

**RECENZJA OSIĄGNIĘCIA ORAZ DOROBKU DOKTORA PAWŁA STROJNEGO  
W POSTĘPOWANIU W SPRAWIE NADANIA STOPNIA DOKTORA HABILITOWANEGO  
W DZIEDZINIE NAUK SPOŁECZNYCH, DYSCYPLINIE PSYCHOLOGIA**

Przedmiotem tej recenzji jest stwierdzenie czy osiągnięcie naukowe oraz dorobek doktora Pawła Strojnego stanowią podstawę nadania mu stopnia doktora habilitowanego nauk społecznych w dyscyplinie psychologia. W szczególności analizie poddaję fakt spełnienia bądź nie warunków wymienionych w Ustawie, tj.: (1) posiadanie stopnia doktora, (2) posiadanie w dorobku osiągnięcia naukowego (monografii lub cyklu publikacji) stanowiącego znaczny wkład w rozwój dyscypliny psychologia oraz (3) wykazywanie się aktywnością naukową realizowaną w więcej niż jednej uczelni, instytucji naukowej lub instytucji kultury, w szczególności zagranicznej.

**Sylwetka Kandydata oraz ogólna charakterystyka dorobku i aktywności naukowej**

Doktor Paweł Strojny jest absolwentem Uniwersytetu Jagiellońskiego, gdzie uzyskał magisterium (2008) i doktorat (2015) w zakresie psychologii. Tak więc spełniony jest pierwszy z warunków (posiadanie stopnia doktora). Obecnie (od 2016) dr Strojny jest zatrudniony jako adiunkt w Instytucie Psychologii Stosowanej UJ (Wydział Zarządzania i Komunikacji Społecznej), wcześniej (między 2015 a 2016) pracował jako asystent w Katedrze Psychologii Uniwersytetu Pedagogicznego KEN w Krakowie. Można więc stwierdzić, że spełniona jest również przesłanka trzecia – habilitant publikował bowiem zarówno z afiliacją UP KEN, jak i UJ, prowadził więc prace naukowe w więcej niż jednej uczelni. Poza pracą naukową dr Strojny przez 5 lat kierował Działem Badań i Rozwoju spółki Nano Games, tworzącej gry i symulatory do tzw. gier poważnych, był więc zaangażowany w praktyczne wdrażanie nowych technologii.

Poza artykułami wchodzącymi w skład osiągnięcia naukowego, które szczegółowo analizuję w dalszej części tej recenzji, opublikowany dorobek dr. Strojnego obejmuje 14 prac naukowych, z czego dwa artykuły: w pismach *Homo Ludens* i *Czasopismo Psychologiczne* ukazały się przed uzyskaniem doktoratu. Wśród 14 wspomnianych wyżej prac, w ośmiu dr Strojny jest pierwszym lub jedynym (2) autorem, a w trzech autorem ostatnim. Doniesienia te ukazywały się w miejscach o bardzo zróżnicowanej jakości – mamy tu bowiem zarówno rozdziały w polskojęzycznych pracach zbiorowych, materiały konferencyjne konferencji krajowych jak i zagranicznych, rozpoznawalne czasopisma polskojęzyczne (*Przegląd Psychologiczny*, *Czasopismo Psychologiczne*, *Roczniki Psychologiczne*), jak też – choć pozostające w zdecydowanej mniejszości – pisma o uznanej renomie międzynarodowej (*Motivation and Emotion*, *Personality and Individual Differences*) oraz pisma wprawdzie nieszczególnie selektywne, ale pozwalające na upowszechnienie wyników w obiegu międzynarodowym (*Psychological Topics*, *Frontiers in Psychology*). Doktor Strojny jest też (współ)autorem dwunastu wystąpień konferencyjnych, jednak wykazuje wśród nich także prezentacje na wewnętrznych sympozjach; jedynie jedno wystąpienie miało miejsce poza granicami kraju (Belgia).

Doktor Paweł Strojny realizował granty finansowane ze środków konkursowych – był laureatem konkursu Preludium NCN, obecnie realizuje zaś grant finansowany w ramach programu Sonata Bis NCN. Uczestniczył jako wykonawca w grantach kierowanych przez innych naukowców. Ma też na swoim koncie trzy granty finansowane przez NCBiR, jednak jego rola (poza grantem poświęconym symulatorowi katastrof zbiorowych, którym kierował) jest opisana w autoreferacie jako funkcja kierownika fragmentu podejmowanych działań badawczych.

Baza Scopus indeksuje 16 artykułów dr. Strojnego, cytowanych dotychczas 233 razy (230 bez autocytowań, indeks  $h = 6$ ), z czego za znaczną część (161 cytowań) odpowiadają cytowania

wchodzącego w skład osiągnięcia artykułu opublikowany we *Frontiers in Psychology* na temat choroby symulatorowej, określanej przez habilitanta jako VIMS. Bardziej inkluzywna baza Google Scholar (GS) pokazuje 447 cytowań publikacji dr. Strojnego ( $h = 10$ ), z czego za 62% cytowań odpowiada praca z *Frontiers in Psychology* (275 cytacji w GS).

Stwierdzam, że habilitant posiada stopień doktora nauk społecznych w zakresie psychologii oraz prowadził badania naukowe w więcej niż jednej uczelni. Kandydat ma też na swoim koncie sukcesy we wnioskowaniu o środki na realizację badań w trybie konkursowym (NCN i NCBiR), a jego prace (zwłaszcza jedna) są cytowane w środowisku międzynarodowym. Jednocześnie jednak dorobek naukowy habilitanta poza osiągnięciem ma w znacznej mierze charakter lokalny – jedynie cztery artykuły zostały opublikowane w czasopiśmie międzynarodowych, rozpoznawalnych w psychologii, jednak o zróżnicowanej jakości (*Motivation and Emotion, Personality and Individual Differences, Frontiers in Psychology, Psychological Topics*). Jest to więc dorobek skromny, w minimalnym stopniu spełniający kryteria mogące świadczyć o wpływie na dyscyplinę.

### Ocena osiągnięcia naukowego

Swoje osiągnięcie naukowe dr P. Strojny zatytułował „Optymalizacją parametrów wirtualnego doświadczenia: pomiar, modyfikacja i społeczne skutki” i włączył do niego dziesięć współautorskich artykułów, które analizuję niżej. Sam tytuł osiągnięcia, zwłaszcza zaś „optymalizacja” wydaje się mocno „stosowany” i techniczny, dodatkowo zaś nie zawsze klarowny: zwłaszcza gdy analizujemy kolejne teksty wchodzące w skład cyklu. Z innego miejsca autoreferatu dowiadujemy się, że w opublikowanych tekstach analizie poddano proces doświadczenia wirtualnego środowiska (bardziej niż same badania nad wirtualną rzeczywistością). Problem jednak w tym, że również owo „doświadczenie” jest w istocie bardzo ogólne i może oznaczać niemal wszystko – każde subiektywne odczucie w zderzeniu z wirtualną rzeczywistością, grą czy symulatorem. Habilitant wyjaśnia, że interesują go zwłaszcza cztery kategorie: poczucie obecności (*presence*) oraz obecności innych (*co-presence*), ocena realizmu środowiska, w którym badany jest umieszczony oraz zaangażowanie w wirtualne doświadczenia. Choć wyróżnienie i charakterystyka tych kategorii w autoreferacie są pomocne dla zrozumienia istoty programu badawczego dr. Strojnego, to nie zmienia to faktu, że program ten ma w znacznej mierze charakter ateoretyczny i eklektyczny. Nie tylko nie odnajdujemy tu prób stworzenia i przetestowania bardziej ogólnego modelu wirtualnego doświadczenia (czy też doświadczenia), ale i odniesienia do ugruntowanych teorii psychologicznych (np.: teoria intensywności motywacji, teoria aktywizacji schematów poznawczych) pojawiają się jedynie sporadycznie i w sposób mało systematyczny i mocno przypadkowy. Poniżej analizuję prace wchodzące w skład osiągnięcia w podziale na trzy grupy, wyróżnione przede wszystkim ze względu na cele i/lub zastosowaną metodologię. Dzielę je na artykuły, które są przeglądami (1), artykuły o narzędziowym-metodologicznym charakterze (2) oraz artykuły o wyraźnie „substancjalnej” orientacji, testujące hipotezy i w założeniu dostarczające nowej wiedzy (3).

**Prace przeglądowe.** Wśród artykułów mających charakter systematycznych przeglądów lub mini-metaanaliz znajdujemy 3 spośród 10 prac włączonych w skład osiągnięcia. Są to prace poświęcone chorobie symulatorowej / VIMS – a więc wspomniany już najczęściej cytowany artykuł habilitanta [Dużmańska, N., Strojny, P., & Strojny, A. (2018). Can simulator sickness be avoided? A review on temporal aspects of simulator sickness. *Frontiers in Psychology*, 9. <https://doi.org/10.3389/fpsyg.2018.02132>], efektywności treningu wirtualnego [Strojny, P., & Dużmańska-Misiarczyk, N. (2023). Measuring the effectiveness of virtual training: A systematic review. *Computers & Education: X Reality*, 2, 100006. <https://doi.org/10.1016/j.cexr.2022.100006>] oraz efektowi facylitacji społecznej w warunkach wirtualnego doświadczenia [Sterna, R., Strojny, P., & Rębilas, K. (2019). Can virtual observers affect our behavior? social facilitation in Virtual Environments: A mini-review. *Social Psychological Bulletin*, 14(3).

<https://doi.org/10.32872/spb.v14i3.30091>]. W dwóch z trzech ww. prac rola habilitanta nie była wiodąca, choć deklaracje autorów wskazują, że była ona znacząca. Jednocześnie trzeba zauważyć, że w artykule z *Frontiers in Psychology* wkład autorów jest zdefiniowany w inny sposób niż w autoreferacie i deklaracjach autorów.

Wszystkie trzy artykuły są zgrabnymi, choć w znacznej mierze deskryptywnymi i głównie narracyjnymi podsumowaniami stanu badań nad poszczególnymi analizowanymi kwestiami. Przegląd z *Frontiers in Psychology* dotyczy problemu ważnego dla wszystkich realizujących badania w VR (odczuwane zawroty głowy czy nudności to częste doświadczenia, zwłaszcza przy niektórych rozwiązaniach technologicznych), stąd nie dziwi duża popularność tego tekstu. Trudno jednak uznać, aby wnosił on nowe ustalenia, hipotezy czy nawet niezadane dotąd pytania – jest dobrze napisanym podsumowaniem, które pokazuje na duże zróżnicowanie metodologiczne dotychczasowych badań oraz trudności z formułowaniem jednoznacznych konkluzji na temat powodów choroby symulatorowej oraz jej przebiegu czy też sposobów jej zapobiegania. Przede wszystkim jednak tekst ten jest w bardzo niewielkim, minimalnym wręcz stopniu, związany z pozostałymi artykułami i badaniami wchodzącymi w skład osiągnięcia.

Podobny charakter ma przegląd badań nad efektywnością treningu realizowanego w warunkach wirtualnych opublikowany w nowym (a więc trudno powiedzieć jak selektywnym) piśmie *Computers & Education: X Reality*. Przegląd ten był niewątpliwie niesłychanie pracochłonny – przeanalizowano ponad 300 artykułów; przegląd poprzedzono przekonującym wprowadzeniem teoretycznym dobrze świadczącym o orientacji autorów w literaturze. Konkluzje płynące z frekwencyjnych podsumowań i zestawień są jednak takie, że „bywa różnie” – zarówno jeśli idzie o metodologię realizowanych badań nad efektywnością takich treningów, czas i metodę pomiaru, jak i same ustalenia na temat efektywności. To dobry przegląd, ale ponownie bardzo luźno związany z pozostałymi pracami: o niejasnym wkładzie do całości cyklu. Doceniam pracochłonność i rzetelność dokonanych podsumowań, jednak nie jest jasne, jakie wnioski płyną z dokonanych analiz.

Z kolei trzeci z przeglądów – z udostępnionymi danymi oraz kodami analiz, za co autorom należy się uznanie – opublikowany w polskim SPB, to niewielka metaanaliza pokazująca znaczenie obecności innych dla wykonywania łatwych i trudnych zadań w warunkach wirtualnych. Ten tekst dla odmiany wyraźnie wiąże się z co najmniej jednym z empirycznych doniesień habilitanta, jednak wkład tego przeglądu do psychologii trudno uznać za znaczący. Autorzy konkludują, że badań pozwalających na sensowną metaanalizę jest po prostu za mało, a te, które są, są obciążone licznymi niedoskonałościami. Prowadzi to autorów do przesunięcia ilościowych wyników analiz do suplementu i pozostania przy narracyjnym przeglądzie. Problem w tym, że jest on niekonkluzywny, bowiem 10 artykułów i 13 badań na niewielkich próbach, dodatkowo jeszcze badań o licznych mankamentach metodologicznych nie jest w stanie wiele pewnego powiedzieć. W konsekwencji trzy artykuły przeglądowe, choć wymagające wysiłku i czasu trudno uznać za wkład do psychologii – paradoksalnie, nawet mimo licznych cytowań jednego z nich – zwłaszcza zaś, że w przypadku dwóch z nich trudno o wskazanie ich znaczenia dla pozostałej części osiągnięcia.

**Prace psychometryczne.** Drugą grupą artykułów dr. Pawła Strojnego są teksty o charakterze psychometrycznym, poświęcone bądź to adaptacji istniejących narzędzi, bądź opracowaniu fizjologicznych wskaźników zaangażowania podczas realizacji zadania w symulatorze (sytuacja wypadku drogowego). W skład tej grupy włączam teksty: Strojny, P., Lipp, N., Strojny, A. (2022). Polish adaptation of the Igroup Presence Questionnaire. *Homo Ludens* 1(15), 177-202, Lipp, N., Sterna, R., Dużmańska-Misiarczyk, N., Strojny, A., Poeschl-Guenther, S., & Strojny, P. (2021). VR realism scale—revalidation of contemporary VR headsets on a Polish sample. *PLOS ONE*, 16(12).

<https://doi.org/10.1371/journal.pone.0261507> oraz Czarnek, G., Richter, M., & Strojny, P. (2020). Cardiac sympathetic activity during recovery as an indicator of sympathetic activity during task

performance. *Psychophysiology*, 58(2). <https://doi.org/10.1111/psyp.13724>. Dwa pierwsze artykuły – opublikowane w *Homo Ludens* i *Plos One* mają bardzo zbliżony charakter, prezentują bowiem proces polskiej adaptacji bardzo prostych skal samoopisowych do pomiaru – odpowiednio – poczucia obecności i postrzeganego realizmu sytuacji VR. Tekstowi w *Homo Ludens* nie towarzyszą dostępne dane i skrypty, brak też informacji na temat ważnych elementów procesu analiz – choćby estymatorów zastosowanych w confirmacyjnej analizie czynnikowej i rzetelności innych wykorzystanych w badaniu narzędzi (a ma to znaczenie dla zrozumienia raportowanych korelacji, w szczególności zaś trafności dyskryminacyjnej w odniesieniu do immersji). Autorzy porównują rozwiązanie trzyczynnikowe z – wcale nie lepiej dopasowanym – rozwiązaniem podwójnego czynnika, które jednak pojawia się *ad hoc*, bez żadnego szczególnego uzasadnienia. Zastrzeżenia można by też mieć do wykorzystanej dość egzotycznej skali odpowiedzi (–3 -- +3). Samo narzędzie wydaje się niespecjalnie wyrafinowane, przede wszystkim zaś trudno wskazać na psychologiczne charakterystyki stojące za wyróżnionymi czynnikami. Zastosowana procedura analityczna jest do zaakceptowania, jednak tłumaczenie na polski istniejącego, prostego narzędzia trudno uznać za wnoszące coś istotnego do nauki (choć niewątpliwie jest to działalność użyteczna).

Artykuł na temat *VR Realism Scale* opublikowany w *Plos One* prezentuje zagregowane rezultaty 6 badań (całkowita próba,  $N = 720$ ), z dostępnymi danymi, lecz niedostępnymi kodami analiz. Sądząc z opisu wkładu autorów, jest to przede wszystkim praca N. Lipp, ale również wkład habilitanta jest tu znaczący. Autorzy analizują strukturę czynnikową (czteroczynnikową) skali realizmu w warunkach wirtualnych. Z artykułem tym i zrealizowanymi analizami wiąże się szereg problemów, które nie świadczą najlepiej o recenzentach *Plos One*. Po pierwsze, w największym z badań włączonych do analiz (badanie 1,  $N = 245$ ) uczestnicy przypominali sobie bardzo różne doświadczenia i sytuacje, nie sposób więc ich porównywać, a nawet zrozumieć, co badani mieli na myśli wypełniając *VR Realism Scale*. Po drugie, nie wiadomo, dlaczego zastosowano taki, a nie inny estymator (DWLS) w confirmacyjnej analizie czynnikowej, jednak biorąc pod uwagę fakt, że sięgnięto po najbardziej liberalny z dostępnych (i co do zasady wymagający dużych prób), pojawia się wątpliwość, jakie byłoby dopasowanie modeli przy typowych w takich sytuacjach estymatorach maksymalnej wiarygodności. Po trzecie, autorzy wspominają, że źródło danych (a więc ich pochodzenie z różnych badań) okazało się istotnie różnicować uzyskiwane wyniki (co nie dziwi w kontekście mojego komentarza dotyczącego charakteru największego z badań), kontrolowali więc ten czynnik w CFA. Nie wiem jednak, jak taka kontrola mogłaby wyglądać poza wielogrupową analizą confirmacyjną, której w tekście brak, a przynajmniej brak dla tej akurat zmiennej. Raportowane korelacje item-skala nie mają sensu przy dostępnych ładunkach czynnikowych CFA. Autorzy pokazują dobre dopasowanie modelu czteroczynnikowego, choć na jeden czynnik latentny składa się tylko jedna pozycja kwestionariusza, co rodzi pytanie o „zbieganie” modelu, nie wiedzieć czemu jednak i tym razem sprawdzany jest model podwójnego czynnika (*bi-factor*). W kolejnych analizach – wydaje się, że również analizie inwariancji, choć nie jest to jasne – autorzy piszą, że używają zarówno ogólnego czynnika, jak i czynników szczegółowych. Problem w tym, że jeśli są to analizy na zmiennych obserwowalnych (*manifest*), to taka analiza jest błędna – wyniki powinny być bowiem wyestymowane przy kontroli ładunków dla czynnika *bi-factor* i poszczególnych skal. Dodatkowo nie jest jasne, w jaki sposób zrealizowana jest analiza inwariancji; na ogół modele podwójnego czynnika w takich sytuacjach cechują się problemami z konwergencją, trudno się jednak zorientować czy rzeczywiście takie właśnie modele były stosowane. Co ważne, już model inwariancji konfiguralnej cechuje się istotnie słabszymi miarami dopasowania niż jednogrupowy model czteroczynnikowy, a kolejne ograniczenia dla modelu wielogrupowego (inwariancja metryczna i skalarna) pokazują problem z ekwiwalencją pomiaru. Trudno również uzasadnić zastosowaną w tekście analizę mediacji – zwłaszcza że została dokonana na danych korelacyjnych – a tłumaczenie, że mediacja jest matematycznie supresją, wydaje się błędne (taką analogię można budować między matematyczną charakterystyką mediacji i wystąpieniem zmiennej

typu *confounder*, ale nie między mediacją a supresją). Autorzy raportują wiele ogólnych efektów pośrednich, bez kontroli liczby porównań, przede wszystkim jednak bez należytego uzasadnienia oczekiwanego przyczynowego charakteru relacji. Ogółem tekst ten jest „przeładowany” pobocznymi analizami, często wątpliwymi, co mu wyraźnie nie służy. Podobnie jak w poprzednim przypadku samo analizowane narzędzie jest proste, trudno jednak wskazać jego psychologiczne podstawy – mówi (i mierzy) bardziej dość oczywiste i ateoretyczne wrażenia, że coś jest realistyczne bądź nie. Choć analizy confirmacyjne – pomijając pytania i wątpliwości wyrażone wyżej – są na ogół zrealizowane w akceptowalny sposób, tekst ten nie wnosi wiele do naszej wiedzy, choć daje badaczom instrument, który może okazać się przydatny w ich studiach.

Trzeci z tekstów o „narzędziowym” czy też – w tym przypadku – „wskaźnikowym” charakterze, to artykuł, którego główną autorką jest Gabriela Czarnek, opublikowany w *Psychophysiology*. Współautorem tego tekstu jest Michael Richter – rozpoznawalny międzynarodowo psychofizjolog i badacz motywacji. Artykuł ten prezentuje reanalizę dwóch badań na łącznej próbie 181 osób, w których udział wzięli strażacy i studenci szkoły pożarniczej. Autorzy stawiają sobie za cel sprawdzenie czy reakcje układu współczulnego dają się przewidzieć parametrami fizjologicznej reakcji organizmu tuż po ustąpieniu wymogów zadania (w okresie odpoczynku). Przeprowadzone analizy wskazują na obiecujący charakter wykorzystanych dynamicznych wskaźników. Biorąc pod uwagę selektywny charakter *Psychophysiology* i jakość recenzji w tym piśmie, nie mam powodów, aby podważać konkluzje tego artykułu, a uzyskanie psychofizjologicznych miar aktywności i zaangażowania w zadanie w warunkach wirtualnego doświadczenia może być faktycznie użyteczne. Jednocześnie, wskaźniki takie w oczywisty sposób wymagają dalszej walidacji, również z wykorzystaniem innych miar, stąd wstępny charakter tych ustaleń i konieczność ich replikacji. Trzeba też zauważyć, że tekst Czarnek i współautorów jest skonceptualizowany jako artykuł wyłącznie metodologiczny – zupełnie abstrahujący od wirtualnego doświadczenia (choć dane zbierano w trakcie badań na strażakach ćwiczących z użyciem symulatora wypadków). W konsekwencji artykuł ten – choć wyrafinowany – jest wyraźnie oderwany od reszty osiągnięcia.

**Oryginalne doniesienia empiryczne.** Ostatnią, najbardziej zróżnicowaną, grupę prac dr. Pawła Strojnego stanowią cztery artykuły, które prezentują oryginalne ustalenia empiryczne o bardziej substancjalnym (w odróżnieniu od artykułów przeglądowych i psychometrycznych) charakterze. Są to teksty: Strojny, P., Strojny, A., & Rębilas, K. (2023). Player involvement as a result of difficulty: An introductory study to test the suitability of the motivational intensity approach to video game research. *PLOS ONE*, 18(3). <https://doi.org/10.1371/journal.pone.0282966>, Strojny, P. M., Dużmańska-Misiarczyk, N., Lipp, N., & Strojny, A. (2020). Moderators of social facilitation effect in virtual reality: Co-presence and realism of virtual agents. *Frontiers in Psychology*, 11. <https://doi.org/10.3389/fpsyg.2020.01252>, Lipp, N., Dużmańska-Misiarczyk, N., Strojny, A., & Strojny, P. (2020). Evoking emotions in virtual reality: Schema activation via a freeze-frame stimulus. *Virtual Reality*, 25(2), 279–292. <https://doi.org/10.1007/s10055-020-00454-6> oraz Czarnek, G., Strojny, P., Strojny, A., & Richter, M. (2019). Assessing engagement during rescue operation simulated in virtual reality: A Psychophysiological Study. *International Journal of Human–Computer Interaction*, 36(5), 464–476. <https://doi.org/10.1080/10447318.2019.1655905>. Trzy spośród czterech tekstów – poza pierwszym z wymienionych, opublikowanym w Plos One – powstały w wyniku *de facto* jednego projektu na tych samych strażakach i studentach służby pożarniczej, co wskazani we wcześniej analizowanym tekście Czarnek i współautorów. Fakt prezentowania danych zebranych w ramach tego samego projektu i na tej samej próbie w szeregu tekstów rodzi pytania (i zastrzeżenia) wobec strategii publikacyjnej, zwłaszcza gdy wyniki te są prezentowane w aż pięciu artykułach.

Pierwszy z tekstów (Strojny, P., Strojny, A., & Rębilas, K. (2023). Player involvement as a result of difficulty: An introductory study to test the suitability of the motivational intensity approach to

video game research. *PLOS ONE*, 18(3)) przynosi bardzo interesujące, lecz – jak już w tytule zauważają autorzy – wstępne, ustalenia na temat możliwości zastosowania teorii intensywności motywacji do analizy ludzkiego zachowania podczas grania w gry. 39 uczestników badania grało w grę polegającą na przedostawaniu się na coraz to wyższe poziomy platformy, a każda z czterech serii była trudniejsza od poprzedniej. Autorzy powołując się na teorię intensywności motywacji, przewidywali, że gdy zadanie stanie się zbyt trudne, zaangażowanie badanych spadnie – oczekiwali więc krzywoliniowej relacji, z wyraźnym obniżeniem się wysiłku w warunku niemożliwym do osiągnięcia, określonym przez nich w tekście dość osobliwie jako „insane”. Badani przeskakiwali na platformy używając spacji i strzałek kierunkowych, liczbę naciśnień spacji na minutę uczyniono wskaźnikiem zaangażowania. Wyniki były zgodne z oczekiwaniami i interesujące. Choć w przypadku miar samoopisowych uczestnicy deklarowali rosnące zaangażowanie i wysiłek, to liczba kliknięć spacji na minutę spadała w warunku „insane”. To ciekawy wynik, jednak wiąże się z nim kilka kłopotów interpretacyjnych, nie zawsze zauważonych przez autorów. Po pierwsze, warunki były prezentowane w identycznej kolejności: rosnącej trudności, co wprawdzie zwiększało trafność ekologiczną (tak wszak jest w grach), ale i mogło powodować zmęczenie i znudzenie badanych. Wprawdzie nie wyjaśnia to różnic między samoopisem a zachowaniem, ale jest czynnikiem, którego rolę należałoby kontrolować w badaniu replikacyjnym. Po drugie, autorzy w tekście informują, że badani mieli maksymalnie 5 minut na każdy z czterech bloków gry (a więc 20 minut łącznie), jednak udostępnione przez nich dane pokazują, że w przypadku pierwszego bloku maksymalny czas spędzony na graniu wynosił 490 sekund (tj. ponad 8 minut), a w przypadku 15 osób (39% całości) przekroczył 5 minut, co rodzi pytania o faktyczną kontrolę w eksperymencie i międzyosobowe zróżnicowanie czasu spędzanego w każdej z czterech faz i jego znaczenie dla wysiłku w ostatniej fazie. Wyniki surowe pokazują również, że badani w ogóle spędzali istotnie mniej czasu na ostatnim z warunków, pytanie więc, czy nie da się tego wzoru wyjaśnić inaczej niż spadkiem inwestowanych zasobów. Wreszcie słabością tekstu – dość oczywistą, biorąc pod uwagę małą próbę – jest brak analiz relacji między subiektywnym a obiektywnym zaangażowaniem; autorzy prezentują osobne wyniki analizy wariancji z powtarzanymi pomiarami, nie wiemy jednak kto i w jakiej mierze przeszacowywał swoje zaangażowanie, kto nie. Tekst ten uznaję za interesujący, jednak o bardzo wstępnym charakterze, wymagającym i replikacji, i lepszej kontroli oraz uwzględnienia alternatywnych wyjaśnień. Sama klasyczna teoria motywacji zastosowana w specyficznym kontekście grania w grę nie stanowi jeszcze znaczącego wkładu w psychologię, choć tekst ten ma heurystyczny potencjał do zapoczątkowania potencjalnie ciekawej linii badań. Słabością jest też brak analizy mocy i ogólnie niewielka próba, która, choć pozwala na główne analizy (zmiany pomiędzy blokami gry przy schemacie z powtarzanymi pomiarami), to jednak nie wystarcza do bardziej pogłębionych analiz.

Drugi z artykułów z oryginalnymi ustaleniami to tekst na temat występowania efektu facylitacji w warunkach wirtualnych (Strojny, P. M., Dużmańska-Misiarczyk, N., Lipp, N., & Strojny, A. (2020). Moderators of social facilitation effect in virtual reality: Co-presence and realism of virtual agents. *Frontiers in Psychology*, 11. <https://doi.org/10.3389/fpsyg.2020.01252>). Artykuł ten jest więc tematycznie związany z przeglądem opublikowanym w *Social Psychological Bulletin*, o którym wspominałem wyżej. Niestety tekst ten z kilku powodów trudno uznać za udany i przekonujący. Pierwszym, zasadniczym problemem artykułu jest jego „niedomocowanie” – 48 strażaków (a finalnie, po wyłączeniach, 44) w schemacie międzygrupowym zostało przypisanych do grupy eksperymentalnej (warunek z gapiami) lub kontrolnej. Nie wiedzieć czemu, zamiast skupić się na kluczowym efekcie głównym: a więc obecności lub nieobecności innych ludzi dla skuteczności działań w warunkach wirtualnych, autorzy prezentują mnóstwo szczegółowych analiz, przekonując ostatecznie, że efekt ten występuje jedynie w warunkach moderowanych przez poczucie realizmu sytuacji i współobecności gapiów. Problem jednak w tym, że analizy moderacji, łączące porządek eksperymentalny z różnicowym (realizm i *co-presence*) oceniane na poziomie każdego badanego, wymagają znacznie większych prób, aby być wiarygodnymi. Dodatkowo tekst ten ma szereg innych – czasem drobnych,

czasem poważniejszych – problemów. Wielość włączonych do analiz zmiennych sprzyja błędowi pierwszego rodzaju. Wykorzystane miary samoopisowe są na ogół kilkuczynnikowe, jednak raportowane są jedynie rzetelności dla globalnych wyników. Kluczowa zmienna zależna (*performance*) jest opisana w sposób uniemożliwiający replikację: brak informacji o jej rzetelności, rozkładzie, sposobie zliczania – zakładam, że była to suma zachowań uznanych za oczekiwane od strażaków, jednak nie jest to jasne. Zupełnie zbędne jest tłumaczenie w tekście, czym jest moderator. Analiza skuteczności manipulacji jest pozbawiona jakichkolwiek testów statystycznych. Relacje prezentowane na wykresach ilustrujących moderacje sugerują znaczenie przypadków odstających, zresztą analiza moderacji przy braku efektu głównego manipulacji oraz wielość zastosowanych miar i analiz nieuchronnie nasuwają zarzut *p-hackingu*. Dodatkowo niedostępność danych oraz inne mankamenty tekstu (np.: Polish pisane małą literą, „to small” zamiast „too small” przy opisie efektów) również nie pozwalają na jego wysoką ocenę. Nie oznacza to oczywiście, że odnotowana moderacja wskazująca na znaczenie obecności innych, ale jedynie przy wysokiej jakości technicznej prezentowanej sytuacji nie ma znaczenia dla wystąpienia efektu facylitacji. Wydaje się to jedynie świadczyć o tym, że realizowany eksperyment (i środowisko VR) był zbyt niedoskonały technicznie, aby badani strażacy byli w stanie w pełni poczuć się tak, jakby sytuacja ta faktycznie miała miejsce. W dobie kryzysu replikacyjnego i wątpliwości związanych z *p-hackingiem* jest praktyką całkowicie naturalną traktowanie tego ustalenia jako wstępnego, prerejestrowanie replikacji, a następnie jej realizacja. W obecnej postaci tekst ten trudno uznać za wnoszący znaczący wkład w psychologię.

Trzeci z artykułów, z Natalią Lipp jako główną autorką (sądząc po kolejności autorów i *author contribution* w tekście), opublikowany w *Virtual Reality*, w trzech badaniach na tych samych strażakach sprawdza, czy i w jakiej mierze proste bodźce (naklejka „dziecko w aucie”, zabawka, fotelik, suknia ślubna) wpływają na odczuwane emocje i stres strażaków. Wyniki tego badania – interpretowane w duchu aktywowania schematów poprzez pokazywanie atrybutów dziecięcości lub ważnej roli (panna młoda) – pokazują, że gdy bodźce te są zauważalne (oczywisty warunek konieczny; bodźce zignorowano w badaniu 1, odnotowano w badaniach 2 i 3) to mają wpływ na stres (badanie z manipulacją elementami dziecięcymi) lub na negatywne emocje (badanie z manipulacją suknią ślubną). Z jednej strony ustalenia te są dość oczywiste (z czego trudno czynić zarzut), zaś schemat badania – międzygrupowy eksperyment w trzech iteracjach, gdzie uczestnicy mieli szansę „nauczyć” się / odgadnąć sens manipulacji nie wydaje się odporne na zarzut niskiej trafności wewnętrznej i zewnętrznej. Małe próby i niestabilność wyników (część znika przy kontroli wielokrotnego testowania) dodatkowo podważają zaufanie do uzyskanych ustaleń. Przede wszystkim jednak wnioskowanie przedstawione w autoreferacie i artykule – zwłaszcza w dyskusji – tj., że jakoby aktywizacja jednego schematu powoduje wzrost stresu, a innego emocji – nie jest właściwie testowana w tym artykule. Taki test wymagałby nie dwóch jednokrotnych porównań między grupą eksperymentalną a kontrolną, ale właśnie testu interakcji. Nie wiedząc, czy w każdej z trzech iteracji badani byli przypisywani do tych samych warunków (eksperymentalnego lub kontrolnego), nie sposób w pełni zinterpretować uzyskanych wyników. Zastosowany schemat w zasadzie uniemożliwia bardziej zaawansowane wnioskowanie, pozostawiając czytelnika z poczuciem dość dużej „oczywistości” efektów zastosowanych manipulacji.

Czwarty i ostatni z tekstów z tej grupy prac zgłoszonych w osiągnięciu dr. P. Strojnego, ponownie z G. Czarnek jako główną autorką (i ponownie M. Richterem jako współautorem) pokazuje spójność między reakcjami fizjologicznymi a deklaracjami strażaków przypisanych do grupy eksperymentalnej i kontrolnej. Tekst ten jest bardzo dobrze napisany, dostępne są dane oraz szczegółowe, dobrze skomentowane skrypty analiz. Ograniczenia tekstu są podobne jak całej serii badań na strażakach – niewielka próba ( $n = 60$ ), przy międzygrupowym schemacie sprawia, że możliwe do wykrycia są jedynie silne efekty. Drugim z ograniczeń jest wspomniany już fakt ciągłego analizowania wyników tego samego eksperymentu, czy też trzech odsłon eksperymentu

prezentowanych jako „badanie podłużne”. Fakt, iż subiektywne zaangażowanie badanych (a przynajmniej średnie różnice między grupą eksperymentalną a kontrolną) oddaje wzór obserwowany dla danych obiektywnych-psychofizjologicznych, można uznać za dobrze świadczący o trafności tych pierwszych (co stoi w opozycji do ustaleń poczynionych w Strojny, P., Strojny, A., & Rębilas, K. (2023). Player involvement as a result of difficulty: An introductory study to test the suitability of the motivational intensity approach to video game research. *PLOS ONE*, 18). Ten artykuł – opublikowany w piśmie o ustalonej renomie – ma szansę na wniesienie wkładu w psychologię, jednak należy zauważyć, że zarówno, jeśli idzie o metodologię, jak i wkład dr. Strojnego, odstaje on wyraźnie od innych prac zgłoszonych w ramach cyklu.

Podsumowując, wśród dziesięciu tekstów zgłoszonych przez dr. Pawła Strojnego jako elementy jego osiągnięcia stanowiącego podstawę starań o nadanie stopnia doktora habilitowanego, trzy prezentują wyniki przeglądów lub metaanaliz, których zakres oraz związek z innymi pracami nie pozwala ich uznać za znaczący wkład, nawet jeśli jeden z nich jest dość często cytowany. Spośród trzech kolejnych artykułów o narzędziowo-wskaźnikowym charakterze, dwa mają dość elementarny charakter adaptacji do polskich warunków niezbyt zaawansowanych skal, które, choć użyteczne dla realizacji badań nie zmieniają naszego rozumienia badanych zjawisk, a same skale zostały stworzone przez innych badaczy. Zaproponowane wskaźniki fizjologiczne są interesujące, ale wymagają dalszych testów. Wreszcie w trzeciej grupie tekstów złożonych do oceny – tym razem czterech – mamy do czynienia z badaniami momentami interesującymi, jednak zrealizowanymi na niewielkich próbach, gdzie trzy teksty powstały na podstawie jednego badania. Spośród tych artykułów najwyżej oceniam tekst Gabrieli Czarnek i współautorów opublikowany w *International Journal of Human-Computer Interaction*. Pozostałe artykuły, nawet gdy zawierają interesujące pomysły, mają też poważne ograniczenia metodologiczne i analityczne, które sprawiają, że można je co najwyżej traktować jako pierwszy krok, wymagający replikacji i dalszych poszukiwań. Co więcej, związek między tymi artykułami – i ogólnie wszystkimi tekstami włączonymi w skład osiągnięcia – jest bardzo luźny: teksty te często bardzo niewiele (poza kontekstem wirtualnego doświadczenia) łączą.

#### KONKLUZJA

Podsumowując, w moim przekonaniu prace przedstawione przez dr. Pawła Strojnego jako osiągnięcie naukowe nie stanowią przekonującej podstawy do poparcia wniosku o nadanie mu stopnia doktora habilitowanego w dziedzinie nauk społecznych w dyscyplinie psychologia. W związku z tym, moja opinia jest negatywna.

A handwritten signature in black ink, reading "Maciej Karwowski". The signature is written in a cursive, flowing style. The first name "Maciej" is written in a slightly larger, more prominent script than the last name "Karwowski". The signature is positioned in the lower right area of the page.