

Recenzja rozprawy doktorskiej mgr Angeliki Olszewskiej  
pt. “Wgląd to coś więcej niż rozwiązywanie zadań. Wykorzystanie wskaźników  
subiektywnego doświadczenia do pomiaru wglądu i opracowania nowego narzędzia”

Przedmiotem recenzowanej rozprawy jest analiza problematyki związanej ze zjawiskiem doświadczania wglądu oraz próba stworzenia nowego, opartego na subiektywnych wskaźnikach, narzędzia do badania charakterystyk wglądu. Autorka przedstawia w niej trzy badania. W pierwszych dwóch prezentuje rozszerzone replikacje niedawnych badań Laukkonena i Tangena (2017), oraz Bar-Hillela i in. (2019). Ostatnie badanie przeprowadzone przez Doktorantkę ma charakter bardziej autorski w tym sensie, że nie stanowi w dużej mierze replikacji wcześniejszych badań, ale jest autorską propozycją rozwiązania konkretnego zagadnienia.

Formalne aspekty przygotowania rozprawy doktorskiej wydają się być bez zarzutu. Praca zawiera 98 stron, opatrzona jest wyczerpującą bibliografią i jest zorganizowana zgodnie ze ogólnymi standardami przygotowywania rozpraw doktorskich. Organizacja rozprawy nieznacznie odbiega w szczegółach od rekomendacji Komitetu Psychologii PAN<sup>1</sup>. Nie stanowi to jednak jakiegoś istotnego uchybienia. W niektórych miejscach zauważyłem nieprecyzyjne sformułowania (np. tytuł Rysunku 3 na stronie 37), które mogą rodzić wątpliwości interpretacyjne i/lub utrudniać podążanie za Autorką, jednak nie jest to problem dotyczący większych fragmentów pracy.

W pierwszej części rozprawy Autorka definiuje i uzasadnia cel pracy wskazując na istotną rolę przyjętej metodologii badań dla rozstrzygnięć teoretycznych. Trudno nie zgodzić się z jej konstatacją, że znikome korelacje między metodami badania tego samego zjawiska (przynajmniej teoretycznie) znacząco utrudniają, a czasami wręcz uniemożliwiają,

---

<sup>1</sup> [https://kompsych.pan.pl/index.php?option=com\\_content&view=article&id=86&Itemid=185](https://kompsych.pan.pl/index.php?option=com_content&view=article&id=86&Itemid=185)

zrozumienie psychologicznych mechanizmów i formułowanie ogólniejszych twierdzeń na jego temat. Zatem, celem pracy w ujęciu Autorki jest analiza tego problemu i wypracowanie metody jego zredukowania poprzez zaproponowanie nowej autorskiej metody pomiaru opisywanego zjawiska, tj. doświadczenia wglądu.

Przedmiotem tej recenzji jest zatem ocena tego, czy cel udało się osiągnąć i ewentualnie w jakim stopniu.

Przegląd teoretyczny dotychczasowych prac dotyczących problematyki wglądu, przedstawiony w pierwszej części rozprawy, jest wyraźnie związany z celem pracy. Autorka skupia się na przeglądzie różnych definicji wglądu oraz porównywaniu dostępnych metod badania tego zjawiska. Przegląd ten jest wystarczająco wyczerpujący, aby zilustrować problemy związane z badaniami nad wglądem i rozumieniem tego zjawiska. Autorka rozpoczyna od przedstawienia siedmiu właściwości charakteryzujących wgląd, wyprowadzonych z analizy dostępnych w literaturze definicji. Robi to w sposób kompetentny i wyczerpujący. Z powodów ogólnych, problematyczne wydaje mi się zdefiniowanie drugiej właściwości przez negację - to tak jakby zdefiniować psychologię, że nie jest matematyką. Nie dowiadujemy się wiele z takiego określenia naszej dyscypliny. Ponadto, Autorka, w mojej opinii, niewystarczającą uwagę przywiązała do analizy relacji między tymi właściwościami oraz ich odniesienia do problemu badawczego zdefiniowanego wcześniej. Pojawia się zatem wiele pytań, tym bardziej, że wymienione zdania podstawowe stanowią pewien drogowskaz w konstruowaniu metody pomiaru wglądu opartej na subiektywnych ocenach. Dla przykładu, czy zjawisko zwane wglądem powinno mieć wszystkie wymienione charakterystyki jednocześnie, czy wystarczy, że posiada jedną, albo jakiś ich podzbiór?

Część teoretyczna rozprawy obejmuje również porównanie i analizę istniejących już metod badania wglądu. W tym fragmencie Autorka uzasadnia konkluzję, że “żadne z wymienionych w Tabeli 2 zadań nie powinno być wykorzystywane do pomiaru wglądu bez uwzględnienia subiektywnego doświadczenia” (s. 25). Rozumiem w zupełności podstawy takiego stwierdzenia, ale równocześnie wydaje mi się, że jego sformułowanie wymagałoby trochę więcej uzasadnienia w rozprawie. Czy miary wglądu, które nie uwzględniają/ły subiektywnego uzasadnienia nie dotyczą wglądu? Czy wyniki badań, w których takie miary zastosowano, nic nie mówią nam o wglądzie? Nie przesądzam odpowiedzi na te pytania, ale sądzę, że powinny one poprzedzać sformułowanie ww. konkluzji.

Wybierając wymiary doświadczenia wglądu Autorka poprawnie wiąże je ze zdaniami podstawowymi V-VII, ale zastanawiające jest pominięcie zdania IV mówiącego o tym, że istotnym elementem istniejących definicji wglądu jest doświadczenie impasu w znajdowaniu rozwiązania zadań. Sugerowałbym ujęcie również tego aspektu w badaniu wymiarów doświadczenia wglądu. Można by spekulować, że pozytywna reakcja następująca po wglądzie jest silnie związana z wcześniejszą frustracją spowodowaną impasem na zasadzie kontrastu, np., że im większy poziom tej frustracji, tym silniejsze pozytywne doznania po wglądzie w rozwiązanie. Mam zatem pytanie do Autorki: czy są i ewentualnie jakie, powody pominięcia tej właściwości wglądu w konstruowaniu miar doświadczeniowych?

Rezygnacja ze skali przyjemności w Badaniu 3 uzasadniona jest koniecznością wyboru między nią a skalą nagłości. Autorka argumentuje, że “(...) *nagłość* ma bardziej ugruntowane podłoże teoretyczne”. Sądzę, że temu twierdzeniu koniecznie powinno towarzyszyć uzasadnienie w postaci analizy literatury.

Wydaje mi się, że uzasadnienie przeprowadzenia Badania 1 powinno być bardziej rozbudowane. Dlaczego Autorka zakłada, że wystąpi transfer z restrukturyzacji percepcyjnej na semantyczną? Nie jest to jasno określone w rozprawie i pozostawia pole dla domysłów czytelnika. Prosiłbym również o zinterpretowanie znaczenia miary spójności wewnętrznej zagadek (alfa Cronbacha), która okazała się zaskakująco wysoka. Moim zdaniem brakuje również opisu kryteriów wyselekcjonowania zagadek słownych w Badaniu 1, które zostało przeprowadzone na podstawie wspomnianych przez Autorkę badań pilotażowych.

Bardzo zastanawiające jest potraktowanie doświadczenia aha! jako stopniowalnego - uczestnicy badania zostali poproszeni o ocenę tego doświadczenia na skali Likerta. Po pierwsze, przynajmniej tradycyjnie rozumiane doświadczenie aha! ma charakter jakościowy, zatem zastosowanie skali porządkowej nie wydaje się tutaj adekwatne. Po drugie, można domniemywać, że doświadczenie to niemal z definicji może przyjąć tylko dwie wartości: albo jest, albo go nie ma. Natężenie odczuwania doświadczenia aha! może być związane z innymi właściwościami osoby lub samego zadania, ale to czy ktoś ma wgląd, czy go nie ma, zdaje się być wydarzeniem zerojedynkowym. Prosiłbym tu o dokładniejsze uzasadnienie pomiaru aha! na skali porządkowej. Podobne wątpliwości mogą dotyczyć również pomiaru nagłości. Literatura sugeruje, że typowe dla wglądu jest nagłe uświadomienie sobie

rozwiązania jakiegoś problemu. Ta nagłość oznacza pewną skokową zmianę w procesie dochodzenia do rozwiązania. Jakie znaczenie ma wobec tego stopniowanie tej nagłości ujęte w postaci skali Likerta? Czy można doświadczyć wglądu raz bardziej, a raz mniej nagle?

Znaczenie i rola parametru szybkości przełączania się percepcyjnego nie zostały wyjaśnione ani w części teoretycznej, ani we wstępie do Badania 1. Z tego względu nie do końca wiadomo, co ewentualnie miałyby oznaczać korelacje tej miary z poprawnością rozwiązywania, czy innymi miarami subiektywnego doświadczenia wglądu. Ponadto, wyjaśnienia wymaga zastosowanie tego parametru w grupie kontrolnej Badania 1, w której prezentowano jednoznaczny percept sześcianu. W oryginalnym badaniu przedstawiano naprzemiennie dwa jednoznaczne percepty zatem w tym wypadku szybkość przełączania byłaby w pewnym sensie miarą koncentracji, czy podtrzymywania uwagi. W przedstawianym Badaniu 1 prezentowano tylko jeden percept w grupie kontrolnej. Wynika z tego pytanie, miarą czego jest szybkość przełączania rejestrowana w grupie kontrolnej w badaniu przedstawionym w rozprawie?

Na stronie 35 zawarte jest stwierdzenie: “Zauważyć jednak można, że w grupie kontrolnej w oryginalnym badaniu osoby badane deklarowały mniej przełączania się perceptu kostki”. Po pierwsze, mniej niż co? A po drugie, dane z tabeli na stronie 36 chyba nie uzasadniają powyższego stwierdzenia - średnia liczba kliknięć oznaczających przełączenie perceptu w oryginalnym badaniu wynosiła 26,10, zaś w próbkach A i B, odpowiednio, 4,50 i 3,75. Ogólnie, dane w Tabeli 1 wydają mi się nieco podejrzane, np., odchylenia standardowe w szybkości przełączania w grupach kontrolnych A i B wynoszą 14,85 oraz 23,11, podczas gdy analogiczna wartość w oryginalnym badaniu wynosi 4. Prosiłbym więc, o szersze wyjaśnienie również tej kwestii.

Jednym z wniosków z Badania 1 jest stwierdzenie, że system monitorowania konfliktu jest zaangażowany przy poprawnym rozwiązywaniu zagadek wymagających wglądu. Na podstawie jakich danych wysunięto ten wniosek? Wszak nie zaobserwowano różnic w poprawności między grupą kontrolną (brak wzbudzenia monitorowania konfliktu) a grupą eksperymentalną (wzbudzenia monitorowania konfliktu) ani w próbce A, ani w B. Ponadto, jeśli sensem manipulacji eksperymentalnej jest wzbudzenie monitorowania konfliktu, to należałoby w badaniu uwzględnić metodę pomiaru skuteczności tej manipulacji w celu potwierdzenia, że była ona skuteczna. Dla jasności dodam, że metoda pomiaru

skuteczności manipulacji eksperymentalnej powinna być niezależna od metody zastosowanej do badania wpływu tej manipulacji na badaną zmienną zależną.

Nie do końca czytelne jest dla mnie przedstawienie we wstępie do Badania 2 teoretycznej zależności między detekcją konfliktu a poprawnością odpowiadania w zadaniach wymagających wglądu oraz w teście refleksyjnego myślenia. Pierwsze pytanie, które się nasuwa brzmi: na czym polega domniemany konflikt w obu rodzajach zadań? O ile pojęcie detekcji konfliktu jest w miarę czytelne w kontekście bimodalnych figur percepcyjnych, o tyle w ww. zadaniach wymagałoby to dalszych wyjaśnień. Niezależnie od powyższego, uważam, że samo sprawdzenie i porównanie subiektywnego doświadczenia towarzyszącego tym rodzajom zadań stanowi wystarczającą motywację do przeprowadzenia Badania 2.

Intrygujące z punktu widzenia osoby, która przez dłuższy czas zajmowała się badaniami z wykorzystaniem RAT, wydaje się również automatyczne kategoryzowanie odpowiedzi w tym teście jako poprawnych lub niepoprawnych. Jakie kryterium poprawności zastosowano w tym algorytmie? Np., jak były kategoryzowane odpowiedzi będące pojęciami bliskoznacznymi do pojęć zdefiniowanych jako poprawne rozwiązania?

Pewnym mankamentem pracy wydaje mi się brak zrelacjonowania w niej zależności między wglądem a kreatywnością. Chociaż związek między nimi wydaje się dobrze ugruntowany w literaturze, to wprowadzenie klasycznego zadania dywergencyjnego w Badaniu 2 wymagałoby choćby szczątkowego uzasadnienia. Na jakiej podstawie (mam na myśli literaturę przedmiotu) Autorka spodziewa się jakiegoś rodzaju związku między poprawnością rozwiązywania zagadek słownych - tudzież jakąkolwiek inną zmienną w badaniu - a płynnością generowania skojarzeń w teście znaczenia wzorów?

Podobnie, głębszego uzasadnienia wymagałoby również wprowadzenie testu szerokości kategorii oraz jego ewentualnych związków z poprawnością rozwiązywania zadań wglądowych oraz wynikami w teście refleksyjnego myślenia. Dla przykładu, w tym drugim teście (szczególnie w wersji słownej) uzyskanie poprawnej odpowiedzi wymaga przecież przełączenia się z intuicyjnie narzucającej się kategorii rozwiązania na inną. Czy szerokość kategorii poznawczych pomaga w takim przełączeniu się? Jeśli tak, to w jaki sposób?

Pewnych trudności interpretacyjnych nastrocza Tabela 13. Rozumiem, że miarą siły doświadczania różnych aspektów wglądu jest suma odpowiedzi dla różnych rodzajów zadań na poszczególnych skalach dla każdej pojedynczej osoby, a Tabela 13 przedstawia medianę tych wartości w grupie osób badanych. Ale czy jest to suma bez rozróżnienia na odpowiedzi poprawne i błędne? Tak się właśnie wydaje (a w każdym razie brakuje informacji na ten temat), a zatem powstaje pytanie o trafność tego wskaźnika. Czego miarą jest np. ocena pewności w sytuacji, kiedy uczestnik badania miał wgląd (czyt. udzielił poprawnej odpowiedzi w zadaniu), a czego miarą jest ocena pewności w zadaniu, w którym wglądu nie było (czyt. nie udzielono poprawnej odpowiedzi)?

W mojej opinii najbardziej nowatorską częścią rozprawy jest próba stworzenia zadania wymagającego wglądu, które opierałoby się na bodźcach percepcyjnych (Badanie 3). Jak podkreśla Autorka, w takim zadaniu dochodziłoby raczej do wyłonienia się struktury (uzyskania w nią wglądu), niż do rekonfiguracji istniejącej przestrzeni problemu. Bardzo doceniam fakt, że Autorka podjęła się tak ambitnego i trudnego zadania. Dostrzegam również to, że przeprowadzenie i opis procedury selekcji i pilotażu wybranych bodźców zostały przeprowadzone bardziej niż poprawnie czy dobrze spełniły swoje cele.

Zauważyłem jednak pewien fundamentalny problem związany ze sposobem rozumowania Autorki na temat ewentualnego znaczenia wyników Badania 3. Na stronie 64 Autorka pisze tak:

*“Udowodnienie, że doświadczenie wglądu podczas rozpoznawania obrazów jest możliwe i równie silne jak przy rozwiązywaniu problemów wskazywałoby na potwierdzenie stanowiska psychologów postaci o wylanianiu się gestaltu, jednocześnie falsyfikowałoby zdanie bazowe V wskazujące, że wgląd jest związany ze zmianą reprezentacji problemu.”*

Po pierwsze, ewentualny brak różnic w doświadczaniu wglądu między zadaniami rozpoznawania obrazów a zadaniami rozwiązywania problemów sam w sobie nie jest potwierdzeniem, że psychologowie postaci mieli/mają rację. Co do zasady, z braku różnic niewiele można wywnioskować. Po drugie, nie dostrzegam w jaki sposób brak ww. różnicy miałby falsyfikować zdanie V. To zdanie dotyczy tylko zadań wglądowych wymagających zmiany reprezentacji problemu. Ponadto, ze zdania V nie wynika logicznie, że wgląd nie może być lub nie jest związany z innymi zadaniami niż zadania wymagające zmiany reprezentacji problemu. W jaki zatem sposób stwierdzenie, że również inne zadania związane są z doświadczeniem wglądu miałyby falsyfikować zdanie V?

Niezależnie od powyższego, w moim przekonaniu samo zbadanie czy i ewentualnie jak, miary doświadczania wglądu funkcjonują w opisanych zadaniach percepcyjnych, oraz która z proponowanych wersji zadania percepcyjnego wywołuje silniejsze doświadczenia związane z wglądem, są wartościowym celem. Chciałbym jednak podkreślić, że postawienie przez Autorkę hipotezy o braku różnic między zadaniem PIT a innymi zadaniami wymagającymi wglądu (H1 w Badaniu 3) jest wysoce nieortodoksyjne, szczególnie jeśli zauważymy, że Autorka wykorzystwała metody statystyczne z repertuaru NHST (null hypothesis significant testing). W tym przypadku należałoby wykorzystać metody bayesowskie, które, w wielkim skrócie, szacują prawdopodobieństwo, że dane potwierdzają hipotezę alternatywną lub zerową. W tym kontekście, stwierdzenie Autorki, że *“Zaobserwowane wyniki jednoznacznie pozwalają na stwierdzenie, że percepcyjne zadania wglądowe charakteryzują się dobrą trafnością zewnętrzną, zarówno zbieżną jak i rozbieżną.”* (str. 78) wydaje się nieco nadmiarowe.

W Badaniu 3B brakuje charakterystyki próby badawczej. Ponadto, sugerowałbym przeformułowanie treści Tabeli 23, w taki sposób, aby czytelnik nie odniósł wrażenia, że zależności przyczynowo skutkowe opisane w Twierdzeniu VI zostały potwierdzone w analizach prowadzonych w modelu korelacyjnym.

Część rozprawy zawierającą dyskusję uzyskanych wyników oraz wskazanie ograniczeń przeprowadzonych badań, uznaję za poprawną i niewymagającą jakiegoś szczegółowego komentarza. W niewielkich fragmentach można dostrzec tendencję do nadmiernego generalizowania, ale nie stanowi to jakiegoś istotnego zarzutu, bowiem widać wyraźnie, że Autorka stara się ograniczyć swoje konkluzje do uzyskanych danych.

Bardzo dużym plusem w mojej opinii jest prerejestracja Badanie 3, oraz udostępnienie szczegółów procedury badawczej i danych w bazie OSF (Badania 1-3).

## **Konkluzja**

Mimo powyższych uwag i wątpliwości, moja ocena rozprawy doktorskiej p. mgr. Angeliki Olszewskiej jest pozytywna. Uznaję, że rozprawa doktorska mgr. Olszewskiej spełnia warunki określone w art. 187 Ustawy z dnia 20 lipca 2018 r. o szkolnictwie wyższym i nauce

(Dz. U. 2018 poz. 1668), dlatego rekomenduję dopuszczenie p. mgr. Angeliki Olszewskiej do dalszych etapów przewodu doktorskiego.

A handwritten signature in black ink, appearing to read 'R. Balas', is positioned above the typed name of the signatory.

Dr hab. Robert Balas, prof. IP PAN