

Krystyna SOCHACKA

Państwowa Wyższa Szkoła Zawodowa w Nowym Sączu

PROBLEMY Z POMIAREM INTELIGENCJI SKALAMI WECHSLERA – WISC-R I WAIS-R (PL)

Streszczenie

Artykuł odnosi się do polskich adaptacji skal inteligencji Davida Wechslera dla dzieci i dorosłych. Praktyka posługiwania się adaptowanymi i starzejącymi się testami nie jest wolna od problemów. Dwa z nich zostaną omówione. Jednym z nich jest zjawisko przyrastania wyników w testach inteligencji wraz z upływem lat, nazywane *efektem Flynna*. Drugi dotyczy porównywalności wyników testów dla dzieci i dorosłych. Szczególną uwagę poświęcono wynikającym z tego porównania i konsekwencjom dla praktyki psychologicznej.

Słowa kluczowe: testy inteligencji, WISC-R i WAIS-R (PL), efekt Flynna.

PROBLEMS WITH MEASUREMENT OF INTELLIGENCE USING WECHSLER SCALES - WISC-R AND WAIS-R (PL)

Summary

The article refers to the Polish adaptation of David Wechsler intelligence scales for children and adults. The practice of using adapted and aging tests is not free from problems. Two of them will be discussed. One of them is the phenomenon of increasing results in intelligence tests over the years, called the *Flynn effect*. The second concerns the comparability of test results for children and adults. Implications for clinical practice are discussed.

Key words: intelligence tests; WISC-R, WAIS-R (PL), Flynn effect.

1. Skale Davida Wechslera w Polsce

Testami powszechnie stosowanymi na całym świecie do pomiaru funkcjonowania poznawczego są skale Davida Wechslera. Zaliczane są do najbardziej znaczących spośród wielu dostępnych testów inteligencji. Mierzą różne zdolności poznawcze, dlatego mogą być użyteczne przy ocenie szkolnych i zawodowych kompetencji badanych. Służą jednak przede wszystkim do pomiaru „ogólnej zdolności do inteligentnego zachowania” (Brzeziński i in., 2007, s. 8). Zdaniem innych autorów (por. Anastasi, Urbina, 1999), skale te służą raczej do pomiaru tylko fragmentu inteligencji, który diagnozowany jest za pomocą tradycyjnych testów i symbolizowany skrótem IQ (*Intelligence Quotient* – Iloraz Inteligencji). Ponieważ inteligencja jest funkcją osobowości, jako całości, testy te mogą być także pomocne przy ocenie zdolności pozaintelektualnych, np. świadomości celu, wytrwałości czy poziomu lęku (Brzeziński i in., 1996; Brzeziński i in., 2007).

Pierwsze wersje tych amerykańskich testów *Wechsler Bellevue Intelligence Scale* (W-B I) z 1939 roku i *Wechsler Bellevue Intelligence Scale-II* (W-B II) z roku 1944/1945 przeznaczone były do badania młodzieży i osób dorosłych (15-60). Stanowiły pierwowzór zarówno skali inteligencji dla dzieci (5-15), określanej skrótem WISC – *Wechsler Intelligence Scale for Children*, powstałej w roku 1949, jak i skali dla dorosłych (15-65) WAIS – *Wechsler Adult Intelligence Scale* z roku 1955. W późniejszych latach powstawały kolejne wersje.

W roku 1974 ukazała się zmodyfikowana wersja dla dzieci (6-16) **WISC-R** – *Wechsler Intelligence Scale for Children – Revised*, a w 1981 roku zrewidowana skala dla dorosłych (15-65) **WAIS-R** – *Wechsler Adult Intelligence Scale – Revised*. Po śmierci D. Wechslera w 1981 roku kontynuatorzy nadal tworzą testy wzorowane na skalach klasycznych i publikują je pod nazwiskiem D. Wechslera, np. WISC-III w 1991, WAIS-III w 1997, WISC-IV w 2003 czy WAIS-IV w 2008 roku. W skład ogólnej puli testów, poza skalami dla dorosłych i dla dzieci w wieku szkolnym, wchodzi także skala do badania dzieci przedszkolnych *Wechsler Preschool and Primary Scale of Intelligence* (WPPSI), których kolejne wersje powstawały w latach 1967, 1989 i 2001¹.

W Polsce pierwsza wersja skali dla dorosłych (WB I) była obecna w diagnostyce psychologicznej od lat 50. XX wieku (Chojnowski, 1951; Grzywak-Kaczyńska, 1949; Polskie Towarzystwo Higieny Psychiczej, 1972; Brzeziński, Hornowska, 1993; Hornowska, 2004). Niestety, bazowała ona na amerykańskich charakterystykach psychometrycznych i normach z 1939 roku, więc pomimo akceptacji skali przez psychologów praktyków, podkreślano jej małą użyteczność. Skalę D. Wechslera dla dzieci z 1949 roku w opracowaniu J. Kostrzewskiego (1970) stosowano w Polsce od lat 70. XX wieku. Skala ta również nie miała polskich norm ani charakterystyki parametrów psychometrycznych i tylko fragmentaryczne informacje na temat jej użyteczności w naszych warunkach kulturowych (Matczak, Piotrowska, Ciarkowska, 1997), w związku z tym kierowano pod jej adresem podobne zastrzeżenia.

W 2 poł. lat 80. XX wieku w ramach Pracowni Testów Psychologicznych Polskiego Towarzystwa Psychologicznego podjęto prace nad pierwszą, pełną polską adaptacją i normalizacją skal D. Wechslera do badania zarówno dzieci, jak też dorosłych: WISC-R z 1974 roku (Matczak, Piotrowska, Ciarkowska, 1991; 1997) oraz WAIS-R z 1981 roku (Brzeziński i in., 1996, 2007). Nie są to wersje nowe, w większości krajów już nie są stosowane, wśród zalet podkreśla się jednak, że spełniają warunki rzetelności, trafności i normalizacji (Krasowicz-Kupis, Wiejak, 2006). Polska wersja *Skali Inteligencji D. Wechslera dla Dzieci – wersja zmodyfikowana* (WISC-R) została opublikowana w 1991 roku, a w latach 1997 i 2008 ulazały się kolejne, poprawione, wydania. Służy do badania dzieci w wieku od 6 do 16 lat. Polska adaptacja *Skali Inteligencji D. Wechslera dla Dorosłych Wersja Zrewidowana* WAIS-R (PL) opublikowana w 1996 roku zawierała normy dla osób od 16. do 54. roku życia. W roku 2004 ukazała się renormalizacja skali WAIS-R (PL), w efekcie zaktualizowano i rozszerzono normy dla osób w wieku od 16 do 79 lat. Wśród powodów renormalizacji autorzy (Brzeziński i in., 1996; Brzeziński i in., 2007) wymieniali: starzenie się norm, brak norm dla starszych grup wiekowych, dezaktualizację treści niektórych zadań oraz potrzebę uszczegółowienia klucza odpowiedzi w testach werbalnych. W wersji renormalizacyjnej uległa zmianie procedura obliczania wyników przeliczonych. Tabele zamiany wyników surowych na wyniki przeliczone opracowano niezależnie dla każdej grupy wiekowej, do jakiej należy osoba badana, a nie – jak poprzednio – w relacji do wyników otrzymanych przez grupę odniesienia, którą stanowiły osoby o najwyższym poziomie intelektualnym w wieku 20-34 lata (Zakrzewska, 2001).

Praktyka posługiwania się adaptowanymi i starzejącymi się testami nie jest wolna od problemów. Dwa z nich zostaną poniżej omówione. Jednym jest zjawisko przyrastania wyników w testach inteligencji wraz z upływem lat, nazywane *efektem Flynna*, zaś drugi dotyczy porównywalności wyników testów dla dzieci i dorosłych. Problemy te wiążą się z pojęciem ciągłości zdolności poznawczych.

¹ Szerzej na temat poszczególnych wersji skal D. Wechslera oraz ich chronologii piszą A. Anastasi i S. Urbina (1999), J. Brzeziński i E. Hornowska (1993), E. Hornowska (2004) oraz G. Krasowicz-Kupis i K. Wiejak (2006, 2011).

2. Ciągłość zdolności poznawczych

Zachowanie cech indywidualnych wraz z wiekiem nazywane jest ciągłością rozwoju (Schaffer, 2006). Uznaje się, że poziom kompetencji intelektualnych stanowi względnie trwałą cechę człowieka, która odróżnia go od innych ludzi (Nęcka, 2003; Strelau, 2001), stąd też wyniki testów badających inteligencję w różnych okresach życia powinny być podobne. Inteligencja jest to cecha indywidualna względnie trwała, gdyż podlega zmianom rozwojowym, degeneracyjnym i cywilizacyjnym, ale w kolejnych pomiarach zachowuje stałą pozycję na tle swojej grupy wiekowej. W testach D. Wechslera iloraz inteligencji 100 jest zawsze równy średniej w danej populacji i w tym sensie jest wartością stałą.

Ciągłość zdolności poznawczych dzieci, młodzieży i osób dorosłych była przedmiotem wielu badań na świecie. W badaniach podłużnych, polegających na badaniu tej samej grupy osób w określonych odstępach czasowych, wykazywano (Anastasi, Urbina, 1999), że w okresie szkoły podstawowej, średniej i wyższej wyniki testowe są dość stabilne. Korelacje pomiędzy wynikami były tym wyższe, im krótszy był odstęp między badaniami oraz im starsze osoby były badane. W sensie statystycznym wyniki te mają znaczenie w przewidywaniach dotyczących grup, natomiast poszczególne osoby mogą wykazywać wzrost lub spadek wyników, co zwykle wiąże się ze zmianami środowiskowymi lub właściwościami emocjonalno-motywacyjnymi badanych. Analizy wyników badań grup populacyjnych (kohort), polegających na badaniu osób urodzonych w różnym czasie ale badanych w tym samym wieku, pokazały z kolei, że w różnych krajach świata wskaźniki liczbowe uzyskiwane w testach inteligencji z biegiem lat systematycznie rosły. Zjawisko przyrastania wyników w testach inteligencji wraz z upływem lat po raz pierwszy opisał J. Flynn w latach 80. XX wieku, dlatego też zostało ono nazywane *efektem Flynna*.

3. Efekt Flynna

J. Flynn (2006) wykazał, że wyniki w testach inteligencji rosną z pokolenia na pokolenie o 5 do 25 punktów, w ciągu dekady zwiększając się przeciętnie o 3 punkty. Dla skal wechslerowskich przyrost ten jest na stałym poziomie – ok. 0,3 punktu rocznie. W konsekwencji, im więcej czasu upływa od momentu normalizacji testu do momentu badania, tym mniej trafnie wynik odzwierciedla pozycję jednostki na tle grupy wiekowej, wskazując zawyżone wartości IQ. Starzenie się norm powoduje wzrost wyników testowych. Im starsze są normy, tym bardziej zawyżone są ilorazy inteligencji, co powoduje, że mniej badanych uzyskuje wynik wskazujący na niepełnosprawność intelektualną oraz więcej osób wykazuje wybitne zdolności. Najbardziej zawyżone wyniki obserwuje się u osób o niskim poziomie inteligencji. Jak temu zaradzić? Zdaniem J. Flynna, niezbędne jest uwzględnianie obserwowanej prawidłowości w ogólnym wyniku testowym. Zaleca on postępowanie według następującej procedury: należy policzyć lata od czasu normalizacji testu do czasu badania (L), pomnożyć je przez 0,3, a następnie uzyskany iloczyn odjąć od otrzymanego w badaniu wyniku ogólnego. Wyraża to następujący wzór: $IQ = IQ \text{ otrzymany} - (L \times 0,3)$.

Na temat efektu Flynna na przykładzie wyników WISC-R w polskiej literaturze pisze A. Jaworowska (2011), przytaczając dodatkowo wyniki polskich badań (Matczak, Piotrowska, Ciarkowska, 1991, 1997, 2008; Bogdanowicz i in., 2008; Jaworowska, Matczak, Stańczak, 2010), z których wynika, że w ciągu ostatnich 20 lat średnie ilorazy inteligencji w *Skali Pełnej* WISC-R wzrosły o 8 punktów u dzieci w wieku 9-10 lat i prawie o 10 punktów u dzieci w wieku 11-12 lat. Średnio wzrastały o 0,4-0,5 punktu rocznie, okazały się więc większe niż w innych krajach. W *Skali Słownej* i *Bezsłownej* przyrosty wyników były podobne.

Wzrost poziomu wyników uzyskiwanych w testach inteligencji nie przekłada się na pozatestowe osiągnięcia badanych, np. nie uwidacznia się w osiągnięciach szkolnych, które raczej ulegają obniżeniu. Z tego względu przyrosty wyników testowych często nie są identyfikowane z przyrostami inteligencji (Matczak, 1994; Anastasi, Urbina, 1999; Jaworowska, 2011). Uznawane są natomiast za artefakt, który może być skutkiem coraz sprawniejszego radzenia sobie z rozwiązywaniem zadań testowych, ponieważ uczniowie ćwiczą te umiejętności w szkole. Badacze z jednej strony próbują wykazać, że efekt Flynna jest artefaktem, a z drugiej poszukują uzasadnienia zwiększających się z czasem możliwości intelektualnych (Neiser, 1997).

Żadna z rozważanych dotychczas hipotez nie wyjaśnia w pełni obserwowanego zjawiska. Przyrost wyników testowych jest zbyt szybki, by mógł być efektem zmian genetycznych. Uznanie go za efekt wpływów środowiskowych też budzi wątpliwości, gdyż w badaniach monozygotycznych bliźniąt okazują się one mało znaczące. Hipotezę uwzględniającą poprawę standardów żywienia trudno zweryfikować ze względu na to, że nie udaje się wykazać bezpośredniego związku pomiędzy dietą a inteligencją (Flynn, 2006). Niezależnie od tego jej zwolennicy twierdzą, że wraz z podwyższaniem się wzrostu o 1 cm w ciągu dekady, powiększa się rozmiar głowy i prawdopodobnie też średnia wielkość mózgu, który w związku z tym ma większe możliwości. Wydłużenie czasu i rozpowszechnienie edukacji nie jest satysfakcjonującym wyjaśnieniem ze względu na to, że zgodnie z powszechnym oczekiwaniem, edukacja bardziej wpływa na zakres wiedzy i tzw. umiejętności szkolne, a nie na rozumowanie, czy sposób rozwiązywania nowych problemów, w czym właśnie obserwowany jest wzrost wyników testowych. Ponadto wyniki w testach rosną nawet w okresie, gdy czas nauki pozostaje taki sam. Podobnie rosły także jeszcze przed erą telewizji, gier video czy komputerów.

Wielu autorów spekuluje na temat przyczyn efektu Flynna. Na ogół zakładają, że jest to efekt wpływów środowiskowych, nawet jeśli nieznana jest jeszcze istota zjawiska. Znaczącą rolę przypisują wzrastającej złożoności życia i stymulującemu środowisku. Podkreślają jednak, że próżno szukać prostych liniowych zależności przyczynowo-skutkowych, znajdują natomiast znaczące interakcje pomiędzy różnymi czynnikami wpływu społecznego. Najczęściej biorą pod uwagę urbanizację, mniej liczne rodziny, lepiej wykształconych i bardziej zamożnych rodziców, zmieniający się stosunek do dzieci, bardziej liberalny styl wychowania, lepszy wgląd w potrzeby dzieci oraz, w większym stopniu niż kiedyś, zwracanie uwagi na ich rozwój poznawczy oraz nacisk na myślenie lateralne w szkołach – niekonwencjonalny sposób korzystania z inteligencji i wiedzy – produktywny i twórczy (Fudala, 2013; de Bono, 2008). Ze względu na to, że jest to proces społeczny, uwzględniane są także takie trendy, jak skłonność do mobilności społecznej i kojarzenie par z szerszej gamy partnerów.

Zdaniem J. Flynna (2006), przyrost wyników testowych nie wskazuje na wzrost inteligencji w ogóle, a świadczy jedynie o zwiększających się możliwościach w zakresie myślenia abstrakcyjnego, które w znacznym stopniu badane są przez testy. Najwyraźniej uwidacznia się w *Podobieństwach* (podtest skali werbalnej testów Wechslera), bazujących na klasyfikacji przy użyciu kategorii abstrakcyjnych, w niewerbalnych procesach poznawczych i przetwarzaniu przestrzennym, a więc w zadaniach, dla których ważna jest umiejętność rozumowania i rozwiązywania nowych problemów niezależnie czy mają one charakter werbalny, czy wykonawczy. Przez długi czas uznawano, że są to umiejętności odporne na trening. Wyjaśnieniem zmian w tym zakresie może być coraz bardziej efektywne wykorzystywanie systemu pamięci operacyjnej (Resing, Tunteler, 2007), co wynika z konieczności przyswajania coraz większej liczby informacji w coraz młodszych wiekach i radzenia sobie z coraz bardziej złożonym środowiskiem.

Mimo wielu hipotez, jak dotąd, nie ma jednoznacznego wyjaśnienia źródeł efektu Flynna.

4. Porównywalność wyników w testach WISC-R i WAIS-R (PL)

Drugi z rozważanych problemów dotyczy porównywalności wyników uzyskiwanych w testach dla dzieci i dorosłych. Wbrew oczekiwaniom, amerykańscy badacze wielokrotnie wykazywali, że osiągnięte w WISC-R i WAIS-R wyniki nie są równoważne (Spitz, 1989).

Odwołania dotyczące tego problemu odnoszą się do literatury z lat 80. XX wieku, ponieważ wtedy skale te były tam jeszcze stosowane, zanim w latach 90. zastąpiły je nowsze wersje (WISC III i WAIS III). W wielu badaniach (Grace, Sweeney, 1986; Vance i in., 1987; Wilcoxon, 1982) u tych samych osób stwierdzano wyższe wyniki przy pomiarze skalą dla dorosłych (WAIS-R) niż skalą dla dzieci (WISC-R), a rozbieżność ta zwiększała się wraz z niższymi ilorazami inteligencji. Niektórzy badacze potwierdzali równoważność obu testów. Przykładowo, J.M. Sattler, J.C. Polifka, S. Polifka i D.E. Hilsen (1984) wykazali brak istotnych różnic pomiędzy wynikami adolescentów badanych 2-krotnie w odstępie 4 lat. Uzyskiwano też wyniki niższe w WAIS-R niż WISC-R (Slate, Frost, Cross, 1990), co uzasadniono zdecydowanie starszymi normami WISC-R w porównaniu z bardziej aktualnymi WAIS-R.

Najczęściej badano 16-latków, ponieważ obie skale zawierają normy dla tego przedziału wiekowego. Oczekiwano, że zadania z testu dla dzieci będą dla nich łatwiejsze, jednak okazywało się często przeciwnie, wyższe wyniki 16-latkowie uzyskiwali w teście dla dorosłych. Różnice te zależne były od poziomu inteligencji, zaś zwiększały się przy niższych wartościach IQ. D. Wechsler (1981) dowodził, że w szerokim zakresie wyników przeciętnych WISC-R i WAIS-R dla 16-latków pozostają porównywalne. Rozbieżności natomiast dotyczą badanych o niskim poziomie umysłowym, u których obserwuje się wyższe wyniki w WAIS-R oraz badanych o najwyższej inteligencji, którzy z kolei wyższe wyniki uzyskują w WISC-R. W swoich opiniach brał pod uwagę badania porównujące także wcześniejsze wersje WISC i WAIS oraz WAIS i WISC-R, pomiędzy którymi obserwowano podobne rozbieżności.

W polskiej literaturze przedmiotu G. Krasowicz-Kupis i K. Wiejak, pisząc o czynnikach, które należy wziąć pod uwagę, podejmując decyzję, którą z polskich wersji skali zastosować w badaniu dzieci z pogranicza grup wiekowych, zaznaczyły: „Należy pamiętać o tym, że dla 16-latków WISC-R jest testem łatwiejszym” (2006, s. 95). Doświadczenia polskich psychologów praktyków i cytowane wyżej badania nie potwierdzają takiego przekonania. Wśród możliwych powodów braku zgodności w wynikach obu skal brano pod uwagę m.in. różne przedziały wieku w grupach normalizacyjnych. Wskazywano na zbyt małą liczbę osób z niskim poziomem funkcjonowania poznawczego w próbach normalizacyjnych. Osób z niepełnosprawnością intelektualną w ogóle nie włączono (Brzeziński i in., 1996; Matczak, Piotrowska, Ciarkowska, 1997). Zdaniem A. Jaworowskiej (2009), powodem są odrębne konwencje obliczania wyników. Badany w wieku 16,5 roku, który rozwiąże po jednym zadaniu z każdego podtestu w WISC-R, otrzyma 10 punktów przeliczonych i IQ poniżej 40, w WAIS-R (PL) uzyska 17 punktów i IQ = 46.

Przeprowadzone w Polsce badania podłużne (Sochacka, 2016) umożliwiły porównanie wyników uzyskanych przez osoby w normie intelektualnej badane w dzieciństwie skalą WISC-R oraz w adolescencji skalą WAIS-R (PL) zarówno w aktualnie stosowanej wersji zrenormalizowanej (2004), jak i w wersji sprzed renormalizacji (1996). W efekcie znaleziono istotne różnice pomiędzy wynikami testów WISC-R i WAIS-R (PL) z 2004 roku na korzyść testu dla dorosłych. Brak jest natomiast takiej różnicy pomiędzy WISC-R i wcześniejszą wersją WAIS-R (PL) z 1996 roku, co wskazuje na zgodność wyników pomiędzy skalą dla dzieci i dorosłych w wersji sprzed renormalizacji. Otrzymany rezultat porównania WISC-R i WAIS-R(PL) w wersji zrenormalizowanej jest zgodny z wieloma podobnymi doniesieniami amerykańskich badaczy. Przeczy jednocześnie domniemaniu równoważności obu skal w zakresie wyników

przeciętnych, które mogło powstać na podstawie amerykańskich analiz D. Wechslera (1981), wykazujących rozbieżności między tymi skalami tylko w przedziałach wyników niskich i wysokich, co mogło tłumaczyć doniesienia praktyków dotyczące problemów w diagnozowaniu osób niepełnosprawnych intelektualnie.

W tym kontekście otwartą kwestią pozostaje pytanie, dlaczego wersja WAIS-R (PL) sprzed renormalizacji nie wykazuje takiego przyrostu możliwości intelektualnych badanych? Różnic pomiędzy wynikami testowymi dzieci i młodzieży, w zależności od tego, czy wyniki opracowane były zgodnie z wersją WAIS-R (PL) sprzed czy po renormalizacji, nie tłumaczy efekt Flynna. Zgodnie z efektem Flynna, starsze normy zawyżają wyniki. W przytoczonych badaniach okazało się odwrotnie, wyniki badanych porównywane ze starszymi normami pierwszej polskiej wersji testu WAIS-R(PL) są znacząco niższe od wyników porównanych do nowszych norm drugiej, zrenormalizowanej wersji testu.

Zawyżone wyniki w skali WAIS-R (PL) u osób w normie intelektualnej, w porównaniu ze skalą stosowaną w dzieciństwie, nie powodują problemów tego rodzaju, jak sygnalizowane wcześniej u osób niepełnosprawnych intelektualnie. W interpretacji łatwiej przyjąć, że zmiany są naturalną konsekwencją rozwoju danej osoby. Testy psychologiczne służą m.in. do rozstrzygania kwestii dotyczących zmian rozwojowych dokonujących się w biegu życia jednostki. Biorąc pod uwagę zrenormalizowaną wersję WAIS-R (PL), pojawia się pytanie, czy przyrost wyników wykazywany przez tę skalę należy traktować, jako efekt zmian w poziomie funkcjonowania poznawczego pomiędzy dzieciństwem i adolescencją? Uzasadnieniem dla odpowiedzi twierdzącej może być z jednej strony dynamika rozwoju w okresie objętym badaniem, a z drugiej ściśle powiązanie wyników testowych ze sprawnościami kształconymi w szkołach. Zwykle badania wykazują wzrastanie wyników u osób, które dłużej kontynuują edukację (Anastasi, Urbina, 1999). W podręczniku do testu (Brzeziński i in., 2007) czytamy, że zmienną, która najsilniej różnicuje wyniki w WAIS-R (PL), jest poziom wykształcenia.

Zdecydowanie trudniej wyjaśnić przyrost wyników testowych w badaniu testem WAIS-R(PL) u osób niepełnosprawnych intelektualnie, z czym borykają się psychologowie praktycy. Niepełnosprawność intelektualna zwykle ma stabilny charakter i nie ulega zasadniczym zmianom. Pojęcie niezmienności IQ odnosi się do populacji, a nie do jednostek, i nie można go odnosić do konkretnej osoby.

5. Konsekwencje dla praktyki

Niezgodne wyniki obu skal stanowią poważny problem dla psychologów praktyków (por. Sochacka, 2016, 2018). Szczególnie wtedy, gdy uczniowie przez wiele szkolnych lat diagnozowani skalą WISC-R, np. jako niepełnosprawni intelektualnie w stopniu umiarkowanym, badani ponownie po ukończeniu szkoły czy uzyskaniu pełnoletności skalą WAIS-R (PL), uzyskują wyniki wskazujące na niepełnosprawność intelektualną lekkiego stopnia, pomimo iż poziom ich funkcjonowania szkolnego ani społecznego nie uległ zmianie. Wyniki testów są uwzględniane przy podejmowaniu decyzji dotyczących osób badanych, dlatego też taka psychometryczna diagnoza może mieć poważne konsekwencje administracyjne, np. może odebrać przysługujące prawo do renty socjalnej. Warto podkreślić, że w polskich poradniach problem ten pojawił się wraz z wersją zrenormalizowaną WAIS-R (PL) z 2004 roku. Praktyka pokazuje jednak, że w kłopotliwej sytuacji dużej rozpiętości wyników pomiędzy WISC-R a WAIS-R (PL) wystarczy wrócić do wersji WAIS-R (PL) z 1996 roku, tej sprzed renormalizacji, by uzyskać porównywalny do mierzonego skalą WISC-R poziom wydolności poznawczej danej osoby.

Nasuwa się jednak pytanie, czy praktyka posługiwania się starszą wersją WAIS-R (PL), z powodu zbieżności jej wyników z WISC-R, może być uzasadniona? Zdaniem A. Anastasi i S. Urbiny (1999), dopóki narzędzie służy do formułowania hipotez, jego stosowanie może być uzasadnione. W diagnozie indywidualnej psycholog dysponuje różnymi danymi, np. z obserwacji, wywiadu, historii przypadku, które łączy z wynikami testowymi w celu uzyskania całościowego obrazu badanej osoby, co chroni go przed nadmiernym generalizowaniem wyników testowych i częściowo wyjaśnia trwanie przy testach niezależnie od właściwości ich parametrów psychometrycznych. Aktywny kontakt w czasie przeprowadzania testu dostarcza doświadczonemu diagnoście znacznie więcej informacji o mocnych i słabych stronach badanego niż pojedynczy wskaźnik, jakim jest IQ.

Testy D. Wechslera od początku swojego istnienia wykorzystywane były do diagnozowania niepełnosprawności intelektualnej, a podstawą diagnozy był najczęściej iloraz inteligencji w *Skali Pełnej*. Opieranie się w diagnostyce psychologicznej tylko na kryterium psychometrycznym, czyli na wartościach ilorazów inteligencji, może prowadzić do uproszczeń, a nawet błędnych konkluzji, na co zwraca uwagę wielu autorów (Brzeziński, Hornowska, 1993; Hornowska, 2004; Krasowicz-Kupis, Wiejak, 2006; Matczak, 1994), którzy przyznają, że testy są pomocne w postawieniu diagnozy, szczególnie na etapie wstępnym, ale rozstrzygające powinno być pełne badanie psychologiczne i kliniczne. Według S. Messicka (2005) odpowiedzialność za trafne wykorzystanie testu spoczywa na osobie, która go interpretuje, a taka interpretacja pociąga za sobą odpowiedzialność za konsekwencje (por. Standardy dla testów psychologicznych i pedagogicznych, 2007). Podobnie A. Anastasi i S. Urbina (1999) uważają, że decydowanie tylko na podstawie testów jest ich nadużyciem. Testy stanowią tylko jedno ze źródeł danych, nie są rozstrzygającymi narzędziami, a decyzje powinni podejmować ludzie.

Z drugiej strony, co podkreślają E. Aranowska i J. Rytel, „wiara urzędników państwowych w nieomyślność diagnozy psychologicznej na podstawie testów jest (...) niepokojąco wysoka” (2013, s. 42), co niejednokrotnie przejawia się domaganiem się od psychologów podania wartości ilorazów inteligencji. W rozumieniu potocznym pojęcie „IQ” utożsamiane jest często ze skrótowym określeniem inteligencji, niesłusznie rozumianej jako jednolita właściwość i nie jest traktowane jako wynik tylko konkretnego testu mierzącego kombinację różnych zdolności potrzebnych do efektywnego funkcjonowania i rozwoju.

Przy założeniu niezmienności IQ w ciągu życia zmiany stwierdzone w powtórnych badaniach testowych, po upływie dłuższego czasu u określonej grupy osób, można przypisywać słabości narzędzi pomiarowych. Udostępnienie polskim psychologom skal do badania inteligencji D. Wechslera ma nieocenione znaczenie dla diagnostyki psychologicznej. Zostały osadzone w polskich realiach kulturowych, zachowały jednak „ducha” oryginału (Brzeziński, Gaul, Hornowska, Jaworowska, Machowski, 1996; Brzeziński i in., 2007; Hornowska, 2004), a wraz z nim problemy sygnalizowane zarówno przez amerykańskich badaczy, jak i polskich psychologów. Nowsze i doskonalsze wersje testów zapewne będą od nich wolne.

Bibliografia

- Anastasi, A., Urbina, S. (1999). *Testy psychologiczne*. Warszawa: Pracownia Testów Psychologicznych PTP.
- Aranowska, E., Rytel, J. (2013). Kontrowersje wokół rzetelności jako pojęcia psychometrycznego. *Przegląd Psychologiczny*, 56(1), 29-43.
- Bogdanowicz, M., Jaworowska, A., Krasowicz-Kupis, G., Matczak, A., Pelc-Pękala, O., Pietras, I., Stańczak, J., Szczerbiński, M. (2008). *Diagnoza dysleksji u uczniów klasy III szkoły podstawowej. Przewodnik diagnostyczny*. Warszawa: Pracownia Testów Psychologicznych PTP.
- Bono de, E. (2008). *Myślenie lateralne. Idee na przekór schematom*. Gliwice: Wydawnictwo Sensus.
- Brzeziński, J., Gaul, M., Hornowska, E., Jaworowska, A., Machowski, A., Zakrzewska, M. (2004, 2007). *Skala Inteligencji D. Wechslera dla dorosłych. Wersja zrewidowana – Renormalizacja, WAIS-R (PL)*. Warszawa: Pracownia Testów Psychologicznych PTP.
- Brzeziński, J., Gaul, M., Hornowska, E., Machowski, A., Zakrzewska, M. (1996). *Skala Inteligencji D. Wechslera dla dorosłych. Wersja zrewidowana WAIS-R (PL)*. Warszawa: Pracownia Testów Psychologicznych PTP.
- Brzeziński, J., Hornowska, E. (red.). (1993). *Skala Inteligencji Wechslera WAIS-R. Polska adaptacja, standaryzacja, normalizacja i wykorzystanie w diagnostyce psychologicznej*. Warszawa: Wydawnictwo Naukowe PWN.
- Choynowski, M. (1951). *Tymczasowy podręcznik do testu Wechsler-Bellevue I*. Kobierzyn: Pracownia Psychometryczna.
- Flynn, J. (2006). Tethering the elephant. Capital cases, IQ and the Flynn effect. *Psychology, Public Policy, and Law*, 12, 170-189.
- Fudala, R. (2013). Myślenie lateralne – pochwała i krytyka twórczości Edwarda de Bono. *Pedagogika Przedszkolna i Wczesnoszkolna*, 1(2), 87-94.
- Grace, W., Sweeney, M. (1986). Comparisons of the $P > V$ sign on the WISC-R and WAIS-R in delinquent males. *Journal of Clinical Psychology*, 42, 173-176.
- Grzywak-Kaczyńska, M. (1949). Metoda Wechsler-Bellevue jako pomoc w diagnozie psychologicznej. *Zdrowie Psychiczne*, 3-4, 53-76.
- Hornowska, E. (2004). *Skale Inteligencji dla dorosłych Davida Wechslera WAIS-R oraz WAIS-III*. Warszawa: Wydawnictwo Naukowe Scholar.
- Jaworowska, A. (2009). *Psychometryczna ocena upośledzenia umysłowego*. Pobrane z: www.practest.com.pl.
- Jaworowska, A. (2011). Starzenie się norm w testach inteligencji. Efekt Flynnna na przykładzie wyników WISC-R. W: K. Wiejak, G. Krasowicz-Kupis (red.), *Kliniczne zastosowania skal inteligencji D. Wechslera*, (s. 17-31). Warszawa: Pracownia Testów Psychologicznych PTP.
- Jaworowska, A., Matczak, A., Stańczak, J. (2010). *Diagnoza dysleksji. Aneks do Przewodnika diagnostycznego. Normalizacja dla uczniów klasy V szkoły podstawowej*. Warszawa: Pracownia Testów Psychologicznych PTP.
- Kostrzewski, J. (1970). *Wprowadzenie do Skali Inteligencji D. Wechslera dla dzieci od 5 do 15 lat*. Warszawa-Łódź: PTHP.
- Krasowicz-Kupis, G., Wiejak, K. (2006). *Skala Inteligencji Wechslera dla dzieci (WISC-R) w praktyce psychologicznej*. Warszawa: Wydawnictwo Naukowe PWN.
- Matczak, A. (1994). *Diagnoza intelektu*. Warszawa: Wydawnictwo Instytutu Psychologii PAN.
- Matczak, A., Piotrowska, A., Ciarkowska, W. (1991, 1997, 2008). *Skala Inteligencji D. Wechslera dla dzieci – Wersja zmodyfikowana (WISC-R)*. Warszawa: Pracownia Testów Psychologicznych PTP.
- Messick, S. (2005). Trafność testu a etyka oceny (diagnozy). W: J. Brzeziński (red.), *Trafność i rzetelność testów psychologicznych*, (s.460-485). Gdańsk: GWP.
- Necka, E. (2003). *Inteligencja. Geneza-Struktura-Funkcje*. Gdańsk: GWP.
- Neisser, U. (1997). Rising Scores on Intelligence Tests. *American Scientist*, 85(5), 440-447.
- Polskie Towarzystwo Higieny Psychiczej (1972). *Instrukcja do Skali Inteligencji D. Wechslera dla młodzieży i dorosłych*. Warszawa: PTHP.

- Resing, W.C.M., Tunteler, E. (2007). Children Becoming More Intelligent: Can the Flynn Effect be Generalized to Other Child Intelligence Tests? *International Journal of Testing*, 7(2), 191-208.
- Sattler, J. M., Polifka, J., Polifka, S., Hilsen, D. (1984). A longitudinal study of the WISC-R and WAIS-R with special education students. *Psychology in the Schools*, 3, 294-295.
- Schaffer, H.R. (2006). *Psychologia dziecka*. Warszawa: Wydawnictwo Naukowe PWN.
- Slate, J.R., Frost, J., Cross, B. (1990). Comparability of WISC-R and WAIS-R scores for a sample of college students with learning disabilities. *Learning Disability Quarterly*, 13, 205-208.
- Sochacka, K. (2016). Porównanie wyników WISC-R i WAIS-R (PL) dzieci i adolescentów w badaniach podłużnych/A comparison of WISC-R and WAIS-R (PL) scores of children and adolescents in a longitudinal study. *Roczniki Psychologiczne/Annals of Psychology*, 19(1), 159-178.
- Sochacka, K. (2018). „Ważne aspekty praktyki posługiwania się polskimi adaptacjami skal inteligencji”. Referat wygłoszony na Konferencji „Sztuka Wychowania”. Organizator: PWSZ w Nowym Sączu, w dniu 11 maja.
- Spitz, H. (1989). Variations in Wechsler Interscale IQ. Disparities at Different Levels of IQ. *Intelligence*, 13, 157-167.
- Standardy dla testów psychologicznych i pedagogicznych. (2007). Gdańsk: GWP.
- Strelau, J. (2001). Różnice indywidualne: opis, determinanty i aspekt społeczny. W: J. Strelau (red.), *Psychologia* (t. 2, s. 653-681). Gdańsk: GWP.
- Vance, H., Brown, W., Hankins, N., Furgerson, S. (1987). A comparison of the WISC-R and the WAIS-R with special education students. *Journal of Clinical Psychology*, 43, 377-380.
- Wechsler, D. (1981). *Manual for the Wechsler Adult Intelligence Scale--revised*. New York: Psychological Corporation.
- Wiejak, K., Krasowicz-Kupis, G. (red.). (2011). *Kliniczne zastosowania skal inteligencji D. Wechslera*. Warszawa: Pracownia Testów Psychologicznych PTP.
- Wilcoxon, K. (1982). *A comparison of the WAIS-R and WISC-R for sixteen-year-old learning disabled and educable mentally handicapped students*. A Thesis Presented to the Department of Psychology Emporia State University: Pobrano z: <https://esirc.emporia.edu/handle/123456789/2120>.
- Zakrzewska, M. (2001). Poziom wykonania testów WAIS-R (PL) w różnych grupach wiekowych. *Roczniki Psychologiczne*, 4, 257-278.