

**Recenzja testu psychologicznego
Skale Inteligencji Stanford-Binet. Piąta Edycja (SB5)
wydana w Polsce przez Pracownię Testów Psycholo-
gicznych i Pedagogicznych
na licencji wydawnictwa PRO-ED (USA)
dla Komisji ds. Testów Psychologicznych KP PAN**

Wstęp

Skala Stanford-Binet jest indywidualnym testem inteligencji, który w 1916 roku stworzył na bazie oryginalnej skali Alfreda Bineta i Theodore Simona psycholog z Uniwersytetu Stanforda Lewis M. Terman. Od czasu pierwszej publikacji pojawiły się cztery zrewidowane wydania Skali Stanford-Bineta, z których pierwsza została opracowana przez Lewisa Termana. Ponad dwadzieścia lat później Maud Merrill wspólnie z Termanem zaczęli pracę nad nową zrewidowaną edycją skali, która ukazała się w 1937 roku (dokładniej: dwie jej wersje równoległe – M i L). Kiedy Lewis Terman zmarł w 1956 r. prace nad trzecią rewizją były już w toku i Maud Merrill opublikowała jej ostateczną wersję w 1960 r. (tzw. wersja L-M). To w niej zastosowano po raz pierwszy skalę dewiacyjnego IQ. Po przejściu Merrill na emeryturę prace nad następną rewizją skali podjął Robert Thorndike, który z pomocą Elizabeth Hagen i Jerome'a Sattlera opublikował w roku 1986 czwarte wydanie skali. Najnowszą wersję Skali Stanford-Bineta opublikował w 2003 roku Gale Roid, który na Uniwersytecie Harvarda był asystentem Davida McClellanda.

Według Wydawcy Skale Inteligencji Stanford-Binet Piąta Edycja (SB5) to test oceniający inteligencję i zdolności poznawcze w populacji ogólnej i w szczególności w grupach specjalnych. Test odpowiedni dla badanych w wieku od 2;0 do 18;11 lat (normy psychometryczne) oraz dla osób dorosłych z niepełnosprawnością intelektualną (tabela norm dla wieków umysłowych; skala CSS). Pełna skala IQ składa się z 10 podskal. Skala skrócona składa się z dwóch testów kierujących (ścieżek dojścia) – Serie/Matryce oraz Słownik. Skali skróconej SB5 można używać m.in. do badań neuropsychologicznych, gdzie dodatkowo stosuje się inne baterie testów uzupełniające pomiar wykonany z zastosowaniem SB5. Test oparty jest na teorii inteligencji CHC.

Polska wersja testu to adaptacja najnowszej wersji oryginalnej, normy opracowano na podstawie badań ogólnopolskich przeprowadzonych w latach 2016-2017.

Niewerbalna skala IQ opiera się na pięciu niewerbalnych podskalach odnoszących się do każdego z pięciu czynników poznawczych badanych za pomocą SB5 (Rozumowanie płynne, Wiedza, Rozumowanie ilościowe, Przetwarzanie wzrokowo-przestrzenne, Pamięć robocza). Używa się jej do diagnozy osób słabosłyszących, a także osób z zaburzeniami komunikacji, autyzmem, trudnościami w nauce, poważnym uszkodzeniem mózgu, władających językiem polskim w spo-

sób ograniczony z powodu pochodzenia, oraz w przypadku afazji lub udaru, które negatywnie wpływają na zdolności językowe jednostki.

Wербalna skala IQ składa się z pięciu werbalnych podskal odnoszących się do każdego z pięciu czynników. Wykorzystuje się ją przeprowadzając badanie skalą pełną, a także w specjalnych przypadkach badania osób z osłabionym wzrokiem, niepełnosprawnych ruchowo, etc.

Wyniki pełnego badania testem obejmują iloraz ogólny, iloraz werbalny, iloraz niewerbalny oraz wyniki poszczególnych czynników oraz wiek umysłowy. Na świecie test stosowany jest od początku alternatywnie do baterii Wechslera, w szczególności tam, gdzie zachodzi podejrzenie, że rozwój pacjenta nie jest w pełni typowy (grupy specjalne, zaburzenia neurorozwojowe, wybitne uzdolnienia) oraz jako jedyny precyzyjnie różnicujący stopnie niepełnosprawności intelektualnej.

Tekst recenzji jest podsumowaniem prac zespołu, powołanego przez Komisję ds. Testów Psychologicznych Komitetu Psychologii PAN, którego członkiem i koordynatorem był prof. dr hab. Władysław Jacek Paluchowski. W skład zespołu wchodził ponadto mgr Dorota Gaul i mgr Mateusz Wiliński, którzy przygotowali niezależne recenzje na podstawie analizy materiałów testowych i przeprowadzonych w tym celu badań. Prof. W. J. Paluchowski przygotował uwagi do części psychometrycznej na podstawie podręczników oraz scalił i uzgodnił z recenzentami i Wydawcą ostateczną wersję tej recenzji. Recenzje sporządzono na podstawie formularza, opracowanego w Komisji ds. Testów Psychologicznych KP PAN.

Dla celów recenzji zbadano 9 osób:

1. chłopiec, 3 lata 8 miesięcy (wyniki przeciętne w stosunku do wieku)
2. chłopiec, 4 lata 8 miesięcy (pogranicze normy)
3. chłopiec, 7 lat 6 miesięcy (wyniki przeciętne w stosunku do wieku)
4. chłopiec, 7 lat 7 miesięcy (wyniki przeciętne w stosunku do wieku)
5. chłopiec, 9 lat 7 miesięcy (pogranicze normy)
6. chłopiec, 10 lat 3 miesiące (niepełnosprawność intelektualna w stopniu lekkim)
7. dziewczynka, 2 lata 5 miesięcy (wyniki przeciętne w stosunku do wieku)
8. dziewczynka, 10 lat 4 miesiące (wyniki przeciętne w stosunku do wieku)
9. dziewczynka, 11 lata 7 miesięcy (wyniki przeciętne w stosunku do wieku)

Recenzenci zwrócili uwagę na fakt, że w "Podręczniku diagnosty" znajdują się odwołania do "Podręcznika do interpretacji" [Roid, G. H., Sajewicz-Radtke, U., Radtke, B. M. (2017) Skale Inteligencji Stanford-Binet, Edycja Piąta. Podręcznik interpretacji. Gdańsk: Pracownia Testów Psychologicznych i Pedagogicznych], którego nie posiadamy. W odpowiedzi Wydawca napisał „Ten podręcznik nie jest integralną częścią zestawu testowego, jest to materiał uzupełniający. Podręcznik jest w tej chwili opiniowany przez amerykańskiego licencjodawcę i ukaże się oficjalnie drukiem gdy tylko uzyskamy akceptację. Wszystkie kluczowe informacje dotyczące testu, walidacji, standaryzacji i normalizacji oraz procedur obliczania wyników znajdują się w standardowych podręcznikach. Jeżeli recenzenci mają pytania dotyczące standardowego zakresu informacji o teście, a nie znaleźli niezbędnych odpowiedzi w podręczniku diagnosty i technicznym

prosimy o ich przedstawienie. Moglibyśmy wówczas przesłać do wglądu fragmenty podręcznika interpretacyjnego (po uzyskaniu stosownej zgody licencjodawcy)”. Tak się też stało.

Konsekwencje użycia recenzowanego testu dla osoby badanej mogą być daleko idące.

Uzasadnienie

- Po pierwsze wyniki testu mogą stanowić przesłanki do podejmowania decyzji orzeczniczych:
 - W systemie edukacji – w zakresie kwalifikacji dzieci do wczesnego wspomaganie rozwoju oraz kształcenia specjalnego z uwagi na niepełnosprawność intelektualną. Zgodnie z deklaracją wydawcy, z uwzględnieniem różnych stopni niepełnosprawności (Roid i in., 2017, s. 159; Radtke, Sajewicz-Radtke, 2017, s. 10).
 - W systemie ubezpieczeń społecznych – w zakresie kwalifikacji osób małoletnich i osiemnastoletnich (aktualnie normy kończą się na poziomie 18;11) jako niepełnosprawnych intelektualnie. Zgodnie z deklaracją wydawcy, z uwzględnieniem różnych stopni niepełnosprawności (*ibidem*).
- Po drugie wyniki testu mogą stanowić przesłanki do formułowania wniosków w opiniach mających konsekwencje prawne, przede wszystkim w systemie edukacji ale również w systemie ubezpieczeń społecznych, sądowno – penitencjarnym oraz ochrony zdrowia. Wydawca odnosi się do tego zagadnienia w podręczniku diagnosty (Roid i in., 2017, s. 3).
- Jako test oceniający zdolności intelektualne może on być wykorzystywany przy kwalifikacji do określonych programów edukacyjnych, czy społecznych.

Obszar diagnozowany przez recenzowany test jest „wrażliwy” i może być powodem stygmatyzacji.

Uzasadnienie

- SB5 jest narzędziem służącym do diagnozy inteligencji oraz funkcjonowania poznawczego (Roid i in., 2017, s. 1). Czynnikiem ten jest powszechnie uznany za istotny dla funkcjonowania osoby a jego wyrażona ocena może wpływać zwrotnie na ustosunkowanie innych osób (np. nauczycieli, lekarzy) do jednostki.
- Na podstawie diagnozy dokonanej Skalą można kategoryzować wyniki osoby badanej, od niepełnosprawności intelektualnej w stopniu umiarkowanym do wybitnych zdolności intelektualnych. Oznacza to potencjalnie daleko idące konsekwencje diagnozy z wykorzystaniem SB5 w zakresie zmiany ustosunkowania innych osób do jednostki badanej. Informacja o wynikach badania tym testem, zwłaszcza o tych niskich, a w niektórych przypadkach przeciętnych, niesie ze sobą konsekwencje etykietowania osoby badanej i ryzyko jej stygmatyzacji.

Recenzowany test wymaga szczególnych kompetencji diagnostycznych i psychometrycznych.

Uzasadnienie

- Test ten obejmuje bardzo szeroką wiekowo grupę osób badanych, Zawiera wiele zadań, wymaga od osób badanych koncentracji uwagi i pozytywnej motywacji zadaniowej przez dość długi czas. Osoby badane mogą czuć to, że ocenia się poprawność wykonywania zadań, co może dla nich stanowić sytuację emocjonalnie trudną. Wszystko to sprawia, że przeprowadzanie tego testu wymaga od badającego dużych kompetencji diagnostycznych dotyczących między innymi tworzenia odpowiednich warunków dających diagnozowanej osobie poczucie bezpieczeństwa, umiejętnego zachęcania do rozwiązywania zadań, koncentrowania na zadaniach itp. A to z kolei wymaga dobrej znajomości zadań testowych i towarzyszących im instrukcji, w takim stopniu, który sprawi, że test zostanie przeprowadzony sprawnie i bez wykraczania poza zasady standaryzacji.
- W konstrukcji testu SB5 wdrożono koncepcję pomiaru kluczowych zdolności do oceny poziomu funkcjonowania osoby badanej („testy kierujące”), koncepcje podziału testów na bloki o zwiększającym się stopniu trudności – o zróżnicowanej treści zadań („poziomy”, „testy cząstkowe”), rotację zadań mierzących różne czynniki zdolności (Roid, Jurek i in., 2017, s. 15). Podejście to jest definiowane jako „poziomy ze zróżnicowaną oceną czynników” i, według autorów podręcznika technicznego (ibidem), koresponduje z nowoczesnymi teoriami testowymi. Ponadto w projektowaniu oryginalnej wersji testu zastosowano teorię odpowiadania na pozycje testu - *item response theory* (ibidem). Wymienione koncepcje, wraz z wykorzystanymi teoriami służącymi projektowaniu testu mogą być dla polskich psychologów praktyków, istotnym *novum*. Dotychczas obecne na rynku testy, w szczególności w zakresie pomiaru intelektu, nie posługiwały się tymi koncepcjami i w niewielkim stopniu odwoływały się do wspomnianych teorii.
- Zgodnie z intencją autorów oraz wydawcy (Roid i in., 2017, s. 1) SB5 może być używany do diagnozy osób głuchych lub słabosłyszących, z zaburzeniami komunikacji, autyzmem, szczególnymi trudnościami w nauce bądź poważnym uszkodzeniem mózgu, osób władających językiem polskim w sposób ograniczony z powodu pochodzenia, w przypadkach afazji lub udaru (w zakresie pomiaru zdolności umysłowych Skalą Niewerbalną IQ) oraz wobec osób z osłabionym wzrokiem, niepełnosprawnych ruchowo i w stanie utrudniającym im ukończenie niewerbalne skali IQ (Skala Werbalna IQ). Przeprowadzenie badań osób z tych grup specjalnych wiąże się z dodatkowymi wymaganiami postawionymi diagnoście zarówno w zakresie tworzenia odpowiednich warunków badania, dostosowywania procedury testowania SB5, jak i elementów postępowania postdiagnostycznego (m. in. w zakresie omawiania wyników). Wytyczne co do wykorzystania testu SB5 są opisane w podręczniku diagnosty (Roid i in., 2017, s. 9-19.) oraz w podręczniku interpretacji (niedostarczonym do recenzji, Roid, Szejewicz-Radtke, Radtke, 2017).

Skale Stanford-Bineta wymagają dogłębnej znajomości psychometrii dla poprawnego stosowania i interpretacji.

Uzasadnienie

- Specjaliści (w szczególności dokonujący interpretacji wyników) stosujący Skale Stanford-Bineta 5 powinni zostać przeszkoleni z zakresu ogólnych pojęć statystycznych niezbędnych do rozumienia wyników oraz w pełni rozumieć pojęcia: błąd pomiaru, znormalizowane wyniki, rzetelność, trafność, normatywne i opierające się na kryteriach interpretacje testów (Roid i in., 2017, s. 6). Powinni również adekwatnie dokonywać wyboru tego testu względem warunków badania oraz zgłaszanego problemu diagnostycznego uwzględniając jego właściwości psychometryczne. Wymogi te są konsekwencją, z jednej strony złożoności testu i jego interpretacji (również psychometrycznej), a z drugiej przeznaczenia testu i obszaru, który on bada.
- Język tłumaczonych podręczników (diagnosty i techniczny) może być niezrozumiały dla potencjalnych użytkowników. W szczególności wykorzystywane pojęcia z zakresu psychometrii (IRT) wykraczają poza standardy kursu tego przedmiotu na studiach psychologicznych.

Poprawne użycie interpretowanego testu wymaga pogłębionej wiedzy psychologicznej i dogłębnej znajomości teorii leżącej u podstaw interpretacji.

Uzasadnienie

- W aktualnych programach studiów psychologicznych przynajmniej najważniejsze teorie inteligencji oraz koncepcje psychometryczne, na których oparty jest test, są nauczane.
- Z jednej strony informacje przedstawione w podręczniku technicznym służą efektywnemu zapoznaniu użytkownika z teoriami, na których oparty jest test, w tym szczególnie z hierarchicznym modelem zdolności intelektualnych CHC (Roid, Jurek i in., s. 1-24). Z drugiej strony model ten, podobnie jak inne teorie psychologiczne, z których korzystano przy konstrukcji testu mają charakter zaawansowany i mogą wymagać bliższego zapoznania się przez użytkowników testów dla jego używania i prawidłowej interpretacji.
- W podręczniku diagnosty (Roid i in. 2017, s. 139-167) zawarto informacje dotyczące podstawowej interpretacji wyników (analizę ilościową, analizę profilową, identyfikację trudności osoby badanej). W tym oraz dostarczonym do recenzji podręczniku technicznym (Roid, Jurek i in. 2017) zawarto również informacje niezbędne do dokonania podstawowej/wstępnej interpretacji psychologicznej wyników testowych.
- Prawdopodobnie, informacje znajdujące się w podręczniku interpretacji (Roid, Sajewicz-Radtke, Radtke, 2017) mogą służyć pogłębieniu interpretacji psychologicznej. Podkreśla się przy tym, że podręcznik ten nie wchodzi w skład sprzedawanego aktualnie zestawu SB5.
- Zaznacza się przy tym, że prawidłowa interpretacja psychologiczna, dokonana tylko na podstawie podręczników do testu (w tym podręcznika interpretacji), bez pogłębionej wiedzy psychologicznej, jest znacznie utrudniona lub niemożliwa.
- Pewnym ułatwieniem dla użytkowników jest możliwość stosowania SYSTEMU WSPARCIA DIAGNOSTY udostępnionego przez wydawcę testu na stronie: <http://id3855.abaks.pl/bateriaBinet/>, który wspiera diagno-

stę w poprawnym obliczaniu wyników i pozwala na wygenerowanie wstępnego opisu wyników do dalszych analiz i opracowań.

Zgodnie z aktualną polityką wydawcy dostęp do testu i posługiwanie się nim wymagają specyficznego treningu bez względu na wcześniejszy poziom kwalifikacji użytkownika.

Uzasadnienie

- Aktualnie, w Polsce, psycholog musi uzyskać certyfikat uprawniający go do posługiwania się testem SB5 i jego zakupu. Wydawca przewiduje możliwość odbycia takiego kursu podczas zajęć dydaktycznych (prowadzonych przez wykładowcę akredytowanego przez Pracownię Testów Psychologicznych i Pedagogicznych w Gdańsku) w czasie jednolitych studiów magisterskich w zakresie psychologii.
- Uznać należy, że stosowanie SB5 stawia przed użytkownikiem wymagania idące dalej niż inne, dostępne na polskim rynku testy inteligencji i zdolności poznawczych (por. wymogi szczególnych kompetencji diagnostycznych i psychometrycznych oraz wiedzy psychologicznej i teorii leżących u podstaw interpretacji) oraz wyrażoną w podręczniku diagnosty (Roid i in., 2017, s. 7) argumentację dotyczącą wagi przeprowadzania testów pod nadzorem dla nabywania biegłości użytkownika w procedurach testowania.
- Test SB5 jest względnie prosty w administracji (w porównaniu np. do testu IDS Skale Inteligencji i Rozwoju), porównywalny do Skal Inteligencji Wechslera (w wersji WAIS-R i WISC-R), także procedura obliczania wyników wydaje się jasna i prosta. Problem stanowi ich interpretacja zważywszy na to, jak wielki jest zakładany przedział wieku osób badanych i do badania jakich różnych grup z zaburzeniami psychorozwojowymi test ten w myśl intencji autorów może być wykorzystywany.
- Oprócz aktualnie wydanych materiałów (podręcznika diagnosty i technicznego) wydawca ma w planie wydanie materiałów ułatwiających korzystanie z tego testu (tj. Roid, Sajewicz-Radtke, Radtke, 2017 – informacja potwierdzona mailowo dnia 27.11.2017 r., na potrzeby recenzji).
- Poprawna ocena i interpretacja wyników testu psychologicznego wymaga dużej wiedzy psychologicznej i doświadczenia w kontakcie z osobami z tej grupy, do której należy konkretna osoba badana. Pomimo, że materiały testowe, za pomocą których przeprowadza się badanie, a także „Podręcznik diagnosty” przygotowane są bardzo starannie, są czytelne, zrozumiałe, przejrzyste, to jednak ze względu na dobro osób badanych, dostęp do testu i posługiwanie się nim wymagają odbycia specyficznego szkolenia i treningu, niezależnie od poziomu kwalifikacji użytkownika.

CZEŚĆ A. Ocena materiałów testowych

A1. Materiały testowe dla osoby prowadzącej badanie

Z perspektywy użytkownika testu, w zasadzie poprawnie opisano, co jest mierzone i uzasadniono dlaczego pomiar w SB5 jest skonstruowany tak, jak jest.

Uzasadnienie:

- W podręczniku diagnosty (Roid i in. 2017, s. 21-27, 47-118) wyjaśniano konstrukcję SB5, wraz ze szczegółowym opisaniem co i jak jest mierzone. W podręczniku technicznym (Roid, Jurek i in. 2017, s. 1-27) opisano konstrukcję testu w odwołaniu do teorii.
- Opis co jest mierzone i wyjaśnienie dlaczego operacjonalizacja i skalowanie mają taką postać (np. dlaczego dobrano takie właśnie zadania, bądź taką kolejność zadań) jest zbyt powierzchowny.
- W związku z tym, że Podręcznik Diagnosty to w zasadzie tłumaczenie, brak jest wskazania zarówno różnicy (podobieństwa) kontekstu kulturowego i jego wpływu na interpretację wyników testu, jak i polskiej literatury i polskich odpowiedników literatury oryginalnej (Roid i in. 2017, s. 167 i dalsze).

Z perspektywy użytkownika testu, nie w pełni opisano polską adaptację testu SB5.

Uzasadnienie:

- W podręczniku technicznym (Roid, Jurek i in. 2017, s. 27-32 i 35-37) opisano procedury zastosowane w procesie aktualizacji (doskonalenia) zadań (pozycji) testu SB5 wydania oryginalnego.
- Brakuje uzasadnienia pozostawienia bez zmiany (lub zmiany) adaptowanych kulturowo pozycji testowych, dlaczego taka, a nie inna jest ich kolejność, jak to się ma do teorii, na podstawie których powstał ten test itp.
- Autorzy zbyt często odwołują się do poprzednich edycji testu, nie podając konkretnych informacji.

Z perspektywy użytkownika testu, poprawnie opisano psychometryczne własności testu.

Uzasadnienie:

- W podręczniku technicznym (Roid, Jurek i in. 2017) znalazły się podrozdziały poświęcone opisaniu rzetelności (Rozdział 3) i trafności SB5 (Rozdział 4) oraz techniczne aspekty interpretacji wyników (Rozdział 5). Opis ten był wystarczający na potrzeby użytkownika testu.

Poprawnie opisano procedurę prowadzenia badania i jej dopuszczalne modyfikacje.

Uzasadnienie

- Opis taki znalazł się w podręczniku diagnosty (Roid i in. 2017, s. 29-122) uwzględniając dopuszczalne modyfikacje i dostosowania dla osób z grup

specjalnych (op. cit., s. 120-122). Wskazano również, że podręcznik do interpretacji (Roid, Sajewicz-Radtke, Radtke 2017) zawierać będzie dalsze propozycje dostosowań.

- Sposób konstrukcji protokołu wykonania SB5 oraz materiałów pomocniczych (w formie zestawów tablic z oddzielnie prezentowanymi bodźcami dla osoby badanej i instrukcjami do odczytania dla diagnosty) ułatwia trzymanie się procedury badania testem bez konieczności używania podręcznika. Podręcznik diagnosty zaś zawiera wskazówki jak radzić sobie w sytuacjach problemowych (np. Roid i in. 2017, s. 31, 37, 38) oraz wskazanie kiedy i w jakim zakresie można odstąpić od warunków standardowych (np. przerywanie testu, dzielenie na sesje – s. 40, dostosowywanie procedur do dzieci w wieku przedszkolnym, s. 42)

Poprawnie wyjaśniono, jak obliczać wyniki.

Uzasadnienie

- W podręczniku zawarto jasną informację, z przykładami słownymi, a sporadycznie również graficznymi, jak obliczać wyniki dla każdej podskali, czynników, wyniku niewerbalnego, werbalnego i ogólnego.

Poprawnie wyjaśniono, jak interpretować wyniki.

Uzasadnienie

- W podręczniku zawarto jasną informację, z przykładami słownymi oraz analizą przypadków, jak interpretować wyniki, zarówno ilościowo (analiza ilościowa, profilowa, różnic) jak i psychologicznie. Korzystnym rozwiązaniem jest również zaproponowanie interpretacji metodą siedmiu kroków oraz uporządkowanie według niej opisu procesu interpretacji testem SB5. Ponadto stwierdza się, że sposób wyrażenia norm (na skali standardowej tzw. wechslerowskiej oraz równocześnie na skali centylowej) jest przystępny i ułatwia dokonywanie bardziej precyzyjnej interpretacji.
- Wyjaśnienia dotyczące interpretacji różnych wyników, rozumienia norm, rozważania relacji pomiędzy poszczególnymi skalami powinny być bardziej rozbudowane i zilustrowane większą liczbą przykładów.

W teście poprawnie opisano i uzasadniono kto może i kto powinien być badany testem.

Uzasadnienie

- Nie wymieniono, kto nie powinien być badany testem, ale dzięki precyzyjnemu opisowi zakresu badanych zmiennych, opisie grup specjalnych, które mogą być badane testem oraz opisanym zakresom dostosowań dla tych grup specjalnych jest czytelne kto powinien, a kto nie powinien być badany testem. Wskazano przy tym, w podręczniku technicznym (Roid, Jurek i in., s. 24), że w przypadku osób „pochodzących ze środowisk zróżnicowanych kulturowo i etnicznie oraz osób z poważnymi niepełnościami lub schorzeniami (...) należy upewnić się, że wymagania

testowe są dla nich sprawiedliwe”. Ponadto wskazano również ograniczenia polskiej adaptacji w zakresie stosowania testu w populacji osób głuchych i słabo-słyszących.

Poprawnie wyjaśniono, jakie kompetencje są niezbędne do właściwego posługiwania się testem.

Uzasadnienie

- W podręczniku diagnosty (Roid i in., 2017, s. 5-8) poświęcono tym zagadnieniom dużo miejsca.
- Jedyne zastrzeżenie budzi brak wyrażenia w podręczniku decyzji wydawcy o konieczności odbycia dodatkowego szkolenia dla psychologów (absolwentów 5-letnich studiów magisterskich) w celu uzyskania certyfikatu uprawniającego do korzystania i dostępu do testu SB5 oraz bezpośredniego uzasadnienia takiej decyzji. Za taką argumentację mogłaby być uznana argumentacja dotycząca zasadności wykonywania testów pod nadzorem, gdyby decyzja odbycia dodatkowego szkolenia dla psychologów była ujawniona w podręczniku.

Podsumowanie oceny, czy wszystkie materiały dla osoby prowadzącej badanie są poprawne:

1. Istnieje zastrzeżenie do sposobu opisu testu Słownik w podręczniku (który jednocześnie stanowi ścieżkę dostępu). W opisie tego testu oraz we wskazówkach do obliczania jego wyniku, nie wskazano jednoznacznie kryterium oceny tych pozycji, które wymagają definiowania słów.
2. Jak pisze Wydawca „Kryteria oceny przedstawiane są podczas obowiązkowego szkolenia kwalifikacyjnego. Wówczas odbywa się również trening w zakresie oceny Słownika. Słusznym wydaje się jednak uzupełnienie podręcznika diagnosty o ten aspekt. Uwaga zostanie uwzględniona w kolejnym wydaniu podręcznika.”
3. W podręczniku diagnosty (Roid i in. 2017, załącznik D.2. Definicje używane do oceny testu kierującego Słownik) i zestawie tablic wskazano właściwe definicje słów oraz przykłady odpowiedzi ocenianych na 1 i 2 pkt. Nie wskazano jednak, że ocenie podlega wyłącznie JAKOŚĆ definicji, co uczyniono w trakcie szkolenia z testu SB5 (Radtke, Sajewicz-Radtke, 2017). Z uwagi na silnie obecną praktykę stosowania testu WISC-R oraz odmienny sposób oceny skali słownikowej w nim zawartej, część psychologów praktyków może mieć inaczej ukształtowane nawyki oceny pozycji słownikowych. W związku z powyższym brak (w podręczniku, w zestawie tablic obszaru werbalnego) precyzyjnie wyrażonego kryterium oceny pozycji słownikowych, wymagających definiowania, niesie, w kontekście polskiej praktyki, wzmożone ryzyko błędu oceny odpowiedzi osoby badanej w tym podteście. W opinii recenzenta podręcznik powinien być skorygowany pod kątem tej uwagi.
4. Informację dotyczącą stosowania testu wyrażoną w podręczniku można uzupełnić o precyzyjne określenie jakie grupy nie powinny być badane testem (np. osoby z całkowitym zaburzeniem widzenia).

5. W Zestawie tablic 2 jest błąd, dwukrotnie pojawiają się strony od 39 do 48. Jak pisze Wydawca jest to „błąd drukarski ujawniony w niewielkiej części nakładu. Zalecono drukarni zwiększenie dokładności kontroli jakości składu.”

A2. Materiały testowe dla osoby badanej

Instrukcje dla osób badanych (z uwzględnieniem poziomu wieku i poziomu startowego testu) są zrozumiałe, poza naprawdę wyjątkowymi przypadkami / sytuacjami nie pozostawiają wątpliwości (dla takich sytuacji, w podręczniku, wskazane są sposoby postępowania).

Uzasadnienie:

- Badane dzieci nie miały problemu ze zrozumieniem zadania. Zrozumienie ułatwia też obecny w wielu początkowych zadaniach danego poziomu pokaz.
- Spośród wszystkich zadań, najtrudniejsze instrukcje (zarówno dla diagnosty – w zakresie ich podawania, jak i dla osoby diagnozowanej) są zawarte w zadaniu dotyczącym pamięci roboczej w obszarze niewerbalnym. Trudność instrukcji wynika tutaj z konstrukcji samego zadania. Zaznacza się przy tym, że po odpowiednim wcześniejszym treningu diagnosta powinien radzić sobie z podawaniem instrukcji (wraz z prezentacją przykładów), tak by nie utrudniać dodatkowo osobie badanej wykonania zadania. W podręczniku diagnosty (Roid i in., 2017, np. s. 70) zawarto uwagi w podobnym duchu, zachęcające do treningu w zakresie przeprowadzania tego podtestu.

Forma graficzna arkusza testowego dla osoby badanej jest czytelna i łatwa do zrozumienia.

Uzasadnienie:

- Forma graficzna nie jest przeładowana graficznie, zawiera tylko niezbędne elementy przedstawione w sposób uporządkowany i estetyczny, właściwie dla zadania.

Podsumowanie opinii, czy materiały dla osoby badanej są poprawne:

1. Materiały i pomoce do badań przygotowano bardzo starannie i ergonomicznie.
2. Nie dostrzeżono znaczących błędów w materiałach dla osoby badanej.
3. Zestaw tablic 2, str. 41 i 42 rzeczy mało znane dzieciom z doświadczenia. Jak pisze Wydawca „Na etapie adaptacji testu, przedstawiony na tablicach testowych materiał został porównany z podstawą programową kształcenia ogólnego MEN. W myśl przywołanej podstawy dzieci na poziomie klas I-III oraz IV-VIII zapoznają się zarówno z prawidłowym sposo-

bem adresowania listu jak i z konceptem wagi szalkowej (treści podstawowe). W związku z tym, że test bada niewerbalną wiedzę wydaje się, że zastosowane bodźce są zasadne.”. Jednak w naszej opinii to, że adresowanie listu, czy koncepcja wagi szalkowej są w podstawie programowej szkoły podstawowej, nie zmienia faktu, że nie są one bliskie doświadczeniu dzieci. Test obejmuje też 5 roczników dzieci przedszkolnych, a zadania te są na 2 i 3 miejscu, jeżeli chodzi o absurdy obrazkowe. Te zadania mogą sprawić, że za chwilę trzeba będzie robić nową adaptację, bo listy znikają z naszego życia, prywatnie ich przecież niemal nie wysyłamy, a te, które dostajemy - urzędowe, czy reklamowe nie mają znaczenia. Waga szalkowa budzi mniej wątpliwości, choć bardziej znana doświadczeniu dzieci również tych przedszkolnych jest huśtawka. Uwaga wydawcy budzi też wątpliwości co do procesu adaptacji test - czy nie powinien on polegać na zbadaniu trafności każdej pozycji dla polskiej populacji, a nie tylko na sprawdzeniu, czy uczy się o nich w naszej szkole?

4. Zestaw tablic 3 str. 2, 3, 4 instrukcja: „opowiedz mi o tym obrazku” wydaje mi się zbyt trudna, dzieci przedszkolne, zwłaszcza te młodsze naturalnie odpowiadają wymieniając obiekty, być może opisanie sytuacji uwzględniające czynności „wymusza” lepiej pytanie „co dzieje się na tym obrazku”. Wbrew pozorom instrukcja "opowiedz mi o tym obrazku" wydaje się większą barierą. Może warto połączyć te dwie instrukcje. Trzeba by to ewentualnie sprawdzić. W opinii Wydawcy „Zastosowana instrukcja wynika ze specyfiki rozwoju mowy najmłodszych i niepełnosprawnych intelektualnie badanych. Uwzględnia ona fakt, że osoby te potrafią jedynie wymienić obiekty na obrazki bez odwoływania się do akcji (relacji między obiektami). Użycie instrukcji koncentrującej na opisie akcji mogłoby stanowić barierę w wykonaniu polecenia przez najmłodszych i niepełnosprawnych intelektualnie. ”.
5. Na str. 11 niejasna instrukcja i kryteria oceny zadania 3, czy chodzi o to, że badana osoba ma położyć klocek gdziekolwiek na rysunku domu, czy na górze? Jeżeli ta pierwsza opcja, to, co z zadaniem 2, jeżeli umieści klocek na obrazku domu, ale na dole? Uwaga Wydawcy jest następująca: „Instrukcja jest precyzyjna: „Połóż klocek na rysunku domu” – czyli gdziekolwiek w obrębie rysunku. Faktycznie trudność może sprawić diagnozę sytuacji, w której dziecko umieszcza klocek w dolnej części rysunku i tym samym pojawia się kłopot w udzieleniu instrukcji do polecenia drugiego: „Przesuń klocek na dół domu”. Na szkoleniu kwalifikacyjnym sytuacja ta jest omawiana, a psychologowie otrzymują szczegółowy instruktaż.”.
6. Na str. 15 może przykład powinien obejmować dwie kategorie, jeden przykład, zwłaszcza kolor, może sprawić, że dziecko, zwłaszcza małe, albo osoba niepełnosprawna intelektualnie zafiksuje się właśnie na tej kategorii. Komentarz Wydawcy: „Wyjaśniono to w podręczniku diagnosty (str. 101)”.

CZĘŚĆ B. Ocena własności psychometrycznych testu

B1. Normalizacja

Dane normalizacyjne są aktualne i poprawie przedstawione.

- W podręczniku technicznym (Roid, Jurek i in. 2017, s. 37- 43) przedstawiono jasno i klarownie zarówno wielkość, jak i pochodzenie prób normalizacyjnych z uwzględnieniem wieku, płci, regionu zamieszkania, miejsca zamieszkania, poziomu wykształcenia rodziców.
- W podręczniku technicznym (Roid, Jurek i in. 2017, s. 27 i dalsze) niepoprawnie używa się terminu „standaryzacja” zamiast – poprawnego w tym kontekście – normalizacja (ibidem str. 37). Zdaniem Wydawcy „W podręczniku użyto dosłownego tłumaczenia z oryginału (*standardization studies*)”, ale na tym właśnie polega problem – w języku polskim jest to termin niepoprawny do opisywania zjawiska prezentowanego w podręczniku.

Trudno ocenić z perspektywy użytkownika testu, czy próby normalizacyjne (referencyjne) odpowiadają przeznaczeniu testu zakładanemu przez jego autora oraz czy normy nadają się do wszystkich zamierzonych zastosowań. Jest tak z pewnością w odniesieniu do populacji ogólnej, z uwzględnieniem podstawowych charakterystyk demograficznych.

Próby normalizacyjne (referencyjne) są reprezentatywne dla populacji.

- W podręczniku technicznym (Roid, Jurek i in. 2017, s. 37-43) wskazano jako podstawę dane GUS.

Uwzględnione zostały dane demograficzne odpowiadające podstawowym zastosowaniom testu (tj. wiek, płeć, region zamieszkania, miejsce zamieszkania, poziom wykształcenia rodziców).

Próby normalizacyjne (referencyjne) zostały opisane (podano liczebności w poszczególnych kategoriach).

Z perspektywy użytkownika testu zastosowana skala wyników przeliczonych (od 1 do 19, ze średnią 10 i SD = 3) pozwala na precyzyjne różnicowanie w zakresie wyników niskich i wysokich dla grupy odniesienia. Również sposób prezentacji wyników znormalizowanych (tzw. skala wechslewska ze średnią 100, SD = 15 oraz skala centylowa) jest adekwatny i pozwalający na różnicowanie w zakresie wyników niskich i wysokich dla grupy odniesienia.

Nie wskazano w podręczniku diagnosty i w podręczniku technicznym ograniczeń interpretacyjnych wynikających ze sposobu doboru prób normalizacyjnych (referencyjnych).

Procedura normalizacyjna jest poprawna.

- W podręczniku technicznym (Roid, Jurek i in. 2017, s. 45-51) opisano przebieg normalizacji.

CZEŚĆ B. Ocena własności psychometrycznych testu B2. Rzetelność

Dostępne są wskaźniki zgodności wewnętrznej wyrażone współczynnikiem alfa Cronbacha oraz współczynniki równoważności międzypółkowej (obliczono współczynnik Spearmana-Browna) zarówno dla ogólnego wyniku, obszarów, czynników jak i poszczególnych skal testu.

Podano metody szacowania współczynników rzetelności testów (zastosowano zarówno współczynniki alfa Cronbacha, półkowania Spearmana-Browna oraz test-retest (patrz rozdział 3. Podręcznika technicznego, Roid, Jurek i in., 2017).

Zastosowano cztery metody szacowania współczynników rzetelności wyników testu, oprócz wyżej wymienionych skorzystano również z oceny funkcji informacyjnej (Roid, Jurek i in., 2017, s. 61 i dalsze; ryc. 3.1-3.9).

- Dla osób bez pogłębionej i wykraczającej poza standardowy program psychometrii ta część podręcznika technicznego może być nieczytelna. Zadaniem Wydawcy „Dlatego obowiązkowe jest szkolenie certyfikujące.”, z czym należy się zgodzić. Nie zmienia to faktu, że ta część podręcznika technicznego można opisać przystępniej (tak, jak to jest prezentowane podczas szkolenia); utrwali to też efekty szkolenia.

Przedstawiono dane na temat SEM zarówno dla wyniku ogólnego, jak i wyników częściowych (Roid, Jurek i in., 2017, s. 60).

Nie przedstawiono wartości półprzedziałów ufności dla wyniku globalnego i/lub dla wyników częściowych, ale przedstawiono możliwy sposób ich określania. W samym narzędziu budowanie przedziałów ufności oparto na szacowaniu prawdziwego wyniku z wykorzystaniem wzoru na standardowy błąd estymacji (SEE – por. Roid, Jurek i in., 2017, s. 59) .

Nie odnaleziono lub przeoczono informacje na temat standardowego błędu pomiaru dla różnic między wynikami w skalach. Wydawca zwraca uwagę, że „Rozdział 5 zawiera metodę badania różnic między wynikami z uwzględnieniem SEM. Tabele w Załączniku B zawierają minimalne wartości różnic.”. Jest tak w istocie.

Przedstawiono wartości minimalnej istotnej różnicy między wynikami danej skali a przeciętną profilu. Wartości te przedstawiono w podręczniku technicznym (Roid, Jurek i in., 2017, s. 129).

Z perspektywy użytkownika testu rzetelność została poprawnie wykazana. Podręcznik techniczny (Roid, Jurek i in., 2017) w sposób przejrzysty prezentuje wszystkie niezbędne informacje dotyczące trafności.

CZEŚĆ B. Ocena własności psychometrycznych testu B3. Trafność

Przedstawiono dogłębnie i względnie szczegółowo charakterystykę konstruktów mierzonych przez test.

Uzasadnienie

- Zagadnieniu temu poświęcony jest rozdział 1 (wstęp i omówienie) podręcznika technicznego (Roid, Jurek i in., 2017). Rozdział ten przedstawia model teoretyczny, na którym oparty jest test oraz historię narzędzia w jego poprzednich wersjach. Uzasadnia on również konstrukcję testu w odniesieniu do konstruktów mierzonych.

Założenia teoretyczne dotyczące mierzonych zmiennych, jak i sposobu konstrukcji testu przedstawiono w sposób ogólny w rozdziale 1 (wstęp i omówienie) podręcznika technicznego (Roid, Jurek i in., 2017). Natomiast w podręczniku diagnostyki (Roid i in., 2017, Rozdział 5, s. 139-167) przedstawiono założenia teoretyczne dotyczące mierzonych konstruktów oraz sposobu ich interpretacji. Przedstawiony opis uznaje się za poprawny.

Przedstawiono wystarczające dane empiryczne dla proponowanej interpretacji testu; w tym: wskaźniki trafności teoretycznej (w tym korelacje między podtestami, wskaźnikami czynników oraz wynikami IQ), strukturę czynnikową SB 5, wskaźniki trafności diagnostycznej (korelacje między SB5, a WAIS-R, CFT 1-R, TRJ). Najślabiej udokumentowane są badania walidacyjne (o tym dalej w recenzji).

Przedstawiono syntetyczne podsumowanie wyników otrzymanych w badaniach walidacyjnych (w tabelach i słowne). Zawarte jest w podręczniku technicznym (Roid, Jurek i in., 2017, s. 84-89).

- Brakuje omówienia wniosków wynikających z badania trafności kryterialnej (np. tab. 4.9). Choć Wydawca komentuje „W podręczniku skoncentrowano się na opisie wyników.”, to nie zmienia to faktu braku omówienia wniosków.
- Wątpliwości budzi brak przywołania, w podręczniku technicznym (Roid, Jurek i in., 2017) wyników badań walidacyjnych testu SB5 w wersji oryginalnej oraz brak dokonania oceny wskaźników trafności prognostycznej (np. wyników szkolnych, zdolności adaptacyjnych). Komentarz Wydawcy „Na stronie 73-74 podręcznika technicznego zreferowano zakres badań nad trafnością SB5 z podaniem źródeł. Faktycznie wyniki te nie były szczegółowo przytaczane.”. Jednak w naszej opinii to zreferowanie ma charakter ogólnikowy (także na następnej stronie, gdzie omawia się trafność teoretyczną), a cytowane źródła pochodzą z lat 1943-2007.
- Zwraca się również uwagę, że do walidacji zastosowano testy CFT1-R oraz test słownikowy (TRJ), które przeznaczone są do badania dzieci w wieku od 4 do 9 lat, oraz skalę wykonaniową WAIS-R (w zakresie porównania wyników osób w wieku 17 i 18), tymczasem test SB5 jest przeznaczony

do badania osób w wieku 2-18 lat. Nie skorzystano z innych skal wykonaniowych ani testów badających zdolności poznawcze i intelektualne, pomimo tego, że niektóre z nich (np. Międzynarodowa Skala Leitera, CFT20-R) są przeznaczone do badania osób w przedziale wiekowym bardziej zbliżonym do przedziału wiekowego osób, które można badać z zastosowaniem SB5.

- Ponadto zwraca się uwagę, w przeciwieństwie do interpretacji zawartej w podręczniku (Roid, Jurek i in., 2017 s. 86), że różnice średnich wartości IQ pomiędzy CFT 1-R oraz SB 5 nie wynikają z efektu Flynna. Polskie normy do testu CFT 1-R pochodzą z 2011 roku i nawet jeśli trafna jest przytoczona argumentacja o wzroście IQ populacji, zgodnie z efektem Flynna (z czym również można dyskutować), to nie tłumaczy ona skali różnicy wartości IQ pomiędzy CFT 1-R oraz SB5. Różnica ta i jej skala wynika prawdopodobnie ze słabości testu CFT 1-R (opisanych w podręczniku i przywołanych w Krasowicz-Kupis, Wiejak, Gruszczyńska, 2015, s. 43), mianowicie test ten jest trudny dla dzieci czteroletnich i łatwy dla dzieci w wieku 8-9 lat, co prowadzi do zaniżania wyników dzieci młodszych i zawyżania wyników dzieci 8-9 letnich. Taka interpretacja różnic wartości IQ, jeśli jest właściwa, wskazuje na słabość zastosowanej procedury walidacyjnej. Rekomendowanym testem do badania dzieci 8-9 letnich (i starszych) jest test CFT20-R, który nie został zastosowany w badaniach walidacyjnych testu SB5. Jak pisze Wydawca „Dziękujemy za tę uwagę. Te wnioski wydają się bardziej trafne niż nasze założenia.”.

B4. Prowadzenie badania

Przedstawiono jasno zasadę przeliczania wyników surowych na wyniki standaryzowane, zarówno szczegółowym i precyzyjnym tekstem, jak i przykładową analizą przypadków w formie grafik oraz tekstu.

Przedstawiono sposób interpretacji wyników zastosowanej skali standardowej:

- a) podstawową kategoryzację skali znormalizowanej, w podręczniku diagnosty (Roid i in., 2017, str. 159).**
- b) właściwości zastosowanej dla podtestów skali standardowej (skala od 1-19 z medianą 10 i SD = 3 są opisane w Roid, Jurek i in., s. 48)**
- c) sposób interpretacji wyników w poszczególnych podskalach, w odniesieniu do konstruktów mierzonych przez podtesty (Roid i in., 2017, str. 144-152).**

- W podręczniku diagnosty, w części dotyczącej interpretacji nie wskazano właściwego sposobu interpretowania wyników zastosowanej w podtestach skali standardowej, jakkolwiek prawidłowo interpretowano wyniki w przywoływanych przykładach (np. str. 157 Roid i in. 2017). Wskazówki do takiej interpretacji znajdują się w protokole wykonanego badania (zaznaczono kolorami wartości z różnych przedziałów skali standardowej).

Przedstawiono przykład normatywnej interpretacji wyników otrzymanych w teście (Roid i in., 2017, str. 159 -162).

Przedstawiono rekomendacje do stosowania testu w badaniach indywidualnych w różnych obszarach praktyki (Roid i in. s. 3).

- Niestety rekomendacje w tym miejscu (a nie znaleziono innych w podręczniku diagnostyki i technicznym) nie uwzględniają wysokiej wartości rzetelności wyników testu. Brak jest wyraźnej i jednoznacznej wzmianki, w podręcznikach diagnostyki i technicznym, że SB5 nie tylko nadaje się do diagnozy indywidualnej, ale spełnia wysokie wymagania dotyczące rzetelności w odniesieniu do skali werbalnej, niewerbalnej i ogólnej w związku z czym może być narzędziem niezwykle użytecznym, w szczególności dla orzecznictwa i opiniowania psychologicznego (tj. podejmowania decyzji na podstawie wyników badań). Jednoznaczne wyrażenie takiej rekomendacji, zaraz na pierwszych stronach podręcznika zwróciłoby uwagę na podstawowe mocne strony testu.
- Zaznacza się przy tym, że powyższe zalety są wyraźnie podkreślane przez wydawcę w innych kontekstach (np. w czasie szkoleń, konferencji).

Nie przedstawiono przykładu ilustrującego zasadę budowania przedziału ufności dla wyniku otrzymanego.

- W podręczniku technicznym (Roid, Jurek i in. 2017, s. 59) zawarto jedynie wzmiankę o możliwości budowania przedziałów ufności na podstawie wartości SEM. Informacje o przedziałach ufności są obliczone z wykorzystaniem wzoru na standardowy błąd estymacji (ibidem), oraz włączone do tabel „Wyników standaryzowanych (ilorazów) dla sum wyników przeliczonych” (Roid i in., 2017, s. 237-249), oddzielnie dla każdego wyniku przeliczonego. Jak komentuje to Wydawca „Na stronie 61 podręcznika technicznego umieszczono przykład liczenia przedziałów ufności z zastosowaniem SEE. Wskazanie użytkownikom „gotowych” przedziałów ułatwia diagnozę. Odpowiedź o przykładzie liczenia przedziałów ufności z zastosowaniem SEE warto uwzględnić.”. Jednak na str. 61 nie ma takiego przykładu i jest tam tylko ogólna informacja o błędzie pomiaru i przedziały ufności.
- Przyjęte rozwiązanie ma zalety - jest wygodne i intuicyjne z perspektywy użytkownika testu. Pozwala na zmniejszenie czasu niezbędnego na opracowanie wyników.

Przedstawiono przykład rekomendowanej interpretacji wyników testu.

- Przede wszystkim wprowadzono procedurę interpretacji testu w siedmiu krokach, omówiono z przykładami słownymi i graficznymi każdy z kroków (Roid i in., 2017, s. 139-152). Dodatkowy podrozdział poświęcono na interpretowanie różnic w wynikach (s. 154-158). Następnie przedstawiono pełną słowną interpretację profilu dziecka, z przywołaniem pierwszej strony jego wyników po analizie ilościowej (s. 159).

Podsumowanie opinii, czy autorzy testu zadbali o jego poprawną standaryzację:

1. Autorzy testu zadbali o jego poprawną standaryzację.
2. Istnieje jedno zastrzeżenie związane z niewystarczającą informacją (zawartą w podręczniku diagnosty) dotyczącą właściwej interpretacji zastosowanej skali standardowej.
3. Poza tym stwierdza się, że autorzy opisali w sposób precyzyjny procedurę analizy ilościowej oraz interpretacji wyników. Zasady tej analizy oraz interpretacji są uporządkowane, opatrzone przykładami i ilustracjami co ułatwia zrozumienie kluczowych punktów postępowania.
4. Ponadto, dla użytkowników testu, przewidziano możliwość stosowania SYSTEMU WSPARCIA DIAGNOSTY udostępnionego przez wydawcę testu na stronie: <http://id3855.abaks.pl/bateriaBinet/>, który wspiera diagnostę w poprawnym obliczaniu wyników i pozwala na wygenerowanie wstępnego opisu wyników do dalszych analiz i opracowań co ułatwia proces interpretacji i minimalizuje ryzyko pomyłek (np. obliczeniowych).
5. W toku lektury podręczników nasuwa się przy tym jedno istotne zastrzeżenie. Z uwagi na wagę analizy interpretacji różnic oraz stopień jej trudności wskazuje się, że decyzja o przeniesieniu tabel istotnych dla obliczania różnic wyników do podręcznika technicznego (Roid i in., s. 155) nie zachęcają użytkownika testu, w sposób wystarczający, do wykonania interpretacji różnic. Zdaniem Wydawcy „Układ treści jest spójny z pierwowzorem.” – jednak można oczekiwać nie tylko tłumaczenia oryginalnych materiałów dla diagnostów, ale – uwzględniając różnice – także spodziewać się ich adaptacji do warunków polskich. Problem jest oczywiście ogólniejszy i nie dotyczy jedynie tego testu.
6. Ponadto uzasadnienie dla przeniesienia tabel do obliczania różnic wyników do podręcznika technicznego (ibidem) budzi wątpliwości. Przywołanie faktu „że niektórzy diagnosty nie decydują się na przeliczanie różnic wyników” oraz zasadności koncentracji na głównym celu badania („skupienie się na przeprowadzaniu testu oraz odpowiednim wypełnieniu *Protokołu*”) dla przeniesienia tych tabel nie jest adekwatne – może zostać odebrane jako uzasadnienie dla rezygnowania z dokonywania analizy różnic w diagnozie indywidualnej. Zdaniem Wydawcy „Samo przeniesienie tabel istotnych dla obliczania różnic wyników budzi wątpliwości w mniejszym stopniu niż uzasadnienie dla tego przeniesienia.”
7. Interpretacja różnic, szczególnie w diagnozie funkcjonalnej, pełni kluczowe znaczenie, a diagnosty bez względu na cel badania i przyjęty model postępowania diagnostycznego powinni być przynajmniej zachęceni do jej wykonania.
8. W związku z powyższym proponuje się włączenie tabel do obliczania wyników do podręcznika diagnosty a przynajmniej położenie większego nacisku na zachęcanie użytkowników do stosowania analizy różnic. Rekomendacja ta jest analogiczna, do nacisków na dokonywanie różnic intra- i interprofilowych wyrażonych w podręcznikach innych skal wykonaniowych dostępnych na polskim rynku (np. IDS Skale inteligencji i Rozwoju, IDS-P Skale Inteligencji i Rozwoju dla Dzieci w Wieku Przedszkolnym, Skala Inteligencji Wechslera w wersji dla Dzieci-Wersja Zmodyfikowana).

Nie zilustrowano przykładem błędów w prezentowanym sposobie interpretacji wyników.

Całościowa ocena testu

Mocne strony testu:

1. Skala SB5 jest powszechnie stosowana na całym świecie.
2. SB 5 Pozwala obliczyć ilorazy inteligencji ogólny, niewerbalny, werbalny oraz dokonywać analizy profilowej, również dla grup specjalnych.
3. SB5 pozwala na dokonywanie diagnozy przesiewowej, z wykorzystaniem dwóch podtestów.
4. Ze względu na bardzo wysokie wskaźniki rzetelności wyniku IQ werbalnego, niewerbalnego i ogólnego SB5 może mieć bardzo dużą użyteczność w orzecznictwie i opiniowaniu.
5. Ze względu na wysokie wskaźniki rzetelności podskal SB5 ma dużą wartość w diagnozie indywidualnej intelektu oraz diagnozie funkcjonalnej zdolności poznawczych.
6. SB5 jest użyteczne do badania dzieci i młodzieży z grup specjalnych, ma opisane możliwości dostosowywania procedury do potrzeb dzieci z różnymi potrzebami wynikającymi z niepełnosprawności.
7. Narzędzie pozwala różnicować wyniki niskie (np. w obrębie niepełnosprawności intelektualnej, pozwala rozróżnić wyniki wskazujące na niepełnosprawność w stopniu lekkim i umiarkowanym) i wysokie, w obrębie szczególnych zdolności poznawczych.
8. Skala SB5 jest łatwa w administracji (ze względu na konstrukcje pomocy i protokołu badania)
9. Procedura badania testem jest przyjazna osobom badanym, dostosowana do ich faktycznych możliwości (dzięki testom kierującym i innym rozwiązaniom zastosowanym), różnorodna i ciekawa dla nich (dzięki testom cząstkowym, podziale na poziomy o różnej strukturze zadań).
10. Możliwości interpretowania (psychologicznego, ilościowego i jakościowego) zebranych za pomocą testu danych są duże.
11. Wydawca wspiera użytkownika testu w dokonywaniu właściwej interpretacji wyników m. in. poprzez platformę SYSTEM WSPARCIA DIAGNOSTY (<http://id3855.abaks.pl/bateriaBinet/>)
12. Dobrze, czytelnie przygotowane materiały, zarówno dla diagnozowanej osoby, jak i prowadzącego badanie które pomagają w sprawnym przeprowadzaniu testem i operowaniu pomocami. Wysoki poziom edytorski. Test umożliwia badanie przedszkolnych dzieci. Jest przyjazny, nie ma presji czasu, instrukcje wzbogacone są o przykłady rozwiązań. Powtórne badanie pozwala też na badanie możliwości badanego. Struktura testu pozwala nie tylko na diagnozę, ale rzeczywiście daje możliwość tworzenia programu dalszych oddziaływań terapeutycznych, czy stymulujących rozwój.

Inne mocne strony, wymagające dalszej oceny:

1. Narzędzie to wydaje się być użyteczne w diagnozie zdolności poznawczych małych dzieci (w tym od 2 r. ż. do 3 r. ż., tj. w grupie dzieci, których diagnoza psychometryczna intelektu przed wydaniem narzędzia była niemożliwa).

Najważniejsze słabe strony testu:

1. Wąski zakres przeprowadzonych badań walidacyjnych.
2. Test (w porównywaniu z innymi skalami wykonaniowymi dostępnymi na polskim rynku) wymagający relatywnie wiele czasu i wprawy w obliczaniu wyników i interpretacji.
3. Wymaga od użytkownika dużej koncentracji i skupienia przy badaniu pamięci roboczej w obszarze niewerbalnym.
4. Pełne badanie w grupie dzieci wczesnoszkolnych (7-12 lat) może trwać zdecydowanie dłużej niż deklarowane przez wydawcę 75 minut.
5. Dość powierzchownie opisane, przynajmniej w dostępnych podręcznikach („Podręcznik diagnosty” i „Podręcznik techniczny”) możliwości interpretacji jakościowej.
6. Brak odniesień do polskiej literatury i kontekstu kulturowo-społecznego. Podobnie jak wyżej, Wydawca pisze w swoim komentarzu do recenzji „Podstawa teoretyczna spójna z oryginalnymi materiałami SB5.”.

Ogólna ocena opisowa narzędzia.

1. Skale Inteligencji Stanford-Binet (Piąta Edycja) oceniamy jako narzędzie zaawansowane technicznie, o dobrych właściwościach psychometrycznych i szerokich możliwych zastosowaniach. Właściwe do psychologicznej diagnozy indywidualnej intelektu oraz psychologicznej diagnozy funkcjonalnej zdolności poznawczych. W tym zakresie pozbawione znaczących wad. Test pozwala na dokonanie szerokiej i zaawansowanej interpretacji ilościowej, jakościowej i psychologicznej mierzonych konstruktów. Nie może być przeznaczony do użytku przez niepsychologów.
2. W szczególności doceniamy opisanie możliwości dostosowań warunków badań do potrzeb różnych grup dzieci (zarówno ze względu na ich wiek, jak i specjalne potrzeby wynikające z niepełnosprawności). Zwiększa to możliwe zastosowania testu. Dostrzegamy również bardzo wysoką rzetelność wyników IQ werbalnego, niewerbalnego oraz ogólnego, która umożliwia stosowanie SB5, z powodzeniem, w orzecznictwie i opiniowaniu.
3. SB5 jest przy tym narzędziem z jednej strony wygodnym w administracji, a z drugiej wymagającym od użytkownika znacznego zaangażowania w zakresie oceny i interpretacji uzyskanych danych.
4. Przeprowadzenie badania testem, a zwłaszcza jego interpretacja w odniesieniu do skrajnych grup wiekowych, różnorodnych grup z zaburzeniami rozwojowymi, takimi chociażby jak spektrum autyzmu czy afazja, wymaga dodatkowych wyjaśnień, przykładów, przeprowadzania testu pod nadzorem, a także doświadczenia.
5. Przychylamy się do argumentacji wydawcy, że użytkownicy testu powinni przypisywać wagę do przeprowadzania analizy i interpretacji wyników pod nadzorem. Uznajemy też zasadność wprowadzenia ograniczeń dostępu do testu psychologom legitymujący się ukończeniem 5-

- letnich studiów magisterskich i wprowadzenia obowiązkowych szkoleń. Z uwagi na swoje mocne strony test ten trafi do poradni i ośrodków diagnostycznych, będzie wykorzystywany w orzecznictwie, jego wyniki mogą mieć daleko idące konsekwencje dla dalszych losów badanych osób i dlatego powinien być testem na swój sposób „elitarnym”, jeżeli chodzi o to, kto może go przeprowadzać.
6. Z przeprowadzonych badań indywidualnych wynika, że materiał testowy SB5 jest dobrze odbierany przez dzieci w wieku przedszkolnym i szkolnym, ze względu na swoją jakość, różnorodność i dostosowany, w procedurze testowania, poziom trudności.
 7. Mierzone konstrukty, konstrukcja testu oraz jego właściwości psychometryczne, z perspektywy użytkownika testu, są poprawnie opisane. Pomimo dostrzeżonych wad w prezentacji narzędzia (opisanych na kartach powyższej recenzji) podkreślić należy ogólnie wysoką jakość samego narzędzia i materiałów do niego dołączonych (w tym podręcznika diagnosty i technicznego).
 8. Wysoko oceniamy wartość tego testu. Wypełnia pewną lukę wśród narzędzi diagnostycznych badających możliwości intelektualne małych dzieci oraz osób niepełnosprawnych intelektualnie dla celów orzecznictwych. Zaletą jest możliwość obliczenia ekwiwalentów wieku dla uzyskanych wyników, co jest bardzo pomocne przy interpretacji i budowaniu programów dalszych oddziaływań. Wydaje się być stosunkowo łatwy do przeprowadzenia, jest przyjazny dla badanego, materiały testowe są czytelne i estetyczne, Zaletą testu jest też stosunkowo krótki czas trwania badania (również to, że można to badanie podzielić na etapy). Dobrze opisane są procedury obliczania wyników i interpretacji ilościowej. Test ten daje też bogate, możliwości interpretacyjne. Słuszna jest sugestia autorów, by dostęp do testu możliwy był po ukończeniu szkolenia, głównie ze względu na dogłębsze niż jest to możliwe na podstawie podręczników poznanie założeń teoretycznych testu i zasad oraz możliwości jego interpretacji.

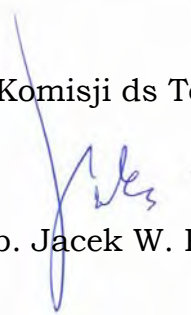
LITERATURA CYTOWANA

- Krasowicz-Kupis, G., Wiejak, K., Gruszczyńska, K. (2015). *Katalog metod diagnozy rozwoju poznawczego dziecka na etapie edukacji przedszkolnej i wczesnoszkolnej. Tom 1. Narzędzia dostępne w poradniach psychologiczno-pedagogicznych i szkołach*. Warszawa: Instytut Badań Edukacyjnych.
- Radtke, B. M., Sajewicz-Radtke, U. (2017). *Skala Inteligencji Stanford Binet 5 w praktyce diagnostycznej. Materiały szkoleniowe*. Gdańsk: Pracownia Testów Psychologicznych i Pedagogicznych.
- Roid, G. H., Jurek, P., Olech, M., Sajewicz-Radtke, U., Radtke, B. M. (2017). *Skale Inteligencji Stanford-Binet, Edycja Piąta. Podręcznik techniczny*. Gdańsk: Pracownia Testów Psychologicznych i Pedagogicznych.

Roid, G. H., Sajewicz-Radtke, U., Radtke, B. M. (2017). *Skale Inteligencji Stanford-Binet, Edycja Piąta. Podręcznik interpretacji*. Gdańsk: Pracownia Testów Psychologicznych i Pedagogicznych.

Roid, G. H., Sajewicz-Radtke, U., Radtke, B. M., Lipowska, M. (2017) *Skale Inteligencji Stanford-Binet, Edycja Piąta. Podręcznik Diagnosty*. Gdańsk: Pracownia Testów Psychologicznych i Pedagogicznych.

W imieniu Komisji ds Testów Psychologicznych Komitetu Psychologii PAN


prof. dr hab. Jacek W. Paluchowski