanie dystrykcyjne, odtwarzali zapamiętane wyrazy, wykonywali dwa zadania wymagające wyszukiwania wzrokowego, zapamiętywali następną listę wyrazów i po zadaniu dystrykcyjnym ją odtwarzali. Lista obejmowała 30 wyrazów wydrukowanych na kartce: 10 o pozytywnym wydźwięku emocjonalnym, 10 o konotacji neutralnej i 10 o negatywnej. Uczenie się każdorazowo trwało 75 sekund i tyle samo czasu trwało odtwarzanie polegające na wypisywaniu wyrazów na kartce.

Te zadania wykonywano w jednym z trzech warunków: w sali komputerowej na uczelni, w komorze hiperbarycznej, w której następowało sprężanie powietrza do 5 ATA, a następnie dekompresja, oraz w komorze hiperbarycznej, w której następowało sprężanie powietrza do 5 ATA i dekompresja, ale uczestnicy oddychali mieszkanką helowo-tlenową przez szczelne maski twarzowe. W komorze hiperbarycznej zapamiętywanie pierwszej listy następowało w normobarii, a odtwarzanie w warunkach 5 ATA, natomiast zapamiętywanie drugiej listy następowało w warunkach 5 ATA, a odtwarzanie – po dekompresji (w normobarii). W każdym z tych trzech ogólnych warunków były dwie grupy różniące się kolejnością list (AB albo BA), a w badaniach w komorze hiperbarycznej – również zadaniami rozdzielającymi dwa cykle: zapominanie – odtwarzanie. Grupy wykonujące zadania na uczelni Autor nazywa normobarycznymi grupami kontrolnymi. Grupy wykonujące zadania w komorze hiperbarycznej będę nazywał „eksperymentalnymi”, w zależności od użytego czynnika oddechowego „powietrznymi” albo „helioksowymi”.

Grupy kontrolne różniły się od eksperymentalnych proporcją kobiet i mężczyzn, a także wiekiem uczestników: w grupach kontrolnych uczestnicy byli młodszy i było wśród nich zdecydowanie więcej kobiet². Obydwa te czynniki mogą przekładać się na przewagę

² W grupach kontrolnych było ponad trzykrotnie więcej osób niż w eksperymentalnych. Autor jednak nie wpadł na pomysł, żeby w analizach uwzględnić ich podgrupę najlepiej dopasowaną do grup eksperymentalnych pod względem liczebności, wieku uczestników oraz proporcji kobiet.

grup kontrolnych nad eksperymentalnymi w wykonaniu zadania pamięciowego. W grupach kontrolnych zadanie dystrykcyjne między zapamiętywaniem a odtwarzaniem trwało każdorazowo 8 minut. W eksperymentalnych odstęp pomiędzy zapamiętywaniem pierwszej listy a odtwarzaniem również wynosił 8 minut, ale odstęp pomiędzy zapamiętywaniem drugiej listy a odtwarzaniem – 11 minut i 40 sekund (wynikało to z wymogów procedury dekompresji). W grupach kontrolnych zadaniem dystrykcyjnym była „gra komputerowa, która angażowała proste funkcje motoryczne” (s. 124), a w eksperymentalnych – „obserwacja i analiza wskaźnika pokazującego ciśnienie i temperaturę w komorze, a w trakcie kompresji – badani musieli także regularnie wyrównywać ciśnienie w uchu środkowym” (s. 137).

W komorze hiperbarycznej odtwarzanie następowało w innych warunkach ciśnienia niż zapamiętywanie, a do tego warunki odtwarzania w drugim cyklu były analogiczne (jeśli chodzi o samo ciśnienie) do warunków zapamiętywania w pierwszym. Na dodatek zadanie dystrykcyjne kierowało uwagę na wskaźnik zmiany ciśnienia! W tych okolicznościach można by przewidywać gorsze odtwarzanie w grupach eksperymentalnych za sprawą zmiany wewnętrznego i zewnętrznego kontekstu przy przejściu od fazy zapamiętywania do odtwarzania, a w drugim cyklu być może również za sprawą nasilonej interferencji asocjacyjnej. W grupach kontrolnych kontekst zapamiętywania i odtwarzania list nie był systematycznie zróżnicowany.

Warunki kontrolne od helioksowych różniło też to, że w tych pierwszych eksperymentator przebywał w trakcie badania z uczestnikami – w grupach helioksowych musiał pozostawać poza komorą hiperbaryczną (zob. niżej).

Grupy powietrzne i helioksove nie różniły się istotnie pod względem wieku uczestników, ale przydział do nich nie miał charakteru losowego. W grupach powietrznych byli pływonurkowie (w tym dwie kobiety), którzy – poza jedną osobą (instruktorem) – znaleźli się w komorze hiperbarycznej po raz pierwszy. W grupach helioksowych znajdowali się żołnierze sił specjalnych (mężczyźni), mający za sobą liczne sesje przebywania w nadciśnieniu w tej komorze (w jednej grupie komandosów średnia szacowanej liczby sprężeń wynosiła 36,87, a w drugiej 16,66).

W grupach powietrznych eksperymentator przebywał w trakcie badania z uczestnikami. Przed każdym zadaniem upewniał się, czy zrozumieli polecenia. W grupach helioksowych eksperymentator nie przebywał w trakcie sprężania w komorze. „Uczestnicy eksperymentu zostali dodatkowo wyposażeni w klarowne plany badawcze, które miały na celu przypomnienie i utrwalenie sposobu wykonania testów (badani zapoznawali się z nimi w normobarii, a także w trakcie presuryzacji jako dodatkowe zadanie dystrykcyjne)” (s. 176).

Jeśli dobrze zrozumiałem, uczestnicy nie mogli liczyć na pomoc badacza przy przechodzeniu do kolejnych zadań, a w każdym razie była ona mniejsza niż w grupach powietrznych (niestety opis procedury pod tym względem nie jest do końca jasny). W zamian mieli wspomagać się dostarczonym planem. Nie jest to niuans, ponieważ oznacza to, że ich procesy zarządcze były obciążone bardziej niż procesy zarządcze uczestników w innych grupach. Wreszcie – uczestnicy z grupy helioksovej wykonywali zadania w maskach, poprzez które podawano mieszanke helowo-tlenową³. Można sobie tylko wyobrazić, że komfort spostrzegania i działania w takich warunkach jest niższy.

Ani w głównym tomie pracy, ani w „Załącznikach”, nie znalazłem dokładnej informacji, jak byli rekrutowani uczestnicy. Można jednak podejrzewać, że te procedury były inne w wypadku pletwonurków cywilnych i żołnierzy. Nasuwa się pytanie, na jakiej zasadzie jedni i drudzy uczestniczyli w badaniach i jakie w związku z nim otrzymywali gratyfikacje. Na przykład, czy żołnierze uczestniczyli w badaniach w godzinach pracy bądź służby czy w czasie wolnym? Zgodnie z tym, co pisze Autor, uczestników informowano, że wyniki badań są poufne (dostępne tylko dla eksperymentatora i promotora) i anonimowe (s. 175), ale jak się dowiadujemy, nagranie wideo pozostało w jednostce wojskowej: „ze względu na ochronę danych osobowych żołnierzy, autor (...) nie otrzymał nagrania z eksperymentu” (s. 176). Czy żołnierze wiedzieli, że filmowa dokumentacja ich zachowania w komorze pozostaje w jednostce? Nie zadaję tych pytań, żeby zacząć rozważać etyczne aspekty sprawy, ale żeby wskazać kolejny potencjalnie ważny czynnik, który mógł różnicować porównywane grupy: rodzaj i siła wzbudzonej motywacji.

Wobec wszystkich tych różnic między grupami powietrznymi i helioksowymi, interpretowanie różnic (czy też braku różnic) między nimi w wykonaniu zadań w kategoriach wpływu (bądź braku wpływu) narkozy azotowej na procesy umysłowe może mieć co najwyżej charakter domysłów.

Analizy statystyczne, które raportuje Autor, mają charakter prostych międzygrupowych porównań średnich w parach, przy czym w każdym ogólnym warunku (grupa kontrolna, powietrzna i helioksova) mamy dwie podgrupy różniące się kolejnością list. Autor dokonuje porównań osobno dla każdego wariantu kolejności list oraz dla

³ Informacja o tym pojawia się nie w opisie procedury, a (w przypisie!) w dyskusji. Autor rozprawy nie rozważa tego, w jakim stopniu oddychanie i patrzenie przez maskę wpływa na wykonywanie zadań, w szczególności zadań wzrokowo-motorycznych. A komponent wzrokowy i motoryczny miały wszystkie zadania wykonywane przez uczestników.

pierwszego i drugiego cyklu zapamiętywanie – odtwarzanie, co daje w sumie 12 porównań. Ponadto powieliła te analizy z rozbiem na trzy kategorie wyrazów. Nie towarzyszy temu refleksja i korekta związana z liczbą wykonywanych porównań, a także analiza konsekwencji tego, że porównywane grupy są bardzo małe bądź grupy małe porównywane są ze znacznie liczeźniejszymi. Brakuje analizy trafności założenia o normalności rozkładów i równości wariancji oraz sprawdzenia, czy występują wartości odstające⁴.

Rezultaty atomistycznych analiz nie układają się w spójny wzór. W formalnie równoważnych warunkach występowały rozbieżne wyniki: na przykład średnie liczby zapamiętanych wyrazów w grupie powietrznej i helioksovej w warunkach zapamiętywania w hiperbarii – odtwarzanie w normobarii były istotnie zróżnicowane w jednym wariancie kolejności list, natomiast różnice pomiędzy nimi były nieistotne (a porządek średnich odwrócony) w drugim wariancie. Analogiczne niespójności wystąpiły w analizach dotyczących sekwencji: zapamiętywanie w normobarii – odtwarzanie w hiperbarii. Jaki wniosek wyciąga Autor?

„(...) mając na uwadze dane pozyskane w trakcie Badań 2b i 3b (hiperbaryczny helioks vs. hiperbaryczne powietrze) stwierdzono zarówno gorsze przypominanie materiału w warunkach narkozy azotowej, który został zapamiętany w normobarii, jak i istotne upośledzenie przypominania słów w warunkach normalnych, które zostały zapamiętane w hiperbarii powietrznej” (s. 206).

W Streszczeniu zaś czytamy: „W toku badań potwierdzono wcześniej uzyskane wyniki o podatności procesu kodowania na zaburzenia spowodowane narkozą azotową” (s. 14).

Takie wnioski nie są uprawnione wobec niespójnych wyników testów. Autor nie relacjonuje analiz wieloczynnikowych, które mogłyby być wykonane na łączonych wynikach grup. Ale nawet gdyby takie analizy pokazały odpowiednie zróżnicowania średnich między warunkami, to nieuprawniona byłaby konkluzja, że przyczyną jest narkoza azotowa. Porównania obejmowały grupy znacznie różniące się pod wieloma względami, a nie tylko składem czynnika oddechowego.

⁴ Gwoli kompletności obrazu trzeba dodać, że wszelkie istotne zależności dotyczące zmiennych zależnych znajdują szczegółową ilustrację w postaci wykresów zawierających wyniki poszczególnych osób (wraz z wartościami numerycznymi w wypadku liczb zapamiętanych wyrazów i liczb błędów) w Załącznikach. Tak więc czytelnik może dokonać oceny na podstawie wykresu lub nawet policzyć wartości statystyk.

Hipotezy dotyczące zróżnicowanego wpływu hiperbarii powietrznej na zapamiętywanie wyrazów różnych kategorii afektywnych wymagałyby testowania efektów interakcji warunków i kategorii wyrazów, ale Autor takich analiz nie przeprowadza. Tym niemniej komunikuje: „zaobserwowano szczególną podatność słów neutralnych na zaburzenia spowodowane narkozą” (s. 14). „Zaobserwowano” w nieformalnym oglądzie konfiguracji „istotnych” i nieistotnych różnic w 36 niezależnych porównaniach międzygrupowych („istotnych” w cudzysłowie, bo nieskorygowanych ze względu na liczbę porównań). Trudno dociec, dlaczego nie wykonano analiz odpowiednich do testowanej hipotezy.

Zadania przeszukiwania wzrokowego

Nie zrównoważenie warunków obejmowanych porównaniami pod względem wartości zmiennych ubocznych nie pozwala też oczywiście na konkluzywną ocenę wpływu narkozy azotowej na wykonywanie zadań wymagających poszukiwania określonych obiektów wśród innych w polu wzrokowym. Autor analizuje wyniki tych zadań podobnie jak wyniki zadań pamięciowych, porównując średnie w parach grup osobno dla każdego wskaźnika i nieuwzględniając w interpretacjach liczby dokonanych porównań. Rozważa więc psychologiczne znaczenie „istotnych” różnic nawet w tych wypadkach, w których należałoby uznać, że nie ma podstaw do odrzucenia hipotezy zerowej. Inny problem nasilony przez osobne porównania dla każdej kategorii błędów wiąże się z charakterem wchodzących w grę rozkładów, które jaskrawo odbiegają od kryteriów normalności – zawierają bardzo dużo zer i nieliczne wartości pozytywne. Do analizy takich wyników potrzebne są specjalne narzędzia statystyczne. ANOVA w tym wypadku nie jest rozwiązaniem poprawnym.

W wypadku dwóch zadań nie było żadnych istotnych różnic międzygrupowych, natomiast „istotne” różnice znaleziono dla niektórych wskaźników poprawności wykonania w dwóch wersjach testu kolorowych połączeń. W obu wersjach łącznie można popełnić błędy ośmiu kategorii. W porównaniach liczb błędów poszczególnych kategorii w grupie powietrznej i kontrolnej znaleziono trzy „istotne” różnice, a w porównaniach dotyczących grupy helioksowej i kontrolnej – dwie. Jeśli za pomocą poprawki Bonferroniego skorygować p ze względu na liczbę dokonanych porównań, to liczby istotnych różnic zmniejszają się odpowiednio do dwóch i jednej. Co szczególnie ważne, nie było żadnych istotnych różnic w liczbie błędów poszczególnych kategorii pomiędzy grupą powietrzną i helioksową. Tym niemniej Autor obwieszcza: „wykazano, że wszystkie funkcje wykonawcze w ujęciu Miyake i in. (2000) mogą być podatne na zaburzenia spowodowane narkozą, w dopuszczalnych

limitach głębokościowych nurkowań powietrznych (tj. do 5 ATA)”. Ze Streszczenia zaś dowiadujemy się, iż „potwierdzono, że wszystkie funkcje wykonawcze są podatne na wpływ narkozy” (s. 14).

Już samo to, że grupy nie były zrównoważone pod względem składu uczestników i różniły się procedurami, nakazywałoby formułować konkluzje bardziej ostrożnie. Ale nawet gdyby tego fundamentalnego problemu nie było, to 1) W badaniu nie było wskaźników powiązanych selektywnie z poszczególnymi funkcjami wyróżnionymi przez Miyakego i in. (2000). Liczby błędów różnego typu w teście kolorowych połączeń nie są takimi wskaźnikami. 2) W grupie powietrznej nie było więcej błędów niż w grupie helioksowej, wyrównanej z tą pierwszą pod względem warunków ciśnienia, a różniące się tym, że mieszanka miała skład nienarkotyczny. Jaką rolę w tych badaniach pełniły grupy helioksowe, jeśli wnioskując o wpływie narkozy azotowej na procesy poznawcze można zignorować to, że pomiędzy grupą helioksową a grupą powietrzną nie było istotnych różnic we wskaźnikach wykonania?

Autor proponuje rozmaite interpretacje tego, co stwierdził w swoich badaniach, odwołując się do różnych założeń i zmiennych, których wyjściowo projekt nie uwzględniał – na przykład wskazuje na prawdopodobną rolę lęku czy stresu wzbudzanego w określonych warunkach, a nawet przywołuje ad hoc specjalne rodzaje stresu (s. 195). Czerwona linia jest już jednak zdecydowanie przekraczana, kiedy spekulacja luźno powiązana z uzyskanymi wynikami pojawia się w formie wyprowadzonego z nich wniosku. Na przykład Autor konstatuje: „w toku badań stwierdzono, że stres (wraz z trudnymi warunkami środowiskowymi) spowodowany ekspozycją hiperbaryczną z użyciem mieszaniny helowo-tlenowej (przy 5 ATA) może negatywnie wpływać zarówno na procesy pamięciowe, jak i na funkcje wykonawcze” (s. 243). Poziom lęku czy stresu u uczestników nie był w tych badaniach kontrolowany!⁵.

Test kolorowych połączeń w warunkach podwodnych

Uczestnicy jednego badania wykonywali test kolorowych połączeń po zanurzeniu na głębokość 30 m (4 ATA) z użyciem powietrza jako czynnika oddechowego. Wskaźniki

⁵ Nawiasem mówiąc, ta propozycja wyjaśnienia wyników badań budzi zdziwienie. W grupie helowo-tlenowej byli wyłącznie pletwonurkowie – żołnierze sił specjalnych, mający doświadczenie w „suchych” nurkowaniach. Wydawałoby się, że w ich zawodzie odporność na stres jest jedną z podstawowych kwalifikacji.

wykonania porównano ze wskaźnikami wykonania w grupie powietrznej w suchej hiperbarii. Grupy były zrównoważone pod względem wieku, proporcji płci i stażu nurkowego uczestników. Jednakże w warunkach podwodnych badanie przeprowadzano przy ciśnieniu niższym niż w komorze hiperbarycznej (5 ATA). W warunkach podwodnych było to jedyne zadanie wykonywane przez uczestników – w komorze hiperbarycznej zadanie wykonywano w trakcie dłuższej procedury, w przerwie pomiędzy jednym i drugim cyklem zapamiętywanie – odtwarzanie. Ponadto badanie w suchej hiperbarii miało charakter grupowy, odbywało się dla wszystkich uczestników w tym samym miejscu i czasie, natomiast w warunkach podwodnych odbywało się w różnych akwenach, zapewne w odmiennych kontekstach społecznych (grupowe wyjazdy na nurkowanie) i prawdopodobnie miało charakter indywidualny (opis procedury nie jest pod tym względem jednoznaczny). Ważne mogą być także inne różnice: jakość oświetlenia w obu warunkach (płetwonurkowie pod wodą używali latarek, komora hiperbaryczna przypuszczalnie była nieźle oświetlona), zróżnicowanie łatwości precyzyjnego prowadzenia linii w wodzie i w suchej hiperbarii, a także różnice w kosztach utrzymywania stabilnej pozycji ciała w obu środowiskach (w warunkach utrudniających utrzymanie równowagi następuje pogorszenie wykonywania zadań poznawczych).

Autor stwierdził, że w grupie podwodnej popełniono istotnie mniej błędów jednej kategorii (tzw. prawie błędów) niż w grupie komorowej. Ten nieoczekiwany wynik (hipoteza przewidywała odwrotny kierunek zależności dla błędów) Autor rozważa w kategoriach przetargu szybkości i poprawności, inaczej rozstrzyganego pod wodą niż w suchej hiperbarii. W mojej opinii są to rozważania o znaczeniu różnicy w sytuacji braku podstaw do odrzucenia hipotezy zerowej. Wykonano osiem porównań, więc konieczna jest odpowiednia korekta p . A w tym wypadku nieskorygowane p wynosi 0,038 (wariancje w porównywanych grupach nie spełniają warunku równości). Nie trzeba tej wartości mnożyć przez osiem, żeby znaleźć się powyżej konwencjonalnego progu istotności.

Ciekawe, że w próbach wyjaśnienia przez Autora tej rzekomo istotnej różnicy nie pojawia się narkoza azotowa. Porównywane warunki różniły się przecież ciśnieniem azotu, więc jego niższy poziom w warunkach podwodnych mógłby tłumaczyć lepsze wykonanie.

Sensowność tego badania wydaje się bardzo wątpliwa. Porównywane grupy różniły się warunkami ciśnienia, a zarazem innymi zewnętrznymi warunkami wykonywania zadań. Co mają nam powiedzieć stwierdzane różnice albo ich brak w wykonaniu pomiędzy tymi warunkami? Autor utrzymuje, że „uzyskane wyniki sugerują istnienie różnic w funkcjonowaniu uwagi i działaniu funkcji wykonawczych w zależności od tego, czy

ekspozycja na nadciśnienie ma miejsce w warunkach komory hiperbarycznej czy w środowisku podwodnym” (s. 14). Jest dla mnie prawdziwą zagadką, co takiego w przeprowadzonych analizach wskazywałoby na istnienie tych różnic. Oczywiście nie dłuższy czas wykonywania zadań w wodzie. A więc różnica w częstości błędów jednej z ośmiu niezależnie analizowanych kategorii uznana przez Autora (niezasadnie) za istotną statystycznie?

Ponownie, na podstawie nie do końca poprawnych analiz statystycznych Autor formułuje wnioski, które nie mają uzasadnienia w otrzymanych wynikach. Towarzyszą temu różne interpretacje lub spekulacje, które jednak w niewielkim stopniu wiążą z tym, co rzeczywiście przyniosło wykonane badanie.

W rozprawie pojawiają się uwagi świadczące o dostrzeganiu przez Autora różnych metodologicznych niedostatków tego przedsięwzięcia badawczego i związanych z nimi problemów interpretacyjnych, a także propozycje przezwyciężenia tych niedostatków w przyszłości. Dobrze to świadczy o Autorze jako badaczu, który nie jest bezkrytyczny, ale nie zmienia ograniczonej poznawczej wartości wykonanych badań. W kontekście tych uwag tym bardziej dziwią wnioski wyprowadzane przez Autora na podstawie wyników. Lektura konkluzji zebranych w Streszczeniu i w Zakończeniu w rozprawie wprowadzi w błąd czytelnika nieznanego badań.

Formalna i językowa strona rozprawy

Praca jest napisana komunikatywnym językiem i zasadniczo starannie przygotowana pod względem formalnym. Gorzej z poprawnością językową tekstu, w którym jest sporo różnego rodzaju błędów. Praca mogłaby być lepsza, gdyby wyeliminować z niej częste powtórzenia, wątki niewiążące się ściśle z podejmowanym problemem, a także konstatacje banalne, niewnoszące istotnej treści. Przede wszystkim jednak rozprawa puchnie w konsekwencji rozbicia prezentacji wyników na segmenty: normobaria – hiperbaria powietrzna, normobaria – hiperbaria helioksowa oraz hiperbaria powietrzna – hiperbaria helioksowa, co wymusza liczne powtórzenia. O wiele bardziej ekonomiczne byłoby umieszczenie w tabelach obok siebie korespondujących ze sobą informacji dotyczących trzech ogólnych warunków. Taka prezentacja również ułatwiałaby czytelnikowi całościowy ogląd i wyrobienie sobie poglądu na temat krytycznych relacji pomiędzy porównywanymi wartościami zmiennych. Podobnie, można by zredukować liczbę hipotez poprzez zredagowanie ich w bardziej ekonomiczny, a zarazem bardziej klarowny sposób.

Lekturę pracy utrudnia stosowana przez Autora niestandardowa terminologia. W dokonywanych porównaniach posługuje się on terminami „grupa eksperymentalna” i „grupa kontrolna” (co sugeruje, że mamy do czynienia z eksperymentem), ale w stosunku do tego, co w eksperymencie normalnie nazwalibyśmy warunkiem bądź grupą, stosuje nomenklaturę „badanie” albo „eksperyment”, mimo że wysunięte hipotezy nie obejmują porównań wewnątrz „badania”. Badania wykonane na uczelni zostały określone jako „pretesty”, co sugeruje analizy podłużne, ale w analizach nie było żadnego posttestu, a odpowiednie grupy/warunki zostały potraktowane jako pełnoprawne grupy kontrolne w porównaniach międzygrupowych. Termin „pretest” pojawia się jeszcze raz – w opisie badania 4, ale tutaj już ewidentnie chodzi o pilotaż (uważam, że lepiej byłoby zamieścić opis tego badania w załącznikach, można by też je w ogóle pominąć).

Sprawozdanie z wykonanych badań w większej części rozprawy jest w czasie przeszłym, ale gdzieś pojawia się czas przyszły, tak jakby przedstawiano projekt przyszłych badań. Niektóre tabele mają niezbyt trafne tytuły. Na przykład tytuł tabeli 12 (i jedenastu innych) sugeruje porównania wewnątrz osób między kategoriami wyrazów, podczas gdy tabela przedstawia porównania między grupami z uwzględnieniem kategorii wyrazów. Wyrażenie „liczba różnych międzygrupowo błędów” (tabele 20 i 40) mylnie sugeruje, że chodzi o błędy, które pojawiły się w jednej grupie albo w drugiej, a nie w obu (choć i w tej interpretacji nie spodoba się miłośnikowi języka polskiego). Problem braku precyzji w pisaniu pojawia się w wielu innych miejscach rozprawy.

Właściwszym odwołaniem do tabeli w tekście jest przywołanie jej numeru niż użycie frazy „poniższa tabela”. Razi użycie w całej pracy dywizu w roli myślnika.

Konkluzja

Jestem zobowiązany do oceny, czy rozprawa spełnia wymogi, które określa ustawa z dnia 20 lipca 2018 r. Prawo o szkolnictwie wyższym i nauce:

1. Rozprawa prezentuje ogólną wiedzę teoretyczną kandydata w dyscyplinie (...) oraz umiejętność samodzielnego prowadzenia pracy naukowej.
2. Przedmiotem rozprawy doktorskiej jest oryginalne rozwiązanie problemu naukowego (...).

Uważam, że recenzowana rozprawa nie spełnia tych kryteriów w wystarczającym stopniu.

Badania wykonane przez Autora – choć potencjalnie nowatorskie, potrzebne, w zamyśle wypełniające luki w dotychczasowej wiedzy – mają niedostatki metodologiczne, które nie pozwalają uznać ich poznawczego rezultatu za „oryginalne rozwiązanie problemu naukowego”. Eksperymenty w psychologii rzadko pozwalają na wyprowadzanie mocnych konkluzji, ale w wypadku badań w opiniowanej rozprawie niekonkluzywność wynika z zasadniczego naruszenia reguł metodologicznych. Trzeba dodać, że ich zachowanie w warunkach ogromnych ograniczeń narzucanych przez przedmiot badań podjętych przez Autora jest znacznie trudniejsze niż w innych obszarach eksploracji w psychologii.

Mniej oczywista jest ocena rozprawy według pierwszego ustawowego kryterium. Uznaję, że rozprawa w dostatecznym stopniu „prezentuje ogólną wiedzę teoretyczną kandydata” w dyscyplinie psychologia, choć oczekiwałbym bardziej specjalistycznej wiedzy o tych procesach, które były przedmiotem badań Autora, i większych kompetencji w zakresie analiz statystycznych. Skoro w badaniach i analizach ich wyników Autor popełnił poważne błędy metodologiczne, a z uzyskanych rezultatów wyprowadził nieuprawnione wnioski, to nie można uznać, że wykazał się umiejętnością samodzielnego prowadzenia pracy naukowej. Zarazem jednak wykazał się specjalistyczną wiedzą i rzadkimi kompetencjami, które są niezbędne, by prowadzić badania nad funkcjonowaniem ludzi w specjalnych warunkach hiperbarii. Uzupełnienie tych kwalifikacji odpowiednio wysokimi kompetencjami psychologa-eksperymentatora mogłoby zaowocować bardzo wartościowymi badaniami.

Moja ocena rozprawy doktorskiej p. mgr. Tomasza Janusa jest negatywna, zatem nie wnoszę o dopuszczenie go do dalszych etapów postępowania w przewodzie doktorskim.

M. Kowalczyk