

JANUSZ KOSTRZEWSKI
Katedra Psychologii UŁ

**TESTY INTELIGENCJI „CULTURE-FAIR”
(powstanie, założenia, rodzaje, ocena)***

1. WSTĘP

Co to są testy inteligencji „culture-free” (tzn. testy wolne od uwarunkowań kulturowych) lub „culture-fair” (tj. testy fair pod względem kulturowym, kulturowo neutralne) lub ‘culture-reduced’ (tj. testy o ograniczonym uwarunkowaniu kulturowym)? Czy rzeczywiście można mówić, iż jakiegokolwiek testy inteligencji mogą być wolne od uwarunkowań kulturowych? Czy nie należy raczej mówić o testach inteligencji, w których występują elementy wspólne różnym, choć nie wszystkim kulturom lub subkulturom? Czy nazwa tych testów nie wyraża tylko życzeń ich konstruktorów, by wyniki badań dokonanych przy zastosowaniu tych testów były możliwe najmniej zależne od takich zmiennych, jak: miejsce zamieszkania (miasto-wieś), środowisko wychowawcze (bardziej lub mniej korzystne), płeć, kraj (przynależność do określonej mniejszości narodowej, np. Meksykanie w USA) itp. Czy sama nazwa: najpierw „culture-free”, później „culture-fair”, obecnie „culture-reduced” wskazujący na ten sam rodzaj testów nie sugeruje, iż ich autorzy doszli do wniosku, iż nie są to w istocie testy wolne od uwarunkowań kulturowych, lecz uwarunkowania te występują tu w mniejszym stopniu niż w tradycyjnych testach inteligencji?

Kiedy powstały testy inteligencji „culture-fair”? Jakie założenia leżą u ich podłoża? Jakie są dzisiaj znane w świecie testy inteligencji „culture-fair” lub „culture-reduced”? Czy potrzebne są dzisiaj testy „culture-fair”? Jeśli

* Fragmenty tego artykułu ukazały się w pracy pod red. Anny Ciechanowicz, *Kulturowa adaptacja testów*, Warszawa 1990.

tak, to do jakich celów są one potrzebne oraz w jakich przypadkach powinny być stosowane? Jaki jest dzisiaj pogląd na wartość tych testów? Na te pytania postaram się odpowiedzieć w tym artykule.

2. POJĘCIE TESTU INTELIGENCJI „CULTURE-FREE” LUB/I „CULTURE-FAIR”

Warto się teraz zastanowić nad tym, co to są testy inteligencji „culture-free” i/lub „culture-fair” oraz „culture-reduced”? Czy pojęcia te są używane zamiennie, czy też nie? Przy okazji warto prześledzić ewolucję w rozumieniu terminu „test inteligencji culture-free”.

Ze względów historycznych należy w tym miejscu wspomnieć, iż nazwy „test inteligencji culture-free” po raz pierwszy użył w 1940 r. Raymond B. Cattell w swoim artykule *A Culture-free Intelligence Test* na określenie skonstruowanych przez siebie testów, które zgodnie z założeniem miały być wolne od uwarunkowań kulturowych.

H. B. English i Ava Ch. English (1958, s. 547), autorzy znanego słownika: *A Comprehensive Dictionary of Psychological and Psychoanalytical Terms*, wyjaśniają, iż „test culture-free jest to test ogólnej inteligencji, z którego wyeliminowano na tyle, na ile to było możliwe, wszystkie zadania zależne od doświadczeń, które występują częściej w jednej kulturze niż innej. Nazwą tą obejmuje się niekiedy grupę testów wolnych od uwarunkowań klasowych; testy te służą do badań osób pochodzących z różnych klas społeczno-ekonomicznych (lub subkultur) w ramach tej samej kultury”. Autorzy ci podają, iż synonimem testów wolnych od uwarunkowań kulturowych („culture-free”) są testy „międzykulturowe” („cross-culture”), bowiem nie są one całkowicie wolne od czynników kulturowych.

R. B. Cattell początkowo posługuje się wyłącznie terminem „test inteligencji culture-free”. W jednej ze swoich prac z 1957 r. podaje następującą definicję tego terminu: „jest to percepcyjny test inteligencji wymagający ujmowania złożonych związków; nie zawiera on materiału obrazkowego, liczbowego oraz wiedzy słownej, stąd też jest stosunowo wolny od różnych wpływów statusu intra- lub interkulturowego” (Cattell 1957, s. 891). Od 1959 r. używa on zamiennie terminu „culture-free” i „culture-fair”.

Od 1971 r. R. B. Cattell używa już trzech terminów na określenie tych samych, skonstruowanych przez siebie testów inteligencji: „culture-free”, „culture-fair” i „culture-reduced”. Stwierdza on, iż „test inteligencji fair pod względem kulturowym (culture-fair), jest wolny od uwarunkowań kulturowych (culture-free) lub też o ograniczonym uwarunkowaniu kulturowym (culture-reduced); jest to taki test, który bazuje na materiale i uwzględnia nawyki wspólne wszystkim kulturom i który stara się zminimalizować wpływ

specjalnych, kulturowo uwarunkowanych sprawności w ujmowaniu złożonych relacji. Nie wymaga eliminacji różnic rasowych i klasowych, jeśli takie różnice istnieją realnie” (Cattell 1971, s. 513).

Jak wynika z faktu nazwania nowych wydań tych samych testów uprzednio nazywanych „testami inteligencji culture-free” – „testami inteligencji culture-fair”, jak też zamienne stosowanie terminu „culture-fair” i „culture-reduced”, R. B. Cattell od 1971 r. podkreśla, iż skonstruowane przez niego testy inteligencji nie są bezwzględnie wolne od uwarunkowań kulturowych, ale są one mniejsze niż w takich tradycyjnych technikach pomiaru inteligencji, jak: skala inteligencji Stanford-Bineta, 1937, 1960, skala inteligencji D. Wechslera dla dzieci (WISC 1949) oraz dla dorosłych (W-B 1939, WAIS 1955) i inne. Rezygnacja z terminu „culture-free” na rzecz terminów „culture-fair” i „culture-reduced” świadczy o pewnej ewolucji poglądów R. B. Cattella na omawiany rodzaj testów inteligencji.

Dość odosobniony jest pogląd M. J. Begaba (1973, s. X) na testy inteligencji „culture-free” i „culture-fair”. Warto o nim wspomnieć, bowiem przedstawił go we wprowadzeniu do znanego *Podręcznika terminologii i klasyfikacji upośledzenia umysłowego* opublikowanego pod redakcją H. J. Grossmana (1973). Otóż testami inteligencji „culture-free” nazywa on te testy, z których usunięto pytania zawierające treść kulturową, natomiast testami „culture-fair”, te, które uwzględniają elementy wspólne, istniejące w każdej kulturze stworzonej przez człowieka. Inne jego poglądy zostaną omówione w dalszej części tego artykułu.

3. Z HISTORII POWSTANIA TESTÓW INTELIGENCJI „CULTURE-FREE”, „CULTURE-FAIR” LUB „CULTURE-REDUCED” ZAŁOŻENIA

Nim przejdę do przedstawienia krótkiej historii powstania testów „culture-fair”, podam krótką historię powstania testów inteligencji. Gdyby bowiem nie było tradycyjnych testów inteligencji, nie byłoby również testów „culture-fair”.

Jak wiemy, w 1904 r. francuskie Ministerstwo Oświaty zwróciło się do psychologa, prof. dr. Alfreda Bineta, z prośbą o skonstruowanie techniki, która ułatwiłaby selekcję dzieci umysłowo upośledzonych ze szkół masowych i skierowanie ich do szkół specjalnych. A. Binet podjął się tego zadania wraz z dr. med. Theodorem Simonem. Owocem tej współpracy była pierwsza w świecie skala pomiaru inteligencji dzieci opublikowana w 1905 r. A. Binet uważał, iż zdolność rozumienia, rozumowania i sądzenia są istotnymi cechami inteligencji. W 1908 r. powstała druga, udoskonalona wersja skali, zaś w 1911 r., jeszcze bardziej udoskonalona, trzecia wersja; służyła ona do badań indywidualnych dzieci od 3 r. ż., młodzieży i dorosłych.

Nieomal natychmiast po ukazaniu się we Francji skali Bineta-Simona na całym świecie zaczęto dokonywać jej przekładów, rzadziej zaś adaptacji i normalizacji na reprezentatywnej próbie populacji danego kraju.

Już w 1908 r. ukazał się w USA angielski przekład tej skali dokonany przez H. H. Goddarda, zaś w 1911 r. rewizja tej skali. W 1912 r. ukazała się rewizja skali Bineta-Simona dokonana przez F. Kuhlmana. W 1915 r. ukazał się również przekład tej skali dokonany przez R. M. Yerkesa, J. U. Bridgesa i R. S. Hardwicka. Jednak stosunkowo najlepszej adaptacji tej skali i jej normalizacji dokonał w USA Lewis M. Terman wraz z grupą współpracowników z Uniwersytetu Stanfordzkiego. W 1916 r. ukazała się pierwsza rewizja skali inteligencji Stanford-Bineta. W 1937 r. ukazała się druga jej rewizja, zawierająca dwie wersje L i M, lepiej znormalizowana. W 1960 r. ukazała się trzecia jej rewizja. Dokonano wyboru najlepszych testów z dwu wersji, tworząc nową wersję L-M. Ponownie ją znormalizowano. W 1972 r. dokonano kolejnej jej normalizacji w USA. Jest to jedna z najbardziej znanych w świecie technik pomiaru inteligencji. W toku jej stosowania zebrano szereg cennych doświadczeń, które przyczyniły się do jej udoskonalenia. Wysuwano pod jej adresem szereg uzasadnionych zarzutów, m. in. iż ma zbyt werbalny charakter (zarzut ten starano się usunąć poprzez wprowadzenie do wersji z 1937 r. szeregu testów wykonawczych dla dzieci od 2 do 5 lat), że na wyniki badań ma wpływ płeć dziecka, miejsce jego zamieszkania (miasto-wieś), zawód rodziców, mniej lub bardziej korzystne środowisko wychowawcze, rasa (ażeby zminimalizować ten zarzut normalizacja skali z 1972 r. objęła nie tylko dzieci reprezentujące rasę białą, ale również inne rasy, których reprezentanci mieszkają w USA), subkultura (dzieci meksykańskie i innych mniejszości narodowych nie znające dobrze języka angielskiego, mieszkające w USA uzyskiwały niskie ilorazy inteligencji).

Warto wspomnieć, iż w 1986 r. ukazało się czwarte, znacznie zmienione i udoskonalone przez R. L. Thorndike'a, Elizabeth P. Hagen, J. M. Sattlera i ich współpracowników wydanie skali inteligencji Stanford-Bineta, które nie przypomina poprzednich.

Poprzednie wydania omawianej skali zostały przełożone na wiele języków świata (m. in. na język polski) i często bez adaptacji i normalizacji (dokonanej na reprezentatywnej próbie populacji) stosowane do celów diagnostycznych, w tym diagnozy upośledzenia umysłowego. Ponadto przy rozpoznawaniu odchyleń od normy nie zawsze uwzględniano dane wywiadu anamnestycznego i środowiskowego, całokształt wiedzy o badanym, jak też całokształt wiedzy klinicznej, nie uwzględniano również niekiedy wyników badań lekarskich, co w efekcie prowadziło do poważnych pomyłek. Takie nadużywanie skali spotykało się z protestami i nawet zakazami jej stosowania.

Po zapoznaniu się z przekładami skali Bineta-Simona na język angielski w USA badacze uświadomili sobie, iż nie może ona być stosowana do badań imigrantów nie znających języka angielskiego. Do badań tych osób

należało skonstruować taką technikę wykonawczą, która nie wymagałaby znajomości języka danego kraju. Technikę taką skonstruował i opublikował w 1914 r. H. A. Knox. Służyła ona do badania pamięci i uwagi. Można ją było stosować tylko do badań indywidualnych.

W 1917 r. USA włączyły się do I wojny światowej. Należało dobrać możliwie szybko ludzi na odpowiednie stanowiska w wojsku, by móc ich szybko przeszkolić. Potrzebne były rzetelne i trafne testy służące do badań zbiorowych. Do tych celów skonstruowano test armii alfa. Zawierał on szereg testów słownych i liczbowych, służących do pomiaru inteligencji. Wkrótce stwierdzono, iż nie nadaje się on do badań osób, nie znających biegle języka angielskiego i nie umiejących pisać. Do badań tych osób skonstruowano test armii beta. Był to test wykonawczy, ale nie test „culture-fair”.

Przez wiele lat posługiwano się skalą inteligencji Stanford-Bineta i testami armii alfa i beta do pomiaru inteligencji osób reprezentujących mniejszości narodowe w USA. Ich znajomość języka angielskiego była dość słaba. Wyniki badań tych osób porównywano z normami opartymi na reprezentatywnej próbie populacji. Z zasady uzyskiwali oni niższe wskaźniki inteligencji; wyniki tych badań były często podstawą niekorzystnych dla nich decyzji dotyczących dalszego kształcenia i zatrudnienia na odpowiednim stanowisku. Takie decyzje budziły sprzeciw nie tylko zainteresowanych i ich rodzin, ale również tych badaczy, którzy widzieli niewłaściwość takich decyzji. Stąd też dla tych osób, które nie znały dobrze języka angielskiego, zaczęto konstruować takie skale wykonawcze, które nie wymagały znajomości tegoż języka.

W roku 1924 S. D. Porteus opublikował zestaw labiryntów służących do pomiaru zdolności przewidywania i planowania u dzieci w wieku od 3 r. ż., młodzieży i dorosłych. Sam test powstał wcześniej. W 1918 r. badacz ten opublikował wyniki wstępnych badań dokonanych tym testem. Był to test wykonawczy, nie wymagający znajomości języka angielskiego. Test ten możemy dzisiaj zaliczyć do kategorii testów „culture-reduced”. Nasunął on R. B. Cattellowi pomysł wykorzystania tego typu testów w swoich badaniach i późniejszego ich włączenia do skonstruowanej przez siebie baterii testów „culture-fair”.

W 1927 r. zrodził się pomysł u R. G. Leitera skonstruowania „międzynarodowej skali czynnościowej”, służącej do pomiaru inteligencji osób mieszkających w różnych krajach. Pierwsza wersja tej skali ukazała się w 1929 r. Skala ta była doskonalona przez wiele lat. Ostatnia jej wersja ukazała się w 1948 r. Zaliczana jest do kategorii testów „culture-fair”, choć jej autor tak jej nie nazwał.

W 1931 r. ukazała się praca W. Lina poświęcona rozwojowi percepcji wzrokowej u dzieci. Wykazał on, iż pewne testy wymagające ujmowania związków między prostymi kształtami geometrycznymi są w wysokim stopniu nasycone czynnikiem „g”. W tym samym czasie M. Fortes wykazał w swojej

pracy doktorskiej, iż testy wymagające ujmowania powiązań istniejących w prostym, nie zawierającym treści materiału percepcyjnym są trafną techniką pomiaru czynnika „g” (Cattell 1940, s. 167). W. Stephenson również w 1931 r. wykazał, iż czynnik „g” występuje zarówno w testach werbalnych, jak i niewerbalnych. Ponadto stwierdził on, iż „testy percepcyjne są jednym z najlepszych sposobów, które służą do mierzenia zdolności myślenia, gdyż są najmniej zależne od doświadczeń życiowych badanej osoby” (Hornowski 1970, s. 21).

W 1933 r. C. Spearman opublikował swój test służący do pomiaru percepcji wzrokowej. Był to test niewerbalny. Instrukcje podawano przy pomocy pantomimiki. W 1936 r. L. Penrose i J. C. Raven opublikowali próbną serię testów percepcyjnych, zwanych „Progressive Matrices”. Ta próbna wersja była podstawą znanej powszechnie techniki J. C. Ravena: „Standard Progressive Matrices”, opublikowanej w 1938 r.

R. B. Cattell przystępując do konstrukcji swoich testów „culture-free” pod koniec lat trzydziestych znalazł istniejące wówczas testy percepcyjne i wiedział z badań, iż są one w stosunkowo dużym stopniu nasycone czynnikiem „g” oraz stosunkowo najmniej zależne od uwarunkowań kulturowych.

Czynnikiem, który spowodował podjęcie się R. B. Cattella konstrukcji testów „culture-fair”, była potrzeba posiadania testów pomiaru inteligencji osób posługujących się różnymi językami, pochodzących z różnych kręgów kulturowych, wychowywanych w różnych warunkach środowiskowych i posiadających różne wykształcenie (Cattell 1950, s. 3). Dobrze on wiedział, iż dzieci murzyńskie, indiańskie lub meksykańskie mieszkające w USA w mniej korzystnych warunkach środowiskowych, gorzej znające język angielski od swoich rówieśników wychowujących się w bardziej korzystnych warunkach środowiskowych uzyskują zazwyczaj niskie ilorazy inteligencji w wyniku badań skalą Stanford-Bineta, 1916, 1937. Wiedział, iż wiele tych dzieci jest bardziej inteligentnych niż na to wskazują tradycyjne testy inteligencji. Bowiem tradycyjne testy inteligencji mierzą inteligencję uwikłaną w uwarunkowania kulturowe. Potrzebne były takie testy, które w miarę możliwości byłyby wolne od uwarunkowań kulturowych. Pragnął posiadać takie narzędzie pomiaru inteligencji, które eliminowałoby wpływ tych wszystkich czynników, które mogą zafałszować prawdziwy obraz tejże inteligencji.

R. B. Cattell konstruując testy inteligencji „culture-fair” zdawał sobie sprawę z tego, iż powinien uwzględnić treści wspólne różnym kulturom. Uważał on, iż wszystkie kultury i subkultury znają ciało człowieka, jego części, ślady stóp ludzkich, schematy drzew, a nie określone rodzaje drzew (wyjątek stanowią Eskimosi), zwierzęta (schematy, a nie różne gatunki zwierząt), Ziemię, niebo, chmury, Słońce, Księżyc, gwiazdy, błyskawicę, ogień, dym, wodę, rodziców, dzieci i proste stosunki rodzinne. Opierając się na tym założeniu w skonstruowanych przez siebie testach percepcyjnych,

uwzględnił to wszystko, co jest wspólne wszystkim kulturom. Pierwszą skonstruowaną serię testów „culture-free” omówił w artykule pt. *A Culture-free Intelligence Test* opublikowanym w 1940 r. Ta pierwsza wersja składała się z siedmiu następujących testów: 1) labirynty, 2) serie, 3) klasyfikacje, 4) macierze I, 5) macierze II, 6) macierze III oraz 7) zwierciadlane odbicia. Przy konstrukcji testu (1) oparł się na pomysły S. D. Porteusa, twórcy testu labiryntów, zaś konstruując test macierzy wzorował się na pomysły J. C. Ravena. Tworząc poszczególne testy, uwzględnił przy ich konstrukcji graficznej schematy liści, zwierząt, ludzi, części ciała ludzkiego, Słońca, gwiazd, a przede wszystkim różnych figur geometrycznych (lini, kół, kwadratów, prostokątów, trójkątów itp.). Część wprowadzająca do badań obejmowała 32 testy, zaś część właściwa – 100 testów. Badania empiryczne pozwoliły mu na ocenę wartości skonstruowanej techniki.

W ciągu następnych lat R. B. Cattell doskonalił pomysł i konstruował odrębne wersje testów „culture-fair” przeznaczone do badań dzieci, młodzieży i dorosłych. Przy konstrukcji poszczególnych wersji testów „culture-fair” autor pominął pytania i zadania wymagające zdolności operowania słowami i liczbami, jak też uwzględniające wiadomości i wiedzę. Wyjątek stanowi tylko skala 1 dla dzieci od 4 do 8 lat i starszych umysłowo upośledzonych, jak to zobaczymy w dalszej części tego artykułu.

R. B. Cattell skonstruował trzy serie testów inteligencji „culture-fair”: Skala 1 służy do pomiaru inteligencji dzieci w wieku od 4 do 8 lat i starszych umysłowo upośledzonych. Skala 2 służy do badań dzieci w wieku od 8 do 13 lat i dorosłych o przeciętnym poziomie inteligencji, zaś skala 3 – do badań młodzieży szkół średnich i dorosłych o wysokim poziomie inteligencji. Opierają się one na wspomnianych założeniach. U ich podłoża leży teoria dwóch czynników ogólnych: czynnika zdolności płynnych i skryzalizowanych, której autorem jest R. B. Cattell. W dalszej części tego artykułu przedstawię zarys tej teorii oraz omówię poszczególne skale pod kątem ich rzetelności i trafności. Ponadto postaram się wykazać na podstawie wyników dotychczasowych badań, iż na ilorazy inteligencji uzyskane w badaniach dokonanych R. B. Cattella testami inteligencji „culture-fair” w mniejszym stopniu wpływają uwarunkowania kulturowe niż na ilorazy inteligencji uzyskane w badaniach tradycyjnymi skalami inteligencji.

4. CHARAKTERYSTYKA ISTNIEJĄCYCH TESTÓW INTELIGENCJI „CULTURE-FAIR” LUB „CULTURE-REDUCED”

Jakie testy pomiaru inteligencji są zaliczane do kategorii testów „culture-fair” lub „culture-reduced”? Które spośród nich są faktycznie testami o ograniczonym uwarunkowaniu kulturowym?

Anne Anastasi (1964, s. 258–268), J. Horrocks (1964, s. 271–296) oraz J. M. Sattler (1982, s. 382) do kategorii testów inteligencji „culture-fair” zaliczają następujące testy: R. B. Cattella testy inteligencji „culture-fair” (4.1), R. G. Leitera międzynarodową skalę wykonawczą (4.2), Johna G. Ravena testy percepcyjne (4.3), Florence Goodeough i D. Harrisa test postaci ludzkiej (4.4), Allison Davis i Kennetha Fellsa test inteligencji ogólnej (4.5) oraz F. L. Wellsa i J. Ruescha semantyczny test inteligencji (4.6). Należy również wspomnieć o Safrana teście inteligencji o ograniczonym uwarunkowaniu kulturowym (Safran Reduced Intelligence Test, SCRIT, 4.7).

4.1. R. B. Cattella testy inteligencji „culture-fair”

4.1.1. Podstawy teoretyczne R. B. Cattella testów inteligencji „culture-fair”

Skonstruowane przez R. B. Cattella trzy serie testów inteligencji „culture-fair”, zgodnie z założeniem ich autora mają mierzyć inteligencję płynną. Opierają się one na jego teorii dwóch czynników ogólnych: czynnika zdolności płynnych (g_f – *fluid*) i skryształizowanych (g_c – *crystalized*). Teoria ta rodziła się w okresie konstruowania pierwszych testów „culture-fair”. Teorię tę zaproponował on w 1940 r., zaś jej zarys opublikował w 1941 r. Ponownie ją przedstawił w 1943 r. W latach 1950 i 1957 została ona bardziej sprecyzowana. Dopiero w 1963 r. zostały opublikowane wyniki badań empirycznych potwierdzających tę teorię.

Warto w tym miejscu przypomnieć, iż w 1904 r. C. Spearman opublikował zarys swojej teorii dwóch czynników. Została ona sprecyzowana w 1927 r. Według tej teorii zdolności umysłowe człowieka współtworzą dwa czynniki: czynnik ogólny (*general factor*) i czynniki specyficzne (*specific*). Tym czynnikiem ogólnym (g) jest ogólna energia umysłowa, która ujawnia się przede wszystkim w wykrywaniu stosunków między przedmiotami i elementami oraz wyprowadzaniu wniosków (Spearman 1927). R. B. Cattell wyszedł od tej teorii, ale na podstawie własnych przemyśleń oraz wyników badań doszedł do wniosku, iż Spearmanowski czynnik ogólny (g) nie stanowi jednorodnego wymiaru, ale mamy do czynienia z dwoma czynnikami ogólnymi. Te dwa czynniki, opierając się na ich właściwościach, nazwał czynnikiem inteligencji płynnej i skryształizowanej.

Inteligencja płynna, którą mierzą skonstruowane przez niego testy inteligencji „culture-fair”, ujawnia się w tych sytuacjach, w których musimy rozwiązywać zupełnie nowe problemy nie mając żadnych gotowych schematów działania, przygotowanego czy zapamiętanego sposobu rozwiązania danego

problemu. Ujawnia się ona w testach wymagających przystosowywania się do wciąż nowych sytuacji. R. B. Cattell (1963, s. 16) uważa, iż inteligencja płynna jest to „zdolność ujmowania związków i wyprowadzania wniosków w znaczeniu zaproponowanym przez C. Spearmana”. Jest ona uwarunkowana biologicznie i fizjologicznie jako funkcja wszystkich komórek kory mózgowej. Nie oznacza to – zdaniem R. B. Cattella – iż jest ona w pełni uwarunkowana czynnikami genetycznymi. Tak nie jest, bowiem od chwili poczęcia się dziecka działają nań różne czynniki środowiskowe. Stąd też takie czynniki, jak choroby zakaźne, środki chemiczne, urazy itp., mogą uszkodzić pełnowartościowe zadatki organiczne, z których ukształtuje się inteligencja płynna. Rozwój inteligencji płynnej u człowieka osiąga stosunkowo wcześniej swój pułap, a mianowicie między 14 a 15 r. ż., zaś około 22 r. ż. zaczyna się stopniowo obniżać aż do wieku starczego (Cattell 1963, s. 3–4).

Inteligencją skryształowaną nazywa on te zdolności poznawcze, w których nabyta uprzednio wprawa w wydawaniu sądów i rozumowaniu została już utrwalona, skryształowana (stąd nazwa tego czynnika) jako wynik wcześniejszego usprawniania w tym zakresie bardziej podstawowej ogólnej zdolności. Przykładami utrwalonych zdolności są: zdolność rozumienia słów, płynność słowna, wyobrażenia przestrzenna, osiągnięcia z zakresu historii, geografii, przyrody itp. Rozwój tej inteligencji zależy od ogólnie panującego w danym kraju modelu kształcenia się. Jeśli model ten zakłada, iż człowiek kształci się do 25–28 r. ż., wówczas do tego czasu rozwija się inteligencja skryształowana. Jeśli człowiek nadal się uczy, to kształtuje się również jego inteligencja skryształowana.

Wszystkie skale pomiaru inteligencji jako przeciętny wskaźnik (Iloraz Inteligencji) przyjmują 100. Mają one jednak różne odchylenia standardowe. R. B. Cattell twierdzi, iż w tych testach, które mierzą inteligencję skryształowaną odchylenia standardowe wahają się w granicach od 12 do 16 punktów, natomiast w testach mierzących inteligencję płynną wahają się one w granicach od 24 do 25 punktów. Dlaczego są tak duże różnice? Otóż zdaniem R. B. Cattella (1963, s. 3) inteligencja płynna jest bardziej zależna od uwarunkowań biologicznych, zaś skryształowana – środowiskowych; zależy bowiem od czasu wyznaczonego w danym kraju na kształcenie się. Twierdzi on, iż obecny system kształcenia szkolnego hamuje rozwój jednostek bardzo zdolnych, zaś siłą stymuluje rozwój intelektualny jednostek mało zdolnych, a nawet tępych. Wskutek takiego systemu kształcenia zmniejszają się bezwzględne różnice między inteligencją płynną i skryształowaną u naszych uczniów. Podkreśla on, iż tradycyjne testy inteligencji (m. in. skala inteligencji Stanford-Bineta, skala inteligencji D. Wechslera) dla dzieci i dorosłych mierzą mieszaninę omawianych dwóch rodzajów inteligencji, zaś jego testy inteligencji „culture-fair” mierzą głównie inteligencję płynną.

4.1.2. Charakterystyka trzech skal inteligencji

Skala 1. Jak już wspomniałem, służy ona do pomiaru inteligencji dzieci w wieku od 4 do 8 lat i starszych umysłowo upośledzonych. Składa się z ośmiu takich testów, jak: 1) podstawianie, 2) klasyfikacja, 3) labirynty, 4) wskazywanie nazywanych przedmiotów, 5) wykonywanie poleceń, 6) niedorzeczności w obrazkach, 7) zagadki oraz 8) podobieństwa. Cała skala składa się z 96 zadań, zaś każdy z testów z 12 zadań. Właściwy czas badania wynosi 22 minuty, zaś całe badanie (podawanie instrukcji, odpowiedzi badanego, krótkie przerwy między poszczególnymi testami) od 40 do 60 min. Wprawdzie cała skala składa się z 8 testów, ale tylko 4 spośród nich tworzą skalę „culture-fair”. Są to: 1) podstawianie, 2) klasyfikacja, 3) labirynty i 4(8) podobieństwa. Dla celów badań międzykulturowych R. B. Cattell zaleca stosowanie wyłącznie tych czterech testów.

Normalizacja tej skali opiera się na próbce obejmującej ponad 400 dzieci w wieku od 3 do 12 lat z Anglii i USA. Autor zapowiedział zwiększenie próbki normalizacyjnej, gdyż ta jest stosunkowo mała. Normy są wyrażone w postaci ilorazów inteligencji, przy czym średni I. I. próbki normalizacyjnej wynosi 100, zaś odchylenie standardowe 20. Uzyskane I. I. można zamienić również na centyle.

Współczynniki wiarygodności poszczególnych testów tworzących skalę wahają się w granicach od 0,57 do 0,71, zaś całej skali 0,80 (Schaie 1967, cyt. za Cattell 1970, s. 14).

Uzyskane przez R. B. Cattella (1970, s. 14) współczynniki korelacji wskazujące na homogeniczność skali (według wzoru Spearmana-Browna 0,91), oraz na zgodność wewnętrzną skali (według wzoru Kudera-Richardsona 0,94) są bardzo wysokie. Współczynniki korelacji poszczególnych testów z całą skalą w badaniach R. B. Cattella (1970, s. 14) wahają się w granicach od 0,67 (podobieństwa) do 0,90 (wskazywanie nazywanych przedmiotów).

Analiza czynnikowa dokonana przez R. B. Cattella (1970, s. 14) wykazała, iż współczynniki trafności czynnikowej poszczególnych testów wahają się w granicach od 0,64 (podstawianie i podobieństwa) do 0,90 (wskazywanie nazywanych przedmiotów). Trafność czynnikowa całej skali wynosi 0,95. Jest ona wysoka.

Jane Lathrop (1966) skorelowała wyniki badań 25 dzieci w wieku od 4 do 8 lat dokonanych skalą inteligencji Stanford-Bineta (1960) z rezultatami badań omawianą skalą 1. Jak należało oczekiwać, uzyskała ona nieco wyższe współczynniki korelacji z testami nasyconymi kulturowo, jak: wskazywanie nazywanych przedmiotów (0,52), wykonywanie poleceń (0,54), niedorzeczności w obrazkach (0,64) i zagadki (0,53) niż z testami „culture-fair”, jak: klasyfikacja (0,19), labirynty (0,11) i podobieństwa (0,47); wyjątek stanowi

test (1) podstawianie (0,58). Między wynikami badań skalą Stanford-Bineta, 1960, a całą skalą 1 uzyskała ona umiarkowany współczynnik korelacji 0,62, wskazujący na istotną współzależność ilorazów inteligencji obu skal, statystycznie istotny. G. Meredith (1965) uzyskał umiarkowany współczynnik korelacji (0,46) między rezultatami badań 72 dzieci dokonanych omawianą skalą 1 oraz testem Florence Goodenough: rysunek postaci ludzkiej. Obie te techniki są zaliczane do testów „culture-fair”. Uzyskany współczynnik korelacji wskazuje na istotną zbieżność wyników badań dokonanych obu tymi technikami (Cattell 1970, s. 14).

W 1990 r. dokonano przekładu omawianej skali na język polski (Szymański 1992). W celu adaptacji tej skali do badań dzieci polskich do testu (7) zagadki dodaliśmy 16 nowych zadań (Kostrzewski i M. Szymański, 1992). Rzetelność i trafność polskiego przekładu tej skali oszacowano na podstawie wyników badań 50 dzieci w wieku 8 lat.

Stabilność bezwzględną skali oszacowano opierając się na wynikach dwukrotnych badań wspomnianej grupy dzieci dokonanych w odstępie od 3 do 4 tygodni. Uzyskane współczynniki stabilności poszczególnych testów wahały się w granicach od 0,86 (wskazywanie nazywanych przedmiotów) do 0,94 (niedorzeczności w obrazkach). Współczynnik stabilności bezwzględnej całej skali wynosi 0,97; wskazuje on na bardzo wysoką stabilność całej skali. Współczynnik stabilności podskali „culture-fair” (którą tworzą testy 1, 2, 3 i 8) wynosi 0,95, zaś podskali nie „culture-fair” (którą tworzą testy: 4, 5, 6 i 7) również 0,95. Oba współczynniki stabilności wskazują na bardzo wysoką stabilność obu wspomnianych podskal.

Współczynniki korelacji poszczególnych testów z całą skalą w naszych badaniach wahały się w granicach od 0,53 (klasyfikacja) do 0,75 (wykonywanie poleceń i zagadki), zob. Szymański 1992. Są one nieco niższe niż w badaniach R. B. Cattella (Cattell 1970, s. 14).

Ażeby oszacować trafność diagnostyczną skali 1, skorelowaliśmy wyniki badań 50 dzieci 8-letnich uzyskane przy jej zastosowaniu z rezultatami badań dokonanych skalą inteligencji D. Wechslera dla dzieci (WISC 1949). Badania wykazały, iż między całą skalą 1 R. B. Cattella a skalą pełną D. Wechslera (0,69), słowną (0,54) i wykonawczą (0,62) zachodzi istotna zależność, na którą wskazują umiarkowane współczynniki korelacji, statystycznie bardzo istotne ($P = 0,001$), zob. Szymański 1992.

M. Choynowski (1977, s. 36) dokonał analizy czynnikowej wyników badań dokonanych skalą inteligencji D. Wechslera dla dzieci. Uzyskał dwa rozwiązania: dwu- i trójczynniki. W rozwiązaniu dwuczynnikowym pierwszy czynnik jest Cattellovskim czynnikiem inteligencji skryzalizowanej, tłumaczącym 56% wariacji wspólnej, drugim zaś czynnikiem inteligencji płynnej, tłumaczącym 44% wariacji wspólnej. Czynnikiem inteligencji skryzalizowanej są nasycone następujące testy skali słownej: wiadomości (0,82), rozumienie

(0,46), arytmetyka (0,49), podobieństwa (0,76) oraz słownik (0,79), natomiast czynnikiem inteligencji płynnej są nasycone takie testy skali wykonawczej, jak: braki w obrazkach (0,64), porządkowanie obrazków (0,59), klocki (0,69) oraz układanki (0,79).

Należy przypuszczać, iż ta część skali 1 R. B. Cattella, która jest faktycznie „culture-fair” (którą współtworzą testy 1, 2, 3 i 8) w większym stopniu będzie korelować ze skalą wykonawczą D. Wechslera, która mierzy inteligencję płynną, niż ze skalą słowną, mierzącą inteligencję skryzalizowaną. Natomiast druga część skali 1, którą tworzą pozostałe testy (tj. 4, 5, 6 i 7), nie wolne od uwarunkowań kulturowych, w nieco większym stopniu będzie korelować ze skalą słowną niż wykonawczą, choć różnice nie muszą być duże. Czy wyniki badań potwierdzają zasadność tej hipotezy?

Jak wykazują wyniki badań, między podskalą „culture-fair” skali 1 R. B. Cattella a skalą słowną D. Wechslera dla dzieci zachodzi zależność wyraźna, lecz mała, na którą wskazuje niski współczynnik korelacji (0,36), statystycznie istotny ($P = 0,01$). Natomiast między omawianą podskalą a skalą wykonawczą zachodzi zależność istotna, na którą wskazuje umiarkowany współczynnik korelacji (0,59), statystycznie bardzo istotny ($P = 0,001$). Natomiast między drugą podskalą Skali 1 R. B. Cattella, którą współtworzą tradycyjne testy, nie wolne od uwarunkowań kulturowych, a skalą słowną (0,60) i wykonawczą (0,56) D. Wechslera dla dzieci zachodzi zależność istotna, na którą wskazują umiarkowane współczynniki korelacji, statystycznie bardzo istotne ($P = 0,001$). Nieco wyższy jest współczynnik korelacji ze skalą słowną niż wykonawczą, choć różnica ta nie jest istotna (Szymański 1992). A zatem zgodnie z wysuniętą hipotezą, istnieje większa współzależność między podskalą „culture-fair” skali 1 R. B. Cattella a tymi testami skali inteligencji D. Wechslera dla dzieci, które mierzą inteligencję płynną, niż z testami mierzącymi inteligencję skryzalizowaną. Natomiast podskala nie wolna od uwarunkowań kulturowych w niemal równym stopniu koreluje z testami mierzącymi inteligencję płynną i skryzalizowaną, choć minimalnie bardziej z testami mierzącymi inteligencję skryzalizowaną.

Jak wiemy, tradycyjne testy inteligencji mierzą oba rodzaje inteligencji. Ponieważ skala pełna inteligencji D. Wechslera dla dzieci mierzy zarówno inteligencję płynną, jak i skryzalizowaną, stąd też należy przypuszczać, iż w podobnym stopniu będą z nią korelować wyniki badań dokonanych osobno podskalą „culture-fair”, jak też podskalą nie wolną od uwarunkowań kulturowych. Badania wykazały, iż zarówno podskala „culture-fair” (0,60), jak też podskala nie wolna od uwarunkowań kulturowych (0,66) w istotnym stopniu korelują ze skalą pełną D. Wechslera dla dzieci. Na zależność tę wskazują umiarkowane współczynniki korelacji, statystycznie bardzo istotne ($P = 0,001$) – Szymański 1992. Zależność ta jest nieco większa w podskali nie wolnej od uwarunkowań kulturowych, bowiem skala pełna D. Wechslera

dla dzieci w nieco większym stopniu jest nasycona czynnikiem inteligencji skryształizowanej, niż płynnej.

Skala 2. Służy ona do badań inteligencji dzieci w wieku od 8 do 13 lat oraz dorosłych o przeciętnym poziomie inteligencji. Posiada ona dwie wersje A i B. Każda z dwu wersji składa się z czterech następujących testów: 1) serie, 2) klasyfikacje, 3) macierze i 4) zależności. Każda wersja zawiera 46 zadań. Czas badania obu wersjami wynosi 25 minut (gdy pominiemy czas przeznaczony na instrukcję wprowadzającą do każdego testu). Skala służy do badań indywidualnych i grupowych.

Normalizacji skali dokonano na podstawie wyników badań 4328 dziewcząt i chłopców w wieku od 7 do 14 lat. Były to dzieci amerykańskie i angielskie. Normy są wyrażone w skali ilorazowej i centylowej. Średni I. I. grupy normalizacyjnej wynosi 100, zaś odchylenie standardowe 24,4.

Współczynniki wiarygodności całej skali wahają się w granicach od 0,82 (badania prowadzone w USA) do 0,85 (badania dzieci angielskich). Takie same wartości współczynników wiarygodności uzyskano we Francji i we Włoszech. Współczynniki homogeniczności poszczególnych testów wahają się w granicach od 0,70 (serie) do 0,92 (zależności). Dotyczą obu wersji łącznie.

Analiza czynnikowa ujawniła następujące nasycenie czynnikiem „g” poszczególnych testów skali: 1) serie – 0,53, 2) klasyfikacje – 0,68, 3) macierze – 0,89 i 4) zależności – 0,99. Badania dokonane we Francji wskazują na podobne nasycenie (od 0,78 do 0,83; zob. Cattell i Cattell 1970, s. 4 i 11).

R. B. Cattell i A. K. S. Cattell (1970, s. 5) podają, iż między wynikami badań dokonanych skalą inteligencji Stanford-Bineta a rezultatami badań omawianą skalą uzyskano podczas jednych badań umiarkowany współczynnik korelacji (0,56), wskazujący na isotonię zbieżność wyników, zaś podczas innych badań – wysoki współczynnik korelacji (0,71), wskazujący na znaczną zbieżność wyników obu testów. Między wynikami badań skalą inteligencji Otisa i tą skalą zachodzi znaczna zbieżność, na którą wskazuje wysoki współczynnik korelacji (0,73), statystycznie bardzo istotny. Ci sami badacze informują, iż między wynikami badań skalą 2 a rezultatami badań dokonanych stanfordzkim testem osiągnięć szkolnych zachodzi zależność wyraźna, lecz mała, na którą wskazuje niski współczynnik korelacji (0,36), statystycznie istotny (Cattell i Cattell 1970, s. 5).

Badania Gabrieli Chojnackiej (1969) dokonane w Polsce wykazały, iż między skalą 2 a skalą pełną inteligencji D. Wechslera dla dzieci (WISC 1949) zachodzi zależność wyraźna, lecz mała (0,23) u 8-latków, zaś istotna (0,42) u dzieci 10-letnich. U 8-latków między wynikami badań tą skalą a średnią ocen z języka polskiego, matematyki i innych przedmiotów szkolnych zachodzi zależność wyraźna, lecz mała, ujemna (–0,27), zaś między średnią

ocen a skalą pełną D. Wechslera dla dzieci zachodzi zależność istotna (0,45). U 10-latków między skalą 2 a średnią ocen szkolnych zachodzi zależność wyraźna, lecz mała, na którą wskazuje niski współczynnik korelacji (0,29), statystycznie istotny. Tymczasem dla skali pełnej WISC uzyskano wysoki współczynnik korelacji (0,72). Stąd należy wnosić, iż w Polsce dla celów prognozy powodzenia szkolnego dzieci 8- i 10-letnich większe zastosowanie ma skala inteligencji D. Wechslera dla dzieci niż skala 2 R. B. Cattella i A. K. S. Cattell.

Badania J. Stachyry (1987, s. 114–115) dokonane omawianą skalą 2 oraz skalą inteligencji Wechsler-Bellevue wśród 32-osobowej grupy studentów pochodzenia arabskiego (0,446) oraz 32-osobowej grupy studentów polskich z Lublina (0,485) wykazały, iż między obu skalami zachodzi istotna zależność, na którą wskazują umiarkowane współczynniki korelacji, statystycznie istotne. Między wynikami badań skalą 2 a ocenami ze studiów zarówno wśród studentów pochodzenia arabskiego (0,354), jak i polskich (0,245) zachodzi zależność wyraźna, lecz mała, na którą wskazują niskie współczynniki korelacji. Tylko dla grupy studentów pochodzenia arabskiego współczynnik korelacji jest statystycznie istotny ($P = 0,05$). Między rezultatami badań skalą pełną Wechsler-Bellevue a ocenami ze studiów zarówno wśród studentów pochodzenia arabskiego (0,324), jak i polskich (0,294) zachodzi zależność wyraźna, lecz mała, na którą wskazują niskie współczynniki korelacji, zbliżające się do granicy istotności statystycznej.

Badania światowe wykazują, iż chłopcy nie różnią się od dziewcząt pod względem wyników osiąganych w badaniach tą skalą, tymczasem istnieją różnice w ilorazach inteligencji chłopców i dziewcząt badanych tradycyjnymi skalami inteligencji.

W. L. Fowler (1955) wykazał, iż nie ma statystycznie istotnych różnic między wynikami badań dzieci rasy białej i ich rówieśnikami – murzynami o takim samym statusie społecznym (Cattell i Cattell 1970).

R. B. Cattell konstruując tę skalę starał się tak dobrać zadania, by zminimalizować wpływ statusu społecznego na ich rozwiązanie. Jednak badania wykazują, iż istnieje współzależność między wynikami badań tą skalą a statusem społecznym, choć jest ona niska. R. B. Cattell i A. K. S. Cattell (1970) twierdzą, iż współzależność taka musi istnieć, bowiem w naszym społeczeństwie jednostki bardziej inteligentne starają się osiągnąć wyższy status społeczny w porównaniu z mniej inteligentnymi.

Skala 3. Służy ona do pomiaru inteligencji dzieci od 13 r. ż., młodzieży i dorosłych o wysokim poziomie inteligencji. Posiada dwie wersje A i B. Każdą wersję tworzą cztery testy: 1) serie, 2) klasyfikacje, 3) macierze i 4) zależności. Każda z wersji zawiera po 50 zadań. Czas zasadniczego badania obu wersjami wynosi 25 minut (bez uwzględniania czasu na wprowadzenie do każdego testu).

Normalizacji skali dokonano na podstawie wyników badań 3140 uczniów szkół średnich i dorosłych. Ponadto zbadano 330 osób w wieku od 16 do 80 lat. Badania te umożliwiły wprowadzenie poprawek na poszczególne poziomy wieku, zaczynając od 28 lat.

Współczynniki wiarygodności skali wahają się w granicach od 0,84 do 0,94. Współczynniki homogeniczności całej skali wynoszą dla trzech badanych grup: 0,82, 0,91 i 0,95. F. G. Rodd (1958) oszacował stabilność omawianej skali. Poddał badaniom młodzież z Taiwanu. Uzyskał on współczynnik stabilności 0,53; jest on niski. Szkoda, iż R. B. Cattell i A. K. S. Cattell (1970) nie podają odstępu czasu między obu badaniami; za nimi omawiam te wyniki badań.

G. Bajard (1955) wykazał, iż nasycenie czynnikiem „g” poszczególnych testów skali waha się w granicach od 0,71 do 0,84. R. B. Cattell i A. K. S. Cattell (1970, s. 16) podają następujące nasycenie poszczególnych testów tym czynnikiem: 1) serie – 0,53, 2) klasyfikacje – 0,68, 3) macierze – 0,89 i 4) zależności – 0,99. Badacze ci twierdzą, iż poprzednie wyniki są zaniżone, zaś ostatnie – zawyżone. R. B. Cattell (1971, s. 485) wykazał, iż obie wersje tej skali są w bardzo wysokim stopniu nasycone czynnikiem inteligencji płynnej (0,96).

G. J. Bansburg i W. Sloan (1955) opierając się na swoich badaniach doszli do wniosku, iż między wynikami badań dokonanych omawianą skalą a skalą inteligencji Stanford-Bineta, 1937, zachodzi znaczna zależność (0,85), na którą wskazuje wysoki współczynnik korelacji, statystycznie bardzo istotny ($P = 0,001$). J. R. Tilton (1949) wykazał, iż między tą skalą a wynikami badań skalą pełną Wechslera-Bellevue zachodzi znaczna zależność, na którą wskazuje wysoki współczynnik korelacji (0,84; cyt. za Cattell i Cattell 1970, s. 8).

Krystyna Głodek (1969) skorelowała wyniki badań dokonanych skalą 3 z rezultatami badań przeprowadzonych w Polsce skalą pełną Wechslera-Bellevue wśród 15- i 18-latków. Między ilorazami inteligencji uzyskanymi w badaniach obu skalami zachodzi zależność istotna zarówno u młodzieży 15-letniej (0,61), jak i 18-letniej (0,45), na którą wskazują umiarkowane współczynniki korelacji, statystycznie bardzo istotne ($P = 0,001$).

Badania R. B. Cattella i współpracowników (1962, s. 34) dokonane wśród grupy uczniów z kl. VII i VIII wykazały, iż między wynikami badań omawianą skalą a Stanfordskim testem osiągnięć szkolnych zachodzi zależność istotna (0,54; $P = 0,01$) u uczniów mieszkających w okolicy Springfield, zaś wyraźna, lecz mała (0,33), ale statystycznie istotna ($P = 0,01$) u uczniów z Paxton. Badania wykazały również, iż między wynikami badań skalą 3 a osiągnięciami z języka chińskiego zachodzi zależność wyraźna, lecz mała (0,30) u młodzieży chińskiej. Natomiast u młodzieży amerykańskiej istnieje zależność istotna (0,40) między rezultatami badań tą skalą a osiągnięciami

z języka angielskiego. Zarówno u młodzieży chińskiej (0,47), jak i amerykańskiej (0,64) zachodzi istotna zależność między wynikami badań skalą 3 a ich osiągnięciami z zakresu matematyki. Wskazują na nią umiarkowane współczynniki korelacji, statystycznie istotne (Cattell 1971, s. 485).

Badania Krystyny Głodek (1969, s. 46) wykazały, iż zarówno u polskich 15-letnich licealistów (0,65), jak i u 18-letnich licealistów (0,54) zachodzi istotna zależność między wynikami badań omawianą techniką a osiągnięciami szkolnymi, mierzonymi średnią oceną ze wszystkich przedmiotów nauczania.

Dotychczasowe badania nie ujawniły statystycznie istotnych różnic między średnimi ilorazami inteligencji uzyskanymi przez dziewczęta i chłopców w badaniach tą skalą (Cattell i Cattell 1970, s. 14).

Badania dokonane skalą 3 wśród młodzieży brytyjskiej, francuskiej i tajwańskiej wykazały, iż nie różni się ona statystycznie od młodzieży amerykańskiej pod względem uzyskiwanych średnich ilorazów inteligencji (Cattell i Cattell 1970, s. 34 oraz Cattell 1971, s. 485).

Wyniki badań porównawczych dokonanych tą skalą, jak też skalą inteligencji D. Wechslera dla dzieci (WISC-R 1974) wśród młodzieży niedostosowanej społecznie, wśród której była młodzież reprezentująca rasę białą, czarną i Meksykańczyków amerykańskich, doprowadziły do wniosku, iż skala „culture-fair jest mniej zależna od uwarunkowań kulturowych niż Skala inteligencji D. Wechslera dla dzieci” (zob. McShane i Cook 1985, s. 771).

4.2. R. G. Leitera Międzynarodowa skala wykonawcza

Skala ta służy do pomiaru inteligencji dzieci i młodzieży w wieku od 2 do 18 lat. Oryginalna wersja skali z 1948 r. składa się z 54 testów wykonawczych. Współczynniki rzetelności wahają się w granicach od 0,80 do 0,90. Trafność tej skali oszacowano poprzez korelację wyników badań grupy dzieci tą skalą z rezultatami badań dokonanych skalą inteligencji Stanford-Bineta i D. Wechslera dla dzieci (WISC 1949). Współczynniki trafności wahają się w granicach od 0,56 do 0,92; średni 0,83. M. E. Tata (1952) w wyniku badań porównawczych doszedł do wniosku, iż ilorazy inteligencji uzyskane w wyniku badań tą skalą nie są mniej zależne od uwarunkowań kulturowych niż w skali Stanford-Bineta (zob. Sattler 1982, s. 250).

W Polsce dokonano adaptacji tej skali i jej normalizacji. Służy ona w Polsce do badań dzieci i młodzieży w wieku od 4 do 18 lat, niesłyszącej. Dane dotyczące rzetelności i trafności tej skali w zastosowaniu do badań dzieci niesłyszących zawiera *Tymczasowy podręcznik...* (Markowska 1977).

4.3. Testy J. C. Ravena

Służą one do pomiaru zdolności spostrzegania i poprawnego myślenia. Znane są następujące wersje testów: Kolorowe matryce (Coloured Progressive Matrices, serie: A, Ab i B, 1947, 1956, 1971); stosuje się do badań dzieci w wieku od 5 do 11 lat i osób w wieku starszym od 65 do 85 lat; mają one dwie postacie: wkładankową, używaną do indywidualnych badań dzieci od 3 r.ż., i zeszytową, do badań indywidualnych i grupowych dzieci od 5 r.ż.; standardowe matryce (Standard Progressive Matrices, serie A, B, C, D i F, 1938, 1956, 1987) służą do badań dzieci, młodzieży i dorosłych oraz matryce dla zaawansowanych, seria I i II (1941, 1962, 1970) służą do badań dzieci, młodzieży i dorosłych o wyższym niż przeciętnym poziomie inteligencji.

Podstawy teoretyczne konstrukcji tych testów omówił B. Hornowski (1970), stąd też pomijam je w tym miejscu. Współczynniki rzetelności poszczególnych testów J. C. Ravena wahają się w granicach od 0,71 do 0,92; są one niższe u najmłodszych dzieci. Współczynniki korelacji wyników badań tymi testami ze skalą inteligencji Stanford-Bineta wahają się od 0,55 do 0,86 (zob. Sattler 1982, s. 249). Współczynniki korelacji wyników badań standardowymi matrycami z rezultatami badań dokonanych skalą pełną dla dzieci wahają się od 0,704 do 0,754 (zob. Barratt 1956, Kostrzewski i Biela 1970). Pomędzy testem standardowych matryc a skalą inteligencji Otisa istnieje zależność istotna, na którą wskazują umiarkowane współczynniki korelacji (0,53 u chłopców i 0,58 u dziewcząt; Hornowski 1970, s. 254).

Zgodnie z przyjętym założeniem skonstruowane przez J. C. Ravena testy służą do pomiaru percepcji i jasnego myślenia (Raven 1971, s. 3). J. H. Court i J. Raven (1977, 1985) wykazali, iż testy te należy zaliczyć do najczystszych miar czynnika „g” (Raven 1985, s. 25). P. E. Vernon (1949) wykazał, iż nasycenie czynnikiem ogólnym standardowych matryc wynosi 0,79; oprócz czynnika „g” wystąpiły jeszcze dwa czynniki grupowe: wyobraźni przestrzennej i praktyczno-mechaniczny (cyt. za Hornowski 1970, s. 69). Niektóre późniejsze analizy czynnikowe wykazały, iż test ten mierzy: 1) myślenie konkretne i abstrakcyjne, 2) uzupełnianie wzorów, ciągłych i nieciągłych i 3) percepcję wzorów (cyt. za Sattler 1982, s. 247). Słusznie się dziś podkreśla, iż omawiane wersje testu J. C. Ravena nie mogą zastąpić badań dokonywanych takimi skalami inteligencji, jak skala Stanford-Bineta i D. Wechslera, nie tylko z tych racji, iż mierzą one mniej zmiennych, ale również dlatego, iż są mniej rzetelne i trafne (zob. Sattler 1982, s. 248).

Warto wspomnieć, iż zarówno kolorowe matryce (Szustrowa i Jaworowska 1992 oraz Woynarowska 1991), standardowe matryce (Jaworowska i Szustrowa 1991) oraz matryce dla zaawansowanych (Jaworowska i Szustrowa 1991)

zostały w Polsce znormalizowane. Oszacowano ich rzetelność i trafność w zastosowaniu do badań naszych dzieci, młodzieży i dorosłych.

W tym miejscu omówię niektóre badania dotyczące trafności prognostycznej poszczególnych wersji matryc J. C. Ravena.

T. Woynarowska (1991, s. 28) skorelowała wyniki badań 1579 dzieci w wieku od 8 do 10–11 lat dokonanych testem – kolorowe matryce J. C. Ravena z ich ocenami szkolnymi z języka polskiego i matematyki, jak też ze średnią oceną z tych dwóch przedmiotów. Uzyskane współczynniki korelacji wahają się w granicach od 0,25 do 0,42; są one statystycznie istotne ($P = 0,01$).

A. Jaworowska i T. Szustrowa (1991, s. 29–30) skorelowały z kolei wyniki badań dokonanych testem standardowych matryc J. C. Ravena 4006 dzieci w wieku od 5;0 do 15;11 lat z ocenami szkolnymi. U dzieci szkolnych uwzględniono średnią ocenę ze wszystkich przedmiotów, średnią ocenę z przedmiotów podstawowych (bez plastyki, wychowania fizycznego, wychowania technicznego, wychowania muzycznego i zachowania), średnią ocen z grup przedmiotów humanistycznych, matematycznych i przyrodniczych (w starszych klasach) oraz oceny z każdego przedmiotu nauczania. Uzyskane współczynniki korelacji najczęściej oscylowały wokół wartości 0,3–0,4, co oznacza, iż tylko około 10–15% (a w najlepszym przypadku 25%) zmienności sukcesów szkolnych, mierzonych ocenami szkolnymi, można wyjaśnić zdolnościami mierzonymi standardowymi matrycami.

T. Stawarz (1984) uzyskała następujące współczynniki korelacji u 260 uczniów kl. VIII między wynikami badań omawianą techniką a ocenami z poszczególnych przedmiotów: 0,16 ($P = 0,01$) geografia, 0,23 ($P = 0,001$) fizyka, 0,24 ($P = 0,001$) język polski i chemia, 0,28 ($P = 0,001$) biologia i 0,29 ($P = 0,001$) matematyka.

A. Jaworowska i T. Szustrowa (1991, s. 24) skorelowały wyniki badań dokonanych testem matryc dla zaawansowanych J. C. Ravena 1402 osób w wieku od 12;6 do 19;5 lat z ich ocenami szkolnymi z języka polskiego, matematyki i średnią oceną z czterech przedmiotów (m. in. z fizyki, biologii, geografii, historii lub innych). Analiza uzyskanych współczynników korelacji wykazuje, iż najbardziej wyraźny jest związek między wynikami badań omawianym testem a ocenami młodzieży z kl. VII i VIII szkół podstawowych i klas I szkół średnich. Uzyskane współczynniki korelacji są niskie lub bardzo niskie. Te współczynniki, które oscylują wokół 0,3, oznaczają, że tylko około 10–15% wariacji ocen szkolnych można wyjaśnić tymi zdolnościami, które mierzy ten test.

J. Raven (1985, s. 28) stwierdza, iż omawiane testy percepcyjne pozwalają w większym stopniu przewidzieć powodzenie życiowe człowieka niż powodzenie w nauce szkolnej.

Badania na ogół nie wykazują statystycznie istotnych różnic między średnimi uzyskanymi przez dziewczęta i chłopców w teście J. C. Ravena.

Testy J. C. Ravena są stosowane do badań osób reprezentujących różne rasy, kultury i subkultury. Badano nimi osoby niesłyszące, upośledzone umysłowo, z zaburzeniami mowy, z mózgowym porażeniem dziecięcym itp.

Badania wykazują, iż na wyniki badań dokonanych tymi testami wpływa środowisko społeczno-ekonomiczne, z którego pochodzi badany, ale wyjaśnia ono stosunkowo mały procent wariacji wyników, natomiast mobilność społeczna w dużym stopniu wpływa na wyniki; w efekcie istnieje istotny związek między wynikami badań omawianymi testami a osiągniętym statusem społeczno-ekonomicznym (Foulds i Raven 1950, cyt. za Raven 1985, s. 28).

4.4. Florence Goodenough i D. Harrisa test postaci ludzkiej

Test postaci ludzkiej Florence Goodenough i D. Harrisa (1963) służy do pomiaru dojrzałości umysłowej dzieci w wieku od 3; 0 do 15; 11 lat. Angażuje on zdolność percepcji (różnicowanie: podobny – różny), abstrakcji (klasyfikacja przedmiotów) oraz uogólniania. D. B. Harris w udoskonalonej przez siebie wersji testu wprowadził jako alternatywną wersję do rysunku mężczyzny – rysunek kobiety i rysunek samego siebie przez dziecko. Ponadto zobiektywizował on system oceny. Współczynniki równoważności oryginalnej wersji: rysunek mężczyzny i rysunek kobiety wahają się od 0,72 do 0,90. Współczynniki stabilności wahają się od 0,50 do 0,70, zaś współczynniki rzetelności konsensusyjnej: od 0,80 do 0,90. Współczynniki korelacji wyników badań oryginalną wersją omawianego testu z rezultatami badań dokonanych różnymi skalami pomiaru inteligencji wahają się w granicach od 0,24 do 0,88; średni współczynnik korelacji wynosi 0,49 dla dzieci białych. Dla różnych grup etnicznych współczynniki te wahają się w granicach od 0,22 do 0,85; średni 0,57 (zob. Sattler 1982, s. 249–250).

Badania prowadzone w całym świecie przy zastosowaniu tego testu wykazały, iż wyniki zależą od uwarunkowań kulturowych w mniejszym stopniu niż tradycyjnych testów inteligencji. Różnice w wynikach badań ujawniają się w tych kulturach, które używają znacznie odmiennych ubiorów i które akcentują pewne części ciała. Test ten powinien być restandaryzowany dla tej grupy etnicznej, która istotnie różni się od pozostałych sposobem ubierania się (Sattler 1982, s. 250). Zalecenia te ograniczają użyteczność omawianej techniki do celów badań międzykulturowych. Słusznie się podkreśla, iż test ten jest niewerbalnym miernikiem zdolności poznawczych i nie może być stosowany jako jedyna technika pomiaru inteligencji dziecka (Sattler 1982, s. 250).

W Polsce test ten został znormalizowany na próbkę obejmującej 1488 dzieci w wieku od 6 do 14 lat (Hornowska i Paluchowski 1987).

4.5. Allison Davis i E. Fellsa test inteligencji ogólnej

Test ten służy do pomiaru zdolności rozwiązywania problemów przez dzieci szkolne z kl. I–VI. Skonstruowano go w tym celu, by móc zmierzyć inteligencję możliwie wolną od uwarunkowań społeczno-ekonomicznych. Został on opublikowany w 1953 r. Zawiera on szereg rysunków ilustrujących różne problemy, które należy poprawnie rozwiązać.

Współczynnik stabilności testu (powtórne badania po upływie 2 tygodni) wahają się w granicach od 0,72 do 0,90 u uczniów kl. od II do IV; u uczniów z kl. I – 0,68. Między wynikami badań tym testem a ilorazami inteligencji uzyskanymi w wyniku badań tradycyjnymi skalami inteligencji zachodzi istotna zależność, na którą wskazują umiarkowane współczynniki korelacji. Trafność prognostyczna tego testu jest mniejsza niż tradycyjnych skal inteligencji, stąd też jest on rzadko stosowany (Anastasi 1964, s. 267).

Badania prowadzone tą techniką wykazały, iż dzieci pochodzące z gorszych warunków społeczno-ekonomicznych uzyskują niższe ilorazy inteligencji, w porównaniu z rówieśnikami pochodzącymi z korzystnych warunków. Testy te nie spełniły oczekiwań (Anastasi 1964, s. 268).

4.6. Semantyczny test inteligencji

Służy on do pomiaru inteligencji młodzieży kierowanej do wojska. Instrukcje są podawane przy pomocy pantomimiki. Zbyt mało jest w piśmiennictwie informacji dotyczących rzetelności i trafności tego testu, jak też jego zastosowań, by móc ocenić jego wartość.

4.7. Safrana test inteligencji o ograniczonym uwarunkowaniu kulturowym (Safran Culture Reduced Intelligence Test, 1960–1969)

Służy on do badań uczniów kl. I–VI. Zbyt mało jest informacji w piśmiennictwie światowym o tym teście, by móc obiektywnie ocenić, czy rzeczywiście różne czynniki społeczno-ekonomiczne i kulturowe mają mniejszy wpływ na wyniki badań dokonanych tym testem, w porównaniu z tradycyjnymi testami inteligencji.

5. ZAKOŃCZENIE

Testy inteligencji „culture-free” i/lub „culture-fair” oraz „culture-reduced” były próbą odpowiedzi na zapotrzebowanie społeczne posiadania techniki pomiaru inteligencji osób posługujących się różnymi językami, pochodzącymi z różnych kręgów kulturowych, wychowywanych w różnych warunkach środowiskowych, niekiedy bardzo niekorzystnych dla rozwoju jednostki, które nie szykanowałyby mniejszości narodowych (np. dzieci murzyńskich, indiańskich, meksykańskich, hiszpańskich i innych mieszkających w USA, nie znających dobrze języka angielskiego), dzieci pochodzących ze wsi itp. Testy te wychodziły również naprzeciw potrzebom poznawczym badaczy; pragnęli oni posiadać obiektywne, rzetelne i trafne narzędzie pomiaru inteligencji osób reprezentujących różne kręgi kulturowe i subkulturowe.

Czy skonstruowane dotąd testy inteligencji „culture-fair” i/lub „culture-reduced” spełniły oczekiwania? Można dzisiaj odpowiedzieć, iż żaden ze skonstruowanych dotąd testów inteligencji „culture-free” i/lub „culture-fair” oraz „culture-reduced” nie zaspokoił w pełni oczekiwań badaczy i praktyków.

Warto w tym miejscu zatrzymać się chwilę nad terminologią. Termin „culture-free” dosłownie oznacza „wolny od uwarunkowań kulturowych”, zaś „culture-fair”, sprawiedliwy, bez skazy, czysty, bezstronny. Test inteligencji „culture-free” i/lub „culture-fair” byłby to taki test, w dosłownym rozumieniu obu tych terminów, na którego wyniki nie mają istotnego wpływu żadne uwarunkowania kulturowe i/lub subkulturowe.

Czy mogą istnieć testy inteligencji całkowicie wolne od uwarunkowań kulturowych? Odpowiedź jest jednoznacznie negatywna. Każde dziecko ludzkie przychodzi na świat w określonym miejscu danego kraju, który należy do określonego kręgu kulturowego, posiada określony poziom rozwoju kultury materialnej i duchowej. Jego rodzice (a w razie ich braku – rodzice zastępczy lub wychowawcy) wcześniej w toku swego rozwoju opanowali określony język, przyjęli określone wzorce zachowań, normy społeczno-moralne oraz zdobyli określony poziom wykształcenia (lub nie zdobyli) w ramach danej kultury i subkultury. I to wszystko przekazują dziecku. Wiemy, iż celem wychowania i nauczania jest przekazywanie dziecku dorobku kulturowego; najczęściej bywa przekazywany dorobek kulturowy tego kręgu kulturowego, w obrębie którego znajduje się kraj, w którym mieszkają rodzice wraz z dzieckiem.

Żaden test – ze swojej istoty – nie może być wolny od uwarunkowań kulturowych, bowiem żadna osoba nie działa w próżni kulturowej. Jak słusznie stwierdza S. Scarr (1978), testy inteligencji nie są nimi w pewien abstrakcyjny, wolny od uwarunkowań kulturowych sposób. One mierzą zdolność funkcjonowania intelektualnego jednostki za pomocą wiedzy

i umiejętności nabytych w tej kulturze, z której pochodzą (Sattler 1982, s. 339). Dobre testy muszą mierzyć reprezentatywną próbkę inteligentnych zachowań określonych jednostek w danym środowisku. Twórcy testów, konstruując testy inteligencji „culture-free” i/lub „culture-fair” w rzeczywistości nie eliminują elementów kulturowych, ale uwzględniają tylko te, które są – ich zdaniem – **wspólne** – wszystkim istniejącym kulturom. Istnieje jednak problem, czy są one istotne wspólne wszystkim, czy też znanym im kulturom? Stąd wniosek, iż żaden test nie jest wolny od uwarunkowań kulturowych. Ponieważ twórcy testów „culture-fair” uwzględniają w swoich testach inteligencji treści wspólne wielu kulturom, stąd też **minimalizują** wpływ określonej kultury na ich rozwiązanie.

Nie należy więc używać nieadekwatnej nazwy na określenie omawianych testów inteligencji, tj. testy inteligencji „culture-free” i/lub „culture-fair”. Należy posługiwać się poprawną nazwą: testy inteligencji „culture-reduced”, są to bowiem takie testy, w których zminimalizowano w mniejszym lub większym stopniu wpływ **niektórych** uwarunkowań kulturowych. Przy ocenie wartości poszczególnych testów inteligencji „culture-reduced” należy zwrócić uwagę na to, wpływ jakich zmiennych kulturowych i w jakim stopniu został zminimalizowany, dzięki odpowiedniej ich konstrukcji. I tak np. oceniając wartość omówionych w tym artykule testów „culture-fair”, należy podkreślić z uznaniem, iż J. C. Ravenowi i R. B. Cattellowi udało się wyeliminować wpływ płci na wyniki badań, miejsce zamieszkania (wieś–miasto, przy jednakowym wykształceniu badanych osób), kraju zamieszkania (jak np. Francja, Anglia, Japonia, Chiny, USA itp.).

Bardzo krytyczny sąd o testach inteligencji „culture-free” i/lub „fair” wypowiadają G. P. Robb, L. C. Bernardoni i R. W. Johnson (1972, s. 93). Twierdzą oni, iż stosowanie testów inteligencji w stosunku do innej grupy niż klasa średnia rasy białej stwarza problemy, które do dzisiaj nie zostały pomyślnie rozwiązane. Poronioną próbą rozwiązania tego problemu – ich zdaniem – była konstrukcja i stosowanie tzw. testów inteligencji „culture-fair” lub „culture-free”. W testach tych posługiwano się zadaniami niewerbalnymi, takimi jak rysunki geometryczne, do pomiaru inteligencji. Ich uwagi krytyczne dotyczą tego, iż testy tego typu: 1) mierzą tylko bardzo wąski aspekt inteligencji, 2) nie są fair pod względem kulturowym dla więcej niż kilku określonych kultur, 3) mają ograniczone zastosowanie praktyczne. Mają ograniczoną trafność prognostyczną, jak też nie dają użytecznych informacji o badanej osobie. Twierdzą oni, iż często kulturowo nasycone testy, które były całkiem niedostosowane do danej populacji miały lepszą trafność prognostyczną, tj. umożliwiały w większym stopniu przewidywanie osiągnięć szkolnych. Przed twórcami testów „culture-fair” stały kilka trudnych problemów wymagających rozwiązania, a mianowicie: 1) trudno jest dobrać reprezentatywną próbkę; 2) takie czynniki jak klasa społeczna, środowisko

wychowawcze, które wpływają na wykonanie testu, są trudne do kontroli; 3) większość naszych testów skonstruowano do badań rasy białej, pochodzących z klasy średniej; oparte są one również na definicji inteligencji charakterystycznej dla tej kultury; 4) motywacja wykonania zadań tworzących dany test może być różna u innych grup kulturowych; 5) trudności językowe często obniżają trafność testu.

Surowy również sąd pod adresem testów „culture-free” wypowiada wspomniany na początku tego artykułu M. J. Begab (1973, s. X). Twierdzi on, iż żadna z prób nie wykazała, iż testy te są mniej stronnicze w zastosowaniu do badań dzieci umysłowo upośledzonych w porównaniu z tradycyjnymi testami pomiaru inteligencji, bowiem dzieci pochodzące z niekorzystnych środowisk uzyskiwały w badaniach tymi testami niższe wyniki niż w testach tradycyjnych. Słusznie on twierdzi, iż wszystkie testy psychologiczne mierzą próbkę określonych zachowań, na które wpływa kultura, w jakiej jednostka kształtuje się w toku swego życia. Uważa on, iż braki, które ujawniają się podczas rozwiązywania przez badanego tradycyjnych testów, są często dostrzegane przez otoczenie podczas nauki szkolnej, podczas pracy oraz w różnych sytuacjach życiowych, a taka wiedza jest niezbędna w celu udzielenia kompetentnej pomocy. Sądzi on, iż zwalczanie tradycyjnych testów inteligencji z racji ich uwarunkowań kulturowych jest ukrywaniem skutków zaniedbań kulturowych u określonych osób, zaś taki sposób postępowania opóźnia akcję pomocy dzieciom pochodzącym z niekorzystnych środowisk oraz opóźnia akcję rozwiązywania ważkich problemów społecznych. Jego zdaniem, te rzekome ograniczenia tradycyjnych testów inteligencji są w rzeczywistości dużą ich zaletą; umożliwiają bowiem one ocenę wykonania określonych testów przez dziecko ukształtowane w danej kulturze. Słusznie on podkreśla, iż stosowanie określonych testów w różnych kulturach bez ich poprawnej normalizacji na reprezentatywnej dla danej kultury próbce populacji prowadzi zazwyczaj do poważnych błędów w diagnozie indywidualnych przypadków i podnosi wskaźnik dzieci, u których błędnie rozpoznaje się upośledzenie umysłowe.

Szkoda, iż ta surowa ocena testów inteligencji „culture-fair” nie została wypowiedziana odnośnie do konkretnych testów stosowanych do badań dzieci umysłowo upośledzonych. Tej oceny nie można uogólnić na wszystkie testy „culture-reduced”.

Czy testy inteligencji culture-reduced powinny być stosowane do badań czy też nie? Odpowiedź jest pozytywna, ale stosujący je powinien wiedzieć, w jakim celu ma dany test zastosować, znać wartość konkretnego testu, jego ograniczenia, jak też to, czy pozwoli na rozwiązanie wysuniętego problemu. Nie są to testy uniwersalne, dobre we wszystkich sytuacjach. Takich testów nie ma. A zatem testy te powinny być stosowane w sposób odpowiedzialny ze świadomością ich ograniczeń.

BIBLIOGRAFIA

- Anastasi A. (1964), *Psychological Testing*, The Macmillan Co., New York
- Barrat F. S. (1956), *The Relationship of the Progressive Matrices (1938) and the Columbia Mental Maturity Scale to the WISC*, „Journal of Consulting Psychology”, 20, 294–296
- Begab M. J. (1973), *Preface*, [w:] *Manual on Terminology and Classification in Mental Retardation 1973 Revision*. American Association on Mental Deficiency, ed. H. J. Grossman et al., Baltimore
- Binet A. et Simon Th. (1905), *Methodes nouvelles pour le diagnostic du niveau intellectuel des anormaux*, „L'Annee Psychologique”, XI, 191–244
- Binet A. et Simon Th. (1908), *Le développement de l'intelligence chez les enfants*, „L'Annee Psychologique”, 14, 1–94
- Cattell R. B. (1940), *A Culture-Free Intelligence Test I*, „The Journal of Educational Psychology”, 31, 161–179
- Cattell R. B. (1941), *Some Theoretical Issue in Adult Intelligence Testing*, „Psychological Bulletin”, 38, 592 i n.
- Cattell R. B. (1943), *The Measurement of Adult Intelligence*, „Psychological Bulletin”, 3, 153–193
- Cattell R. B. (1957), *Personality and Motivation Structure and Measurement*. World Book Company. Yonkers-on-Hudson, New York
- Cattell R. B. (1963), *Theory of Fluid and Crystallized Intelligence. A Critical Experiment*, „The Journal of Educational Psychology”, 54, 1–22
- Cattell R. B. (1965), *The Scientific Analysis of Personality*, Baltimore, Penguin Books
- Cattell R. B. (1950, 1970), *Handbook for the Individual or Group Culture-Fair (or Free) Intelligence Test (A Measure of „g”). Scale 1. Children of 4 to 8 Years and Mentally Defective Adults*, IPAT, Champaign, Ill.
- Cattell R. B. (1971), *Abilities, Their Structure, Growth and Action*, Houghton Mifflin Co., Boston
- Cattell R. B. and Bristol H. (1933), *Intelligence Tests for Mental Ages of Four to Eight Years*, „The British Journal of Educational Psychology”, 3, 142–169
- Cattell R. B., Butcher J., Connor D., Sewney A. B. and Tsujioka B. (1962), *Prediction and Understanding of the Effect of Children's Interest upon School Performance*, University of Illinois, Urbana
- Cattell R. B., and Cattell A. K. S., (1949, 1957, 1960, 1970), *Culture Fair Intelligence Test. Scale 2. Forms A and B (Children 8–13 Years and Average Unselected Adults)*, IPAT, Champaign, Ill.
- Cattell R. B. and Cattell A. K. S., (1959, 1970), *Culture Fair Intelligence Test. Scale 3. Forms A and B (High School Pupils and Adults of Superior Intelligence)*, IPAT, Champaign, Ill., 1959, 1970
- Chojnacka G. (1969), *Problem trafności diagnostycznej testu inteligencji abstrakcyjnej (Culture Fair Intelligence Test) R. B. Cattella. Badania eksperymentalne dzieci w wieku 8 i 10 lat*, niepublikowana praca magisterska, Archiwum Biblioteki Głównej KUL, Lublin
- Chojnacka-Szałowska G. (1987), *Sprawność umysłowa dzieci i młodzieży z zaburzeniami motorycznymi*, Gdańsk
- Choynowski M., (1977), *Podręcznik do 'Krótkiej Skali Inteligencji'*, [w:] *Testy psychologiczne w poradnictwie wychowawczo-zawodowym*, Warszawa, 8–96
- Court J. H. and Raven J. (1977), *Reliability, Validity and Normative Studies a Carried out Using Raven's Progressive Matrices and Mill Hill Vocabulary Scales*, [w:] J. C. Raven, J. H. Court and J. Raven, *A Manual for Raven's Progressive Matrices and Mill Hill Vocabulary*, H. K. Lewis, London

- Court J. H. and Raven J. (1982), *Additional Reliability, Validity and Normative Studies Carried out Using Raven's Tests: 1982 Update. Research Supplement No 2 to the Manual for Raven's Progressive Matrices and Mill Hill Vocabulary Scales*, H. K. Lewis and Co., Ltd., London
- Davis A. and Fells K. (1953), *Davis-Fells Games*. Yonkers, World Book, New York
- English H. B. and English A. B. (1958), *A Comprehensive Dictionary of Psychological and Psychoanalytical Terms. A Guide Usage*, Longmans Green and Co., New York
- Głodek K. (1969), *Problem trafności diagnostycznej Testu Inteligencji Abstrakcyjnej (Culture Fair Intelligence test) R. B. Cattella. Badania eksperymentalne młodzieży w wieku 15 i 17 lat*, niepublikowana praca magisterska, Archiwum Bibl. Głównej KUL, Lublin
- Goddard H. H. (1908), *The Binet and Simon Tests of Intellectual Capacity*, „Training School”, 5, 3–9
- Goddard H. H. (1911), *The Binet-Simon Measuring Scale for Intelligence. Revised*, „Training School”, 8, 56–62
- Harris D. B. (1963), *Children's Drawings as Measures of Intellectual Maturity: A Revision and Extension of the Goodenough Draw-a-Man Test*, Haracourt, Brace and World, New York
- Horn J. L. and Cattell R. B. (1966), *Refinement and Test of the Theory of Fluid and Crystallized General Intelligence*, „The Journal of Educational Psychology”, 57, 253 i n.
- Hornowska E. Paluchowski W. J. (1987), *Rysunek postaci ludzkiej według Goodenough-Harrisa (DAMT)*, Poznań
- Hornowski B. (1970), *Analiza psychologiczna skali J. C. Ravena*, Warszawa
- Horrocks J. E. (1964), *Assessment of Behavior*, Charles E. Merrill Books, Inc., Columbus, Ohio
- Jaworowska A. Szustrowa T. (1991), *Podręcznik do Testu Matryc Ravena. Wersja Standard (1956). Polska standaryzacja 1989 (5; 11–15; 11)*, Warszawa
- Jaworowska A. Szustrowa T. (1991), *Podręcznik do Testu Matryc Ravena. Wersja dla zaawansowanych (1962). Polska standaryzacja: uczniowie 12; 6–19; 5 (1991), studenci (1990)*, Warszawa
- Knox H. A. (1914), *A Scale Based on the Work of Ellis Island for Estimating Mental Defect*, „Journal of American Med. Assoc.”, 62, 741–747
- Kostrzewski J. (1990), *Testy inteligencji culture-fair*, [w:] *Kulturowa adaptacja testów*, red. Ciechanowicz A., Warszawa, 49–58
- Kostrzewski J. i Biela F. (1970), *Problem współzależności standardowych matryc J. C. Ravena oraz skali Wechslera do badania inteligencji dzieci*, „Przegląd Psychologiczny”, 19, 81–94
- Kuhlmann F. (1912), *A Revision of the Binet-Simon System for Measuring the Intelligence of Children*, „Journal of Psycho-Asthenics”, Monograph Supplement, 1, 1–41
- Leiter R. G. (1948), *Leiter International Performance Scale*, Stoelting, Co., Chicago
- Leiter R. G. (1959), *Part I of the Manual for the 1948 Revision of the Leiter International Performance Scale: Evidence of the Reliability and Validity of the Leiter Tests*, „Psychological Service Center Journal”, 11, 1–72
- Lewicka E. (1981), *Inteligencja płynna i skryształizowana. Koncepcja R. B. Cattella*, „Psychologia Wychowawcza”, 24(38), 2, 223–229
- Line W. (1991), *The Growth of Visual Perception in Children*, „The British Journal of Psychol.”, Monograph Supplement, 15
- Markowska B. (1977), *Tymczasowy podręcznik do polskiej adaptacji skali R. G. Leitera dla dzieci głuchych*, Warszawa
- Martin A. W. and Wiechers J. E. (1954), *Raven's Coloured Progressive Matrices and the Wechsler Intelligence Scale for Children*, „Journal of Consulting Psychology”, 18, 143–145
- McShane D. and Cook V. (1985), *Transcultural Intellectual Assessment*, [w:] *Handbook of Intelligence. Theories, Measurements and Applications*, ed. B. B. Wolman, John Wiley & Sons, New York, 737–785

- Penrose L. S. and Raven J. C. (1936), *A New Series of Perceptual Tests: Preliminary Communication*, „The British Journal of Medical Psychology”, 18, 1
- Piotrowska A. (1988), *Dwa rodzaje inteligencji ogólnej i sposoby ich pomiaru*, „Psychologia Wychowawcza”, 31(45), 3, 347–355
- Porteus S. D. (1918), *The Measure of Intelligence 653 Children Examined by the Binet and Porteus Tests*, „Journal of Educational Psychology”, 9, 13–31
- Porteus S. D. (1924), *Guide to the Porteus Maze Tests*. Vineland, New York Training School
- Raven J. C. (1938), *Progressive Matrices*. Lewis H. K., London
- Raven J. C. (1947, 1970), *Advanced Progressive Matrices*. H. K. Lewis, London
- Raven J. C. (1947, 1971), *Coloured Progressive Matrices*. H. K. Lewis & Co., Ltd. London
- Raven J. (1985), *Raven's Progressive Matrices: Their Origin and Contribution to Society*, „Psychodiagnostyka w socjalistycznych stranach”, 1(IX), 27–36
- Robb C. P., Bernardoni L. C. and Johnson R. W. (1972), *Assessment of Individual Mental Ability*. Intext Educational Publishers, Scranton
- Sattler J. M. (1982), *Assessment of Children's Intelligence and Special Abilities*, Second Edition, Allyn and Bacon, Inc., Boston
- Sattler J. M. (1988), *Assessment of Children*, Third Edition, Jerome M. Sattler, Publisher, San Diego
- Spearman C. (1933), *The Spearman Visual Perception Test*, London
- Spearman C. (1927), *The Abilities of Man*, Macmillan, New York
- Stachyra J. (1987), *Inteligencja i jej pomiar w świetle teorii Cattella*, [w:] *Z badań nad zdolnościami i uzdolnieniami specjalnymi młodzieży*, red. S. Popek, Lublin, 105–116
- Stawarz T. (1984), *Psychologiczne i społeczne uwarunkowania osiągnięć szkolnych uczniów klasy ósmej*, nieopublikowana praca doktorska, Archiwum Biblioteki Głównej UŁ, Łódź
- Szustrowa T. i Jaworowska A. (1992), *Podręcznik do Testu Matryc Ravena. Wersja kolorowa (1956). Polska standaryzacja, 1991 (3; 11–9; 11)*, Warszawa
- Szymański M. (1992), *Rzetelność i trafność R. B. Cattella „Culture Fair Intelligence Test (Skala 1)” w zastosowaniu do badania dzieci 8-letnich*, nieopublikowana praca magisterska, Archiwum Katedry Psychologii UŁ, Łódź
- Terman L. M. (1916), *The Measurement of Intelligence*, Houghton Mifflin Co., Boston
- Terman L. M. and Merrill M. A. (1937), *Measuring Intelligence*, Houghton Mifflin Co., Boston
- Terman L. M. and Merrill M. A. (1960), *Stanford-Binet Intelligence Scale. Manual for the Third Revision. Form L–M*, Houghton Mifflin Co., Boston
- Terman L. M. and Merrill Maud A. (1973), *Stanford-Binet Intelligence Scale. Manual for the Third Revision. Form L–M. 1972 Tables of Norms by R. L. Thorndike*, Houghton Mifflin Co., Boston
- Thorndike R. L., Hagen E. P., Sattler J. M. (1986), *The Stanford-Binet Intelligence Scale. Fourth Edition*, The Riverside Publishing Company, Chicago
- Wechsler D. (1949), *Manual for the Wechsler Intelligence Scale for Children*. Psychological Corporation, New York
- Wechsler D. (1955), *Manual for the Wechsler Adult Intelligence Scale*. Psychological Corporation, New York
- Wechsler D. (1967), *Manual for the Wechsler Preschool and Primary Scale of Intelligence*. Psychological Corporation, New York
- Wojnarowska T. (1991), *Test Barwnych Matryc J. C. Ravena. Wyniki badań dzieci polskich*, Warszawa
- Yerkes M., Bridges J. W. and Hardwick R. S. (1915), *A New Point Scale for Measuring Mental Ability*, Warwick and York, Baltimore
- Wechsler D. (1939), *The Measurement and Appraisal of Adult Intelligence*, The Williams & Wilkins Co., Baltimore

Janusz Kostrzewski

**CULTURE FAIR INTELLIGENCE TESTS
(The Origin, Foundations, Types and Evaluation)**

The Autor presents in this article the origin and history of the culture-fair intelligence tests and theoretical foundations they are based on. The results of the world's studies including the Polish ones on the reliability and validity of such tests as: R. B. Cattell's Culture Fair Intelligence Test - Scale 1, R. B. Cattell's and A. K. S. Cattell Culture Fair Intelligence Test

and their construction minimized to a greater or lesser degree the influence of certain cultural factors. While estimating the value of particular kinds of culture-reduced intelligence tests, attention should be given not only to their reliability and validity, but also cultural variables that have been minimalized and to the degree their influence have been minimalized by an appropriate construction.