

Maciej Karwowski  
Instytut Psychologii  
Uniwersytet Wrocławski

Warszawa 2023-05-31

**RECENZJA ROZPRAWY DOKTORSKIEJ ANGELIKI OLSZEWSKIEJ  
„WGLĄD TO COŚ WIĘCEJ NIŻ ROZWIĄZYWANIE ZADAŃ. WYKORZYSTANIE WSKAŹNIKÓW  
SUBIEKTYWNEGO DOŚWIADCZENIA DO POMIARU WGLĄDU I  
OPRACOWANIA NOWEGO NARZĘDZIA”**

Praca doktorska magister Angeliki Olszewskiej została przygotowana pod kierunkiem profesora Czesława Nosala i doktor Agaty Sobków na Wydziale Psychologii we Wrocławiu Uniwersytetu SWPS. Jak czytamy we wstępie, praca ta stawia sobie za cel „rozwiniecie wiedzy na temat wglądu i rozwiązanie niektórych problemów związanych z tym zjawiskiem, takich jak brak spójnej definicji i nieliczne narzędzia badawcze”. Cel ten realizowany jest w serii badań, wśród których znajdują się zarówno konceptualne (rozszerzające) replikacje wcześniejszych ustaleń, jak i w pełni autorskie studia empiryczne.

Praca A. Olszewskiej jest studium dojrzałym i wskazującym na znakomitą orientację Autorki w analizowanej tematyce. Wydaje się, że to właśnie świetna znajomość literatury sprawiła, że praca jest też stosunkowo zwięzła – bez bibliografii i suplementu liczy sobie bowiem 90 stron, na których opisane zostały i kluczowe kwestie o charakterze teoretycznym i dość szczegółowo zaprezentowane procedury i rezultaty prowadzonych badań.

Bardzo wysoko oceniam teoretyczną część doktoratu, *de facto* składającą się z dwóch krótkich rozdziałów: „Czym jest wgląd” oraz „Czym mierzyć wgląd”. Autorka zastosowała tu bardzo oryginalny a przy tym trafny zabieg. Zamiast przydługich podręcznikowych relacji dotyczących tego, kto, kiedy i jak wgląd rozumiał bądź badał – co nieuchronnie prowadziłoby do przepisywania podręczników historii psychologii – podjęła ambitną próbę wyizolowania siedmiu kluczowych właściwości wglądu, będących efektem krytycznej analizy wcześniejszych definicji. Właściwości te, to uznanie, że wgląd: (1) jest rezultatem procesu poznawczego, (2) nie jest efektem rozwiązywania problemów metodą prób i błędów (bądź jakkolwiek metodą analityczną), (3) jest skutkiem mentalnej restrukturyzacji, (4) jest poprzedzony impasem, (5) pojawia się nagle, (6) wzbudza pewność co do poprawności rozwiązania oraz (7) jest związany z pozytywnym afektem. Te, tak zwane, „zdania podstawowe”, mogą stanowić coś na kształt formalnej teorii wglądu, a przynajmniej podejścia mającego zadatki takiej teorii.

Choć nie wszystkie wyrażone wyżej postulaty wydają mi się równie uprawnione – na przykład o ile nagłość czy też odrębność wglądu od analitycznego myślenia są niekontrowersyjne, o tyle już konieczność poprzedzania wglądu impasem czy związek z pozytywnym afektem są mniej uniwersalne, to doceniam elegancję oraz rygor w dekomponowaniu wcześniejszych ujęć i przełożenie ich na szereg falsyfikowalnych twierdzeń. Co istotne, w programie badań opisanym na kartach pracy większość z tych „zdań podstawowych” jest poddana dość skrupulatnym – choć nie zawsze w pełni rozstrzygającym – testom.

W warstwie empirycznej praca A. Olszewskiej składa się z trzech badań, choć biorąc pod uwagę pilotaże, trzeba by powiedzieć, że mamy tu raczej do czynienia z trzema liniami badań realizowanymi w ramach tego programu badawczego. Zaletą wszystkich badań jest transparentność Autorki w raportowaniu procedur oraz udostępnianiu danych (choć niestety już nie skryptów analiz, te jednak a dużej mierze są zawarte w opisach procedur) – wszystkie są dostępne w archiwach OSF – a ostatnie z badań zostało również prerejestrowane.

### **Analiza zrealizowanych badań**

Badanie pierwsze stawiało sobie za cel replikację badania Laukkonena i Tangena. Autorka zmodyfikowała oryginalną procedurę, prezentując badanym w schemacie międzygrupowym sześcian Neckera, tj. figurę w warunku konfliktu (dającą możliwość dostrzeżenia dwóch różnych ułożeń sześcianu) oraz standardowy sześcian w warunku bez konfliktu. Uczestnicy dwóch badań – mniejszego i przeprowadzonego na studentach badania 1A oraz większego, zrealizowanego na panelu Prolific badania 1B – przyglądali się sześcianowi, poprzez naciśnięcie spacji odnotowując zmianę percepcji obiektu, następnie zaś rozwiązywali szereg zadań powszechnie uważanych za wymagające wglądu i raportowali swoje wrażenia związane z doznaniem wglądu (pewność, przyjemność, nagłość, doświadczenie olśnienia). Subiektywne wrażenia mierzone były na pięciostopniowych Likertowskich skalach, co jest zgodne z procedurami stosowanymi we wcześniejszych badaniach, choć czasem można by mieć wrażenie (na przykład przy nagłości czy „doświadczeniu aha”), że bardziej trafne i łatwiejsze dla osób badanych byłyby skale zerojedynekowe – coś albo się pojawiło, albo nie.

Badanie 1A – jak słusznie zauważa Autorka – było statystycznie „niedomocowane”: przy 69 uczestnikach (po wyłączeniu 18 osób) i w schemacie międzygrupowym szanse wykrycia relatywnie słabych efektów – a takie powinny być oczekiwane – były niewielkie.

Badanie to pokazało, że w warunku z konfliktem generowanym przez wieloznaczną figurę badani przełączali się znacznie szybciej niż w warunku bez konfliktu (co jest miarą skuteczności manipulacji), brak było natomiast różnic międzygrupowych w zakresie poprawności rozwiązyanych zadań wglądowych prezentowanych po ekspozycji figury. Grupy nie różniły się również subiektywnymi wrażeniami dotyczącymi wglądu. To badanie ze względu na małą próbę jest niekonkluzywne, z czego zresztą Autorka zdaje sobie sprawę.

W badaniu 1B wzięły udział 184 osoby (po wykluczeniach), tak więc tym razem próba pozwalała na wykrycie mniejszych różnic międzygrupowych. Jednak również to badanie nie miało dostatecznej mocy, aby wykryć efekt z oryginalnego badania ( $d = 0,26$ ). Również i tym razem przełączanie się badanych było istotnie szybsze w warunku z konfliktem, jednak i tym razem nie odnotowano różnic w zakresie poprawności problemów wglądowych. Okazało się natomiast, że badani oglądający „niejednoznaczny sześcian” deklarowali większe nasilenie „doświadczenia aha” przy rozwiązywaniu prezentowanych po nim zadań. To subiektywne wrażenie nie przekładało się jednak na rzeczywistą poprawność rozwiązań.

Osobne analizy korelacji dla wyników obu badań pokazały dość spójny wzór relacji – fenomenologiczne relacje były ze sobą silnie pozytywnie skorelowane – na poziomie, który wręcz sugerowałby posługiwanie się jednym wskaźnikiem, co zresztą Autorka czyni w jednym z kolejnych badań. Co ciekawe, w obu badaniach pewność, przyjemność i nagłość doświadczeń wglądu były pozytywnie związane z trafnością odpowiedzi w zadaniach wglądowych. Mniej systematyczne były związki między trafnością odpowiedzi a „doświadczeniem aha” – w badaniu 1A relacja ta była istotna i silna ( $r = 0,48$ ), w badaniu 1B nie różniła się istotnie od 0 ( $r = 0,02$ ).

Ogółem oba badania pokazują, że subiektywnie raportowane wrażenia doznawane w trakcie rozwiązywania zadań potencjalnie wymagających wglądu są pozytywnie związane z poprawnością rozwiązań. Nie udało się natomiast wykazać różnic we wzbudzaniu i transferze wglądu między sytuacją konfliktu i braku takiego konfliktu a kolejnymi zadaniami. Niewystarczająca moc badań, zwłaszcza badania 1A, skłania do pytania, dlaczego Autorka nie zdecydowała się na ich połączenie – czy to w postaci reanalizy zagregowanych danych surowych (nawet przy sprawdzeniu ewentualnych moderujących efektów zmiennej „badanie”) czy też metaanalizy. Pozwoliłoby to na uzyskanie większej mocy, a więc i konkluzyjności uzyskanych rezultatów. Statystyki opisowe w obu badaniach są zbliżone, nie

wskazując na znaczący efekt próby (studenci versus panel), tak więc warto by taką procedurę przeprowadzić.

Innym pytaniem, jakie rodzi się pod wpływem lektury obu pierwszych badań, jest kwestia relacji poprawności rozwiązania zadań wglądowych i raportowanej fenomenologii procesu. Korelacje te bywają dość silne i o ile w przypadku relacji pewność-poprawność nie jest to specjalnie zaskakujące, to już pozytywne związki poprawności z przyjemnością i nagłością są znacznie ciekawsze teoretycznie. Szkoda, że bazując na swoich danych, Autorka nie spróbowała analitycznie przetestować jednego ze swoich „zdań podstawowych”, sprawdzając, czy i jak często problemy wglądowe mogły być rozwiązywane, sięgając po metody *stricte* analityczne lub metodę prób i błędów. Dysponując danymi z badań 1a i 1B stosunkowo łatwo byłoby bowiem stworzyć zmienną, gdzie wysokiej poprawności rozwiązania zadań towarzyszyłoby niskie (np.: poniżej 3 na skali 1-5) „poczucie aha” (lub innych fenomenologicznych wskaźników wglądu). A jeśli badani byli w stanie problem rozwiązać, a nie towarzyszyły im przy tym wrażenia właściwe dla wglądu, to można by wnioskować (pośrednio, ale jednak) o „niewglądowym” charakterze samego procesu w ich przypadku. Taka analiza jest możliwa na danych z badania 1A i 1B.

Druga z serii zrealizowanych badań eksplorowała związki między skutecznością rozwiązywania zadań wglądowych (głównie zagadek) a zadań mierzących myślenie analityczne. W badaniu 2A ( $N = 132$ ), uczestnicy panelu Prolific rozwiązywali 12 zagadek w założeniu wymagających wglądu, polską wersję Testu Odległych Skojarzeń (RAT), test kreatywności (*pattern line meaning task*), test szerokości kategorii, Test Refleksyjnego Myślenia w wersji klasycznej (mocno matematycznej) oraz werbalnej, a także test inteligencji płynnej (ICAR) i cztery sylogizmy logiczne. Biorąc pod uwagę liczbę zastosowanych miar oraz ich charakter badanie trzeba uznać za niezwykle wymagające dla osób badanych. Brak informacji o czasie trwania badania, ale wolno podejrzewać, że trwało 40-60 minut. Tak długi czas rodzi oczywiście pytania o jakość danych, rzetelność miar i uważność badanych, choć wątki te nie są dyskutowane w rozprawie. Autorka wspomina również, że niektóre rezultaty były skalowane zgodnie z metodologią probabilistycznej teorii testu: a ściślej – modelu Rascha, choć niestety nie podaje żadnych szczegółów na temat stworzonych modeli, ich dopasowania i szczegółowych parametrów. Podobnie nie zawsze raportowane są miary rzetelności zastosowanych narzędzi, a ma to duże znaczenie dla interpretacji uzyskanych

współczynników korelacji – wiemy bowiem, że są one zaniżone ze względu na niepełną rzetelność miar, nie wiemy jednak w jakim stopniu.

Wyniki analizowane są w postaci macierzy korelacji, następnie zaś porównań współczynników korelacji pomiędzy różnymi, teoretycznie uzasadnionymi kategoriami. Patrząc na dane zaprezentowane w Tabeli 10, wydaje się – choć brak tu formalnego testu – że w istocie widzimy tu dwa czynniki. Pierwszy to właśnie wgląd – z relatywnie silnie skorelowanymi miarami zagadek, RAT oraz dwóch wariantów testu refleksyjnego myślenia. Z kolei miary kreatywności i inteligencji były słabiej związane z zadaniami wglądowymi, choć od tej reguły są wyjątki (np.: korelacja  $r = 0,40$  między ICAR i CRT). Zapewne niewielka próba była powodem, dla którego wyniki nie zostały przeanalizowane w schemacie zmiennych latentnych – a szkoda, bo pozwoliłoby to i na bardziej elegancką, i statystycznie odporną analizę. Samo porównywanie współczynników korelacji jest wprawdzie akceptowalne (i świetnie, że Autorka sięgnęła po pakiet cocor, który kontroluje fakt, że różne korelacje pochodzą z tej samej próby), jednak nie pozwala na daleko idące wnioski.

W Badaniu 2B ( $N = 137$ , studenci), A. Olszewska sprawdzała, czy i w jakiej mierze badanym towarzyszy doświadczenie wglądu podczas rozwiązywania różnego rodzaju problemów i zagadek – tak „wglądowych”, jak i sylogizmów logicznych. Tym razem, w odróżnieniu od badania 1B, posłużono się ogólnym wskaźnikiem wglądu (choć nie wiemy, jaka jest jego rzetelność). Nieparametryczny wariant ANOVY z powtarzanymi pomiarami pokazał wprawdzie istotny statystycznie efekt doświadczanego wglądu między zadaniami, kierunek różnic nie był jednak zgodny z oczekiwaniami. Wrażenia badanych pokazały, że sylogizmy są równie wglądowe, jak CRT, ten zaś nie różnił się od zagadek wglądowych. Paradoksalnie zagadki: w założeniu wymagające wglądu, okazały się go wzbudzać w najmniejszym stopniu. Autorka interpretuje ten wynik, odwołując się do badań Weisberga i hipotezy, że wszystkie wykorzystane zadania mogą pierwotnie angażować myślenie analityczne. To możliwe, choć nie wiemy, czy tak jest. Wydaje się jednak, że powodem uzyskanych wyników może być nieoptymalny pomiar fenomenologii wglądu. Pięciostopniowa skala wykorzystana w często zerojedynekowej sytuacji odczuwania bądź nieodczuwania czegoś może wprowadzać wariancję błędu i szum. Potrzebujemy kolejnych badań, które pozwoliłyby przetestować zmodyfikowane procedury eksperymentalne.

Trzecia linia badań Angeliki Olszewskiej poświęcona jest stworzeniu autorskiego narzędzia (test PIT) mierzącego percepcyjny wgląd, nawiązując tym samym z jednej strony do klasycznych teorii psychologów postaci, z drugiej zaś do badań Malcolma Westcotta nad naturą intuicji. W tej linii badań wgląd nie jest ujmowany jako restrukturyzacja sytuacji problemowej, ale właściwość percepcji, zgodnie z klasycznymi pracami psychologów postaci. Olszewska opisuje proces konstruowania testu, gdzie sięgając po procedurę rozmazywania obrazów oraz odkrywania ich kawałków (jak u Westcotta), dobrano obiekty do ostatecznej wersji testu (dwa badania pilotażowe). Ta procedura jest klarownie opisana i przeprowadzona zgodnie z regułami sztuki.

W badaniu 3A ( $N = 201$ , panel Prolific) sprawdzano, czy odczucia wglądu („doświadczenie aha”, nagłość, pewność, przyjemność, pobudzenie) różnią się między wariantami wykorzystanych sposobów „ukrywania” obiektów (odkrywające się pola vs. zamazany obraz). Okazało się, że kolejno odkrywające się pola skutkowały wyższym nasileniem „doświadczenia aha”, nagłości, przyjemności i pobudzenia, natomiast niższym wrażeniem pewności niż rozmazany obraz.

W prerejestrowanym badaniu 3B (zakładam, że zrealizowanym na tej samej próbie, co badanie 3A, nie jest to jednak jasne), Autorka sprawdzała korelacje nowostworzonego testu z wcześniej stosowanymi zadaniami wymagającymi i niewymagającymi wglądu oraz porównywała nasilenie doświadczeń wglądowych podczas rozwiązywania tych różnych zadań. Tym razem okazało się, że „zadania wglądowe” – w tym nowy test figuralny – wzbudzają wyższy poziom wrażeń wglądowych niż zadania analityczne. Trzeba też zauważyć, że niektóre hipotezy tego badania – zakładające, że nowy test będzie wzbudzał podobny poziom wrażeń wglądowych jak test RAT i problemy „zapałczane” miały charakter hipotez zerowych. Na odnotowanie zasługuje sięgnięcie przez Autorkę po test ekwiwalencji, który pozwalał stwierdzić, czy hipotezy zerowe są zgodne z danymi. Na podstawie wyników tego badania można uznać, że nowy test wzbudza podobne odczucia wglądu jak klasyczne zadania wglądowe, co dobrze świadczy o jego trafności – przynajmniej w odniesieniu do fenomenologii wglądu. Nieco mniej przekonująco wygląda trafność, gdy wyniki nowego testu skorelowano z poprawnością rozwiązywania innych zadań – tak wglądowych, jak i niewglądowych. Wyniki w PIT okazały się niezwiązane z rezultatami uzyskanymi w teście kreatywności, inteligencji i efektywnością rozwiązywania analogii werbalnych. Z kolei związki z zadaniami wglądowymi były bardziej zróżnicowane: od zerowych przy zagadkach ( $r =$

0,058), po słabe, istotne korelacje z problemami zapalczanymi ( $r = 0,19$ ) i testem RAT ( $r = 0,25$ ). I w tym miejscu nieco brakuje wyjścia poza proste analizy korelacyjne (nie wspominając o braku wszystkich współczynników rzetelności). Model strukturalny z latentnym czynnikiem PIT skorelowanym z pozostałymi miarami mógłby lepiej pokazać faktyczne związki, skorygowane o błąd pomiaru. Autorka jest przy tym nieco nazbyt optymistyczna (albo niedostatecznie samokrytyczna) pisząc w podsumowaniu, że: „zaobserwowane wyniki jednoznacznie pozwalają na stwierdzenie, że percepcyjne zadania wglądowe charakteryzują się dobrą trafnością zewnętrzną, zarówno zbieżną jak i rozbieżną”. Owszem, pewne argumenty za tym stwierdzeniem wyraźnie widać w danych, trudno jednak powiedzieć, że korelacje rzędu 0,19 czy 0,25 pozwalają na „jednoznaczne” potwierdzenie trafności zbieżnej.

W Dyskusji Angelika Olszewska nie tylko podsumowuje i krytycznie analizuje wyniki swoich badań w kontekście literatury przedmiotu, ale – przede wszystkim – wraca do sformułowanych we wstępie siedmiu „zdań podstawowych”. Dwa z nich („wgląd nie jest efektem rozwiązania problemu metodą prób i błędów” oraz „wgląd jest poprzedzony impasem”) nie były w Jej programie badawczym wprost testowane, choć jak pisałem wyżej, można sobie wyobrazić pośredni test pierwszego twierdzenia na bazie zebranych danych. Jedno ze zdań: „wgląd jest efektem mentalnej restrukturyzacji” zostało sfalsyfikowane poprzez wykazanie, że wglądowe zadanie percepcyjne powoduje podobne nasilenie wrażeń wglądowych jak rozwiązywanie zadań wymagających wglądu. Trzy spośród zdań podstawowych: „wgląd pojawia się nagle”, „wzbudza pewność co do poprawności zadania” oraz „jest związany z pozytywnym afektem” zostały potwierdzone. Potwierdzenia te uznałbym jednak za bardzo wstępną korroborację tych twierdzeń, to zaś za sprawą faktu, że ich test oparto wyłącznie na deklaracjach badanych; dodatkowo deklaracjach i ocenach udzielanych w sposób blokowy (tj. kilka kryteriów ocenianych jedno po drugim), co mogło skutkować efektem wariancji metody i stylu odpowiadania. Stąd też wciąż konieczne są dalsze badania, zwłaszcza eksperymentalne, które pozwolą na bardziej odporne metodologicznie testy tego rozumowania.

Wreszcie w odniesieniu do pierwszego z twierdzeń: „wgląd jest rezultatem procesu poznawczego”, wyniki przynieszone przez poszczególne badania zostały przez Autorkę uznane za niejednoznaczne. Wydaje się jednak, że poznawczy charakter wglądu jest trudny do kontestacji – ewentualne wątpliwości pojawiają się chyba jedynie wówczas, gdy nada się

bardzo specyficzną (tj. wąską bądź szeroką) szeroką konotację kategorii „procesu poznawczego”.

### **KONKLUZJA**

Podsumowując to, co zostało powiedziane wyżej, praca doktorska magister Angeliki Olszewskiej jest dojrzałym i ambitnym studium wglądu. Trzy linie badań, poprzedzone szeregiem pilotaży, oryginalne pomysły badawcze, przede wszystkim jednak ścisłe powiązanie realizowanych badań z szeregiem wydedukowanych z literatury założeń na temat natury wglądu, pokazują dużą dojrzałość teoretyczną i metodologiczną Autorki. W związku z tym stwierdzam, że praca A. Olszewskiej spełnia wymogi ustawowe i opowiadam się za dopuszczeniem Doktorantki do dalszych etapów procesu doktorskiego.

Maciej Karwowski