

**Uchwała Komisji habilitacyjnej
z dnia 29 września 2022 r.
powołanej przez Radę Naukową Dyscypliny Psychologia
Uniwersytetu im. Adama Mickiewicza w Poznaniu
w postępowaniu w sprawie nadania stopnia doktora habilitowanego
w dziedzinie nauk społecznych w dyscyplinie psychologia
wszczętym na wniosek dr. Pawła Kleki**

§ 1

Komisja habilitacyjna, powołana przez Radę Naukową Dyscypliny Psychologia Uniwersytetu im. Adama Mickiewicza w Poznaniu uchwałą z dnia 10 maja 2022 r., działając na podstawie art. 221 ust. 10 ustawy z dnia 20 lipca 2018 r. Prawo o szkolnictwie wyższym i nauce (Dz. U. z 2022 r. poz. 574 z późn. zm.), po zapoznaniu się z recenzjami i dokumentacją wniosku, stwierdza że aktywność naukowa oraz osiągnięcia naukowe stanowią istotny wkład w rozwój dyscypliny naukowej psychologia i wyraża pozytywną opinię w sprawie nadania dr. Pawłowi Klece stopnia doktora habilitowanego w dziedzinie nauk społecznych w dyscyplinie psychologia, uznając spełnienie przesłanek warunkujących nadanie stopnia doktora habilitowanego, o których mowa w art. 219 ust. 1 pkt. 1-3 wskazanej ustawy.

UZASADNIENIE

Podstawa prawna

Postępowanie w sprawie nadania dr. Pawłowi Klece stopnia doktora habilitowanego w dziedzinie nauk społecznych w dyscyplinie psychologia prowadzone było zgodnie z przepisami ujętymi w:

- ❑ *Ustawie z dnia 20 lipca 2018 r. Prawo o szkolnictwie wyższym i nauce* (Dz. U. z 2022 r. poz. 574 z późn. zm.);
- ❑ *Uchwale nr 134/2020/2021 Senatu Uniwersytetu im. Adama Mickiewicza w Poznaniu z dnia 28 czerwca 2021 r. w sprawie sposobu postępowania w sprawie nadania stopnia doktora habilitowanego.*

Komisja habilitacyjna

Komisja powołana została w trybie przedstawionym w przywołanej powyżej Ustawie i Uchwale Senatu Uniwersytetu im. Adama Mickiewicza z Poznaniu. W jej skład wchodziły następujące osoby:

- prof. dr hab. Bogdan Zawadzki (Uniwersytet Warszawski) – przewodniczący,
- dr hab. Łukasz Kaczmarek (Uniwersytet im. Adama Mickiewicza z Poznaniu) – sekretarz,
- dr hab. Oleg Gorbaniuk (Katolicki Uniwersytet Lubelski Jana Pawła II) – recenzent,
- dr hab. Paweł Jurek (Uniwersytet Gdański) – recenzent,
- prof. dr hab. Romuald Polczyk (Uniwersytet Jagielloński w Krakowie) – recenzent,
- dr hab. Marcin Zajenkowski (Uniwersytet Warszawski) – recenzent,
- dr hab. Maciej Hanczakowski (im. Adama Mickiewicza z Poznaniu) – członek Komisji.

Przebieg postępowania

Komisja habilitacyjna wzięła pod uwagę: I. Spełnienie wymogów ustawowych, tj. 1. Posiadanie doktoratu z określonej dziedziny; 2. Osiągnięcie naukowe; 3. Aktywność naukową. II. Omówiła także: dodatkowo osiągnięcia dydaktyczne i organizacyjne.

Podstawą sformułowanych przez komisję ocen w.w. dokonań były:

- ✓ osiągnięcie naukowe Habilitanta, którym jest monografia: „*Trafność pomiaru w praktyce psychometrycznej. Badanie empiryczne*” (2021, Poznań: Wydawnictwo Nauk Społecznych i Humanistycznych Uniwersytetu im. Adama Mickiewicza, ss. 124);
- ✓ zestawiony przez Habilitanta wykaz publikacji i pozostałych form aktywności naukowej, dydaktycznej i organizacyjnej;
- ✓ sporządzony przez Habilitanta autoreferat;
- ✓ sporządzone przez Recenzentów oceny osiągnięć naukowych, dydaktycznych i organizacyjnych Habilitanta.

Posiedzenie komisji odbyło się dnia 29 września 2022 r. – zdalnie, za pomocą środków elektronicznych (platformy MSTeams). Posiedzenie komisji przebiegało w dwóch fazach. W pierwszej fazie [kolokwium habilitacyjne] komisja poprosiła Habilitanta, aby przedstawił swoje osiągnięcie habilitacyjne oraz dorobek naukowy, zaś później członkowie komisji zadawali pytania, na które Habilitant udzielił odpowiedzi. Ta faza postępowania została przez sekretarza komisji nagrana i zaprotokołowana. Po dyskusji nad wynikami badań, przedstawionych w monografii, komisja zaprosiła Habilitanta na fragment drugiej części posiedzenia, w celu przedstawienia danych z dodatkowych analiz wyników badawczych.

Dalsza część drugiej fazy postępowania komisji nie była dostępna Habilitantowi. Był jednak poinformowany przez przewodniczącego, że po zakończeniu pracy komisji uzyska informację o wynikach głosowania nad uchwałą. W drugiej fazie postępowania komisja dokonała: (a) szczegółowej oceny osiągnięcia naukowego oraz dodatkowego dorobku Habilitanta, (b) omówiła dodatkowo Jego dorobek organizacyjny, dydaktyczny i związany z popularyzacją nauki, a następnie podjęła uchwałę w przedmiocie wystąpienia do Rady Dyscypliny Naukowej Psychologia Uniwersytetu im. Adama Mickiewicza w Poznaniu o nadanie Habilitantowi stopnia naukowego doktora habilitowanego. Głosowanie komisji przeprowadzono w trybie jawnym - Habilitant nie wnioskował o przeprowadzenie głosowania w trybie tajnym. Głosowanie w sprawie nadania stopnia doktora habilitowanego odbyło się w trybie większości bezwzględnej, zgodnie z przepisami Uchwały Senatu Uniwersytetu im. Adama Mickiewicza w Poznaniu. Podjęta przez Komisję habilitacyjną uchwała ma dla Rady Dyscypliny Psychologia Uniwersytetu im. Adama Mickiewicza status opinii. Ta faza postępowania również została przez sekretarza komisji nagrana oraz zaprotokołowana.

Sylwetka Habilitanta

Dr Paweł Kleka uzyskał stopień doktora w roku 2013 na Uniwersytecie im. Adama Mickiewicza w Poznaniu na podstawie rozprawy „*Zastosowanie teorii odpowiadania na pozycje testowe (IRT) do tworzenia skróconych wersji testów i kwestionariuszy psychologicznych*”. W latach 2001-2013 był zatrudniony na UAM na stanowisku specjalisty naukowo-badawczego, a od 2013 roku - na stanowisku adiunkta. Od 2021 roku pracuje na 1/2 etatu na stanowisku adiunkta w Instytucie Psychologii Wydziału Nauk Społecznych Uniwersytetu Zielonogórskiego. Współpracuje również z innymi

instytucjami naukowymi, realizując projekty badawcze i uczestnicząc w programach współpracy międzynarodowej. Zainteresowania naukowe dr. Pawła Kleki koncentrują się na zagadnieniu psychometrii i pomiaru w psychologii.

Recenzje

Osiągnięcie i dorobek naukowy Habilitanta uzyskały jedną negatywną oraz trzy pozytywne oceny. Wszystkie recenzje zawierały elementy krytyczne, jak też opinie pozytywne, wystawione przez Recenzentów, których kompetencje naukowe odpowiadają kryteriom Ustawy *Prawo o szkolnictwie wyższym i nauce*.

Dr hab. Oleg Gorbaniuk (wyciąg z recenzji)

Ocena osiągnięcia naukowego: Jako główne osiągnięcie naukowe, stanowiące podstawę wniosku habilitacyjnego, dr Paweł Kleka wskazał monografię „*Trafność pomiaru w praktyce psychometrycznej. Badanie empiryczne*” (124 strony). [...] Praca składa się z czterech rozdziałów, każdy z których realizuje inny cel, tworząc na poziomie celów logiczną a zarazem różnorodną/bogatą całość. W pierwszym rozdziale Habilitant przedstawił problematykę pomiaru, w szczególności krytykę klasycznego podejścia oraz analizę warunków, których spełnienie pozwala na pomiar ilościowy konstruktów psychologicznych. Kolejne trzy rozdziały - po sformułowaniu problemu badawczego - zawierają badania/analizy własne Habilitanta. Cechą wyróżniającą tych analiz jest odmienny charakter danych stanowiących ich podstawę. Są one odpowiednio efektem symulacji komputerowej, metaanalizy publikacji oraz badań eksperymentalnych. Świadczy to z jednej strony o przemyślanej koncepcji pracy, z drugiej o kompetencjach analitycznych Autora. Faktycznym celem badań przedstawionych w drugim rozdziale było przetestowanie wrażliwości wskaźników trafności na trzy rodzaje deformacji danych w stosunku do danych rzeczywistych na bazie sztucznie wygenerowanych danych. [...] Problem wpływu samowiedzy / zdolności do introspekcji na jakość udzielanych odpowiedzi i trafność narzędzia sam w sobie jest ważny z punktu widzenia metodologicznego. Jednak zapowiedź jego zgłębienia w postaci tytułu rozdziału jest nieadekwatna w stosunku do treści, która opiera się na sztucznych danych, jedynie symulujących różne możliwe sytuacje, a sama symulacja bazuje na czysto hipotetycznych przewidywaniach, jak zdolność do introspekcji może zniekształcać zebrane dane. [...] Powstaje zatem problem trafności operacjonalizacji (jej specyficzności dla badanej zmiennej), czyli czegoś, co sam Autor niejednokrotnie krytykuje. Biorąc pod uwagę specyfikę użytych danych i wykonanych analiz bardziej adekwatne byłoby odwrócić kierunek rozumowania: przetestowano kilka możliwych deformacji zebranych danych w stosunku do prawdziwych, których wystąpienie może zmniejszać trafność narzędzia i na które tradycyjne wskaźniki trafności mogą być niewystarczająco wrażliwe. Jednym z wielu czynników, które może powodować takie deformacje, jest poziom samowiedzy lub zdolność do introspekcji. Tu przydałby się porządny a zarazem zwięzły przegląd literatury przedmiotu uzasadniający postulowany charakter deformacji danych przez samowiedzę/introspekcję. Przeprowadzenie własnych badań empirycznych (eksperymentalnych) i zebranie realnych danych poprzedzony systematycznym przeglądem literatury na temat deformującego wpływu czynników sprzężonych z metodami samoopisowymi na wyniki badań psychologicznych byłoby znaczącym posunięciem uzupełniającym zaprezentowane symulacje. Podsumowując drugi rozdział można stwierdzić, że zawiera interesujące symulacje komputerowe pokazujące ograniczenia często stosowanego wskaźnika trafności, ale obietnica dostarczenia wiedzy na temat wpływu metod samoopisowych na wyniki badań psychologicznych wynikająca z tytułu rozdziału została niezrealizowana, zatem tytuł w

mojej ocenie jest nieadekwatny do treści. Trzeci rozdział pracy zawiera analizę i ocenę stanu wykorzystywanych wskaźników trafności na bazie metaanalizy recenzowanych publikacji w latach 1999-2018, dokonaną przez Habilitanta. [...] Pomysł zrobienia metaanalizy w podanym odcinku czas jest godny pochwały, aczkolwiek jego realizacja budzi już zastrzeżenia natury metodologicznej. Analiza zebranych danych ma charakter opisowy i daje fragmentaryczny wgląd w analizowany odcinek czasu, ukazując dominację współczynnika korelacji i analizy czynnikowej jako głównych metod weryfikacji trafności narzędzi psychologicznych. [...] Aby zwiększyć wiarygodność takich analiz należałoby sięgnąć po znacznie większą próbę publikacji (Habilitant uwzględnił tylko 101) zdecydowanie dbając jednocześnie o jej reprezentatywność: w obecnej próbie 64% ukazało się 2018 roku, co samo w sobie powinno wzbudzić wątpliwości Autora odnośnie reprezentatywności materiału; nie uwzględniono płatnych artykułów, które raczej reprezentują ponadprzeciętną jakość, a do których można łatwo dotrzeć zwracając się wprost do autorów. [...] Ogólnie w całej pracy Habilitant wyraża przekonanie, że dominujące aktualnie w badaniach psychologicznych wskaźniki trafności są mało wiarygodnymi jej miarami. Aby potwierdzić zasadność swojego krytycznego/sceptycznego nastawienia (wspieranego przez cytowanie metaanaliz literatury przedmiotu innych autorów) Habilitant zaprojektował badania opisane w rozdziale czwartym, które w znacznym stopniu sprzyjają potwierdzeniu takiego oczekiwania poprzez wybór metody badań obniżającej najczęściej jakość zbieranych danych: (1) jako metodę zbierania danych wybrano ankietę internetową; (2) uzyskano bardzo niski stopień zwrotu wypełnionych kwestionariuszy; (3) kwestionariusz, którego treścią manipulowano, był jednowymiarowy (wielowymiarowy pozwoliłby na znacznie bogatsze analizy, szczególnie w kontekście wieloczynnikowej CFA); (4) pozycje kwestionariusze nie były wzajemnie odwrócone; (5) respondenci byli zmuszani do udzielania odpowiedzi (nie było możliwości udzielenia odpowiedzi „nie wiem”, „trudno powiedzieć”, „nie dotyczy”); (6) poprzez umieszczenie pytań metryczkowych na początku badania respondent, nie znając jeszcze treści pytań, był przynajmniej częściowo pozbawiany anonimowości, co wpływa na jego postawę wobec badań; (7) opis procedury badań nie wskazuje, aby zastosowana metoda badań motywowała respondentów do rzetelnego wypełniania kwestionariuszy; (8) zabrakło definicji populacji, (9) liczebność próby nie była ustalona w oparciu o analizę mocy, (10) nie przedstawiono dowodów na spełnienie warunków II zasady randomizacji (szczególnie w kontekście przerwania wypełnienia kwestionariusza przez 1/4 respondentów), (11) analizowany jest efekt główny w układzie niezrównoważonym itd. Wypunktowałem ważniejsze słabe punkty i wady metodologiczne zrealizowanych przez Autora badań empirycznych. [...] Habilitant w swoich wnioskach/podsumowaniu do nich nie ogranicza się, traktując zrealizowane badania raczej jako standard w psychologii rozszerzając tym samym w sposób nieuzasadniony zakres generalizacji wniosków z badań własnych. Warto odnotować, że w dyskusji wyników Habilitant nie przyjmuje postawy samokrytycznej i nie przedstawia w zasadzie ograniczeń badań własnych. [...] W szczególności jednym z problemów, który należałoby zbadać, jest odpowiedź na pytanie, w jaki sposób jakość metody badań (operacjonalizacja, próba, procedura) wpływa na wskaźniki trafności i ich wartość informacyjną. W moim przekonaniu podjęty przez Habilitanta problem z uwagi na swój potencjał wymaga szerszej zaplanowanego programu badań empirycznych, szczególnie jeżeli wynik tych badań ma stanowić główne osiągnięcie Habilitanta i wносить znaczący wkład w psychologię/metodologię odpowiadający randze przewodu habilitacyjnego. [...] Bez względu na moją krytyczną ocenę badań przedstawionych w rozdziale czwartym trzeba zaznaczyć ponownie, że one zachowują

swoją wartość naukową - w pewnym zakresie przedstawionych analiz - w odniesieniu do badań o niskiej jakości metodologicznej, w których autorzy mogą także stawiać sobie za cel walidację narzędzi psychologicznych. Badania eksperymentalne Habilitanta wyraźnie dowodzą, że przy metodzie badań o niskiej jakości możemy uzyskiwać przyzwoite klasyczne wskaźniki psychometryczne narzędzi, które są pozbawione sensu. Prawdopodobnie zawdzięczamy to bezrefleksyjnemu zaznaczaniu tej samej odpowiedzi przez respondenta, który nie ma motywacji lub merytorycznego powodu (w przypadku bezsensownych pozycji) do ich różnicowania. Swoją drogą warto byłoby wykonać analizy umożliwiające wgląd w style zaznaczania, analizując profile odpowiedzi i porównując średnie wyniki w grupach eksperymentalnych, które stanowiłyby dowód, że respondenci udzielają tych samych odpowiedzi bez względu na wymianę kluczowego pojęcia. [...] Recenzowana praca habilitacyjna objętościowo nie jest obszerna i z pewnością nie jest przeciążona przeglądem literatury przedmiotu ukazującego aktualny stan wiedzy w zakresie omawianych zagadnień. Osobiście w kilku miejscach, gdzie Autor wprost deklaruje, że nie zamierza rozwijać zagadnienia lub dokonywać przeglądu literatury, oczekiwałem takiego przeglądu. W różnych miejscach pracy (wprowadzeniu teoretycznym, interpretacji wyników badań a szczególnie w dyskusji/podsumowaniu wyników) możemy ponadto spotkać obszerne fragmenty, w których Habilitant dzieli się obserwacjami (w tym na bazie własnych badań) lub doświadczeniem ze swojej praktyki analitycznej/badawczej, z którymi osobiście często się zgadzam. Brakuje tu jednak odwołań do literatury, gdzie tego typu prawidłowości oparte na wynikach badań są dobrze opisane i usystematyzowane. Przykładem może być szereg opisywanych efektów związanych z procesem gromadzenia danych, które są typowe dla różnych ankiet (czyli technik pośredniej komunikacji z respondentem), a które stanowią przedmiot systematycznych badań od co najmniej 40 lat i są bardzo dobrze opisane w literaturze przedmiotu poświęconej metodologii badań ankietowych. Habilitant trafnie wypunktowuje prawidłowości na przykładzie swoich badań, nie robi jednak odniesienia do literatury, przez co czytelnik nie jest w stanie ocenić, na ile są to obserwacje odkrywcze, a na ile są potwierdzeniem znanych od dawna prawidłowości. [...] Ostatnia część pracy habilitacyjnej (zakończenie), zawiera szereg dojrzałych refleksji, będących z jednej strony pochodną problematyki analizowanej w rozdziale pierwszym, z drugiej zaś strony są one kontynuacją wniosków wynikających z uzyskanych wyników badań przez Habilitanta. W szczególności Autor proponuje niektóre rozwiązania zidentyfikowanych problemów związanych z pomiarem i badaniem trafności narzędzi psychologicznych. Te rozważania, podobnie jak treści zamieszczone we wcześniejszych rozdziałach, niewątpliwie wzbudzą refleksję nad cudzą i własną praktyką psychometryczną i dydaktyczną zainteresowanego czytelnika, zmuszając do bardziej krytycznej i powściągliwej postawy. [...] Oceniając główne osiągnięcie naukowe, mimo zgłoszonych uwag, uważam wciąż, że stanowi ono wystarczający ustawowo wymagany wkład Habilitanta w psychologię jako naukę.

Ocena pozostałego dorobku naukowego: Dorobek towarzyszący Habilitanta obejmuje 43 publikacje naukowe, w tym 16 publikacji w czasopismach z listy JCR, 23 publikacje w czasopismach z listy B i C MNiSW oraz czterech rozdziały w pracach zbiorowych. Sumaryczny Impact Factor, liczba uzyskanych punktów wg MNiSW, indeks h są na przyzwoitym poziomie. W zdecydowanej większości przypadków Habilitant występuje w roli drugiego lub kolejnego autora (36 z 43), z kolei w roli pierwszego autora występuje w 5 artykułach opublikowanych w czasopismach i 2 rozdziałach w pracach zbiorowych. W publikacjach wieloautorskich, gdzie Habilitant nie występuje w roli pierwszego autora, tylko w ok. 15% przypadków były zamieszczone dane odnośnie

wkładu każdego autora. W przypadku Habilitanta typową rolą było dbanie o stronę metodologiczną badań lub samej publikacji, w szczególności wykonanie analiz statystycznych i napisanie prawdopodobnie odpowiednich części artykułu. Trudno określić, na ile trafna jest generalizacja tej roli na pozostałe publikacje, w których nie określono wkładu autorów. Tematyka publikacji współautorskich jest bardzo szeroka, ale w znacznej części przypadków obejmowały one opracowanie nowych testów lub adaptacja istniejących, gdzie rola Habilitanta, jak można przypuszczać, była kluczowa. Również publikacje, w których Habilitant jest pierwszym autorem, dotyczą w 4-5 z 7 przypadków problematyki metodologii badań i psychometrii. Ranga tych czasopism jest zróżnicowana, dwa najwyżej punktowane mają 70 i 140 pkt. Podsumowując, w ramach dorobku towarzyszącego Habilitant z jednej strony zdobył duże doświadczenie metodologiczne i psychometryczne, z drugiej strony dzięki jego kompetencjom w tym obszarze teksty skutecznie przeszły proces recenzji i ukazały się w wysoko punktowanych czasopismach reprezentujących różne dyscypliny naukowe. [...] W okresie po doktoracie Habilitant był autorem lub współautorem 29 prezentacji (w połowie przypadków jako pierwszy autor) na konferencjach odbywających w Polsce i 2 – na zagranicznych (tylko w roli drugiego autora). Habilitant był także członkiem komitetu organizacyjnego 4 konferencji krajowych. O ile aktywność konferencyjną na rynku krajowym można uznać za wystarczającą, o tyle w przestrzeni międzynarodowej, szczególnie w kontekście tematyki pracy habilitacyjnej, aktywność jest niewystarczająca – głos Habilitanta nie jest obecny w tej formie na poziomie międzynarodowym. [...] W okresie po doktoracie nie występował w roli kierownika grantów, ale był członkiem zespołu badawczego realizującego m.in. granty NCN. Habilitant nie określił swojej roli od strony formalnej, ale od strony merytorycznej wykonane prace dotyczyły analiz statystycznych i psychometrycznych. [...] Habilitant był lub jest członkiem dwóch stowarzyszeń (krajowego i międzynarodowego) po 2013 roku. W tym okresie kilkanaście razy występował w roli recenzenta artykułów skierowanych do czasopism polskich i zagranicznych. Kilkakrotnie wystąpił w roli eksperta, wykonując ekspertyzy z zakresu metodologii badań dla podmiotów zewnętrznych. [...] W okresie po doktoracie Habilitant systematycznie doskonalił swoje kompetencje poprzez zaoczny udział w warsztatach organizowanych przez czołowe zagraniczne uniwersytety (ukończył 12 kursów z zakresu m.in. metodologii, statystyki, e-learningu).

Ocena dorobku dydaktycznego i organizacyjnego: Habilitant ma wieloletnie doświadczenie w prowadzeniu różnych zajęć o charakterze metodologicznym, w tym z zakresu statystyki i psychometrii w UAM oraz Uniwersytecie Zielonogórskim dla studentów studiów magisterskich i doktoranckich. Od 2015 roku występuje w roli promotora prac dyplomowych, promując w tym okresie 12 prac licencjackich i 10 magisterskich (psychologia i kognitywistyka). [...] W ramach UAM przez lata Habilitant pełnił funkcję pełnomocnika ds. e-learningu. Ogólnie dorobek organizacyjny Habilitanta w świetle dostarczonych przez niego materiałów jest dość skromny.

Konkluzja końcowa Recenzenta: Oceniając dorobek i aktywność naukową dr. Pawła Kleki na podstawie kryteriów zawartych w art. 219 ust. 1 pkt. 2 Ustawy z dnia 20 lipca 2018 r. należy stwierdzić, że Habilitant posiada znaczącą część osiągnięć, które wymienia ustawa. Posiada on duże kompetencje w zakresie metodologii badań, w tym w zakresie analiz statystycznych i psychometrii. W wieloautorskich przedsięwzięciach Habilitant najczęściej zadowala się jednak drugoplanową rolą. Analiza dorobku badawczego i dydaktycznego wskazuje na wybór przez Habilitanta raczej ścieżki wąsko specjalizacyjnej w nauce. Posiadane kompetencje zostały udokumentowane/potwierdzone w postaci licznych publikacji i są przekazywane

kolejnym pokoleniom studentów. W mniejszym stopniu te kompetencje zostały ujawnione w pracy habilitacyjnej w zakresie zrealizowanych przez Habilitanta badań własnych, w których można było pokusić się o bogatszy/szerszy program badawczy, aby w pełni wykorzystać posiadany potencjał. Oceniając główne osiągnięcie naukowe, mimo zgłoszonych uwag, uważam wciąż, że stanowi ono wystarczający ustawowo wymagany wkład Habilitanta w psychologię jako naukę. Wnoszę na tej podstawie o nadanie dr. Pawłowi Klece stopnia naukowego doktora habilitowanego nauk społecznych w zakresie psychologii.

Dr hab. Paweł Jurek (wyciąg z recenzji)

Ocena osiągnięcia naukowego: Dr Paweł Kleka jako swoje osiągnięcie naukowe będące podstawą ubiegania się o nadanie stopnia doktora habilitowanego wskazał monografię pod tytułem „*Trafność pomiaru w praktyce psychometrycznej*”, której jest jedynym autorem. [...] Wskazana monografia składa się z czterech rozdziałów, w tym trzech zawierających sprawozdania z działań badawczych przeprowadzonych przez Autora. Cała praca dotyczy wybranych aspektów jakości pomiaru w psychologii. Podjęta przez Autora tematyka leży w polu zainteresowań przede wszystkim psychometrów oraz naukowców mocno zanurzonych w metodologii badań psychologicznych, rzadziej – co należy stwierdzić z ubolewaniem – przebija się do standardowych dyskusji w kontekście planowania i realizowania projektów badawczych. Być może właśnie dlatego, że stawiane przez Habilitanta pytania mogą wydawać się niszowe, ich waga jest szczególnie duża. [...] Monografia dr. Pawła Kleki stanowi istotny wkład w realizację tego zadania, przede wszystkim w zakresie zdiagnozowania i wyjaśnienia wagi problemu. Przedstawione w monografii wyniki prac badawczych Autora trafnie ilustrują wyzwania związane z trafnością pomiaru w badaniach psychologicznych. [...] W rozdziale poświęconym problemom pomiaru w psychologii, dr Paweł Kleka naświetla dominujące w badaniach psychologicznych podejście klasyczne reprezentowane przez S. Stevensa i zestawia je z aksjomatyczną teorią pomiaru. Autor przytacza krytykę operacjonalizacji zmiennych wg reguł pozwalających na zaklasyfikowanie zebranych danych do jednej z czterech skal pomiarowych. Przekonuje, wspierając się argumentacją innych badaczy, że dopiero zastosowanie mocnego pomiaru w psychologii bazującego na odkrywaniu ilościowych właściwości mierzonych konstruktów, może ugruntować psychologię jako naukę empiryczną. W mojej ocenie, rozdział ten trafnie zarysowuje dylematy i problemy, które mają już swoją historię w psychometrii. Jednocześnie stanowi głos Autora w tej dyskusji, w którym da się odczytać apel o nieignorowanie problemów z pomiarem w psychologii. Ciekawym zabiegiem jest przedstawienie analizy danych zebranych w badaniach z zastosowaniem jednego z dobrze znanych narzędzi psychometrycznych i ich ocenę w kontekście aksjomatycznej teorii pomiaru. Konkluzja, że wykorzystanie do pomiaru inteligencji jednego z dostępnych na rynku testów prawdopodobnie nie pozwala dostarczyć wyników ilościowych w rozumieniu aksjomatycznej teorii pomiaru, pobudza do refleksji. [...] Niewątpliwą zaletą tego rozdziału jest skupienie uwagi na problemie, którego praca ta dotyczy. Czytając tę część monografii nabyłem przekonania, że Autor jest ekspertem w swojej subdyscyplinie, zna wagę tego o czym pisze i jest świadomy wyzwań, jakie stoją przed współczesnymi naukowcami w psychologii w kontekście projektowania pomiaru badanych zmiennych. [...] W rozdziale drugim dr Paweł Kleka podejmuje temat obciążenia wyników pomiaru z zastosowaniem metod samoopisowych. Opisuje fazy formułowania odpowiedzi osoby badanej na pozycję kwestionariusza, akcentuje jak ten proces jest podatny na zniekształcenia. W dalszej części rozdziału Autor koncentruje się na zdolności do introspekcji (samowiedzy), jako potencjalnym źródle obciążenia wyników testu. W

przedłożonej monografii przedstawia wyniki badań symulacyjnych, w których przetestował kilka modeli ilustrujących związek pomiędzy samowiedzą a uzyskiwanymi wynikami pomiaru zadanej cechy. [...] Za pewne ograniczenie uważam porzucenie na badaniach symulacyjnych i nie podjęcie próby zaprojektowania i przeprowadzenia analiz (choćby na danych zastanych), które wsparłyby poczynione ustalenia. Być może wątek ten został „porzucony” przez Autora na rzecz badań ilustrujących kolejne aspekty jakości badań, niemniej pogłębienie tematu roli samowiedzy w trafności pomiaru z zastosowaniem metod samoopisowych uważam za niezwykle istotny dla współczesnych badań w psychologii. [...] Rozdział trzeci przedłożonej monografii poświęcony jest zagadnieniu trafności pomiaru. Habilitant przytacza w nim jak na przestrzeni dekad zmieniano się podejście do ustalania trafności pomiaru i jednocześnie jak rozbudowywał się warsztat analityczny pozwalający na zbieranie różnego rodzaju dowodów (argumentów) przemawiających za trafnością narzędzia, czy też interpretacji i wykorzystania wyników narzędzia. W rozdziale trzecim dr Paweł Kleka zamieścił także rezultaty systematycznego przeglądu literatury, którego celem było ujawnienie powszechnie stosowanego modelu ustalania trafności narzędzi psychometrycznych. Autor przeanalizował 101 artykułów raportujących łącznie 178 wyników badań zaprojektowanych i przeprowadzonych w celu walidacji narzędzi psychometrycznych. Zastosowana procedura oraz kryteria włączenia i wykluczenia artykułów do przeglądu oceniam jako właściwe dla sformułowanego problemu badawczego. Za pewnego rodzaju niedopatrzenie uznaję brak bibliografii dla artykułów uwzględnionych w przeglądzie – sama charakterystyka pozycji nie pozwalała na zweryfikowanie ustaleń w kontekście dodatkowych nasuwających się pytań, np. czy występują różnice w standardach ustalania właściwości psychometrycznych w walidacji narzędzi stosowanych w różnych obszarach psychologii. Z otrzymanych rezultatów wynika, że w ustalaniu trafności dominuje badanie struktury czynnikowej narzędzia oraz analizowanie korelacji wyników nowego testu z wynikami uzyskiwanymi z wykorzystaniem innych, walidowanych wcześniej narzędzi psychometrycznych. Wyniki te dają dobry punkt odniesienia na temat strategii psychometrycznych stosowanych w praktyce konstruowania nowych narzędzi. Zaletą wykonanego przeglądu jest również dokonana przez Autora ocena poprawności stosowanych metod analitycznych – z ustaleń wynika, że w artykułach psychometrycznych często brak jest kluczowych informacji o zastosowanych rozwiązaniach, a nie rzadko zdarzają się również błędne interpretacje otrzymanych wyników badań. Zaprezentowane w rozdziale trzecim doniesienia uważam za bardzo cenne i inspirujące zarówno z perspektywy tworzenia narzędzi psychometrycznych, recenzowania artykułów z tego obszaru, a także uczenia psychometrii przyszłych i obecnych psychologów. [...] W rozdziale czwartym dr Paweł Kleka przedstawił wyniki przeprowadzonego przez siebie eksperymentu mającego na celu przetestowanie wrażliwości powszechnie stosowanej procedury walidacji narzędzi badawczych w psychologii w detekcji wadliwych, czy nawet bezsensownych kwestionariuszy. Pomysł na badanie uważam za bardzo ciekawy, a dobór warunków eksperymentalnych przemyślany. Generalnie, zgadzam się również z interpretacją otrzymanych wyników, choć w pewnych fragmentach mam wątpliwości, które dałoby się łatwo rozwiązać wprowadzając drobne, lecz istotne zmiany w zaprojektowanym badaniu. Autor monografii przeprowadził analizę rzetelności oraz konfirmacyjną analizę czynnikową na danych z badania, w którym zastosował 10-itemowy kwestionariusz w czterech wersjach. Trzy z wykorzystanych wersji zostały na tyle zmienione, że treściowo nie badały oryginalnej zmiennej, a w przypadku dwóch wersji nawet żadnej sensownej zmiennej, ponieważ zmodyfikowane itemy nie zawierały żadnej sensownej treści.

Zdaniem Autora, dowodem na to, że standardowe procedury psychometryczne nie są w stanie wychwycić błędnych narzędzi jest fakt, że uzyskane wyniki – dla wszystkich czterech wersji narzędzia – wskazują na satysfakcjonującą rzetelność pomiaru i dobre dopasowanie modeli jednoczynnikowych. Pewnym ograniczeniem tego badania jest właśnie zastosowanie narzędzi o jednoczynnikowej strukturze. [...] Wątpliwości te można byłoby rozwiązać, gdyby struktura czynnikowa testowanych narzędzi była bardziej złożona i z góry zadana. Gdyby podobne wyniki uzyskać dla bezsensownej wersji kwestionariusza o zadanej trzy-czynnikowej strukturze, byłbym zaskoczony. Niemniej, moje uwagi nie zmieniają podstawowej konkluzji z tego eksperymentu, że wykonywane analizy statystyczne – choćby bardzo zaawansowane – pozostają wyłącznie analizami statystycznymi i wykorzystywanie ich wyników bezrefleksyjnie w oderwaniu od przyjętych procedur postępowania (np. związanych z walidacją treści jednostek testowych na odpowiednim etapie konstruowania narzędzia) może doprowadzić do absurdalnych wniosków. [...] Ostatnia część monografii dr. Pawła Kleki to zakończenie zawierające syntezę wcześniejszych treści, omówienie ograniczeń przeprowadzonych badań oraz – co najcenniejsze – rekomendacje w postaci pewnych propozycji rozwiązania zidentyfikowanych w pracy problemów. Ten fragment oceniam szczególnie wysoko, stanowi bowiem wyraźne potwierdzenie wkładu, jaki praca Habilitanta ma w rozwój dyscypliny psychologia. Podstawową konkluzją Autora jest przekonanie, że jakość pomiaru w psychologii może być poprawiona dzięki zastosowaniu procedur eksperymentalnych poprzez „poszukiwanie takich manipulacji, które oddziałują na psychologiczną zmienną ukrytą i obserwowanie, czy mierzone zmienne zmieniają się zgodnie z zakładanym modelem powiązań między teorią a empirią”. [...] Co do zasady, trudno się z tym postulatem nie zgodzić. Jednakże w literaturze nadal brakuje przykładów badań pokazujących jakie wymierne benefity można osiągnąć zastępując metody kwestionariuszowe metodami eksperymentalnymi, jaka jest efektywność takiego działania. Sądzę zatem, że temat ten wymaga dalszych, przede wszystkim empirycznych prac i doniesień. W tym kontekście należy docenić zainteresowanie i wysiłki dr. Pawła Kleki w rozwój psychometrii, w promowanie tworzenia narzędzi, które w większym stopniu będą czułe na ujawnianie ilościowej struktury mierzonych właściwości psychologicznych. Biorąc pod uwagę powyższe oraz wcześniej przytoczone argumenty uważam, że wskazane osiągnięcie naukowe dr. Pawła Kleki stanowią znaczący wkład w rozwój dyscypliny psychologia.

Ocena pozostałego dorobku naukowego: Dr Paweł Kleka jest współautorem ponad 40 prac naukowych opublikowanych po uzyskaniu stopnia doktora. Wykazany dorobek należy uznać za znaczący (sumaryczny *Impact Factor* = 45,746). Na dorobek naukowy Habilitanta składają się prace z różnych subdyscyplin psychologii (psychologia zdrowia, neurobiopsychologia, psychologia społeczna), a także prace interdyscyplinarne (np. w kooperacji z Uniwersytetem Medycznym oraz Akademią Wychowania Fizycznego). Dr Paweł Kleka w istotny sposób przyczynił się do powstania tych prac, najczęściej poprzez zaprojektowanie i wykonanie analiz statystycznych pozwalających na zweryfikowanie testowanych modeli lub oszacowanie parametrów tworzonych lub adaptowanych narzędzi psychometrycznych. W tym kontekście dr Paweł Kleka doskonale odnajduje się w pracy zespołowej, wprowadzając do każdego zespołu unikatowe i niezbędne dla pracy naukowej kompetencje. Wiele artykułów, których Habilitant jest współautorem, dotyczy narzędzi psychometrycznych i szerzej pomiaru w naukach społecznych i naukach o zdrowiu. Całość stanowi spójne osiągnięcie wskazujące na to, że dr Paweł Kleka posiada bogate doświadczenie, wiedzę i umiejętności w zakresie psychometrii. [...] Dr

Paweł Kleka pracował w Uniwersytecie im. Adama Mickiewicza w Poznaniu oraz w Uniwersytecie Zielonogórskim. Uczestniczył w projektach badawczych realizowanych we współpracy z naukowcami z Akademii Medycznej w Poznaniu, Akademii Wychowania Fizycznego w Poznaniu, a także Uniwersytetu Gdańskiego. Habilitant jest również współautorem licznych publikacji opracowanych we współpracy z przedstawicielami wielu ośrodków naukowych w Polsce i z zagranicy.

Ocena dorobku dydaktycznego i organizacyjnego oraz popularyzacyjnego: Habilitant posiada doświadczenie w prowadzeniu zajęć z obszaru statystyki, psychometrii i metod badawczych w psychologii; pełnił funkcję promotora prac licencjackich i magisterskich; angażował się w prace na rzecz instytutu i wydziału; był członkiem komitetu organizacyjnego kilku konferencji o zasięgu krajowym.

Konkluzja końcowa Recenzenta: Podsumowując, uważam, osiągnięcie naukowe dr. Pawła Kleki ma znaczący wkład w rozwój dyscypliny psychologia. W swojej recenzji pokazałem, że monografia stanowiąca osiągnięcie naukowe zarówno pod względem ilościowym, jak i jakościowym odpowiada wymaganiom określonym w art. 219 ust. 1 pkt 2 Ustawy z dnia 20 lipca 2018 r. Prawo o szkolnictwie wyższym i nauce, stawianym kandydatom do uzyskania stopnia doktora habilitowanego. Ponadto uważam, że dr Paweł Kleka prowadzi działalność naukową, która przejawia się zarówno w dobrych publikacjach, jak i w aktywności konferencyjnej, aktywności organizacyjnej i recenzenckiej. Dr Paweł Kleka wykazał się istotną działalnością naukową w więcej niż jednej uczelni. Konkludując, moja opinia na temat osiągnięć naukowych Habilitanta jest pozytywna i tym samym wnioskuję o Jego dopuszczenie do dalszych etapów postępowania oraz przyznanie dr. Pawłowi Klece stopnia doktora habilitowanego.

Prof. dr hab. Romuald Polczyk (wyciąg z recenzji)

Ocena osiągnięcia naukowego: Dr Kleka ubiega się o stopień dra habilitowanego na podstawie osiągnięć naukowych zawartych w monografii „*Trafność pomiaru w praktyce psychometrycznej. Badania empiryczne*”. [...] Zacznę od tego, iż samą tematykę tej książki oceniam jako bardzo ważną i potrzebną. O wiele zbyt często widujemy w pracach dotyczących stworzenia nowych narzędzi czy adaptacji istniejących przewagę analiz dotyczących rzetelności, niż trafności. Wynika to być może z tego, że rzetelność jest stosunkowo łatwo zbadać i policzyć. Badanie trafności wymaga w wielu przypadkach znacznie bardziej zaawansowanego programu badawczego. [...] *Rozdział 1. Problematyka pomiaru w psychologii.* W rozdziale 1 Autor trafnie opisuje trudności, jakie spotykały pierwszych badaczy uprawiających analizy empiryczne, oraz problemy z określeniem, czy pomiar w psychologii jest ilościowy. Jednak pewne rozstrzygnięcia, jakie proponuje, są w najlepszym razie niejasne, w najgorszym nieprzekonujące. Przykładem może być stwierdzenie: „Ponieważ przedmiot badań w psychologii znajduje się między pojęciem a rzeczywistością, do badań i analiz w tej dziedzinie niezbędny stał się termin konstrukt psychologiczny (s. 16-17). Nie rozumiem ani, co to znaczy być między pojęciem a rzeczywistością, ani co ma do tego termin „konstrukt psychologiczny”. Podzielam argumentację Autora, że istniejące definicje pojęcia „konstrukt” są niezadowolające, wydaje mi się jednak, że Jego własne ujęcie nie jest jaśniejsze. [...] Wątpliwości budzi też stwierdzenie „Większość badań empirycznych w psychologii polega na jednym z dwóch sposobów pomiaru. Zgodnie z pierwszym zliczana jest częstotliwość reakcji osoby uczestniczącej w badaniu na określony zestaw bodźców. Dla każdego bodźca istnieje predefiniowana reakcja poprawna [...] Drugi sposób pomiaru [...] to liczba zadań (dziś powiedzielibyśmy - pozycji testowych) poprawnie rozwiązanych” (s. 17-19). Nie wiadomo, jak w tym podziale mieści się lwią część pomiaru w psychologii, czyli kwestionariuszowy samoopisowy

miar cech indywidualnych. Ani nie są w nim zdefiniowane reakcje poprawne, ani nie jest on zbiorem zadań z określonym poprawnym rozwiązaniem. [...] Wobec raczej pesymistycznej wymowy tego rozdziału, dotyczącej możliwości prawdziwie ilościowego pomiaru w psychologii, nieco zaskakująco brzmi stwierdzenie w podsumowaniu „Na podstawie zaprezentowanych rozważań można uznać, że przy spełnieniu warunków wstępnych jest możliwy pomiar atrybutów (zmiennych) psychologicznych, takich jak postawy, przekonania, zdolności poznawcze i afekty” (s. 34). Na wielu stronach wywodu wcześniej Autor pokazuje trudności w dokonywaniu pomiaru w psychologii, przez nierzadko dość radykalne opinie, np. metody IRT i SEM nazwane są „ucieczką do przodu”, stosowaną pomimo braku dowodów na ilościowy charakter zmiennej. Pesymistyczny wydźwięk ma też analiza empiryczna testu Omnibus, która wskazała, iż wątpliwe jest, że pomiar w tym teście ma charakter ilościowy. W zasadzie wyłącznie paragraf poprzedzający to stwierdzenie (o możliwości pomiaru atrybutów psychologicznych) wnosi jakiś powiew optymizmu; mimo to podsumowanie rozdziału ma przesłanie optymistyczne. Zaskakuje też nazwanie wspomnianego podejścia IRT nadzieją na mocny pomiar w psychologii, choć wcześniej Autor napisał „Dopasowanie modeli IRT lub SEM nie świadczy o ilościowości konstruktów, wręcz przeciwnie - ilościowość badanego konstruktów pozwala dopiero na stosowanie tych modeli” (s. 32). [...] Niezależnie od powyższych zastrzeżeń, omawiany rozdział dowodzi dużych kompetencji Autora i ma znaczną wartość poznawczą. Czytelnik może poznać zaawansowane podejścia, nieczęsto omawiane w typowych podręcznikach psychometrii, na przykład aksjomatyczną teorię pomiaru, wraz z jej krytyką. Opis jest na tyle precyzyjny, że czytelnik może pokusić się o nauczenie się samodzielnego stosowania tej metody. [...] *Rozdział 2. Wpływ metod samoopisowych na wyniki badań psychologicznych.* Tytuł tego rozdziału jest cokolwiek niezręczny; to nie metody samoopisowe mają wpływ na wyniki badań, tylko pewne zjawiska psychologiczne wiążące się z tymi metodami mogą na coś wpływać (w tym rozdziale chodzi głównie o wpływ samowiedzy). [...] Zawarty w tym rozdziale jest opis symulacji obliczeniowej i jej wyników dotyczącej efektów samowiedzy. Idea takiej analizy jest bardzo interesująca, jednak wymowę uzyskanych wyników osłabiają w moim odczuciu pewne kwestie techniczne. Na przykład, dla potrzeb symulacji konieczne było określenie minimalnego wpływu samowiedzy na wyniki obserwowane. Niezależnie od tego, zarówno w tym artykule, jak i w tym rozdziale, przyjęte wartości są raczej arbitralne. [...] Wątpliwości nie rozwiewa fragment: „Założony przeze mnie pierwszy model odwzorowywał to, że osoby o niskim poziomie samowiedzy będą zawyżać swoje odpowiedzi, ale tylko przy niskim poziomie badanej cechy. Efekt ten byłby podobny do wpływu czynnika aprobaty społecznej. Gdy badana jest społecznie pożądana cecha - osoby z niskimi wynikami będą „dodawały” sobie trochę wartości, osoby z wynikami wysokimi nie mają zaś takiej potrzeby, dlatego można założyć, że w tym modelu wraz ze wzrostem poziomu cechy efekt zanika.” (s. 42). Rozumowanie jest przekonujące dla aprobaty społecznej właśnie; nie wiadomo, dlaczego Autor widzi analogiczny mechanizm w przypadku zdolności do introspekcji. W ogóle wiązanie zdolności do introspekcji z jakąkolwiek motywacją, czy do zawyżania czy do zaniżania wyników, jest niejasne i zupełnie brak dla niego uzasadnienia w tekście. Jest to rozczarowujące zwłaszcza na tle deklaracji Autora, iż w psychologii ważna jest teoria zbudowana na przesłankach empirycznych (s. 35). Odnośnie do analogii czy związku między zdolnością do introspekcji a aprobatą społeczną brak w tekście i teorii, i odniesień do wyników empirycznych. Budzi to niestety wątpliwości dotyczące tego, o czym właściwie mówią przeprowadzone symulacje i jaki jest ich sens aplikacyjny. [...] Największy jednak

mój sprzeciw budzi konkluzja: „Czy zatem psychometria powinna się zajmować problemem obciążenia wyników „czynnikiem ludzkim” mimo potencjalnie niewielkiego obciążenia nim realnych wyników? Moim zdaniem warto [...]” (s. 53). Wyniki przeprowadzonej przez Autora symulacji nie dotyczą obciążenia wyników „czynnikiem ludzkim”, a tylko obciążenia jednym wymiarem czynnika ludzkiego, spośród setek potencjalnie możliwych - zdolnością do introspekcji. Zacytowane zdanie jest absolutnie na wyrost, wykracza poza rzeczywiste znaczenie przeprowadzonej analizy i jego wyników, i jest po prostu nieuzasadnione. [...] Jednak również w tym rozdziale, jego zalety przeważają nad zastrzeżeniami. Opis przeprowadzonej symulacji matematycznej odznacza się właściwą Autorowi precyzją; czytelnik zyskuje solidny grunt, by samodzielnie nauczyć się wykonywać podobne analizy dla innych zmiennych zakłócających. Nasza wiedza o jakości pomiaru dzięki temu może się znacznie powiększyć. [...] *Rozdział 3. Jakość pomiaru. Od trafności jedno- do wieloaspektowej.* W rozdziale tym Autor najpierw przeprowadza analizę, w tym krytyczną, różnych pojęć i ujęć dotyczących trafności, a następnie przedstawia wyniki badań własnych, dotyczących tego, jak wyglądają analizy trafności w istniejących opublikowanych artykułach. Początkowe podrozdziały, dotyczące samego pojęcia trafności, oceniam wysoko, doceniając nierzadko bardzo wnikliwie analizy Autora dotyczące zalet i wad różnych podejść i stanowisk. To powiedziawszy, bardzo mi jednak przeszkadzało wrażenie braku należytego rozróżnienia między trzema pojęciami: (1) trafnością jako taką, (2) rodzajami trafności i (3) metodami badania trafności. Przykładem ogólnej definicji trafności jako takiej jest np. ujęcie Borsbooma i in. (2004): „test jest trafny, jeśli spełnione są dwa warunki: 1) istnieje konstrukt, który możemy mierzyć, oraz 2) istnieje przyczynowa relacja między poziomami (natężeniem) konstruktów a wynikami testu” (s. 67); albo np. najstarsza klasyczna definicja trafności: dotyczy ona tego, CO test mierzy. Oprócz takich definicji ogólnych, mówimy o rodzajach trafności, najczęściej o ujęciu trynitarnym: trafność kryterialna, treściowa i teoretyczna. Wreszcie, analiza korelacji to jest metoda, która może służyć do badania trafności, i to różnych jej rodzajów. Podobnie, macierz WCWM (*MTMM*) to jest metoda badania trafności zbieżnej i różnicowej. Brak precyzyjnego rozróżniania między definicją trafności, jej rodzajami oraz metodami jej badania powoduje wrażenie silnego niedosytu dotyczącego precyzji definiowania kluczowego wszak w pracy pojęcia. W dalszej części tego rozdziału Autor przedstawia przeprowadzone badania. Nie ukrywam, że oceniam je najslabiej z wszystkich badań i analiz dr. Kleki przedstawionych w recenzowanej monografii. Pierwszy „zgrzyt” pojawia się już na samym początku opisu. W pierwszym paragrafie znajdujemy deklarację: „Chciałbym się zatem przyjrzeć temu, jak w praktyce psychologowie i psychometryści definiują i badają trafność konstruowanych przez siebie narzędzi.” (s. 69). Potem następuje podrozdział „Metoda”, i czytamy w nim znowu o celu badania: „Celem systematycznego przeglądu literatury było określenie dominującego modelu badania trafności narzędzi do badań psychologicznych.” (s. 69). Jest to drobiazg, ale kontrastujący ze zwykłą precyzją Autora. Cel przeprowadzonych badań zasługiwał na osobny krótki podrozdział, nie tylko na pojedyncze zdania w pierwszym paragrafie, i na pewno nie w „Metodzie”. Autor zlokalizował blisko 40 tysięcy artykułów możliwych do przeanalizowania, lecz analizie zdecydował się poddać zaledwie 101 z nich. Jest to raczej mała próba - niezrozumiałe jest, dlaczego zdecydowano się na przeanalizowanie zaledwie 101 artykułów, skoro wyszukano ich prawie 40000. Na jakiej podstawie Autor uznał, że próba ta jest „reprezentatywna”? Nie przeprowadzono żadnych analiz mocy, ani oszacowań, jak szerokie byłyby

przedziały ufności dla różnych procentów przy takiej próbie. Co do przedziałów ufności, to zdziwienie budzi, że Autor w ogóle ich nie policzył, podając wskaźniki procentowe. Brak przedziałów ufności prowadzi do stwierdzeń ocierających się o nieprawdę, np. zdanie „[...] i tylko raz użyto rekomendowanej przez Campbella i Fiskego (1959) metody wielu cech wielu metod!” (s. 74) jest przecież mylące - w rzeczywistości metoda WCWM pojawiła się raz wśród 101 przeanalizowanych przez Autora artykułów, ale próba ta pochodziła z około 40 000 artykułów. Można więc oszacować, że metody tej użyto około 400 razy, a nie jeden raz. Dokładniej mówiąc, na podstawie jednego przypadku z próby 101 można oszacować, że metoda pojawiła się między około 20 a 1844 razy (90% przedziały ufności). Brak przedziałów ufności dla prostych danych procentowych jest niezrozumiały, tym bardziej że Autor był w stanie wyinferować je, jeśli ich brakowało w danym artykule, dla bardzo niebanalnych parametrów, np. RMSEA. Znacznie gorsze jednak jest to, że przeprowadzone analizy nie odpowiadają na zadeklarowane przez Autora pytanie badawcze, które dotyczyło tego, „jak w praktyce psychologowie i psychometryści definiują i badają trafność konstruowanych przez siebie narzędzi” (s. 69). Nie przedstawiono żadnych wyników dotyczących tego, jak definiowana była trafność w analizowanych artykułach. Wyniki dotyczą metod badania trafności, jak analiza czynnikowa, IRT, analiza regresji, korelacji i inne. To nie są ani definicje, ani rodzaje trafności, to są metody jej badania. Ponownie pojawia się tu problem braku rozróżnienia między definicjami, rodzajami i metodami badania trafności. Pytanie o to, jak definiowana była trafność w analizowanych artykułach, po prostu nie ma odpowiedzi. Podobnie, nie wiadomo, ilu spośród autorów analizowanych publikacji określało siebie jako psychologowie, a ilu jako psychometryści. Znajdziemy za to wyniki, które w ogóle nie mają nic wspólnego z postawionymi pytaniami badawczymi - mam na myśli analizę korelacji między rokiem wydania publikacji a wartościami wskaźników dopasowania RMSEA i CFI, czy między rokiem publikacji a wartością korelacji test-test. Po prostu nie wiem, po co to zostało policzone i o czym miało by świadczyć. Podsumowując, uważam, że badania przedstawione w tym rozdziale nie odpowiadają na zadeklarowane pytania badawcze i zawierają poważne błędy metodologiczne, w tym za mała próba oraz brak przedziałów ufności dla głównych analiz. [...] *Rozdział 4. Walidacja narzędzi badawczych w psychologii.* W rozdziale tym dr Kleka przedstawia kolejne przeprowadzone przez siebie badania. Polegały one, najogólniej mówiąc, na poddaniu analizom walidacyjnym narzędzia (kwestionariusza samoskuteczności GSES) celowo „uszkodzonego”: w jednej wersji uszkodzenia, kluczowe stwierdzenia zastąpiono słowem „poczuciowość”; w drugiej - wygenerowano zdania absurdalne, w trzeciej - same słowa były absurdalne. Szokującym rezultatem analiz było to, że wprawdzie wersje uszkodzone miały większość wskaźników jakości niższych niż wersja oryginalna, lecz wiele z tych wskaźników, w tym korelacje z kryterium zewnętrznym, było na akceptowalnym poziomie! W rezultacie dr Kleka musiał napisać: „Wartości indeksów dopasowania modelu do danych dla wersji narzędzia wykorzystującej w badaniu pozycje składające się z bezsensownych wyrazów, zarówno po skorelowaniu trzech błędów, jak i po zmianie metody estymacji na nieparametryczną DWLS, spełniają wymagane progi, by uznać narzędzie za trafne. Jest tylko jeden problem - to narzędzie niczego nie mierzy!” (s. 101). Wyniki te uważam za najważniejsze osiągnięcie dra Kleki, które przesądzają o tym, iż wniośł On znaczący wkład do psychometrii i psychologii. [...] Bardzo zabrakło mi prostych badań dodatkowych zbliżonych do procedury analizy procedury rozwiązywania testu, polegających po prostu na wywiadach z osobami badanymi, co właściwie sobie myślały, widząc absurdalne słowa i stwierdzenia. Być może dowiedzielibyśmy się

wtedy czegoś, co mogłoby jakoś objaśnić fakt zadowalających wskaźników dla absurdalnego kwestionariusza. Jest na przykład możliwe, że dużo osób badanych odpowiadało w absurdalnym narzędziu w jakiś jednolity sposób, np. wybierając opcję „zdecydowanie nie”, albo na przykład opcję pośrednią. Jeśli dużo osób badanych odpowiadało tak samo, mogłoby to tłumaczyć dobrą spójność wewnętrzną, a nawet wyniki analiz czynnikowych.. Nie da się w ten sposób wyjaśnić korelacji z kryterium zewnętrznym - to można np. próbować tłumaczyć skłonnością części osób badanych do wybierania raczej górnych, czy też raczej dolnych odpowiedzi. Obecnie w monografii nie ma danych, które pozwoliłyby to rozstrzygnąć. Nie ma też prób interpretacji uzyskanych wyników, poza ogólnymi wątpliwościami co do wartości stosowanych obecnie metod walidacji testów i ich podstaw teoretycznych. Nie zmniejsza to jednak wartości odkrycia podstawowego - możliwości uzyskiwania dobrych wskaźników psychometrycznych dla bezsensownych narzędzi. [...]

Uważam, że taki znaczący wkład został przez dra Pawła Klekę do psychologii wniesiony, tym samym moja ocena Jego osiągnięć jest pozytywna. Za najważniejsze uważam empiryczne zademonstrowanie przez Autora, iż można uzyskać zadowalające wyniki różnych analiz trafności, nawet jeśli badane narzędzie jest pozbawione wszelkiego sensu. Wyniki te były dla mnie absolutnym wstrząsem poznawczym i uważam, że wprost nie sposób przecenić ich znaczenia. Mam liczne i w moim odczuciu poważne zastrzeżenia wobec omawianej monografii, lecz nie zmieniają one faktu, że Autor moim zdaniem dokonał prawdziwego i ważnego odkrycia naukowego w dziedzinie psychometrii, tym samym wniósł istotny wkład w rozwój psychologii.

Ocena pozostałego dorobku naukowego: Jeśli idzie o dorobek towarzyszący, to dla publikacji po doktoracie Autor podaje sumaryczny Impact Factor dla wynoszący 45.746, a sumaryczną punktację według MEiN: 2773 pkt. Indeks Hirscha wyniósł od 6 do 10, w zależności od źródła. Ogółem liczba publikacji po doktoracie wyniosła 43, w tym 16 w czasopiśmie z listy JCR. Ogólna liczba cytacji wyniosła 423. [...]

Komentując osiągnięcia Autora można stwierdzić, że od strony liczbowej i naukometrycznej przedstawia się on dobrze, natomiast odznacza się ogromnym rozproszeniem tematyki. Znajdziemy w nim artykuły dotyczące na przykład postaw wobec lesbijek, systemów CAPTCHA, posługiwania się pismem u dzieci, komunikacji w negocjacjach, wpływu hałasu turbin wiatrowych na stres, pomocy pacjentom z demencją, schematów płci, integracji osobowości w psychopatii, predyktorów statusu związku u studiujących kobiet, metapoznawczego Ja, gadatliwości kobiet, organizacji języka u osób leworęcznych, komunikacji internetowej, związku między poczuciem humoru a postawą rodzicielską, i wiele, wiele innych tematów. Jest oczywiste, że nikt nie jest w stanie być specjalistą w tylu dziedzinach. Można się domyślać, choć Autor tego wyraźnie nie pisze, iż Jego rola polegała w większości tych tekstów na konsultacji metodologicznej i wykonywaniu analiz. Nie ma w tym oczywiście niczego złego, przeciwnie, mieć w swoim zespole naukowym eksperta od metodologii, statystyki i psychometrii jest wprost błogosławieństwem, i z pewnością wkład dr. Kleki w liczne artykuły przyczynił się znaczenie do sukcesu tych artykułów. Nie zmienia to jednak faktu, że jedynym spójnym wątkiem, jaki można znaleźć w dorobku Autora, są adaptacje testów i teksty dotyczące metodologii i psychometrii, i tylko w pojedynczych wypadkach był On pierwszym (lub ostatnim, co też sugerowałoby ważną rolę) Autorem. [...]

Oczywiste jest też, że dorobek ten ma charakter interdyscyplinarny i powstał we współpracy z różnymi ośrodkami (współpraca z różnymi ośrodkami wynika też jednoznacznie z Autoreferatu, a nawet z spisu miejsc zatrudnienia), a tym samym

spełniony jest ustawowy wymóg wykazywania się istotną aktywnością naukową realizowaną w więcej niż jednej uczelni czy instytucji naukowej.

Ocena dorobku dydaktycznego i organizacyjnego: wychodząc z założenia, że zgodnie z Ustawą opinia sporządzona przez recenzenta nie musi zawierać elementu oceny aktywności dydaktycznej lub organizacyjnej Habilitanta, kwestiom tym nie poświęcił dużo miejsca w recenzji, choć zaznaczył, że ocena tych aspektów jest w przypadku dr. Kleki jednoznacznie pozytywna.

Konkluzja końcowa Recenzenta: Stwierdzam, iż moim zdaniem monografia dr. Pawła Kleki „*Trafność pomiaru w praktyce psychometrycznej. Badania empiryczne*” spełnia warunki zapisane w ustawie z dnia 20 lipca 2018 r. „Prawo o szkolnictwie wyższym i nauce” (z późniejszymi nowelizacjami i zmianami, art. 219). W szczególności, odkrycia zawarte w tej monografii stanowią znaczny wkład w rozwój psychologii i psychometrii. W związku z tym, w mojej opinii, dr Paweł Kleka zasługuje na przyznanie stopnia doktora habilitowanego.

Dr hab. Marcin Zajenkowski (wyciąg z recenzji)

Ocena osiągnięcia naukowego: Doktor Paweł Kleka ubiega się o uzyskanie stopnia doktora habilitowanego na podstawie monografii „*Trafność pomiaru w praktyce psychometrycznej. Badania empiryczne*” wydanej przez Wydawnictwo Naukowe Wydziału Nauk Społecznych Uniwersytetu Adama Mickiewicza w Poznaniu. Praca składa się z czterech rozdziałów (124 strony), w których zaprezentowano koncepcje teoretyczne pomiaru w psychologii, analizy symulacyjne, systematyczny przegląd literatury oraz jedno badanie z udziałem osób badanych. [...] Monografia dr. Kleki skupia się na tematyce pomiaru w psychologii. W rozdziale pierwszym autor przedstawia rys historyczny tego zagadnienia, sięgając do prac badawczych z XIX w., a więc początków psychologii, kończąc na współczesnych nurtach. Autor zastanawia się nad samą istotą empirycznego badania zjawisk psychicznych a co za tym idzie, przypisywania liczb atrybutom psychologicznym. Przedstawia kontekst, w którym zrodziło się najbardziej rozpowszechnione podejście do pomiaru w psychologii jakim jest koncepcja skal pomiarowych Stevensa. Habilitant przedstawia krytykę tego podejścia i wskazuje na największy grzech współczesnej psychologii jakim jest lekkomyślne założenie, że wszystkie atrybuty psychiczne są zasadniczo ilościowe oraz brak prób udowodnienia ich ilościowego charakteru. Dr Kleka zauważa, że psychologia skupia się na liczbach, które wcale nie muszą reprezentować empirycznych ilości. Jako alternatywę dla podejścia Stevensa, przywołuje aksjomatyczną teorię pomiaru Krantz i wsp. (2014). Jak zauważa jednak sam Habilitant, ta ostatnia jest nieco zapomniana i nie zyskała szerokiego zainteresowania wśród badaczy. Pewną nadzieję Autor upatruje natomiast w używanym szerzej w psychologii teorii odpowiedzi na pozycje testowe (*item response theory*) i powiązania wyników w kontekście probabilistycznego powiązania między empiryczną właściwością a latentnym konstruktem. Na końcu rozdziału, Autor wskazuje również na ogólniejszą teorię pomiaru – teorię reprezentacji – jednakże stwierdza, że wykracza ona poza ramy obecnych rozważań. Szkoda, bo wydaje się, że po krytyce różnych podejść, warto wskazać na potencjalne rozwiązania. Rozdział kończy konkluzja, że pomimo pewnych problemów, pomiar w psychologii jest możliwy, choć badacze powinni wykazywać więcej refleksji na temat przedmiotu badania. Cały rozdział jest interesującym przedstawieniem historii pomiaru, jednakże trudno stwierdzić w jakim stopniu łączy się z następnymi częściami książki. Jest to jedyny rozdział czysto teoretyczny, spodziewać by się można na wskazanie w nim potencjalnych zagadnień, które zostaną podjęte w dalszej części, to jednak nie następuje. [...] Drugi rozdział podejmuje zagadnienie oddziaływania wglądu i samowiedzy osób badanych na wyniki miar

samoopisowych. Autor zauważa, że poziom zdolności do introspekcji może mieć znaczenie dla wyników narzędzi samoopisowych. Autor postanowił przyjrzeć się obciążeniu wyników badań samoopisowych zdolnością do introspekcji. Postanowił przeprowadzić analizy symulacyjne związków pomiędzy kwestionariuszem a obiektywnym kryterium zakładając trzy scenariusze. Założone modele różniły się tym, w jakim stopniu poziom samowiedzy wpływa na zniekształcanie odpowiedzi. Następnie testował jak obciążenie wyników „czynnikiem ludzkim” (zdolnością do wglądu) wpływało na korelację pomiędzy kryterium a narzędziem samoopisowym. Wyniki ujawniły znaczący wpływ obciążenia samowiedzą na współczynniki korelacji. Przedstawione analizy wydają się interesującym podejściem, jednakże, jak wskazuję poniżej, w rozdziale brakuje osadzenia teoretycznego, co utrudnia interpretację wyników. [...] Rozdział trzeci zgłębia zagadnienie trafności pomiaru. W pierwszej części rozdziału dokonano historycznego przeglądu podejść do trafności w psychometrii, by następnie przejść do badań własnych. Te ostatnie to, jak utrzymuje Habilitant, systematyczny przegląd literatury, którego celem jest ustalenie jakich metod używają badacze do analizy trafności. W pracy dokonano analizy 100 artykułów podejmujących to zagadnienie. Ogólne wnioski z tych badań wskazują na confirmacyjną analizę czynnikową (CFA) jako główną metodę testowania trafności, następne w kolejności są korelacje oraz eksploracyjna analiza czynnikowa (EFA). [...] Czwarty rozdział wydaje mi się najbardziej twórczy. Autor podejmuje w nim zagadnienie tego, na ile stosowane w psychometrii metody są w stanie „przepuścić” złe narzędzie. W tym celu, Habilitant postanowił zbadać właściwości psychometryczne używanego w psychologii kwestionariusza do badania samoskuteczności, modyfikując jego treść. Stworzył kilka jego wersji, m.in. taką, w której kluczowe słowo zostało zamienione na „poczuciowość”, wersję z bezsensownymi zdaniami i wersję z bezsensownymi słowami. Następnie korelował wszystkie wersje z kwestionariuszem osobowości, analizował ich rzetelność i strukturę czynnikową. Wniosek płynący z tych badań był taki, że uznane w psychometrii narzędzia statystyczne nie pozwalają dobrze odróżnić narzędzi sensownych od celowo bezsensownych. [...] Opisana powyżej monografia z pewnością pobudza do zastanowienia się nad istotą pomiaru, tego, w jakim stopniu liczby oddają rzeczywistość psychiczną. Skłania również do krytycznej oceny stosowanych w psychometrii metod. Jednakże w mojej opinii, monografia jako osiągnięcie naukowe ma dość sporo niedoskonałości, jak na dzieło będące podstawą do ubiegania się o stopień doktora habilitowanego. Po pierwsze, monografia nie przedstawia spójnego projektu a raczej kilka, dość luźno powiązanych, rozdziałów podejmujących różne zagadnienia w obrębie tematyki pomiaru psychologicznego. Zarówno w samej monografii nie znajdujemy syntezy całości, ani we wprowadzeniu, ani w zakończeniu. Również w autoreferacie nie znalazło się miejsce na podsumowanie i pokazanie jak poszczególne aspekty badane w monografii mają się do siebie. Rozdział pierwszy przedstawia ogólne rozważania historyczne i krytykę niektórych podejść, jednak nie jest to wstęp do całego projektu, w którym przedstawione zostałyby założenia, pytania badawcze czy hipotezy. Szkoda, bo forma monografii (w przeciwieństwie do cyklu artykułów) umożliwia takie podejście. Kolejne rozdziały traktują o obciążeniu wyników samoopisowych zdolnością do introspekcji, praktyce badaczy oraz niedoskonałościach metod psychometrycznych. [...] Wydaje mi się, że praca zyskałaby znacznie, gdyby Autor skupił się na jednym z nich i pogłębił dane zagadnienie. Przykładowo, można by przyjrzeć się tematyce samowiedzy, rozpoczynając od symulacji danych a następnie przeprowadzić badania na ludziach. Albo pogłębić badanie dotyczące analiz bezsensownych narzędzi. To bardzo interesujące badanie przedstawione przez Habilitanta aż się prosi o rozwinięcie w serii

kolejnych badań, które odpowiedzą na bardziej szczegółowe zagadnienia. Tak się jednak nie dzieje, w zamian zaś otrzymujemy kilka odrębnych badań i analiz. Niestety cierpi na tym również poziom zagłębienia w temat. W wielu miejscach Habilitant stwierdza, że dane zagadnienie wykracza poza ramy monografii, po czym przyjmuje różne, dość arbitralne, rozwiązania nie osadzając ich w literaturze, czy nie uzasadniając wyborów badawczych. [...] Szczegółowa analiza rozdziału pierwszego skłania do opinii, że jest to niezły przegląd historyczny podejść do pomiaru w psychologii. Sam w sobie byłby zapewne ciekawym esejem (został zresztą częściowo opublikowany wcześniej), jednakże jako wprowadzenie do całej monografii pozostawia niedosyt. Autor krytykuje używane w psychologii pragmatyczne podejście operacjonistyczne, słusznie wskazuje na brak refleksji nad stosowaniem liczb do zjawisk psychicznych. Wszystko to jednak już wiemy, ta krytyka była już podejmowana wśród badaczy w przeszłości. Jedną z propozycji to odświeżenie teorii aksjomatycznej, co mogłoby być ciekawe, gdyby w dalszej części pracy do niej nawiązywał, prowadził wokół niej badania, pokazując jej użyteczność. Tak się jednak nie dzieje, czytelnik zostaje z krytyką i konkluzją, że pomimo problemów, jednak zjawiska psychiczne da się badać. Rozdział pobudza do myślenia, ale na tym pozostaje. [...] Drugi rozdział wzbudza najwięcej wątpliwości. Habilitant przygląda się w nim samowiedzy i testuje pewne założone modele za pomocą analiz symulacji. We wprowadzeniu wspomina, że co prawda istnieją badania na temat samowiedzy, ale stwierdza, że nie chce tutaj „prowadzić szerokich rozważań stricte na temat konstruktów samowiedzy” (s. 41). Uważam takie podejście za spory błąd, który ma konsekwencje dla kolejnego kroku Autora. Habilitant przechodzi bowiem do przedstawienia proponowanych przez siebie modeli. Te ostatnie nie są jednak osadzone w żadnej literaturze, nie odwołuje się do żadnych innych badań, nikogo nie cytuje. Po prostu stwierdza, że przyjął następujące trzy sytuacje wpływu samowiedzy na wyniki w kwestionariuszu. [...] Pogłębienie tematu samowiedzy pozwoliłoby zatem na osadzenie teoretycznych testowanych modeli. Innym wątkiem, który warto by tu również podjąć jest kwestia tego, czym jest kryterium. Z przyjętych przez Autora założeń wynika, że ma ono oddawać badaną cechę, jednakże wiemy, że istniejące kryteria są również jedynie przybliżeniem określonej cechy. Na przykład wynik w teście inteligencji nie oddaje idealnie samej cechy, tak samo średnia ocen szkolnych nie oddaje idealnie stanu wiedzy uczniów. [...] Trzeci rozdział dokonuje przeglądu literatury, aby odpowiedzieć na pytanie jakich metod używają współcześnie badacze do oceny trafności narzędzi psychologicznych. Jest to niewątpliwie interesujący wątek, jednak lektura tego rozdziału znów nasuwa kilka problemów. Po pierwsze, trudno powiedzieć co ma na celu takie badanie, tzn. do czego może nam się przydać wiedza na temat stosowanych w praktyce metod. Po drugie, to wiąże się znów z brakiem osadzenia teoretycznego tych analiz. Autor nie przedstawia żadnych pytań badawczych, hipotez ani przypuszczeń. [...] Sam systematyczny przegląd też ma pewne niedociągnięcia. Przykładowo, dlaczego Autor wybrał jedynie 100 artykułów, dlaczego odrzucił artykuły o płatnym dostępie (podejrzewam, że było ich sporo), czy analizował jakość artykułów i czasopism? Wreszcie, jak sam zauważa, w próbie uzyskał przewagę artykułów z roku 2018, ponieważ pojawiały się w bazie jako pierwsze. To wskazuje na to, że nie brał pod uwagę czynnika daty publikacji, który wydaje się tu ważny. [...] Jak już wspominałem, ostatni rozdział wydaje mi się najciekawszy i najbardziej twórczy. Przedstawia również jedyne badanie z udziałem ludzi. Wydaje mi się, że rozwinięcie tematu sensorycznych/bezsensorycznych narzędzi mogłoby sporo wnieść do lepszego rozumienia metod psychometrycznych. Analiza dokonana przez Habilitanta na to wskazuje. Pokazał, że bezkrytyczne wykorzystywanie popularnych metod może prowadzić do

poważnych błędów. Dodatkowo, rzuca to również cień na standardowe narzędzia, wszakże popularny kwestionariusz do badania osobowości korelował zarówno z prawdziwym kwestionariuszem oraz całkiem podobnie korelował z jego zmienioną wersją, która celowo nie mierzyła nic sensownego. Te wyniki sprawiają, że można być sceptycznym co do wielu danych psychologicznych, zwłaszcza takich, które analizują związki samoopisu z samoopisem. Pomimo, że pomysł na to badanie uważam za bardzo dobry, prezentacja całego rozdziału czwartego ma jednak sporo niedociągnięć. Ponownie, jak w poprzednich rozdziałach, brakuje odniesienia teoretycznego, jasnych pytań badawczych lub hipotez. Opis metody jest bardzo skrótowy i brakuje w nim najważniejszych danych. [...] Autor nie tłumaczy również swoich decyzji metodologicznych, np. dlaczego zdecydował się na użycie terminu „poczuciowość” albo w jaki sposób tworzył bezsensowne słowa. [...] Wreszcie, wydaje się, że Habilitant wyciąga zbyt pochopnie niektóre wnioski. Przykładowo, pisze, że rodzaj użytego kwestionariusza nie wpływa na współczynniki korelacji między osobowością a zmienionym kwestionariuszem GSES, podczas gdy nie przedstawia żadnych statystyk na potwierdzenie. Na pierwszy rzut oka, analiza statystycznych różnic w korelacjach wskazuje, że np. korelacje zwykłego kwestionariusz GSES różnią się od tego z bezsensownymi wyrazami. Trudno jednak poddać je analizie, gdyż Autor nie podaje dokładnych liczebności w poszczególnych grupach. Kolejna sprawa dotyczy interpretacji współczynników rzetelności. [...] Autor nie próbuje jednak dociekać przyczyn wysokiej spójności wewnętrznej tych kwestionariuszy. Niemniej jednak, samo pokazanie tego, że bezsensowny kwestionariusz ma wysoką rzetelność uważam za wartościowe i ważne dla praktyki psychometrycznej. Badacze z pewnością zbyt lekkomyślnie używają metod analiz, uznając, że jeżeli przekracza pewną wartość to nie trzeba się nią dalej przejmować. Badanie dr. Kleki pokazuje, że to błąd. [...] Ostatnia kwestia, o której chciałbym wspomnieć to brak wskazania ograniczeń własnych badań oraz wskazania ewentualnych dalszych kierunków. W podsumowaniu: moja ocena przedstawionego osiągnięcia naukowego jest zróżnicowana. Doceniam wiele jego elementów takich jak spora erudycja Habilitanta związana z historią pomiaru, wiele bardzo interesujących pomysłów na badania, czy też zwrócenie uwagi na fundamentalne zagadnienia dla całej naukowej psychologii. Habilitant zauważa w autoreferacie, że jego „monografia ma na celu wzbudzić refleksję nad własną praktyką psychometryczną i dydaktyczną w zainteresowanym tą problematyką czytelniku”. Uważam, że dr Kleka osiąga tak postawiony cel znakomicie – rzeczywiście lektura jego książki każe pomyśleć o własnych praktykach badawczych, skłania ku refleksji bardziej fundamentalnej – co właściwie badamy. Jednocześnie jednak, trudno odwrócić uwagę od niedociągnięć całego projektu. W mojej opinii, monografia wymaga sporego dopracowania, rozwinięcia pewnych wątków, przedstawienia planu badania, pytań badawczych, hipotez, osadzenia w teorii poszczególnych badań własnych, jasnej syntezy uzyskanych wyników, czy wskazanie na ograniczenia badań. Moja konkluzja jest zatem negatywna. Stwierdzam to z przykrością, ponieważ widzę ogromny potencjał w opisanych przez Habilitanta pomysłach, które mogłyby być rozwinięte i pogłębione.

Ocena pozostałego dorobku naukowego oraz aktywności organizacyjnej: Poza monografią zgłoszoną jako osiągnięcie habilitacyjne, Dr Kleka posiada w dorobku pokaźną liczbę publikacji, bo aż 43, po uzyskaniu stopnia doktora. Spora część prac opublikowana została w prestiżowych międzynarodowych czasopiśmie. Publikacje dotyczą różnych zagadnień w tym psychologii zdrowia, metapoznania, pomiaru oraz właściwości psychometrycznych narzędzi. Sumaryczny Impact Factor tych prac jest wysoki i wynosi 45.746. Habilitant wykazuje również wysoki wskaźnik cytowań – wg

Google Scholar wynosi on 479 a indeks H = 10. Od strony ilościowej ten dorobek jest bardzo duży jak na etap kariery – przed habilitacją. W większości prac, zwłaszcza tych opublikowanych w prestiżowych czasopismach, Habilitant nie jest jednak wiodącym autorem, nie przedstawia również informacji o jego wkładzie w każdą publikację. Na podstawie autoreferatu można przypuszczać, że Habilitant jest specjalistą i ekspertem w dziedzinie metodologii, statystyki i psychometrii. [...] Za bardzo ważne uznać należy jego prace nad tworzeniem nowych narzędzi, ich skracaniem oraz adaptacją istniejących testów. Przykładem jest tu Test Sortowania Kart dla Dzieci. Habilitant współpracuje w tym zakresie z wieloma badaczami z Polski, z różnych ośrodków i uczelni. Są to prace nieocenione w kontekście praktyki psychologicznej. [...] Dr Kleka raportuje również udział w 31 konferencjach krajowych oraz 6 międzynarodowych. Wykazuje również udział w sześciu projektach badawczych krajowych. Niestety nie określa charakteru udziału (kierownik, wykonawca), natomiast wskazuje na zakres prac: analizy statystyczne oraz konstrukcja narzędzi. W jednym projekcie uczestniczył w komisji rekrutacyjnej, co trudno uznać za merytoryczny wkład. [...] Dr Kleka brał udział w dwóch warsztatach statystycznych prowadzonych w zagranicznych ośrodkach, jest członkiem trzech towarzystw naukowych, przygotowywał recenzje dla wielu czasopism naukowych zarówno polskich, jak i zagranicznych. Brał również udział w organizacji pięciu polskich konferencji oraz wykonał kilka ekspertyz statystycznych i metodologicznych różnych projektów naukowych i komercyjnych. W podsumowaniu: dodatkowy dorobek naukowy dr. Kleki jest dość bogaty i zróżnicowany. Habilitant we współpracy z wieloma badaczami z różnych ośrodków opublikował szereg prac w uznanych międzynarodowych czasopismach, ma spory sumaryczny Impact Factor oraz wskaźniki bibliometryczne adekwatne do etapu kariery naukowej, na której się znajduje. Jednocześnie jego rola w większości publikacji nie jest wiodąca, co utrudnia ocenę na ile jest w stanie samodzielnie, z sukcesem, opublikować wyniki własnych badań w uznanych czasopismach. Dr Kleka uczestniczy aktywnie w życiu naukowym poprzez udział w konferencjach, współpracuje w projektach naukowych, choć sam nie kierował projektem finansowanym ze źródła zewnętrznego.

Konkluzja końcowa Recenzenta: Przeprowadzona ocena dorobku naukowego dr. Pawła Kleki wskazuje na jego mocne i słabe strony. Doceniam oryginalne pomysły badawcze oraz poziom erudycji zawarte w monografii stanowiącej podstawę ubiegania się o stopień doktora habilitowanego. Habilitant posiada również ilościowo spory dorobek dodatkowy, współpracuje z wieloma badaczami z różnych ośrodków, jest aktywny naukowo. Jednocześnie, szczegółowa ocena, zwłaszcza głównego osiągnięcia naukowego, przedstawiona wyżej wskazuje na spore niedostatki monografii Habilitanta. Moja ocena jest więc negatywna, co czynię z przykrością, ponieważ widzę spory potencjał prac Habilitanta. Na tym etapie stwierdzam jednak, że oceniane osiągnięcie naukowe nie spełnia wymogów ustawowych niezbędnych do uzyskania stopnia doktora habilitowanego.

Opinie członków komisji habilitacyjnej innych niż recenzenci

Dr hab. Maciej Hanczakowski: ocenił pozytywnie zarówno główne osiągnięcie naukowe, jak też dorobek naukowy Habilitanta. W stosunku do monografii wskazał na dwa zasadnicze elementy krytyczne: brak próby wyjaśnienia uzyskanych efektów w badaniach własnych Habilitanta, co stanowi istotny wymóg prac empirycznych oraz przyjęcie monografii za podstawę głównego osiągnięcia naukowego, a nie cyklu artykułów, które byłyby poddane szerszej ocenie recenzentkiej. W stosunku do dorobku naukowego zaznaczył, iż jest to w zasadzie dorobek o charakterze analitycznym, pochodzący z badań innych autorów i mało jest w dorobku Habilitanta

prac jednoautorskich, prezentujących wyniki własnych badań. W konkluzji stwierdził jednak, że poziom monografii, jak też dorobku pozwalają na stwierdzenie, iż spełnione zostały warunki ustawowe.

Dr hab. Łukasz Kaczmarek: podkreślił znaczenie monografii, która podejmuje kluczowe kwestie pomiaru w psychologii – jest ciekawa i ważna, choć nie stanowi przełomu. W stosunku do dorobku naukowego wskazał na ważny wkład Habilitanta w publikacje, podkreślając Jego kreatywność w tworzeniu nowych metod oraz kreowaniu nowych podejść analitycznych. Wskazał jednak także, iż w dorobku brakuje prac, stanowiących uogólnienie osiągnięć badawczych Habilitanta. W konkluzji stwierdził, że monografia stanowi znaczny wkład w rozwój psychologii, zaś Habilitant wykazuje się istotną aktywnością naukową.

Prof. dr hab. Bogdan Zawadzki: zaznaczył, że podziela uwagi krytyczne, wyrażone zarówno w recenzjach, jak też opiniach członków komisji. Habilitantowi należy się jednak uznanie za podjęcie kluczowego problemu pomiaru w psychologii oraz wyrażenie krytycznego stanowiska wobec stosowanych metod badawczych, sformułowanego na podstawie wyników własnych badań. Kwestie te wymagają pogłębienia, ale monografia stanowi wkład w rozwój psychologii, zaś Habilitant - bez wątpliwości - wykazał się istotną aktywnością naukową.

Przebieg dyskusji oraz oceny osiągnięcia i dorobku naukowego Habilitanta

W dyskusji opinie dotyczące osiągnięcia naukowego, aktywności naukowej, jak też kolokwium były podzielone. W stosunku do monografii habilitacyjnej wskazywano zarówno na jej istotny wkład do psychologii, jak też konieczność pogłębienia teoretycznego podjętych problemów badawczych oraz pogłębienia analiz uzyskanych wyników badań własnych. W stosunku do przebiegu kolokwium habilitacyjnego wskazywano zarówno na uzyskanie wyjaśnienia wątpliwości teoretycznych, jak też pogłębienie niejasności uzyskanych wyników i konieczność podjęcia dalszych analiz. W stosunku do aktywności naukowej wskazywano, iż spełnia ona warunek ustawowy i była realizowana w więcej niż jednej jednostce naukowej. Podkreślano, iż ilościowo dorobek publikacyjny jest obszerny i opublikowany w dobrych czasopismach, ale treściowo - mocno rozproszony (co wynika z faktu, iż Habilitant wykonywał w nich głównie zadania analityczne). Wskazywano także na małą liczbę samodzielnych prac prezentujących własne dokonania badawcze Habilitanta oraz stosunkowo małą liczbę zagranicznych prezentacji konferencyjnych (przy bardzo licznych wystąpieniach krajowych). Podkreślono jednak, że Habilitant wykazuje znakomite umiejętności metodologiczne, psychometryczne i analityczne, ma duży potencjał badawczy i już obecnie może być uznany za samodzielnego badacza. Przy pozytywnej ocenie aktywności naukowej Habilitanta (oraz dodatkowo – Jego osiągnięć dydaktycznych, organizacyjnych i popularyzacyjnych) kluczowym aspektem była ocena osiągnięcia naukowego, która uzyskała (jednoznacznie albo umiarkowanie) pozytywną ocenę u większości członków komisji habilitacyjnej. Opinie te znalazły swoje odzwierciedlenie w głosowaniu uchwały habilitacyjnej.

Uchwała komisji

Biorąc pod uwagę sporządzone w postępowaniu recenzje oraz opinie i stanowiska członków komisji, uwzględniając ocenę osiągnięcia naukowego kandydata oraz aktywność naukową, komisja habilitacyjna na posiedzeniu przeprowadzonym zdalnie w głosowaniu jawnym, w obecności wszystkich siedmiu członków podjęła uchwałę zawierającą pozytywną opinię w sprawie nadania dr. Pawłowi Klece stopnia doktora habilitowanego w dziedzinie nauk społecznych w dyscyplinie psychologia. Uzyskano

następujące wyniki głosowania: głosowało: 7, głosów za: 5, głosów przeciwnych: 1, głosów wstrzymujących się: 1.

§ 2

Na niniejszą uchwałę nie przysługuje zażalenie. Uchwała wchodzi w życie z dniem jej podjęcia.

PODPISY:

Przewodniczący/a Komisji habilitacyjnej

Prof. dr hab. Bogdan Zawadzki



.....

Treść uchwały została zaaprobowana przez wszystkich członków komisji habilitacyjnej