



UNIwersytet
Warszawski

Wydział Psychologii



Warszawa, 02.12.2023

Dr hab. Maciej Stolarski, prof. ucz.

Uniwersytet Warszawski

Wydział Psychologii

**Recenzja rozprawy doktorskiej mgr Justyny Józefowicz
pt. „*The role of player's flow in the complex skill learning*”,
napisanej pod kierunkiem dr hab. Anety Brzezickiej, prof. Uniwersytetu SWPS,
oraz prof. dr. hab. Adama Wierzbickiego**

Rozprawa doktorska mgr Justyny Józefowicz dotyczy zagadnienia interesującego teoretycznie i istotnego praktycznie: Doktorantka podjęła bowiem próbę pogłębionej empirycznej analizy roli stanu *flow* w procesie nabywania nowych, złożonych umiejętności, na przykładzie popularnej gry strategicznej czasu rzeczywistego – Starcraft II. Mgr Józefowicz zrealizowała dwa badania empiryczne, z których pierwsze ma wyrażnie charakter instrumentalny względem drugiego – jego celem było bowiem stworzenie polskich wersji językowych dwóch kwestionariuszy: Skali Dyspozycyjnego Przepływu-2 (ang., *Dispositional Flow Scale-2*) oraz Skali Stanu Przepływu-2 (ang. *Flow State Scale-2*). Narzędzia te zostały następnie z powodzeniem wykorzystane w badaniu drugim, które bez wątpienia należałoby określić mianem właściwego lub głównego, jako że to właśnie w nim poddano weryfikacji empirycznej główne hipotezy sformułowane przez Autorkę rozprawy.

Już na wstępie pragnę zaznaczyć, iż bardzo nie lubię najbardziej rozpowszechnionego, polskiego tłumaczenia pojęcia *flow*, czyli „przepływu”. Termin ten w mojej ocenie brzmi sztucznie, a przede wszystkim nie do końca oddaje istotę doświadczenia opisanego przez Csikszentmihalyi’ego. Uważam, że najbliższe oryginału jest słowo *pochłonięcie* (ew. absorpcja), jednak i ono nie oddaje w pełni tego, co niesie ze sobą jego angielski

odpowiednik, podobnie jak spotykane niekiedy w literaturze polskojęzycznej *uskrzydlenie* lub *uniesienie*. Dlatego też ucieszyłem się gdy zobaczyłem, że Doktorantka zdecydowała się na napisanie swojej rozprawy w języku angielskim, co pozwala uniknąć czytelnikowi pracy kilkusetkrotnego spotkania z niefortunnym „przepływem”. Z serca zachęcam Autorkę do rozważenia zastąpienia go w innych pracach i wypowiedziach w języku polskim terminem angielskim, lub jedną z zasugerowanych przeze mnie opcji polskojęzycznych (używa go bowiem w streszczeniu w języku polskim).

Praca ma formę rozprawy o charakterze monograficznym, pisanej w języku angielskim. Warto w tym miejscu zaznaczyć, że rezultaty pierwszego badania zostały częściowo opublikowane w czasopiśmie *Frontiers in Psychology* (IF=3.8; 70 pkt. MEIN). Uważam, że fakt ten zasługuje na podkreślenie i docenienie. Doktorantka sama bowiem wskazuje, iż cel badania pierwszego był dwójaki – po pierwsze, było nim przygotowanie rzetelnych polskich wersji narzędzi pozwalających na zbadanie nasilenia dyspozycji do doświadczania stanu *flow* (DFS-2) oraz sytuacyjnego doświadczenia *flow* (FSS-2), co miało umożliwić przeprowadzenie badań własnych. Jak widać, został on w pełni zrealizowany, czego świadectwem pozostaje recenzowana przeze mnie praca i opisane w niej badanie 2. Drugim celem, wskazanym przez Autorkę, było zapewnienie polskim badaczom dostępu do solidnych narzędzi psychometrycznych, pozwalających na prowadzenie badań nad zjawiskiem *flow*. Z przyjemnością mogę stwierdzić, że ów drugi cel również jest już realizowany, a mój zespół był jego pierwszym beneficjentem – rzeczne skale wykorzystaliśmy już bowiem w badaniach dotyczących doświadczenia *flow* u biegaczy długodystansowych, opublikowanego w tym roku na łamach *Personality and Individual Differences* (Błachowska, Waleriańczyk i Stolarski, 2023).

Pracę rozpoczyna niezwykle bogate wprowadzenie teoretyczne, zakotwiczące zaplanowane przez Autorkę analizy w literaturze przedmiotu. W mojej ocenie w jednoznaczny sposób dowodzi w nim swojej znajomości literatury w analizowanym obszarze. Zarówno liczba, jak i dobór pozycji dowodzą, że mgr Józefowicz poświęciła dużo czasu na systematyczny i pogłębiony przegląd dostępnych prac naukowych – zarówno dotyczących zjawiska *flow*, jak i analizujących rozmaite powiązania pomiędzy graniem w gry video a funkcjonowaniem poznawczym. Ów rozległy i aktualny przegląd literatury przeczytałem z przyjemnością, a moja wiedza, nawet z zakresu zjawiska *flow*, które z racji swojego temporalnego charakteru znajduje się w kręgu moich zainteresowań badawczych, zdecydowanie się powiększyła.

Procedura adaptacji narzędzi DFS-2 i FSS-2 została przeprowadzona zgodnie z regułami sztuki i obejmowała zarówno weryfikacji zgodności z wykorzystaniem metody tłumaczenia zwrotnego (*back-translation*), jak i oceny właściwości psychometrycznych obydwu narzędzi, z uwzględnieniem konfirmacyjnej analizy czynnikowej i analizy rzetelności z wykorzystaniem wskaźnika *omega* McDonalda. Nawiasem mówiąc decyzja o wykorzystaniu *omegi* w miejsce wciąż najpopularniejszej *alfy* zasługuje w mojej ocenie na pochwałę. Wiele analiz dowodzi bowiem przewagi nowego wskaźnika, a mimo to przyzwyczajeni do *alfy* badacze często pozostają niechętni wobec nowego (lepszego!) rozwiązania. Wielkości obydwu prób badawczych wydają się wystarczające (reguła minimum 10 osób badanych na pozycję testową została zachowana). Zastanawiające, dlaczego wskazania analizy normalności z wykorzystaniem testu Shapiro-Wilka wykazała tak wyraźne odchylenia od normalności dla wszystkich skal obydwu narzędzi, podczas gdy zarówno wartości skośności, jak i kurtozy, mieszczą się w powszechnie akceptowanych granicach ± 1 . Chętnie zobaczyłbym, jak w rzeczywistości wyglądają rozkłady tych zmiennych – być może pozwoliłoby to zrozumieć, co takiego zadecydowało o owych rozbieżnościach (może rozkłady były np. dwumodalne?).

Jeśli chodzi o mankamenty strony statystycznej badania 1, to z pewnością za takowy należy uznać brak raportowania wielkości efektów w analizach, w których Autorka porównuje średnie wyniki w adaptowanych skalach i podskalach. Raportowanie *d*-Cohena (dla testu *t*) oraz η^2 (dla ANOVY) to współcześnie kwestia absolutnie podstawowa, dużo więcej mówiąca o rzeczywistej różnicy niż poziom *p* (przy małych próbach nawet wobec całkiem znaczącej różnicy *p* może być większe od 0,05, z kolei przy próbach dużych – takich jak próby w badaniu mgr Józefowicz – nawet bardzo niewielkie różnice o marginalnym znaczeniu mogą okazać się statystycznie istotne). Autorka podaje co prawda wielkości efektu w badaniu drugim, ale tutaj niestety ich zabrakło.

Drugim problemem są w mojej ocenie niewystarczające dane dotyczące trafności nowych skal. Doktorantka wydaje się co prawda świadoma tego, co nazywa „trzema możliwymi kierunkami walidacji” narzędzi do pomiaru *flow* (patrz s. 80), jednak w mojej ocenie badania w nich sugerowane nie tyle stanowią opcję, ile konieczność – i przynajmniej część z nich (np. walidacja krzyżowa, polegająca na analizie związków dyspozycja-stan) powinna być wykonana już na samym początku prac nad narzędziem (znajdujemy ją w badaniu drugim, ale jest to analiza przeprowadzona na bardzo małej, jak na procedury walidacyjne, próbie). Nie wiem jaka część obydwu prób badawczych z badania pierwszego się pokrywała (Autorka wspomina w pracy, że część osób brała udział w badaniu obydwoma narzędziami) –

jeśli osób takich było choćby kilkadziesiąt, i istniała możliwość powiązania danych pochodzących od jednej osoby, przedstawienie takiej analizy byłoby zdecydowanie uzasadnione. Doktorantka wskazuje ograniczenia badania, podkreślając wśród nich przede wszystkim fakt prowadzenia badania online. Jednak świetle analiz porównujących jakość badań stacjonarnych, prowadzonych metodą papier-ołówek z badaniami online, nie wydaje się to jednak większym problemem. Dużo istotniejszym ograniczeniem jest fakt braku analiz trafności zbieżnej (np. poprzez przeprowadzenie wspomnianych badań, zasygnalizowanych przez Autorkę).

Badanie drugie zostało przeprowadzone z wykorzystaniem ambitnej metodologii – było to badanie podłużne, z wielokrotnym pomiarem. Jego zasadniczy cel zasadniczo pokrywa się z głównym celem całej rozprawy – chodzi mianowicie o odpowiedź na pytanie o rolę doświadczenia *flow* przez gracza w nabywaniu złożonych umiejętności poznawczych podczas grania w Starcraft II. Cele specyficzne obejmowały 1) określenie zależności pomiędzy nasileniem *flow* a wynikami uzyskiwanymi w grze, 2) identyfikację „predyspozycji” (może lepiej byłoby użyć określenia „charakterystyk”) gracza, które pozwalają przewidywać nasilenie *flow*, a także 3) określenie zależności pomiędzy nasileniem *flow* a wynikami uzyskiwanymi w zadaniach poznawczych. Tym samym Autorka podjęła próbę uchwycenia zarówno pewnych mechanizmów (czy raczej dyspozycji podmiotowych), które leżą u podstaw *flow*, jak i jego konsekwencji – zarówno w odniesieniu do poziomu wykonania w grze, której swoisty trening badany przechodzi, jak i w odniesieniu do ogólnych procesów poznawczych (jak np. aktualizacja pamięci roboczej, czy przełączania się pomiędzy zadaniami).

Oceniając samo badanie, jak również, szerzej, jego ideę i sposób konstruowania go, chciałbym rozróżnić pomiędzy dwoma jego aspektami. Po pierwsze, jak wspomniałem już na początku, sam pomysł na badanie uznaję za wartościowy poznawczo i poszerzający wiedzę na temat znaczenia doświadczenia *flow* dla poziomu wykonania poznawczego (*cognitive performance*) i dla rozwoju w tym zakresie. Z drugiej jednak strony, czytając tekst miałem dość wyraźne poczucie ateoretyczności – Doktorantka stawia pytania badawcze, formułuje hipotezy, ale nie przedstawia ich klarownego uzasadnienia teoretycznego. Hipotezy te wydają się co prawda jak najbardziej zasadne, jednak od pracy naukowej wymagamy zazwyczaj argumentacji – dlaczego przewidujemy tak, a nie inaczej. Na przykład, warto byłoby określić, co takiego jest w stanie *flow*, że możemy oczekiwać jego pozytywnego wpływu na poziom wykonania oraz rozwój poznawczy? Pośrednią odpowiedź na to pytanie znajdziemy co prawda w przeglądzie

literatury, ale uzasadnienie dla hipotezy powinno być wyraźnie wyartykułowane przed jej podaniem.

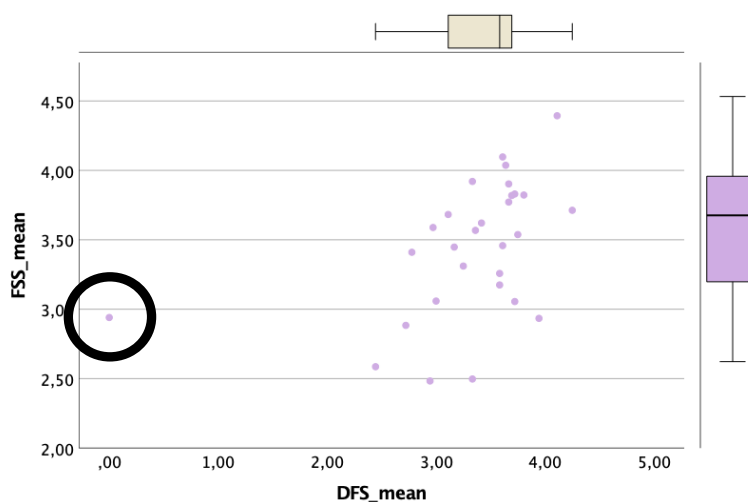
Podobnego uzasadnienia oczekiwałbym w odniesieniu do analizowanych przez Autorkę potencjalnych uwarunkowań podmiotowych *flow* – dlaczego na przykład jako miarę osobowości Doktorantka wybrała model Wielkiej Piątki osobowości? Wydaje mi się, że bardziej biologicznie zorientowane teorie osobowości lub temperamentu mogłyby być nawet lepszymi „kandydatami” do wyjaśnienia częstotliwości doświadczania *flow*. W tej części znajdziemy też drobne pomyłki – np. w hipotezie H.2.3. ewidentnie brakuje odniesienia do inteligencji (której dotyczy umieszczone obok pytanie badawcze), zaś sama hipoteza nie ma sensu: Autorka bowiem przewiduje w niej, że *flow* nie koreluje ze stanem *flow*... Ogólnie rozdział przedstawiający pytania badawcze i hipotezy umieściłbym znacznie wcześniej, przed opisem metody, jak aby przedstawiane wskaźniki mogły być przez czytelnika od razu odnoszone do postawionych hipotez.

Analizy statystyczne w badaniu drugim zostały generalnie dobrane adekwatnie. Pewnym problemem jest spadek liczebności w zestawieniu z przeprowadzonymi analizami mocy. Wygląda na to, że pierwotna liczebność próby była wystarczająca dla uzyskania zakładanej mocy statystycznej, (0.8), jednak następnie, wskutek szeregu dość powszechnie spotykanych w badaniach przyczyn, opisanych klarownie przez Doktorantkę, liczebność ta wyraźnie się zmniejszyła. Czy Doktorantka (lub szerzej – Autorzy projektu, w ramach którego badanie zostało zrealizowane) nie zakładali *drop-outu*? Kwestia ta powinna zostać dość szczegółowo poddana analizie w kontekście ograniczeń badania. Zasadnym wydaje się też przeprowadzenie analiz *post-hoc* rzeczywistej mocy uzyskanej w badaniu.

Zastanawia mnie sensowność prowadzenie analiz nieparametrycznych (mam tu na myśli analizę korelacji z wykorzystaniem współczynnika *rho*-Spearmana). Rozumiem, że było to spowodowane brakiem normalności większości rozkładów (choć w mojej ocenie można było ją przyjąć (zachowując ostrożność w interpretacji wyników), opierając się o statystyki skośności i kurtozy). Jednocześnie przy takiej decyzji nie do końca zrozumiała była dla mnie inna decyzja - o przeprowadzeniu analiz regresji przy analizach, w których predyktorami były zmienne osobowościowe. Skoro zakładamy brak normalności, to powinniśmy być w tym konsekwentni i zrezygnować już całkowicie z analiz parametrycznych (a do takich należy przecież i analiza regresji). Nie uważam tego za absolutny błąd (jak wspomniałem było to względnie uprawnione w oparciu o statystyki skośności i kurtozy), jednak problemem jest tu brak konsekwencji odnośnie tej kwestii. Tymczasem uważam, że analizy regresji przydałyby

się już przy testowaniu hipotez zw. z pierwszym celem badania - przy raportowaniu tak licznych korelacji ze zmiennymi, które są ze sobą dodatnio skorelowane (a jak wiemy z badania 1 podskale *flow* generalnie wiążą się ze sobą dodatnio) wiele z tych wyników może w istocie być redundantnymi. Analiza regresji pozwoliłaby na wychwycenie owej redundancji (lub dowiedzenie jej braku). Rozumiem, że z racji skorelowania niektórych z owych podskal istniałoby ryzyko współliniowości – jednak nawet gdyby takowa zaistniała, możliwe byłoby obejście tego problemu – np. z wykorzystaniem zabiegu statystycznego, polegającego na obliczeniu ogólnej wartości *flow*, a następnie zapisaniu reszt dla poszczególnych podskal po „wyłączeniu” wariancji dzielonej z czynnikiem ogólnym. Wtedy możliwe byłoby uchwycenie wysoce specyficznych efektów poszczególnych aspektów *flow*.

Moją uwagę zwróciła konkluzja, dotycząca wysokiej, dodatniej korelacji pomiędzy stanem a cechą *flow* (s. 120). Otóż jeśli spojrzymy na wykres rozrzutu punktowego, już na pierwszy rzut oka widać, że przed przeprowadzeniem tej analizy należało usunąć z bazy *outliera*, który uzyskał zero (0, **sic!**) w skali Likerta 1-5. Osobę tę pozwoliłem sobie nawet zaznaczyć na rysunku:



Korelacja Pearsonowska, jako statystyka parametryczna, jest wskaźnikiem bardzo podatnym na wartości skrajne. Tutaj mamy do czynienia z gigantycznym wręcz odchyleniem od średniej, przez co uzyskana przez Autorkę wartość $r = .63$ jest bardzo silnie zawyżona. Choć rzut oka na pozostałe punkty dowodzi, iż mamy tu faktycznie do czynienia z zależnością dodatnią, to jej wielkość jest najpewniej w rzeczywistości znacząco niższa. Rodzi się pytanie, czy raportowane dalej korelacje, obliczane dla poszczególnych podskal – jeśli znalazła się wśród nich ta sama osoba badana z tym konkretnym „zerowym” wynikiem, wówczas i te dane należałoby przeliczyć raz jeszcze.

Z racji moich zainteresowań badawczych związanych z osobowością i różnicami indywidualnymi, z dużym zainteresowaniem zapoznałem się z wynikami analiz regresji dotyczących predykcji *flow* z wykorzystaniem cech Wielkiej Piątki. Wykazały one wyraźną rolę Neurotyczności, jako czynnika ograniczającego *flow*. Oprócz tego wyniku uwagę zwracają już nieco mniej systematyczne rezultaty wskazujące na związki poszczególnych aspektów *flow* i cech Ugodowości i Otwartości na Doświadczenie. Po analizach regresji następuje szereg interesujących, choć niestety znów dość słabo uzasadnionych teoretycznie, analiz mediacji. Ich wyniki wydają mi się istotne i same w sobie mogłyby stanowić podstawę interesującej publikacji. Problemem może być jednak niewielka liczebność próby. I jeszcze jedna uwaga dotycząca tej części pracy: zważywszy na liczebność, zasadnym byłoby raportowanie tzw. *ratio of indirect to total effect*. Statystyka ta pozwala dobrze uchwycić, w jakiej części testowany mediator wyjaśnia efekt głównego predyktora na zmienną zależną, stanowiąc swoisty substytut miar wielkości efektu w modelach mediacyjnych.

Jeśli chodzi o dalsze analizy zmiennych podmiotowych, to w mojej ocenie Doktorantka zdecydowanie za szybko zrezygnowała z analiz dotyczących inteligencji, stwierdziwszy, że nie wiąże się ona ze stanami *flow*. Tymczasem modele mediacyjne to przecież nie jedyne modele, jakie możemy poddać analizie. Inteligencja z definicji powinna przewidywać szybkość uczenia się i poziom wykonania wszelkich zadań poznawczych. Tym samym, jako miara ogólnej sprawności poznawczej, pozostaje ważną zmienną w kontekście prowadzonych analiz. Wydaje mi się na przykład, iż dość łatwo można było uzasadnić model, w którym inteligencja byłaby predyktorem wykonania w Starcraft II, zaś *flow* byłoby moderatorem owej zależności – jako czynnik facylitujący, ułatwiający przekucie potencjału intelektualnego mierzonego testem IQ na poziom wykonania w tak angażującej poznawczo grze.

Z kolei przy analizach dotyczących testowania hipotezy H.2.4b (s. 141) sugerowałbym zastąpienie zastosowanych analiz (*rho*-Spearmana porównane z wykorzystaniem „reguły kciuka” dla warunków „win” i „lost”) analizą moderacji (z dychotomicznym moderatorem w postaci zmiennej win/lost) lub ostatecznie obliczenie istotności różnic pomiędzy parami korelacji (są do tego opracowane określone metody analiz).

Ostatnia, największa grupa wyników dotyczy analiz weryfikujących efekty udziału w treningu gry RTS dla poziomu wykonania niezależnych od samej gry zadań poznawczych, reprezentujących określone funkcje poznawcze. Autorka porównuje w nich m. in. efekty treningu w stałym vs. zmiennym środowisku gry w interakcji z poziomem *flow* charakterystycznym dla danego zawodnika. Ta część pracy dostarcza ogromu danych, liczb

analiz, z których każdą z osobna można by uznać za interesującą i wartościową. Jednak ich przytłaczająca wielość sprawia, że czytelnikowi nie jest łatwo śledzić narrację i dokonywać samodzielnie jakiegokolwiek syntezy. Z pomocą przychodzi sama Autorka, po każdej grupie analiz dokonując pewnego podsumowania i uwypuklając najważniejsze stwierdzenia empiryczne. Mimo tej pomocy ze strony Doktorantki, lektura tej części pracy wywołała u mnie niespotykane wysokie poczucie *cognitive load* – obawiam się, że dla wielu potencjalnych odbiorców ta część pracy może okazać się zbyt dużym wyzwaniem. I choć przedstawione w niej analizy zostały generalnie przeprowadzone i opisane poprawnie – jeśli spojrzymy na każdą z osobna – to ich wielość, a razem z nią, przytłaczająca liczba wyników, nasuwają mi jedną niepokojącą konkluzję. Otóż już na dość wczesnym etapie edukacji statystycznej przekazujemy studentom wiedzę dotyczącą **problemu porównań wielokrotnych, związanego z ryzykiem popełnienia tzw. błędu I rodzaju**, polegającego zasadniczo na zwiększonym ryzyku omyłkowego przyjęcia fałszywej hipotezy alternatywnej przy wykonywaniu wielu porównań / testów statystycznych, dotyczących tej samej grupy hipotez jednocześnie. Tutaj mamy do niestety do czynienia z przypadkiem wręcz ekstremalnym – tak przytłaczająca liczba analiz wydaje się wręcz „wołać” o zastosowanie poprawki na porównania wielokrotne (np. poprawki Bonferroniego, lub zdecydowanie mniej kosztowną pod względem mocy statystycznej poprawki Bonferroniego-Holma, ew. jednej z procedur kontrolujących tzw. *false discovery rate, FDR*). Niestety zważywszy na ograniczoną liczebność próby obawiam się, iż poprawka taka mogłaby dość wyraźnie zmienić obraz wyników, pozostawiając istotnymi wyłącznie najsilniejsze z raportowanych przez Autorkę efektów. Jest to zdecydowanie największy mankament recenzowanej rozprawy. Najprostszym rozwiązaniem tego problemu mogłoby być znaczące ograniczenie liczby analiz (np. przez wybór najważniejszych, lub przez skoncentrowanie się na wskaźnikach ogólnych, tudzież prowadzenie analiz wyłącznie w odniesieniu do zmiennych, co do których istnieją silne przesłanki teoretyczne). Muszę przyznać, że dość długo zastanawiałem się, czy pozwala on uznać wyniki pracy za uprawniające do wyciągania jakichkolwiek wiążących wniosków, a w konsekwencji – czy uznać pracę za spełniającą kryteria ustawowe, stawiane rozprawom doktorskim. Zważywszy jednak na dużą pracochłonność opisanych w rozprawie badań, a także wartość poznawczą i teoretyczną niektórych spośród opisanych przez mgr Józefowicz stwierdzeń empirycznych (tj. tych, które cechuje relatywnie duża siła efektu oraz duży poziom „wewnętrznej replikowalności”, tj., systematyczna obecność efektu w odniesieniu do dużej liczby analiz -jak to miało miejsce np. w przypadku neurotyczności) zdecydowałem się rozstrzygnąć tę kwestię na korzyść Doktorantki.

Dyskusja wyników jest zwięzła i prowadzona raczej płynnie. Czyta się ją z łatwością. Autorka sygnalizuje najważniejsze wątki, zarysowuje główne konkluzje i, w większości przypadków, stara się odnosić uzyskane wyniki do innych rezultatów i stwierdzeń znanych z literatury. Wydaje się, że zważywszy na rozległość części teoretycznej owych odniesień mogłoby być nieco więcej, jednak ogólnie tą część pracy oceniam raczej pozytywnie.

Pracę zamyka rozdział dotyczący wniosków płynących z pracy. Uważam, że wyznaczone przez Autorkę dalsze kierunki badań mają zdecydowanie zbyt ogólny charakter – ich wskazanie byłoby równie zasadne niezależnie od przeprowadzonych przez Autorkę badań i płynących z nich wniosków oraz wskazanych na kolejnych stronach ograniczeń. Zdecydowanie bardziej wartościowy i uzasadniony wydaje się podrozdział *Implikacje*, w którym Doktorantka bezpośrednio odnosi najważniejsze ustalenia empiryczne do zastosowań praktycznych, a także formułuje już znacznie bardziej precyzyjne wskazania dotyczące kwestii metodologicznych i dalszych kierunków badawczych. Całkiem interesujących wniosków dostarcza podrozdział dotyczący konkluzji o charakterze logistycznym i organizacyjnym.

Za wyraźnie słabszy aspekt pracy muszę również niestety uznać jej stronę formalną. Autorka nie ustrzegła się niefortunnnych gramatycznie lub składniowo sformułowań (np. „*Supervision of the study was largely supervised by computer*”, s. 86), braku konsekwencji w używaniu niektórych terminów (np. *extraversion* vs *extroversion*) oraz bardzo licznych niespójności w formatowaniu pracy. Na przykład cytując prace dwojga autorów, Doktorantka niekiedy rozdziela je łącznikiem „and”, innym razem wykorzystują tzw. etkę (&), a w pozostałych przypadkach rozdziela je przecinkiem. Podobną niespójność łatwo zauważyć w odniesieniu do statystyk – raportując setne i dziesiętne części liczby, Autorka oddziela je raz kropką (co byłoby zgodne z oryginalnymi, amerykańskimi wytycznymi APA), a raz przecinkiem (jak sugerują autorzy polskich „adaptacji” standardów APA). W części cytowań znajdziemy przecinek przed podaniem roku, ale w wielu z nich przecinka brakuje. W bibliografii końcowej błędów formatowania jest jeszcze więcej, począwszy od braku kursywy w tytułach czasopism, poprzez brak podawania redaktorów monografii, z których pochodzą cytowane rozdziały, a skończywszy na stosowaniu skrótów przy niektórych tytułach czasopism oraz podawaniu kursywą bądź inną czcionką całych pozycji. Przy niektórych tekstach (np. Kongmee i in, 2011) nie podano w ogóle skąd dany tekst pochodzi – zamieszczono wyłącznie tytuł. Do tego dochodzą błędy rozdzielaniu akapitów (np. na s. 8 nowy akapit zaczyna się... w połowie zdania), a także nadmiernie długie akapity (wg standardów APA długość akapitu nie powinna przekraczać strony, tutaj taką sytuację znajdujemy kilkakrotnie).

Oczywiście aspekt formalny nie ujmuje pracy wartości naukowej, ale wydaje mi się, iż od osoby składającej do recenzji jedną z najważniejszych prac naukowych w swoim życiu, można wymagać większej dbałości i chociażby przeczytania pracy „na spokojnie” już po jej napisaniu.

Jeszcze jedna uwaga formalna – hipotezy dotyczą istnienia ogólnej zależności, a nie wyników uzyskanych w danym badaniu. Z kolei uzyskane wyniki dotyczą wyłącznie danej próby w danej sytuacji i czasie. Dlatego najbardziej uzasadnione jest formułowanie hipotez w czasie teraźniejszym („istnieje dodatnia zależność...”), zaś opisywanie wyników z użyciem czasu przeszłego („zmienna A dodatnio korelowała ze zmienną b...”). Użycie czasu przyszłego w odniesieniu do hipotez lub pytań badawczych (np. s. 177: „*The second hypothesis assumed that we will see this relationship on the session by session level.*”) sugeruje, że hipoteza dotyczy wyłącznie danego badania (jeśli tak – to po co ją weryfikować?). Z kolei użycie czasu teraźniejszego do opisu rezultatów, w sposób nieuzasadniony może sugerować ich nadreprezentatywność (np. „*there is a weak, but statistically significant negative correlation between time since the last session and the autotelic experience*”). Dużo lepiej byłoby takich przypadkach użyć czasu przeszłego („*there was...*”). Nie twierdzę, że Autorka systematycznie popełnia tego rodzaju błędy – przeciwnie w większości przypadków zastosowana forma jest adekwatna (czas przeszły dla uzyskanych wyników, a teraźniejszy dla hipotez), jednak ze względu na pojawienie się kilku, może kilkunastu odstępstw od tej zasady, postanowiłem uwypuklić tą kwestię.

Konkludując, pomimo opisanych powyżej wątpliwości – zwłaszcza związanych z problemem ryzyka błędu typu I – moja ocena recenzowanej rozprawy jest ostatecznie pozytywna. Autorka podjęła interesujący problem badawczy, przeprowadziła wystarczająco zaawansowane, złożone i pracochłonne badanie empiryczne, a także dostatecznie dobrze poradziła sobie z opracowaniem oraz interpretacją ich wyników. Recenzowana rozprawa spełnia warunki określone w Rozdz. 2. Ustawy z dn. 20 lipca 2018 Prawo o szkolnictwie wyższym i nauce: stanowi oryginalne rozwiązanie problemu naukowego, wskazuje na ogólną wiedzę teoretyczną Doktorantki w dyscyplinie psychologia oraz na umiejętność samodzielnego prowadzenia pracy naukowej. Stwierdzając, że opiniowana praca doktorska zasługuje na pozytywną ocenę, wnioskuję o dopuszczenie mgr Justyny Józefowicz do dalszych etapów przewodu doktorskiego.



Dr hab. Maciej Stolarski, prof. ucz.