

WŁADYSŁAW JACEK PALUCHOWSKI

DIAGNOZA
OPARTA NA DOWODACH EMPIRYCZNYCH
– CZY POTRZEBNY JEST „POLSKI BUROS”?

Celem artykułu jest opis diagnozowania w praktyce psychologicznej. Artykuł skupia się na dwóch istotnych problemach diagnozowania opartego na dowodach empirycznych. Pierwszy z nich to zastosowanie EMB w psychologii. Drugi to opis działań i osiągnięć Buros Institute of Mental Measurements. Poddano też dyskusji implikacje dla przyszłości diagnozowania w polskiej praktyce psychologicznej.

Słowa kluczowe: diagnoza oparta na dowodach empirycznych – czy potrzebny jest „polski Buros”? diagnozowanie, EBA, EBM.

Podczas XXXIII Zjazdu Naukowego Polskiego Towarzystwa Psychologicznego miałem zaszczyt organizować i uczestniczyć w dyskusji panelowej na temat „Testy psychologiczne w świetle prawa”¹. Dla mnie najważniejsze były następujące wątki naszej dyskusji:

(1) niska kultura prawna, mająca swe źródło głównie w braku wiedzy (też – na studiach), powodująca naruszanie (głównie) majątkowych praw autorskich;

DR HAB. WŁADYSŁAW JACEK PALUCHOWSKI, Instytut Psychologii UAM, ul. Szamarzewskiego 89, 60-568 Poznań; e-mail: jwpprof@gmail.com

¹ Brały w niej udział także: Aneta Fita (SHL Polska Sp. z o.o.), Ewa Goryńska (Laboratorium Technik Diagnostycznych, Wydział Psychologii UW), Magdalena Kaczmarek (Centrum Technik Pomiaru Psychologicznego SWPS), Ewa Okręglicka-Forysiak (Pracownia Metod Badań Psychologicznych, Uniwersytet Wrocławski), Katarzyna Smyła (Sedlak & Sedlak Personnel Consulting), Joanna Stańczak (Pracownia Testów Psychologicznych Polskiego Towarzystwa Psychologicznego Sp. z o.o.) oraz Tatiana Szostok (ALTA Sp. z o.o.). Wszystkim Im dziękuję za krytyczne uwagi dotyczące pierwszej wersji tego artykułu.

(2) niski poziom standardów zawodowych użytkowników testów (psychologów), a także niski poziom oczekiwań co do standardów, jakie mają spełniać narzędzia psychologiczne;

(3) oraz – jako konsekwencja tej diagnozy stanu testowania – a) konieczność powołania instytucji zajmującej się propagowaniem standardów używania testów i formułującej ogólne zalecenia w tym zakresie; b) konieczność powołania instytucji zbierającej informacje o konkretnych testach rynkowych i wydającej o nich niezależne opinie oraz je upowszechniającej; c) wspieranie powstawania większej liczby prac akademickich (w tym magisterskich) na temat testów komercyjnych.

Problemy poruszane podczas tej dyskusji dotyczyły głównie niskiej jakości naukowej bazy praktyki testowania oraz etycznych i prawnych aspektów tej praktyki. Niniejszy artykuł jest podsumowaniem pierwszego wątku tej dyskusji.

Powszechne jest błędne przekonanie, że tylko w fazie przygotowywania testu i jego komercjalizacji ważna jest poprawność metodologiczna procedur i – co za tym idzie – że po opublikowaniu testu psychologiczny jest skończonym produktem (narzędziem). Jednak równie ważne jest to, co dzieje się po opublikowaniu narzędzia – ustawiczna aktualizacja norm oraz badania naukowe nad obszarami jego trafności². O wartości testu psychologicznego decyduje więc zarówno przestrzeganie standardów psychologicznych podczas jego tworzenia oraz podczas korzystania z niego w praktyce, jak i liczba prac naukowych, których jest on przedmiotem. Problem oparcia diagnozowania na dowodach naukowych oraz swoistej kontroli rynku komercyjnych narzędzi diagnostycznych jest ważny z jednej strony dlatego, że coraz silniejszy jest nacisk na analizowanie wiarygodności diagnostycznych narzędzi psychologicznych, oraz dlatego, że – z drugiej strony – niewiedza potencjalnych klientów sprzyja swoistej „tabloidyzacji” testów. Wzorcowe może być przedsięwzięcie Oskara K. Burosa – utworzenia organizacji zajmującej się zarówno inicjowaniem badań naukowych dotyczących testów, jak i krytyczną refleksją nad istniejącym instrumentarium diagnostycznym psychologów praktyków.

² Na przykład czy test sprawdzony „w klinice” ma sens podczas selekcji (co jest w polskiej praktyce błędem nagminnym).

OD TEGO SIĘ ZACZEŁO
– MEDYCYNA OPARTA NA DOWODACH EMPIRYCZNYCH

Rola badań naukowych nad procesem diagnozowania i stosowanych w nim narzędzi rośnie w ciągu ostatnich lat. Jest to z jednej strony spowodowane środowiskowym naciskiem na to, by decyzje dotyczące ludzi oparte były na najlepszej i najnowszej wiedzy naukowej (dowodach, przesłankach), oraz – z drugiej – nacisku instytucji dysponujących publicznymi pieniędzmi, by racjonalnie wykorzystać ograniczone środki przeznaczone na służbę zdrowia i wydawać je tylko na te procedury, których trafności i skuteczności dowiedziono naukowo³. W efekcie nie wystarcza już wiedza zdobyta w czasie studiów, ale konieczne jest ciągle jej wzbogacanie przez całe życie zawodowe.

Ruch EBM wychodzi poza to, co można nazwać medycyną naukową, i kładzie nacisk na sposób gromadzenia naukowych danych użytecznych klinicznie (praktycznie). Jest on nie tylko walką z medycyną intuicyjno-kazuistyczną, opartą na przesądach i niezweryfikowanych empirycznie metodach, ale także przeciwwagą dla medycyny akademickiej, nastawionej na badania podstawowe. Ruch ten rozpoczął się w medycynie, bo to tam powstało pojęcie Evidence Based Medicine (EBM) – czyli strategii zdobywania wiedzy opartej na poprawnie metodologicznie przeprowadzonych i aktualnych badaniach naukowych⁴ i jednocześnie potrzebnej w codziennej praktyce do podejmowania decyzji klinicznych w zakresie diagnostyki i leczenia. Istotą podejścia EBM jest więc integracja wniosków płynących z doniesień naukowych z doświadczeniem klinicznym oraz preferencjami pacjentów (Guyatt, 2003; Sackett i in., 1996). Te wnioski stają się bazą do wypracowania standardów postępowania, które mają pomagać w podejmowaniu optymalnych decyzji (wytycznych; por. Gajewski i in., 2007; Jaeschke, Cook, Guyatt, 1999b, 1999c; Jaeschke i in., 1999a, 1999b; Jaeschke i in., 1999).

Główny zarys systemu EBM powstał w początkach lat dziewięćdziesiątych XX wieku, w zespole naukowców Uniwersytetu McMasters w Ontario, kierowanym przez Davida L. Sacketta, początkowo jako model teoretyczny. Medycyna oparta na faktach została określona jako „dokładne, jasne i rozważne użycie aktualnie najlepszych danych przy podejmowaniu decyzji terapeutycznych w sto-

³ Najczęściej oznacza to wykorzystywanie badań eksperymentalnych w modelu losowych badań kontrolowanych (*randomized controlled trials*, RCT). Grupa CONSORT (Consolidated Standards of Reporting Trials) opracowała wytyczne co do sposobu prezentowania wyników takich badań (por. Łuszczynska, Doliński, 2008). Warto jednak zaznaczyć, że badania realizowane w takim modelu dotyczą głównie farmakoterapii i (rzadziej) procedur terapeutycznych (np. zabiegów operacyjnych), sporadycznie natomiast procesu czy następstw diagnozy.

⁴ Co nie wyklucza, że dzisiejszy dowód stanie się jutro mitem!

sunku do poszczególnych pacjentów, łączące doświadczenia kliniczne z najlepszymi zewnętrznymi danymi klinicznymi, pochodzącymi z systematycznych badań” (Sackett i in., 1996). EBM to podejście do rozwiązywania problemów klinicznych integrujące (Sackett i in., 1996) wiedzę, umiejętności i doświadczenie kliniczne, dane z badań naukowych oraz wartości pacjenta i jego preferencje (w pediatrii często też rodziców lub opiekunów dziecka; Szajewska, 2008). Dziedziną medycyny, która w największym stopniu przyczyniła się do przyjęcia i do rozwoju naukowych metod badawczych, była nowoczesna epidemiologia (Sackett i in., 1991). Autorem oryginalnego pojęcia jest profesor medycyny Gordon Henry Guyatt, pracujący w Zakładzie Epidemiologii Klinicznej i Biostatystyki na uniwersytecie McMaster w Hamilton w Kanadzie (Guyatt, Cairns, Churchill, 1992; Guyatt, 2003).

W polskiej literaturze nie ma zgody co do tłumaczenia tego angielskiego terminu – tłumaczony jest on jako „medycyna faktów”, „medycyna wiarygodna”, „medycyna oparta na dowodach empirycznych”, „krytyczna ewaluacja badań medycznych” (proponycja Centrum Monitorowania Jakości w Ochronie Zdrowia), „medycyna oparta na dowodach naukowych – MODNa”, „praktyka oparta na wiarygodnych i aktualnych publikacjach (POWAP)” (Jaeschke, Cook, Guyatt, 1998a; Jaeschke i in., 2000). Jednak generalnie w większości publikacji autorzy używają skrótu EBM. Kiedy w 2003 roku w Krakowie został powołany Polski Instytut EBM, to także pozostano przy skrócie EBM – jak piszą jego twórcy (Gajewski, Jaeschke, Mrukowicz, 2003, s. 33): „[...] zdecydowaliśmy się nie forsować polskiej nazwy i posługiwać się oryginalnym określeniem”.

Zgodnie z pierwotnymi założeniami termin EBM miał opisywać (i wspomagać) proces podejmowania decyzji w medycynie praktycznej. Podejście to zmieniło także stosunek do badań naukowych – wymogi ogólnometodologiczne dotyczyły wcześniej przede wszystkim prac naukowych, pisanych w celach poznawczych, a nie praktycznych (choć są też głosy sceptyczne wobec innowacyjności tego podejścia – por. Doroszewski, 2005). Jednym z odrzucanych jest założenie, że wiarygodne wnioski kliniczne powinny być oparte na przedklinicznych badaniach podstawowych (głównie *in vitro* i na zwierzętach), pozwalających na zrozumienie anatomicznych, fizjologicznych i patologicznych podstaw obserwowanych objawów (Jaeschke, Cook, Guyatt, 1998a). Nie oznacza to oczywiście, że w EBM ignoruje się wiedzę ogólną (oraz doświadczenie i intuicję); jest wręcz przeciwnie – uważa się je za konieczne, choć niewystarczające, jednak nacisk kładzie się na jasne wyartykułowanie przesłanek wnioskowania dotyczącego danego, konkretnego przypadku i zwraca uwagę na to, że wnioski oparte na dedukcji z nauk podstawowych mogą być mylne („oświecony sceptycyzm”).

Z drugiej strony tylko wiedza ogólna (dotycząca tzw. modeli eksplanacyjnych, przyczynowych) pozwala dostrzec i zinterpretować wskaźniki.

Innym odrzucanym założeniem (Jaeschke, Cook, Guyatt 1998a; por. też Jaeschke i in., 2000) jest to, że obserwacje pochodzące z praktyki klinicznej (własnej lub innych osób) stanowią wiarygodne i wystarczające źródło praktycznej wiedzy medycznej. Osobiste doświadczenie i praktyka są przeceniane co do ich wartości, zawsze bowiem mają one charakter lokalny, ograniczony do przypadków niereprezentatywnych dla wszystkich pacjentów, ale – kontekstowo – dobranych (próba obciążona, stronicza), a ich ocena zależy od specyficznych cech podmiotowych lekarza (rodzaj wykształcenia czy preferencja poznawcza). Ponadto bazowanie na doświadczeniu autorytetów częściej ma za podstawę wiarę niż mądrość, która zastępowana bywa przez utrwalone stereotypy i modę (Sułkowski, 2007). Jednocześnie zwraca się uwagę na to, że – na przykład – wyniki postępowania diagnostycznego zależą w dużej mierze od poświadczenia diagnosty (Łanda, Walczak, 2005).

Ruch EBM jest formą realizacji bardziej ogólnego paradygmatu – standaryzacji procedur i kontroli jakości. Przyjęcie podejścia EBM wymagało także ujednoczenia metodyki badań klinicznych (co krytycy nazywają McDonaldyzacją medycyny). We wspomnianym Zakładzie Epidemiologii Klinicznej i Biostatystyki na uniwersytecie McMaster zaproponowano odpowiednią strukturę pytania badawczego o akronimie PICO lub PICO(T) – powinno ono zawierać informacje o populacji, której ma dotyczyć wynik badań, rodzaju badanej interwencji, o grupie kontrolnej (porównawczej) i stosowanych wobec niej oddziaływaniach (interwencji alternatywnej) oraz o sposobie mierzenia efektu badanej interwencji; niekiedy dodatkowo wskazuje się na konieczność podania przedziału czasu, w którym obserwowano badane grupy (Sackett i in., 1997; Mrukowicz, 2005). Przykładowo (Mrukowicz, 2005), niepoprawne pytanie badawcze brzmi: „Czy tamoksyfen jest skuteczny w raku sutka?”; natomiast jego poprawna wersja ma postać: „Czy u kobiet w wieku 35-70 lat obciążonych dużym ryzykiem zachorowania na raka sutka (P) tamoksyfen stosowany w dawce 20 mg na 24 h przez 5 lat (I) zmniejsza w porównaniu z placebo (C) ryzyko wystąpienia raka sutka (O)?”

Punktem wyjścia koncepcji EBM była też konstatacja, że wiedza nabywana podczas studiów medycznych często już na studiach jest anachroniczna, a po nich szybko się dezaktualizuje – im dalej od studiów, tym większe jest posiadane doświadczenie z jednej strony i tym mniej aktualna wiedza medyczna z drugiej (Jaeschke, Cook, Guyatt, 1998a). W przeszłości przyjmowano, że wiedza przekazywana podczas studiów jest wiedzą aktualną nie tylko w momencie jej po-

wstawania, ale także wiele lat po studiach. Odrzucenie tego założenia powoduje konieczność ciągłego uzupełniania wiedzy (doświadczenia) – zazwyczaj za pomocą takich środków, jak szukanie w podręcznikach medycznych (często z czasów studiów), zbieranie opinii bardziej doświadczonych współpracowników, udział w konferencjach i szkoleniach itp. (tamże).

Jednak nawet mimo posiadania ogólnej wiedzy metodologicznej, potrzebnej do czytania prac naukowych (o czym od lat pisze Brzeziński – 2004), lekarze praktycy nie są w stanie ocenić i dokonać całościowej analizy publikowanych danych klinicznych, zdobywanych przecież w różnych warunkach, na różnych próbach i za pomocą różnych metod (Jaeschke, Cook, Guyatt, 1998a), nawet jeśli korzystają oni z artykułów przeglądowych (zawierających metaanalizy; Jaeschke, Cook, Guyatt, Bochenek, 1998). Zdaniem zwolenników EBM realizacja postulatu stałego aktualizowania wiedzy jest możliwa tylko wówczas, gdy istnieje możliwość łatwego pozyskania informacji o wynikach badań z baz danych, zawierających publikacje wyselekcjonowane pod względem znaczenia klinicznego i wiarygodności. Konieczne jest też kształcenie w zakresie korzystania z nich, czyli tzw. umiejętności informacyjnych (Lasoń i in., 2005): posiadania wiedzy o źródłach informacji i umiejętność ich wyszukiwania oraz zadawania pytań w odpowiedniej formie (por. opis: Jaeschke, Cook, Guyatt, 1998a, s. 7), a także umiejętności krytycznej oceny znalezionych informacji, ich interpretacji oraz ich porządkowania i syntezy.

Koncepcja EBM ma swoje konsekwencje także dla kształcenia (EBL – Evidence Based Learning, EBE – Evidence Based Education, także BEME – Best Evidence in Medical Education). Określa się je jako kształcenie skoncentrowane na problemach i studencie (*problem based study & student oriented study* – Łanda, Walczak, 2005; Lasoń i in., 2005; Del Mar, Glasziou, Mayer, 2004). Polega ono na tym, że stawia się studentów przed konkretnym problemem klinicznym, który wymaga samodzielnego rozwiązania przez wskazanie najlepszego w tym przypadku sposobu diagnozowania zakończonego prawidłowym rozpoznaniem oraz najlepszego sposobu leczenia, wyprowadzonym z aktualnego stanu wiedzy medycznej. Niezbędnym zapleczem do poszukiwania rozwiązań jest dostęp do medycznych baz danych.

DIAGNOZA PSYCHOLOGICZNA OPARTA NA DOWODACH EMPIRYCZNYCH

W omawianym obszarze psychologia niewiele się różni od medycyny. I tu, w interesie coraz lepiej wyedukowanych klientów, leży posługiwanie się takimi psychologicznymi metodami diagnozy i terapii, których skuteczność udokumentowana jest naukowo. Także dla psychologii ważne jest, by badania były zgodne z jakimś ogólnie przyjętym i akceptowanym standardem. Ważne jest to, by gromadzone dane oparte na poprawnie metodologicznie przeprowadzonych i aktualnych badaniach naukowych miały nie tylko charakter poznawczo-akademicki, ale by były też użyteczne praktycznie. Również w psychologii badania podstawowe trudno przekładają się na praktykę psychologiczną i stanowią słabe uzasadnienie stosowanych metod diagnozowania i interwencji. Jednocześnie wartość naukowa obserwacji własnej lub cudzej praktyki klinicznej także i w psychologii są przeceniane co do ich wartości.

Duch EBM przeniósł się najpierw do badania psychoterapii (EST – Empirically Supported Therapies, por. Westen, Nowotny, Thompson-Brenner, 2004⁵), potem diagnozy (EBA – Evidence-Based Assessment, por. Hunsley, Mash, 2005; Mash, Hunsley, 2005), a następnie ogarnął całość praktyki psychologicznej (EBPP – Evidence-Based Practice in Psychology, Levant i in., 2006; Hunsley, 2007). Recepcja w amerykańskiej psychologii koncepcji EBM jest specyficzna i w jakiejś mierze megalomańska: w raporcie grupy zadaniowej APA⁶ (Levant i in., 2006) wspomina się o tym, że w psychologii postulat empiryczności ma długą historię (datująca się od 1896 roku! – s. 271), co świadczy o dobrym samopoczuciu autorów; że zawsze była wrażliwa na różnice indywidualne⁷ (co skądinąd jest prawdziwe, ale użyte w tym kontekście sugeruje jakieś specjalne w tym względzie zasługi psychologii – s. 274, 278-279); że wśród dowodów empirycznych najważniejsze są te, które czerpane są z obserwacji, analizy przypadku oraz badań podstawowych (a uważane za najbardziej poprawne i wiarygodnie randomizowane badania kliniczne z podwójnie lub potrójnie ślepą próbą znalazły się na miejscu przedostatnim, a na ostatnim metaanalizy – s. 274); że zadaniem psychologów praktyków jest przystosowywanie ogólnych prawidło-

⁵ W innej pracy Westen, Nowotny, Thompson-Brenner (2005) piszą, że EST jest częścią EBPP.

⁶ Notabene pisze się tam (s. 273), że diagnozowanie jest formą interwencji, co pozwala stosować te same (co w ocenianiu terapii) kryteria.

⁷ Z naciskiem na „indywidualne”. Na marginesie warto zaznaczyć, że podnoszenie teoretycznego zakotwiczenia narzędzi ma wyzwolić psychodiagnostykę z dziedzictwa poszukiwania różnic indywidualnych.

wości do konkretnej sytuacji (s. 275, 277, 279). Ten ostatni postulat – uwzględniania indywidualnych właściwości klienta – jest jednak inaczej rozumiany w koncepcji EBM. Jak pisze Mrukowicz (2005), owo uwzględnianie indywidualnych właściwości pacjenta powinno mieć miejsce w czasie badań (zbierania dowodów empirycznych), a nie podczas aplikowania wykrytej w badaniach naukowych prawidłowości ogólnej i uzasadniania odstępowaniem od niej odwoływaniem się do indywidualizacji.

W 2005 roku ukazały się dwa monograficzne numery międzynarodowych czasopism psychologicznych⁸, poświęcone diagnozowaniu opartemu na dowodach empirycznych (EBA). Na czym polega specyfika tego podejścia i co wnosi ono nowego do diagnozy psychologicznej (por. też Stemplewska-Żakowicz, Paluchowski, 2008)? W propozycji EBA podkreśla się konieczność wykorzystywania badań i wiedzy do świadomego wyboru celów diagnozowania, metod i konkretnych narzędzi i/lub procedur diagnostycznych, a także integrowania „twardych” danych naukowych i „miękkich” danych klinicznych. Stosowanie się do standardów EBA to nie tylko respektowanie wymogu psychometrycznej dobroci stosowanych narzędzi, ale pójście dalej⁹.

Podkreśla się (Hunsley, Mash, 2005; Kazdin, 2005) przede wszystkim, że nie ma wiarygodnej interwencji bez wiarygodnej diagnozy. Diagnozowanie oparte na dowodach empirycznych dotyczyć może (powinno) zarówno oceny efektów interwencji, jak i tradycyjnie rozumianego diagnozowania, w tym psychometrycznych parametrów stosowanych narzędzi w odniesieniu do konkretnego celu ich użycia¹⁰ oraz dodatkowo klinicznej i diagnostycznej użyteczności procedury diagnostycznej, kosztu diagnozy, czułości i specyficzności stosowanych wskaźników itp. (Hunsley, Mash, 2005; por. też Paluchowski, 2007, s. 53-57, s. 82; Stemplewska-Żakowicz, Paluchowski, 2008, s. 50). Towarzystwo Psychologii Pediatrycznej (54 Oddział APA) opracowało zestaw kryteriów do oceny narzędzi diagnostycznych, uzupełniony później przez Diane L. Chambless i Thomasa H. Ollendicka (za: Cohen i in., 2008), i obejmujący następujące wysokie wymagania co do metod diagnozy: narzędzie powinno być opisane w co najmniej

⁸ Oba numery monograficzne wydano w 2005 roku, są to: *Psychological Assessment* (tom 17) oraz *Journal of Clinical Child and Adolescent Psychology* (tom 34).

⁹ Co widać w zmianie sposobu podejścia w nowych „Standardach dla procedur stosowania testów w psychologii i edukacji” (AERA, APA, NCME, 2007; Hornowska, 2007).

¹⁰ Szczególnie trafności kryterialnej diagnostycznej oraz aspektu różnicowego trafności i przyrostu (*incremental validity*) trafności predykcji dla konkretnych sytuacji diagnostycznych. Warto to podkreślić – nie mówi się bowiem o cechach psychometrycznych narzędzi „w ogóle”, ale wyników (i decyzji na nich opartych) uzyskanych w konkretnej sytuacji (czy ich klasie).

dwóch recenzowanych czasopismach przez niezależnych badaczy lub zespoły¹¹, muszą być dostępne informacje w celu ewentualnego powtórzenia badań oraz narzędzie musi mieć wystarczające parametry psychometryczne przynajmniej w jednej z tych publikacji¹².

John Hunsley i Eric J. Mash (2005) zwracają uwagę na konieczność – ze względu na stosunkowo mało czasu, który można przeznaczać na diagnozowanie – wykorzystywania tych narzędzi, które dotyczą rdzeniowych aspektów problemu (aspekt użyteczności diagnostycznej), który jest badany (a także takich, które są mniej czasochłonne). Mówi się (Hunsley, Mash, 2005), że ogólne baterie testów powinny ustąpić miejsca testom specyficznych funkcji, co oczywiście wymaga bardziej precyzyjnych hipotez co do źródeł i mechanizmu zaburzeń. Podkreśla się także konieczność empirycznego dokumentowania sposobu integrowania informacji diagnostycznych oraz ich całościowej interpretacji (Hunsley, Mash, 2005). Zwraca się także (Hunsley, Mash, 2005; Hunsley, 2007) uwagę na konieczne zmiany kształcenia psychologów, wynikające z postulatów EBA – większy udział takich metod, jak arkusze samoobserwacji objawów specyficznych dla zaburzenia, ich monitoringu i ustrukturuowanych wywiadów.

W diagnozowaniu zgodnym z postulatami EBM ważne są też takie charakterystyki wskaźników diagnostycznych, jak dokładność, wartość predykcji na podstawie wyników dodatnich (zwana też mocą prognozy pozytywnej), wartość predykcji na podstawie wyników ujemnych, krzywe ROC itp. (Jaeschke, Cook, Guyatt, 1998b, 1998c). Poszukuje się takich testów wzorcowych (test referencyjny, tzw. złoty standard), które mogłyby w danym obszarze stanowić kryterium wiarygodności. Szczegółowe pytania, jakie stawia się w tym obszarze, to (Szajewska, 2004; por. też Stemplewska-Żakowicz, Paluchowski, 2008): jakie są wskaźniki wiarygodności? (czy test jest użyteczny, tzn. czy ma potencjalne znaczenie dla praktyki klinicznej; czy wyniki testu porównano z jakimś testem referencyjnym)? czy wyniki badań będą pomocne w praktyce klinicznej (czy powtarzalność wykonania samego testu i powtarzalność jego interpretacji są wystarczające w danych warunkach; czy charakterystyka danego testu potwierdzi się wśród moich pacjentów; czy wyniki testu zmienią leczenie pacjenta; czy wykonanie testu pomoże mojemu pacjentowi; czy badanie testem określi stopień zaawansowania terapii)?

¹¹ Por. kryteria Dauberta, opisane dalej.

¹² W analizie, w jakim stopniu stosowane faktycznie metody spełniają te wysokie kryteria, okazało się (Holmbeck i in., 2008), że w tym obszarze jest ich aż 92% wśród analizowanych metod. Jednocześnie niewiele z nich spełnia kryteria wymionione wyżej (diagnostyczna i kliniczna użyteczność).

Warto może też wspomnieć o tzw. kryteriach Dauberta (Widacki, 2007; por. też Stemplewska-Żakowicz, Paluchowski, 2008). Stosuje się je w Stanach Zjednoczonych do oceny „naukowości” przywoływanego podczas procesu sądowego dowodu lub opinii biegłego (także psychologa)¹³. Chodziło o to, by zapewnić oskarżonemu poczucie bezpieczeństwa, że badania będące podstawą opinii (dowodu, ekspertyzy) wykonywane są według jasno określonych i sprawdzonych eksperymentalnie kryteriów. Generalnie przyjęto sześć takich kryteriów: falsyfikacji, recenzji, publikacji, wartości diagnostycznej (poziom błędu), standaryzacji i powszechnej akceptacji. Wszystko zaczęło się w roku 1993, gdy przed Sądem Najwyższym Stanów Zjednoczonych toczyła się sprawa *Daubert vs Merrell Dow Pharmaceuticals Inc.*, w której William Daubert i inni wystąpili przeciwko firmie farmaceutycznej o odszkodowanie za upośledzenia noworodków, spowodowane przez zażywanie leku o nazwie *Bendectin* przez matki w ciąży. Sąd wówczas orzekł, że brak jest naukowych dowodów na związek między zażywaniem leku a uszkodzeniem płodu. Kryteria, którymi się kierował, oceniając wiarygodność opinii biegłych, stały się znane jako „kryteria Dauberta”. Były one później uszczegóławiane i rozwijane w sprawach *General Electric Co. vs Joiner* oraz *Kumho Tire Co. vs Carmichael*. Kryteria Dauberta to w gruncie rzeczy – jak piszą William Grove i R. Christopher Barden (1999) – podstawowe zasady naukowego poznania, sformułowane jako pytania o wskaźniki naukowej rzetelności (tamże, s. 226):

1. Czy teoria lub technika, stanowiąca podstawę dowodu, jest sprawdzalna? Czy da się ją empirycznie weryfikować?¹⁴
2. Czy została ona sprawdzona za pomocą trafnych i rzetelnych procedur i czy wynik powtórnych badań był pozytywny?
3. Czy teoria lub technika, stanowiąca podstawę dowodu, została opisana w recenzowanej (*peer review*) publikacji w literaturze fachowej? Czy była faktycznie przedmiotem środowiskowej dyskusji naukowej?
4. Czy znany jest poziom błędów uzyskiwanych przy stosowaniu tej metody? Jaki jest błąd pomiaru techniki? Jak często wynik metody jest błędem pierwszego rodzaju (wynik fałszywie pozytywny, polegający na błędnej identyfikacji), a jak często drugiego (wynik fałszywie negatywny, polegający na błędnej eliminacji)? Jak duże jest ryzyko, że przywoływana w opinii teoria jest błędna?
5. Czy zostały ustanowione standardy dla stosowania danej techniki?

¹³ Swoją drogą, to trochę zabawne, że sąd ustala, iż wiedza naukowa to taka, która zdobyto metodologicznie poprawnym sposobem, oraz jakie są kryteria naukowej (metodologicznej) poprawności.

¹⁴ Jest to odwołanie się wprost do krytycznego racjonalizmu Karla Poppera i tzw. kryterium demarkacji, czyli odróżniania nauki od pseudonauki.

6. Czy dana teoria lub technika jest akceptowana (*general acceptance*) przez odpowiednie środowisko naukowców¹⁵, więc nie tylko specjalistów, ale także przez przedstawicieli ośrodków naukowych zajmujących się daną problematyką? Czy jest ona zgodna z aktualnym stanem wiedzy w danej dziedzinie lub ogólniej – czy jest zgodna z obecnym stanem nauki?

Źródłem tego, jak mogłaby wyglądać praktyka testowania, można więc poszukiwać zarówno w wypracowanych w medycynie postulatów praktyki opartej na dowodach empirycznych, jak i na tzw. standardach Dauberta.

INSTYTUT BUROSA

Przykładem źródła danych o narzędziach psychologicznych może być działalność Buros Institute of Mental Measurements¹⁶. Oscar Krisen Buros (1905-1978) założył swoiste forum wymiany informacji na temat testów dostępnych na rynku (komercyjnych), na którym praktycy i naukowcy mogli publikować recenzje dostępnych narzędzi i polemizować ze sobą, a gdzie ich potencjalni klienci mogli tych informacji poszukiwać. Przyświecały mu ważne cele: poprawa statusu narzędzi pomiarowych w psychologii oraz poprawa praktyki testowania. Forum służy interesom nie tylko psychologów praktyków (pozwalając dokonać wyboru wiarygodnych narzędzi i stymulując prace badawcze), ale także pozwala na kontrolę społeczną. Stworzył on dwie serie publikacji (*Mental Measurements Yearbook* i *Tests in Print*), które stały się źródłem standardów w tym zakresie. Po śmierci Burosa jego dzieło kontynuowała żona, Luella Gubrud Buros. Archiwa Buros Institute zostały przekazane do Teachers College Hall at the University of Nebraska-Lincoln, który teraz zajmuje się publikacjami Instytutu (szerzej o działalności BIMM piszą Claiborn, 1991 oraz Plake, Conoley, Kramer, Murphy, 1991).

Oskar K. Buros był z wykształcenia pedagogiem – w roku 1925 ukończył studia na University of Minnesota (Educational Administration and Supervision)

¹⁵ Jest to pierwotna wersja standardu oceny wiarygodności dowodu czy opinii – tzw. *Frye test* (lub *Frye standard* – pochodzącego ze sprawy *Frye vs United States* z 1923 roku i dotyczącego wiarygodności badań poligraficznych).

¹⁶ Testowanie wiele zawdzięcza też Usługom Testów Edukacyjnych (ETS, Educational Testing Service). Jest to organizacja non-profit, założona w 1947 roku przez Henry’ego Chaunceya (kierował ETC do 1970 roku), by przejąć działalność testową od Amerykańskiej Rady Edukacji. Inaczej niż w przypadku Buros Institute, jej celem jest nie opiniowanie, lecz tworzenie narzędzi pomiaru kompetencji (głównie szkolnych i językowych). Odwołuje się przy tym do standardów Evidence-Centred Design (ECD; Mislevy, Almond, Lukas, 2003). ETS stworzyło popularne testy: SAT, TOEIC, TOEFL i GRE®. Jego polska filia nosi nazwę ETS Global Polska.

i prowadził w latach 1929-1965 zajęcia z psychometrii i metod statystycznych na Rutgers University. Tam też otrzymał tytuł profesora pedagogiki. Stworzony przez niego (i dzięki pomocy żony, z początku jedyne go pracownika) Institute of Mental Measurements pierwotnie koncentrował się na testach edukacyjnych, a Buros zamierzał krytycznie analizować nieuzasadnione obietnice składane przez autorów testów w celu zwiększenia ich sprzedaży. Jednak w opublikowanym w 1938 roku pierwszym *Mental Measurements Yearbook* (MMY) znalazł się przegląd ponad 300 testów wykorzystywanych nie tylko w edukacji, a także psychologii i biznesie. Kolejno ukazywały się nowe roczniki poświęcone nowym testom na rynku – za życia Burosa ukazało się ich osiem, w 2007 roku opublikowano *The Seventeenth Mental Measurements Yearbook* (w latach 1941-1985 MMY ukazywał się nieregularnie, od 1989 co dwa lata). Do 1989 roku opisano 8779 testów stosowanych w edukacji, psychologii, poradnictwie, doradztwie, prawie i zarządzaniu (Plake i in., 1991, s. 451), a łącznie we wszystkich MMY znajduje się 10775 opisów testów. Okazjonalnie pojawiały się też publikacje poświęcone testom w danym obszarze, jak na przykład *Personality Tests and Reviews*, *Science Tests and Reviews*, *Social Studies Tests and Reviews*, *Vocational Tests and Reviews* lub *Reading Tests and Reviews*. Typowa pozycja w MMY zawiera przegląd testu komercyjnego adresowanego do angielskojęzycznych użytkowników i obejmuje opis testu (nazwa, cel, grupa odniesienia, data publikacji, sposób badania, czas badania, autor i wydawca, krótkie omówienie, historia testu, szczegóły techniczne), jedną lub dwie krótkie recenzje (1000-1500 słów) oraz wskazówki literaturowe i bibliografię (por. EFPA/Bartram, Lindley, Kennedy, 2008). Od lat siedemdziesiątych autorzy lub wydawcy testu muszą dostarczyć jego pełną dokumentację dotyczącą własności psychometrycznych, by sprawdzić, czy spełniają podstawowe wymagania w tym zakresie. Recenzenci dobierani są ze względu na tematykę (obszar, grupę docelową) testu oraz doświadczenie w jego stosowaniu, stopień doktora odpowiednio do obszaru testu i znajomość psychometrii. Recenzenci głównie odwołują się do „Standardów dla testów stosowanych w psychologii i pedagogice” (APA 2007), a ich celem jest nie tylko diagnoza stanu rzeczy i jego krytyka, ale także proponowanie kierunku zmian. Recenzowanie jest bezpłatne. Płaci się za opracowania MMY w wersji papierowej lub za korzystanie z recenzji za pośrednictwem Internetu.

W roku 1961 pojawiła się nowa publikacja, uzupełniająca MMY – *Tests in Print* (TIP); dotychczas ukazało się siedem tomów. Jak wskazuje pełna nazwa tej pozycji (*Tests in Print: An Index to Tests, Test Reviews, and the Literature on Specific Tests*), jest to wyczerpujący spis aktualnie dostępnych komercyjnych

testów, zawierający podstawowe o nich informacje (opis testu) oraz bibliografię recenzji o testach zamieszczonych w publikacjach MMY.

W 1994 roku, dzięki wsparciu żony Burosa, powstało Centrum Testowania (Oscar and Luella Buros Center for Testing), obejmujące swym działaniem dwa obszary ważne dla testów: kontynuujący swoją pracę Instytut (Buros Institute of Mental Measurements – BIMM), oferujący także publikacje elektroniczne (w tym *Test Reviews Online*, zawierający informacje na temat testów publikowane w MMY od 1985 roku), oraz Buros Institute for Assessment Consultation and Outreach (BIACO), zajmujący się nie tylko testami spoza poprzednich obszarów zainteresowania, ale szerzej rozumianymi procedurami ewaluacji w nauczaniu i wszelkiego rodzaju certyfikacji i prowadzący na wniosek wydawcy (*test provider*) audyt ich zgodności z *Standards for Proprietary Testing Programs*.

JAK BYWA I CO MOŻNA ZROBIĆ, BY BYŁO LEPIEJ?

Obserwuje się różne naganne zjawiska związane z upowszechnianiem testów i dotyczące głównie łamania zasad etyki zawodowej: publikowanie zadań testowych¹⁷, przykładów rozwiązań lub testów z kluczem¹⁸, szkolenia w rozwiązywaniu testów selekcyjnych¹⁹ itp. Innym nagannym zjawiskiem jest nie tylko kopiowanie metod zamiast ich kupowania, ale także wykorzystywanie bez zgody autora testu z publikacji naukowej do celów komercyjnych lub naukowych czy nawet wykorzystywanie fragmentów (zadań, pytań). Nie wystarczy ustawa o zawodzie psychologa, kodeks (PTP, 2000) czy standardy (AERA, APA, NCME, 2007²⁰), by wyeliminować z rynku tego rodzaju patologie. Nie pomoże w tym i „polski Buros”.

¹⁷ Ciekawa jest sprawa techniki Rorschacha – oficjalnie od 1992 roku plansze testowe nie podlegają już ochronie (po 70 latach od śmierci twórcy utwory stają się własnością publiczną). Zostały one w związku z tym opublikowane m.in. w Wikipedii, co spowodowało protesty Amerykańskiego Stowarzyszenia Psychologicznego oraz Brytyjskiego Towarzystwa Psychologicznego, które twierdzą, że udostępnienie testu jest nieetyczne (Cohen, 2009).

¹⁸ Tu niesławnym przykładem może być czasopismo *Forum Psychologiczne*, wydawane przez Europejską Akademię Psychologii Integracyjnej Sp. z o.o. w Łodzi.

¹⁹ Jak pisze jeden z użytkowników Internetu: „Nawet psychologowie ogłaszają się w sieci z propozycją przygotowania kandydata «pod test»: 4 wizyty – 600 zł”. W kontekście przestrzegania praw autorskich zabawna jest notka zamieszczona w Internecie przez autora zestawu rozwiązań do testu psychologicznego do policji, który powołuje się na prawo autorskie do nich i zabrania innym jego sprzedawania!

²⁰ W których zalecenia w tych sprawach są jasno sformułowane – por. AERA, APA, NCME, 2007, s. 193-205 czy s. 126-132.

Istnieją Wytyczne Międzynarodowej Komisji ds. Testów (International Test Commission – ITC) dotyczące stosowania testów, w tłumaczeniu Teresy Szustrowej (PTP, 2000; por. ITC, 2000), zamieszczone na stronie Pracowni Testów Psychologicznych. Opisuje się tam postępowanie zgodne z zasadami dobrej praktyki: w tym, jak oceniać potencjalną użyteczność badania testowego w danej sytuacji diagnostycznej i jak je adekwatnie do celu badania dobierać, na co zwracać uwagę oceniając technicznie parametry jakości testu, oraz jak sprawdzać przydatność testu w sytuacji zmian społecznego kontekstu ich stosowania. Także Europejska Federacja Towarzystw Psychologicznych (The European Federation of Psychologists Association – EFPA) powołała komisję do opracowania standardów stosowania testów w obszarze psychologii pracy i organizacji (EFPA, EAWOP 2005). Być może zawarte tam postulaty powinny również być poddane dyskusji i być może powinna powstać analogiczna komisja koordynująca ustalenia z polskim porządkiem prawnym i propagująca je w Polsce. Pod jej patronatem opracowano także standardy opisu testów (EFPA/Bartram, Lindley, Kennedy, 2008). Wytyczne nie zastąpią jednak informacji.

Psychologowie swoimi decyzjami wpływają na wiele dziedzin życia: oświatę, wymiar sprawiedliwości (zarówno w sprawach karnych, jak i cywilnych, w tym sprawach dotyczących nieletnich, władzy rodzicielskiej lub prawa opieki nad dzieckiem, zdolności do zawarcia małżeństwa itd.), zdrowie i opiekę społeczną, różne służby medycyny pracy, służby mundurowe itp. (por. Szustrowa, 2003). I tak jak orzekanie w sprawach ściśle regulowanych przez przepisy (takich jak wydawanie zezwoleń różnego rodzaju, badania kierowców²¹ itp.) raczej nie spotyka się z negatywnym odbiorem społecznym, to psychologiczne opinie i ekspertyzy sądowe (Stanik, Majchrzyk 1995) bywają efektem zarówno niewiedzy, jak i stereotypizacji oraz podlegania emocjonalnym przesądom (Stemplewska-Żakowicz, Paluchowski, 2008, s. 72; Słysz, 2008, s. 52-55), co powoduje, że są częstym tematem złych opinii w Internecie (np. opisuje się, że w orzeczeniach co do ograniczenia władzy rodzicielskiej w RODK badania bazowały na „skróconej wersji MMPI”²², i na tej podstawie psycholog orzekał,

²¹ I tu jednak zdarzają się przepisy, które stwarzają dogodne okoliczności do obniżania standardów profesjonalnych – np. w uchylonym w 1999 r. Zarządzeniu Ministra Transportu i Gospodarki Morskiej z dnia 12 stycznia 1995 r. w sprawie kierowania osób ubiegających się o uprawnienia do kierowania pojazdami (MP 1995 nr 3 poz. 41) zalecano, że „w przypadku uzyskania przez osobę badaną w skali introwersji-ekstrawersji, stabilności emocjonalnej-neurotyczności wyniku poza obszarem normy psychometrycznej, psycholog stosuje test MMPI-WISKAD”, a więc odsyłało do narzędzia nie mającego debitu w Polsce oraz ewidentnie nieaktualnego.

²² Jak pisze się o jednym z takich przypadków: „Biegli stwierdzili, że zastosowali wersję skróconą, opracowaną przez psychologa z Poznania. [...] Skontaktowałam się więc z Pracownią Testów Psychologicznych Polskiego Towarzystwa Psychologicznego. Potwierdziły się moje przypuszczenia

„kto daje większe gwarancje prawidłowego wychowywania dziecka”, lub że badano iloraz inteligencji za pomocą testu Rorschacha lub „wybranymi podtestami Wechslera”, uzasadniając to faktem, iż „zastosowano eksperyment kliniczny”, albo w przypadku podejrzenia molestowania badano dziecko testem Drzewa, Zwierzyńiec i CAT²³). Tu więc problem leży – w braku wiedzy opartej na poprawnie metodologicznie przeprowadzonych i aktualnych badaniach naukowych, a nie na lekceważeniu standardów etycznych zawodu.

Od ponad 20 lat mamy do czynienia z komercyjnym rynkiem metod psychologicznych (diagnostycznych), a głównym wydawcą testów jest spółka Pracownia Testów Psychologicznych, której jedynym właścicielem jest Polskie Towarzystwo Psychologiczne²⁴. Możliwość zakupu pewnej grupy testów psychologicznych mają tam wyłącznie osoby legitymujące się dyplomem ukończenia magisterskich studiów psychologicznych²⁵, inne przeznaczone są dla innych profesjonalistów (bez jednoznacznego zdefiniowania tej grupy na stronie internetowej Pracowni), a niektóre dostępne są dla wszystkich chętnych (por. też Jaworowska, 2000). Studenci, którzy chcą w swojej pracy magisterskiej prowadzić badania za pomocą testów, mogą je kupić, ale pod warunkiem przedstawienia upoważnienia swojego promotora, który musi być psychologiem. Odpowiedzialne stosowanie to przede wszystkim kwestia osoby z niej korzystającej, a brak wiedzy i świadomości konsekwencji opartych na ich wynikach decyzji jest groźny. Za to, co dzieje się po nabyciu testu, odpowiadają więc psychologowie. Niepsycholog stosujący (w sposób nieuprawniony) testy nie jest do niczego zobligowany, a psycholog – jest ograniczony przez kodeks etyczno-zawodowy (PTP, 2000) i standardy (AERA, APA, NCME, 2007). Jest też w Polsce jeszcze kilka innych firm, zajmujących się sprzedażą testów lub całych procedur diagnostycznych. Wykorzystuje się także wiele testów znanych z innych krajów

– niezależnie od wersji (skrótowa czy długa) testu tego NIE MOŻNA stosować w Polsce, bo nikt w Polsce nie ma praw autorskich do niego. A [...] inni «psychologowie» go stosują”.

²³ Jak pisze internauta: „Sądy skazują mężczyzn za molestowanie dzieci na podstawie agresywnej kreski i pokolorowanej na czerwono spódniczki” (por. też Stemplewska-Żakowicz, Paluchowski, 2008, s. 51).

²⁴ Jej historia przedstawiona jest krótko na stronie <http://www.practest.com.pl/html/o-nas> (23.11.2009).

²⁵ Czemu towarzyszą (niesprawiedliwe) wypowiedzi, takie jak na jednym z forów: „[...] ustawa o zawodzie psychologa zabrania niepsychologom używania w Polsce testów psychologicznych, wyłącza konstytucyjne prawo do prowadzenia badań nad testami psychologicznymi i ogólnie zakłada, że najlepiej, jakby pozostawały tajne, to nie będzie trzeba wkładać wysiłku w ich ewaluację, a ciemnogród pozostanie ciemnogrodem ;-)”. Z drugiej strony w Internecie są i takie głosy: „[...] bez podyplomowej specjalizacji i praktyki zawodowej nie będziesz ich [testów] umiała prawidłowo zinterpretować. Jeżeli masz co do tego jakieś wątpliwości, to przeczytaj kilka podręczników neurologii i spróbuj zinterpretować parę wykresów EEG :)”.

(głównie Stanów Zjednoczonych), które nie mają opublikowanych badań ich własności psychometrycznych i które w związku z tym powinny być nazywane psychozabawą lub testem pseudopsychologicznym (choć często reklamowane i sprzedawane są pod nazwą „testy psychometryczne”)²⁶. Zdarza się też, że upowszechniane są i wykorzystywane (nawet w opiniach sądowych) testy bez odpowiedniego debitu, czyli prawa do rozpowszechniania ich na terenie Polski (a więc zgody autora czy oryginalnego wydawcy).

Aczkolwiek deklarowanym przedmiotem działalności Pracowni Testów Psychologicznych jest „szerzenie etycznych i profesjonalnych standardów diagnozy opartej na narzędziach psychometrycznych”, to jednak wydawca testów nie może być w ich sprawie bezstronnym opiniodawcą. Trudno też pogodzić postulat wolnego rynku z kontrolą nad wydawaniem testów. Ponadto inicjowanie badań naukowych nad testami nie jest zasadniczym celem firmy komercyjnej. Czym innym jest także propagowanie standardów (co często bywa deklarowanym celem korporacji zawodowych), a czym innym dostarczanie informacji (recenzji, publikacji) na temat konkretnych technik diagnostycznych. Czego więc można oczekiwać, by praktyka testowania w Polsce uległa poprawie?

Kształcenie w czasie jednolitych magisterskich studiów psychologicznych nie ma charakteru zawodowego, więc nie dostarcza kompetencji wystarczających do prawidłowego interpretowania wyników testów psychologicznych²⁷. Ukończenie psychologii (a nawet specjalności klinicznej) nie pociąga za sobą ani późniejszej pracy w zawodzie psychologa, ani – tym bardziej – bycia w przyszłości psychodiagnostą i/lub psychoterapeutą. Podobnie jak w opisywanym przypadku medycyny opartej na dowodach empirycznych tak i w psychologii (dokładniej: psychodiagnostyce) jest prawdą, że często nabywana podczas studiów wiedza szybko się dezaktualizuje, natomiast nie poddane refleksji doświadczenie zawodowe zamienia je w zmitologizowaną wiedzę (por. np. badania Kim

²⁶ W ich opisach spotkać można takie sformułowania, jak: „Wyniki nie różnicują osób badanych ze względu na płeć lub pochodzenie etniczne” (co usprawiedliwia brak polskich norm), że procedura diagnostyczna jest „[...] przekazywana do firmy w formie swoistego *know-how*, którym przedsiębiorstwo samo dysponuje po odbytych warsztatach certyfikujących. Jest to bardzo wygodna forma współpracy, ponieważ nikt nie narzuca gotowych rozwiązań i każdy użytkownik decyzje podejmuje sam” (co uzasadnia brak nadzoru i usprawiedliwia brak kompetencji osób posługujących się daną procedurą), czy że test „przekłada język psychologii [...] na język kolorów, symbolizujących określony zespół kompetencji psychologicznych” (co ma uzasadniać jego wyższość nad narzędziami „akademickimi”).

²⁷ Trzeba odróżnić kompetencje w zakresie przeprowadzania badania testowego (zwanej z angielska administracją; najczęściej dość dokładnie opisanego w instrukcji do testu dla osoby badającej) od kompetencji potrzebnych do interpretowania (rozumienia znaczeń) wyników testu w danym obszarze i wydawania na tej podstawie opinii (orzeczeń).

i Ahn: Stemplewska-Żakowicz, Paluchowski, 2008, s. 72). Być może zbieranie doświadczeń w zakresie diagnozowania powinno być podobnie nadzorowane (podlegać superwizji), jak w zakresie terapii; powinna też towarzyszyć temu analogiczna oferta studiów podyplomowych (kursów certyfikujących). Z dezaktualizacji wiedzy diagnostycznej wynika postulat konieczności ciągłego jej uzupełniania. Oczywiście źródłem takiej wiedzy mogłyby być naukowe doniesienia z badań nad istniejącymi na rynku komercyjnymi procedurami diagnostycznymi, publikowane w odpowiednich działach w istniejących czasopismach²⁸. Jednocześnie psychologowie praktycy nie otrzymują danych do podejmowania decyzji o wyborze testu, gdyż publikacji takich ewidentnie brakuje.

Zgodzić się należy, że praktycy nie są w stanie ocenić i dokonać syntezy publikowanych danych i że potrzebne są w tym względzie bazy danych na temat metod wraz z syntetycznym omówieniem wyników badań – coś na kształt publikowanych przez Instytut Burosa MMY. Wzorem opisanych wyżej rozwiązań warto rozważyć powołanie organizacji, zajmującej się publikowaniem bibliografii prac naukowych na temat testów oraz recenzji narzędzi dostępnych na rynku (komercyjnych) i gdzie potencjalni klienci mogliby poszukiwać informacji. Publikacja taka zawierać by mogła zarówno informacje formalno-psychometryczne – ogólnodostępne, jak i szczegółowe opinie psychometryczne oraz merytoryczne – dostępne tylko dla psychologów czy innych uprawnionych osób. Pozwalałaby ona na ugruntowane na wynikach badań empirycznych wybory narzędzi oraz zdynamizowałaby prace badawcze nad nimi. Jednym z celów takiej organizacji mogłoby też być propagowanie powstawania prac akademickich (w tym magisterskich) na temat testów komercyjnych. Wzorem Buros Institute of Mental Measurements (BIMM) recenzje byłyby efektem działalności *non profit* recenzentów, a zawartość „polskiego Burosa” – skutkiem współpracy wydawców testów, wyższych uczelni prowadzących studia psychologiczne oraz psychologów praktyków.

BIBLIOGRAFIA

- AERA, APA, NCME (2007). *Standardy dla testów stosowanych w psychologii i pedagogice*. Gdańsk: Gdańskie Wydawnictwo Psychologiczne.
- Brzeziński J. (2004⁵). *Metodologia badań psychologicznych*. Warszawa: Wydawnictwo Naukowe PWN.

²⁸ Pełniłyby te doniesienia funkcje krytyki naukowej i jakiejś formy społecznej kontroli, podobnie jak recenzje opublikowanych książek. Być może warto nawet powołać nowe czasopismo, np. „Diagnoza psychologiczna”.

- Claiborn, C. D. (1991). The Buros Tradition and the counseling profession. *Journal of Counseling and Development*, 69, 456-457.
- Cohen, N. (2009). A Rorschach Cheat Sheet on Wikipedia? *The New York Times* (July), 29, A1.
- Cohen, L. L., La Greca, A. M., Blount, R. L., Kazak, A. E., Holmbeck, G. N., Lemanek, K. L. (2008). Introduction to special issue: Evidence-based assessment in pediatric psychology. *Journal of Pediatric Psychology*, 33, 9, 911-915.
- Del Mar, Ch., Glasziou, P., Mayer, D., (2004). Teaching evidence based medicine. *British Medical Journal*, 329, 7473, 989-990.
- Doroszewski, J. (2005). Evidence-based medicine: New wine in old bottles? *Roczniki Akademii Medycznej w Białymstoku*, 50, 15-17.
- EFPA, EAWOP (2005). *European Test User Standards for test use in Work and Organizational settings*; <http://www.efpa.eu/download/2d30edd3542f33c91295487b64877964>] (22.12.2009).
- EFPA/Bartram, D., Lindley, P., Kennedy, N. (2008). *EFPA Review Model for the Description and Evaluation of Psychological Tests. Test Review Form and Notes for Reviewers version 3.42*; <http://www.efpa.eu/download/9044bd41c7953b956876e06c797f8c9f> (20.12.2009).
- Gajewski, P., Jaeschke, R., Brożek, J., Schünemann, H. (2007). Czy wytyczne praktyki klinicznej mogą sprowadzić na manowce? *Polskie Archiwum Medycyny Wewnętrznej*, 117, 4, 12-15.
- Gajewski, P., Jaeschke, R., Mrukowicz, J. (2003). Evidence based medicine (EBM) współczesną sztuką lekarską. Cele Polskiego Instytutu Evidence Based Medicine. *Medycyna Praktyczna*, 3, 31-34.
- Grove, W., Barden, R. C. (1999). Protecting the integrity of the legal system. The admissibility of testimony from mental health experts under Daubert/Kumho analyses. *Psychology, Public Policy, and Law*, 5, 224-242.
- Guyatt, G. (2003). Praktyczne aspekty EBM. *Medycyna Praktyczna, Ginekologia i Położnictwo*, 3, 72-73.
- Guyatt, G., Cairns, J., Churchill, D. [i in.] (1992). Evidence-based medicine. A new approach to teaching the practice of medicine. *Journal of the American Medical Association*, 268, 2420-2425.
- Holmbeck, G. N., Thill, A. W., Bachanas, P., Garber, J., Miller, K. B., Abad, M., Bruno, E. F., Carter, J. S., David-Ferdon, C., Jandasek, B., Mennuti-Washburn, J. E., O'Mahar, K., Zukerman, J. (2008). Evidence-based assessment in pediatric psychology: Measures of psychosocial adjustment and psychopathology. *Journal of Pediatric Psychology*, 33, 9, 958-980.
- Hornowska, E. (2007). *Testy psychologiczne w świetle Standardów dla procedur stosowania testów w psychologii i edukacji*. W: W. J. Paluchowski (red.), *Diagnoza psychologiczna. Proces – narzędzia – standardy* (s. 198-210). Warszawa: Wydawnictwa Profesjonalne i Akademickie.
- Hunsley, J. (2007). Training psychologists for evidence-based practice. *Canadian Psychology/Psychologie Canadienne*, 48, 1, 32-42.
- Hunsley, J., Mash, E. J. (2005). Introduction to the special section on developing guidelines for the Evidence-Based Assessment (EBA) of adult disorders. *Psychological Assessment*, 17, 3, 251-255.
- ITC (2000). *International Guidelines for Test Use Version 2000*; <http://www.intestcom.org/Downloads/ITC%20Guidelines%20Download%20Version%204.doc> (23.12.2009).
- Jaeschke, R., Cook, D. J., Guyatt, G. H. (1998a). EBM – Wprowadzenie. *Medycyna Praktyczna*, 1, 143-150.
- Jaeschke, R., Cook, D. J., Guyatt, G. H. (1998b). Ocena artykułów na temat testów diagnostycznych. Cz. I: Ocena wiarygodności i uwagi ogólne. *Medycyna Praktyczna*, 10, 156-159.
- Jaeschke, R., Cook, D. J., Guyatt, G. H. (1998c). Ocena artykułów na temat testów diagnostycznych. Cz. II: Metody określania przydatności testu. *Medycyna Praktyczna*, 11, 184-191.

- Jaeschke, R., Cook, D. J., Guyatt, G. H. (1999a). Evidence based medicine (EBM), czyli praktyka medyczna oparta na wiarygodnych i aktualnych publikacjach (POWAP). *Medycyna Praktyczna* (wydanie specjalne), 1.
- Jaeschke, R., Cook, D. J., Guyatt, G. H. (1999b). Analiza ekonomiczna. Cz. II: Ocena i zastosowanie wyników. *Medycyna Praktyczna*, 9, 169-174.
- Jaeschke, R., Cook, D. J., Guyatt, G. H. (1999c). Analiza decyzyjna. Cz. I: Ocena wiarygodności. *Medycyna Praktyczna*, 5, 176-181.
- Jaeschke, R., Cook, D. J., Guyatt, G. H. (1999d). Analiza decyzyjna. Cz. II: Ocena wyników i ich zastosowanie. *Medycyna Praktyczna*, 6, 195-203.
- Jaeschke, R., Cook, D. J., Guyatt, G. H., Bochenek, T. (1998). Ocena artykułów przeglądowych i metaanaliz. *Medycyna Praktyczna*, 9, 177-184.
- Jaeschke, R., Cook, D. J., Guyatt, G. H., Gajewski, P. (2000). EBM po Y2K. *Medycyna Praktyczna*, 1, 225-229.
- Jaeschke, R., Cook, D. J., Guyatt, G. H., Gajewski, P., Łanda, K., Wcisło, J. (1999a). Wytyczne postępowania. Cz. I: Ocena wiarygodności, czyli którym wytycznym można wierzyć. *Medycyna Praktyczna*, 10, 161-166.
- Jaeschke, R., Cook, D. J., Guyatt, G. H., Gajewski, P., Łanda, K., Wcisło, J. (1999b). Wytyczne postępowania. Cz. II: Sposób przedstawiania wytycznych, czyli siła zaleceń. *Medycyna Praktyczna*, 11, 153-157.
- Jaeschke, R., Guyatt, G. H., Cook, D., Gajewski, P., Łanda, K., Wcisło, J. (1999). Evidence based medicine (EBM), czyli praktyka medyczna oparta na wiarygodnych i aktualnych publikacjach (POWAP). Wytyczne postępowania – ocena wiarygodności, czyli którym wytycznym można wierzyć. *Medycyna Praktyczna*, 10, 110-114.
- Jaworowska A. (2000). Dostęp do testów psychologicznych i kompetencje zawodowe ich użytkowników. W: J. Brzeziński, M. Toeplitz-Winiewska (red.), *Etyczne dylematy psychologii* (s. 269-278). Poznań: Wydawnictwo Fundacji Humaniora.
- Kazdin, A. E. (2005). Evidence-based assessment for children and adolescents: Issues in measurement development and clinical application. *Journal of Clinical Child and Adolescent Psychology*, 34, 3, 548-558.
- Lasoń, W., Walecki, P., Sarapata K., Roterman-Konieczna, I. (2005). *Nauczanie medycyny z wykorzystaniem paradygmatu Medycyny Opartej na Dowodach Naukowych (Evidence Based Medicine)*; <http://www.ap.krakow.pl/ptn/ref2005/lason.pdf> (05.01.2009).
- Levant R. F. [i in.] (2006). Evidence-based practice in psychology. APA presidential task force on evidence-based practice. *American Psychologist*, 61, 4, 271-285.
- Łanda, K., Walczak, K. (2005). *EBM – medycyna oparta na dowodach naukowych*; <http://www.hta.pl/dok/EBM.pdf> (05.01.2009).
- Łuszczczyńska, A., Doliński, D. (2008). Jak przeprowadzać i opisywać eksperymentalne badania psychologiczne: wytyczne CONSORT. Tekst instruktażowy. *Studia Psychologiczne*, 46, 1, 57-64.
- Mash, E. J., Hunsley, J. (2005). Evidence-based assessment of child and adolescent disorders: Issues and challenges. *Journal of Clinical Child & Adolescent Psychology*, 34, 3, 362-379.
- McFall, R. M. (2005). Theory and utility-key themes in evidence-based assessment: Comment on the special section. *Psychological Assessment*, 17, 3, 312-323.
- Mislevy, R. J., Almond, R. G., Lukas, J. F. (2003). *A brief introduction to evidence-centered design (Research Report RR-03-16)*. Princeton, NJ: Educational Testing Service; <http://www.ets.org/Media/Research/pdf/RR-03-16.pdf> (6.06.2005).
- Mrukowicz, J. (2005). *Podstawy „evidence based medicine” (EBM), czyli o sztuce podejmowania trafnych decyzji w opiece nad pacjentami*. „*Medycyna Praktyczna On-line*”; <http://www.mp.pl/artykuly/?aid=25574> (23.11.2009).

- Paluchowski, W. J. (2007). *Diagnoza psychologiczna. Proces – narzędzia – standardy*. Warszawa: Wydawnictwa Profesjonalne i Akademickie.
- Plake, B. S., Conoley, J. C., Kramer, J. J., Murphy, L. U. (1991). The Buros Institute of mental measurements: Commitment to the tradition of excellence. *Journal of Counseling and Development*, 69, 449-455.
- Polskie Towarzystwo Psychologiczne (2000). *Kodeks etyczno-zawodowy psychologa*. W: J. Strelau (red.), *Psychologia. Podręcznik akademicki* (t. 3, s. 836-844). Gdańsk: Gdańskie Wydawnictwo Psychologiczne.
- PTP (2000). *Wytyczne Międzynarodowej Komisji ds. Testów (International Test Commission – ITC) z 2000 roku dotyczące stosowania testów: wersja polska*; http://www.practest.com.pl/files/ITC-Stosowanie_testow.pdf (23.12.2009).
- Sackett, D. L., Haynes, R. B., Guyatt, G. H., Tugwell, P. (1991). *Clinical epidemiology-a basic science for clinical medicine*. London: Little Brown.
- Sackett, D. L., Richardson, W. S., Rosenberg, W. M. C., Haynes, R. B. (1997). *Evidence-based medicine: How to practice and teach EBM*. New York: Churchill Livingstone.
- Sackett, D. L., Rosenberg, W. M. C., Gray, J. A. M., Haynes, R. B., Richardson, W. S. (1996). Evidence-based medicine: What it is and what it isn't? *British Medical Journal*, 312, 71-72.
- Słysz, A. (2008). *Typy diagnostów. Preferencje poznawcze psychologów a strategie diagnozowania*. Poznań: Wydawnictwo UAM.
- Stanik, J. M., Majchrzyk, Z. (red.) (1995). *Etyczno-zawodowe problemy biegłego sądowego psychologa i psychiatry w praktyce sądowej*. Katowice: Katedra Psychologii Klinicznej Uniwersytetu Śląskiego i Sekcja Psychiatrii Sądowej Polskiego Towarzystwa Psychiatrycznego.
- Stemplewska-Żakowicz, K., Paluchowski, W. J. (2008). Podstawy diagnozy psychologicznej. W: J. Strelau, D. Doliński (red.), *Psychologia. Podręcznik akademicki* (t. 2, s. 25-94). Gdańsk: Gdańskie Wydawnictwo Psychologiczne.
- Sułkowski, H. (2007). *Co lekarz praktyk o medycynie opartej na dowodach wiedzieć powinien* <http://www.umed.lodz.pl/kmz/doc/ebm.pdf> (17.08.2008).
- Szajewska, H. (2004). Świadome korzystanie z literatury medycznej, czyli o jednym z elementów evidence based medicine (EBM). *Magazyn Otolaryngologiczny*, 3, 11, 71-78.
- Szajewska, H. (2008). Praktykowanie pediatrii w duchu EBM. *Medycyna Praktyczna. Pediaatria*, 1, 15-20.
- Szustrowa, T. (2003). *Raport o stanie orzecznictwa psychologicznego w Polsce*. Warszawa: PTP; http://www.ptp.org.pl/teksty/raport_orzecznictwo.doc (11.03.2005).
- Westen, D., Novotny, C. M., Thompson-Brenner, H. (2004). The empirical status of empirically supported psychotherapies: Assumptions, findings, and reporting in controlled clinical trials. *Psychological Bulletin*, 4, 631-663.
- Westen, D., Novotny, C. M., Thompson-Brenner, H. (2005). EBP ≠ EST: Reply to Crits-Christoph [et al.] (2005) and Weisz [et al.] (2005). *Psychological Bulletin*, 3, 427-433.
- Widacki, J. (2007) (red.) *Kryminalistyka*. Warszawa: Wydawnictwo C. H. Beck.

EVIDENCE-BASED ASSESSMENT
– DO WE NEED “POLISH BUROS”?

S u m m a r y

The aim of this paper is to describe the practice in psychological assessment and diagnosis in Poland. This article focuses on two key themes in the on evidence-based assessment. The first one is application of EBM in psychology. The second is activity and achievement of the Buros Institute of Mental Measurements. The implications of these themes for the future of psychological assessment practice are discussed.

Key words: evidence-based assessment – do we need “polish Buros”?, assessment, EBA, EBM.