

Opinia

w postępowaniu o nadania stopnia doktora habilitowanego

Panu doktorowi Pawłowi Klece

Przedstawienie podstawowych danych o kandydacie

Pan dr Paweł Kleka uzyskał stopień doktora w roku 2013, na Uniwersytecie im. Adama Mickiewicza w Poznaniu, na podstawie rozprawy zatytułowanej „Zastosowanie teorii odpowiadania na pozycje testowe (IRT) do tworzenia skróconych wersji i kwestionariuszy psychologicznych”. Nie ubiegał się wcześniej o nadanie stopnia doktora habilitowanego. Od 2001 do 2013 roku był zatrudniony na Uniwersytecie im. Adama Mickiewicza w Poznaniu na stanowisku specjalisty naukowo-badawczego, a od 2013 roku na tymże uniwersytecie na stanowisku adiunkta. Od 2021 roku pracuje także na 1/2 etatu na stanowisku adiunkta w Instytucie Psychologii Wydziału Nauk Społecznych Uniwersytetu Zielonogórskiego.

Przedstawienie informacji o obowiązujących przepisach prawa na dzień wszczęcia ocenianego postępowania habilitacyjnego, w tym obowiązujących kryteriach oceny

Na dzień wszczęcia ocenianego postępowania habilitacyjnego obowiązywała ustawa z dnia 20 lipca 2018 r. „Prawo o szkolnictwie wyższym i nauce” (z późniejszymi nowelizacjami i zmianami). Kryteria nadania stopnia dra habilitowanego precyzuje art. 219, pkt.1, 1-3:

1. Stopień doktora habilitowanego nadaje się osobie, która:

- 1) posiada stopień doktora;

2) posiada w dorobku osiągnięcia naukowe albo artystyczne, stanowiące znaczny wkład w rozwój określonej dyscypliny, w tym co najmniej:

- a) 1 monografię naukową wydaną przez wydawnictwo, które w roku opublikowania monografii w ostatecznej formie było ujęte w wykazie sporządzonym zgodnie z przepisami wydanymi na podstawie art. 267 kryteria ewaluacji jakości działalności naukowej ust. 2 pkt 2 lit. a, lub
- b) 1 cykl powiązanych tematycznie artykułów naukowych opublikowanych w czasopismach naukowych lub w recenzowanych materiałach z konferencji międzynarodowych, które w roku opublikowania artykułu w ostatecznej formie były ujęte w wykazie sporządzonym zgodnie z przepisami wydanymi na podstawie art. 267 kryteria ewaluacji jakości działalności naukowej ust. 2 pkt 2 lit. b, lub
- c) 1 zrealizowane oryginalne osiągnięcie projektowe, konstrukcyjne, technologiczne lub artystyczne;

3) wykazuje się istotną aktywnością naukową albo artystyczną realizowaną w więcej niż jednej uczelni, instytucji naukowej lub instytucji kultury, w szczególności zagranicznej.

Jednocześnie, art. 221, ust. 8 wspomnianej ustawy stanowi:

8. Recenzenci, w terminie 8 tygodni od dnia doręczenia im wniosku, oceniają, czy osiągnięcia naukowe osoby ubiegającej się o stopień doktora habilitowanego odpowiadają wymaganiom określonym w art. 219 warunki nadania stopnia doktora habilitowanego ust. 1 pkt 2, i przygotowują recenzje.

Z powyższego wynika, że opinia sporządzona przez recenzenta nie musi zawierać aktywności dydaktycznej lub organizacyjnej Habilitanta, w związku z czym kwestiom tym

nie poświęcam w niniejszej recenzji dużo miejsca (aczkolwiek ocena tych aspektów jest w przypadku dra Kleki jednoznacznie pozytywna).

Przedstawienie informacji o ocenianych osiągnięciach naukowych

Dr Kleka ubiega się o stopień dra habilitowanego na podstawie monografii zatytułowanej „Trafność pomiaru w praktyce psychometrycznej. Badania empiryczne” (2021, Poznań, Wydawnictwo Nauk Społecznych i Humanistycznych Uniwersytetu im. Adama Mickiewicza). 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 czasopismach 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 dra 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 pisu miejsc zatrudnienia), a tym samym spełniony jest ustawowy wymóg wykazywania się istotną aktywnością naukową realizowaną w więcej niż jednej uczelni czy instytucji naukowej.

Ocena monografii „Trafność pomiaru w praktyce psychometrycznej. Badania empiryczne”

Uwagi ogólne

Dr Kleka ubiega się o stopień dra habilitowanego na podstawie osiągnięć naukowych zawartych w monografii „Trafność pomiaru w praktyce psychometrycznej. Badania empiryczne”. Poniżej przedstawiam moją ocenę tej monografii.

Zacznę od tego, iż samą tematykę tej książki oceniam jako bardzo ważną i potrzebną. O wiele zbyt często widzimy 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: wystarczy kilkaset czy nawet kilkadziesiąt osób badanych i można wyznaczyć alfę Cronbacha i wiele innych

wskaźników. Badanie trafności wymaga w wielu przypadkach znacznie bardziej zaawansowanego programu badawczego.

W drugiej uwadze wstępnej pozwolę sobie streścić i zapowiedzieć konkluzję dotyczącą wymogu ustawowego dotyczącego stanowienia znacznego wkładu w rozwój psychologii. **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 szczegółowa

Rozdział 1. Problematyka pomiaru w psychologii

Autor podaje, że fragment tego rozdziału ukazał się już wcześniej w pracy zbiorowej „Mity współczesnej psychologii” (red. M. Obrębska i A. Pankala, 2020). Niestety bez sięgnięcia do wspomnianej pracy zbiorowej nie wiadomo, o który fragment chodzi.

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 (str. 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” (str. 17-19). Nie wiadomo, jak w tym podziale mieści się lwią część pomiaru w psychologii, czyli kwestionariuszowy samoopisowy pomiar 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” (str. 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” (str. 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. Autor przyjął tutaj wartości opisywane w innych badaniach (Paunonen i LeBel, 2012), dotyczących jednak nie samooceny, tylko aprobaty społecznej. Samowiedza i potrzeba aprobaty społecznej to przecież zupełnie odrębne zjawiska. 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, co ilustruje rys. 2.2.” (str. 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 (str. 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 [...]” (str. 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 (o tych ostatnich recenzent też ma obowiązek pisać). 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 wnikliwe 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) konstruktu a wynikami testu” (str. 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 dra 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.” (str. 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.” (str. 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!” (str. 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” (str.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 „uszkodzanego”: w jednej wersji uszkodzenia, kluczowe stwierdzenia zastąpiono słowem „poczuciowość” (np. „Z łatwością

potrafię się trzymać swojej poczuciowości"; w drugiej - wygenerowano zdania absurdalne (np. „Zmartwione doświadczenie projektuje lot”), w trzeciej - same słowa były absurdalne (np. „Pochan enifru midal sapiks trge huze"). 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!” (str. 101). Wyniki te uważam za najważniejsze osiągnięcie dra Kleki, które przesądzają o tym, iż wniósł On znaczący wkład do psychometrii i psychologii.

Jednak i tu muszę dodać nieco dziegiu do miodu - 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. Dalsze badanie tego zjawiska z pewnością się dokona.

Konkluzja

Stwierdzam, iż moim zdaniem monografia dra Pawła Kleki „Trafność pomiaru w praktyce psychometrycznej. Badania empiryczne” (2021, Poznań, Wydawnictwo Nauk Społecznych i Humanistycznych Uniwersytetu im. Adama Mickiewicza) 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, Pan dr Paweł Kleka zasługuje na przyznanie stopnia doktora habilitowanego.

prof. dr hab. Romuald Polczyk