

Susan K. Johnsen
Linda Brown
Rita J. Sherbenou

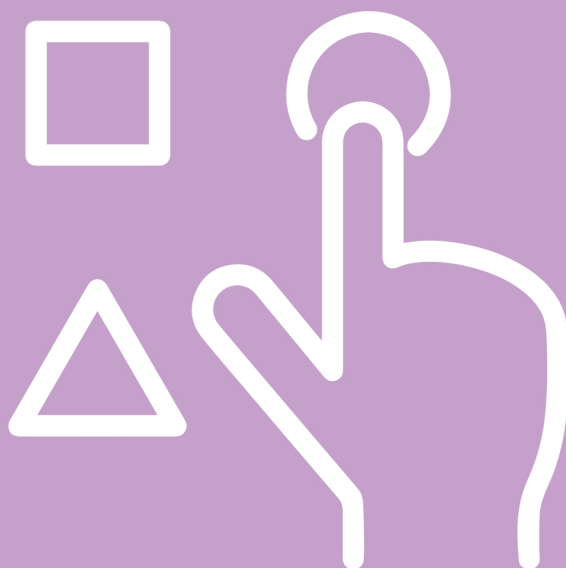
5–69



Urszula Sajewicz-Radtke
Bartosz M. Radtke
Ariadna Łada-Maśko

Niewerbalny Test Inteligencji TONI

przegląd badań 1982–2020



Susan K. Johnsen

Linda Brown

Rita J. Sherbenou

5–69



Urszula Sajewicz-Radtke

Bartosz M. Radtke

Ariadna Łada-Maśko

Niewerbalny Test Inteligencji TONI

przegląd badań 1982–2020



Pracownia Testów Psychologicznych i Pedagogicznych

Gdańsk 2021

Autorzy:

Linda Brown

Rita J. Sherbenou

Susan K. Johnsen

dr Urszula Sajewicz-Radtke

Pracownia Testów Psychologicznych i Pedagogicznych,
Specjalistyczna Poradnia Psychologiczno-Pedagogiczna
„Tęcza” w Gdańsku

dr Bartosz M. Radtke

Pracownia Testów Psychologicznych i Pedagogicznych,
Specjalistyczna Poradnia Psychologiczno-Pedagogiczna
„Tęcza” w Gdańsku

Ariadna Łada-Maśko

Instytut Psychologii, Uniwersytet Gdański

Recenzent:

dr hab. Bernadetta Izydorczyk, prof. UJ

Projekt graficzny:

Damian Chomątowski

dr Bartosz M. Radtke

Skład:

Agata Gruszczyńska

Pracownia Register

Redakcja językowa:

Olga Hollek

Wszystkie prawa zastrzeżone. Żadna część tego materiału chronionego prawem autorskim nie może być powielana w jakiegokolwiek formie lub za pomocą jakichkolwiek środków elektronicznych, czy mechanicznych, w tym poprzez kserowanie lub nagrywanie, ani przez jakiegokolwiek systemy przechowywania lub wyszukiwania informacji bez uprzedniej pisemnej zgody właściciela praw autorskich.

© Copyright of the original English edition 2010 by PRO-ED, Inc., USA.

© Copyright of the Polish edition 2021 by Pracownia Testów

Psychologicznych i Pedagogicznych – Bartosz M. Radtke,

Gdańsk, Poland, and PRO-ED, Inc., USA.

All rights reserved.

Wydanie pierwsze ♦ ISBN 978-83-65453-74-7

Wydrukowano w Rzeczypospolitej Polskiej

Wydawca:

Pracownia Testów Psychologicznych i Pedagogicznych

ul. Czarnieckiego 5A poziom 1, 80-239 Gdańsk

tel./faks: 58 340 38 29 ♦ poczta@pracowniatestow.pl

pracowniatestow.pl ♦ facebook.com/ptpip

Spis treści

Wstęp • 5

1. Przegląd recenzji • 7

Recenzje TONI • 7

Recenzje TONI-2 • 9

Recenzje TONI-3 • 11

2. Rzetelność • 17

Błąd losowy treści • 17

Błąd losowy czasu • 23

Błąd wynikający z różnic pomiędzy osobami
przyznającymi punkty • 25

Podsumowanie wyników rzetelności • 25

3. Trafność • 29

Trafność kryterialna/prognostyczna • 29

Trafność konstruktu • 43

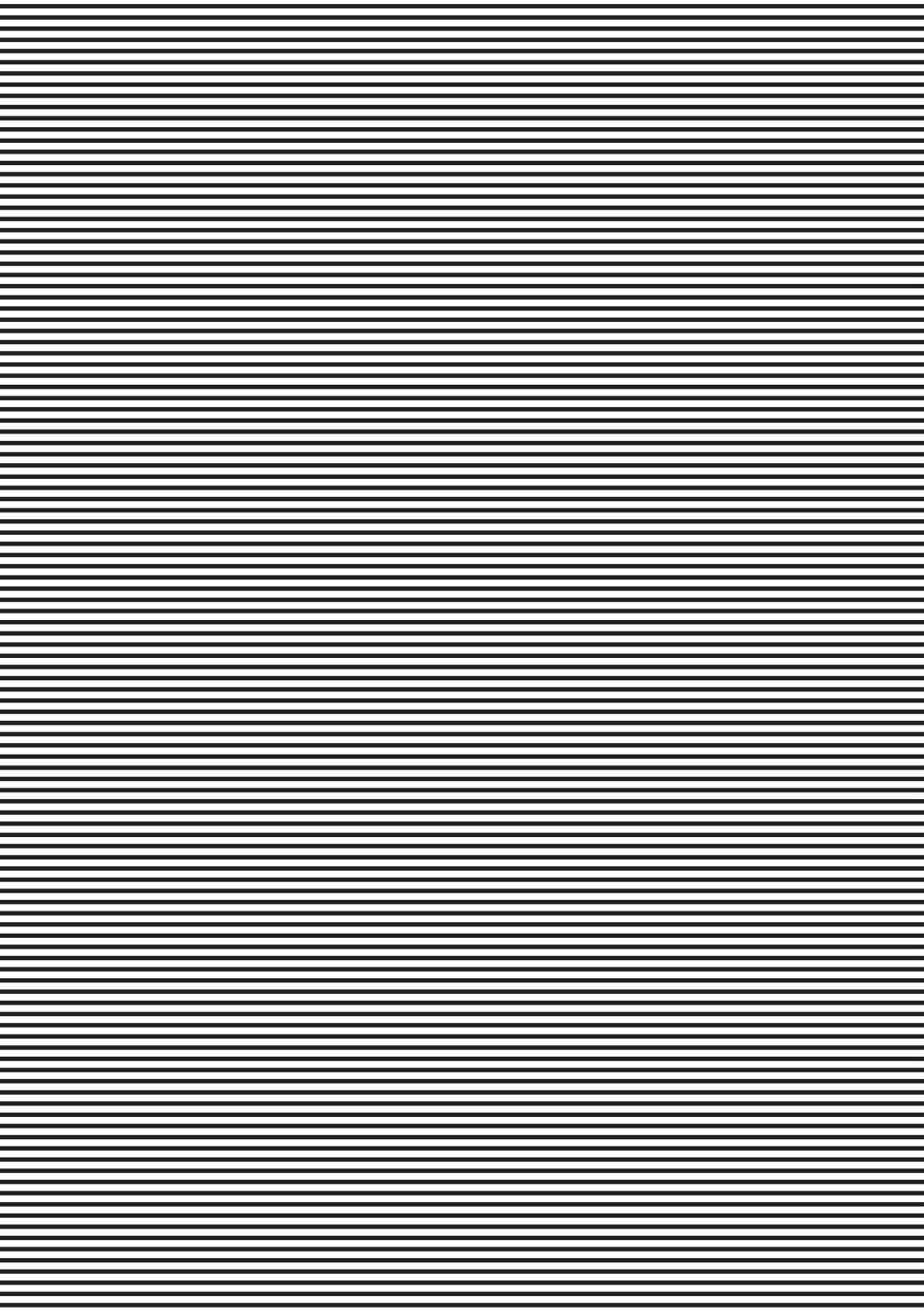
4. Nowe amerykańskie wydanie TONI-4 • 73

Zalety TONI-4 • 75

Ograniczenia TONI-4 • 76

Bibliografia • 79

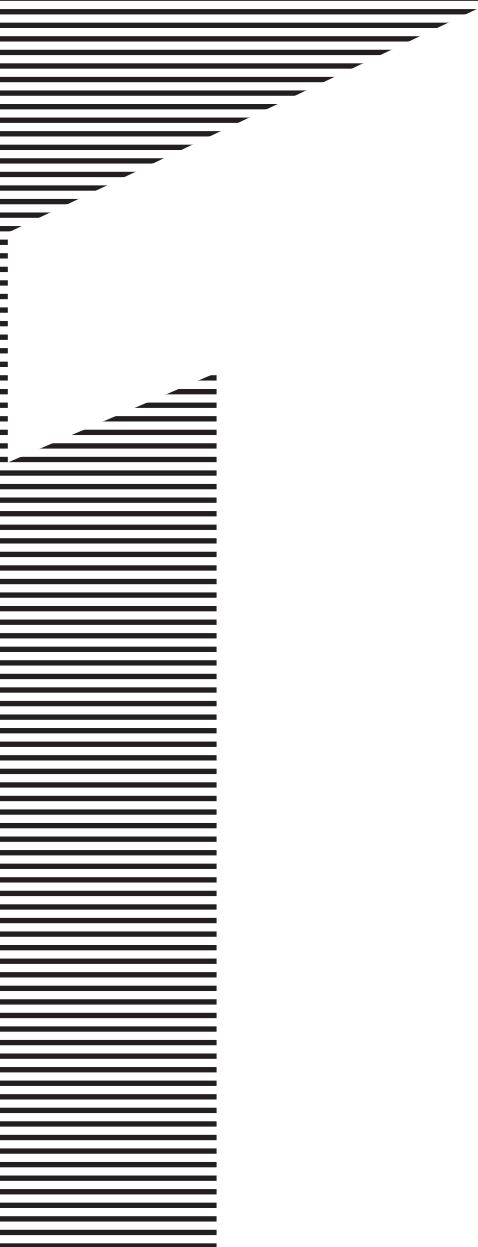
Załączniki • 83



Test Inteligencji Niewerbalnej (TONI) opracowano w 1982 roku w celu oceny inteligencji dzieci i dorosłych, których zdolności językowe i motoryczne mogłyby negatywnie wpłynąć na wykonanie przez nich tradycyjnych testów inteligencji lub przeszkodziłyby im w zrozumieniu zadań testowych i udzieleniu na nie odpowiedzi. Na przestrzeni lat zgromadzono znaczną liczbę uwag, które przeanalizowano w niniejszym kompendium.

Wyniki badań odnoszących się do obecnego wydania testu TONI można znaleźć w podręczniku do polskiej wersji metody TONI-4 (Brown, Sherbenou, Johnsen, Sajewicz-Radtke, Radtke, 2021). Dane dotyczące poprzednich wydań zebrano tutaj. Wcześniejsze badania mają duże znaczenie dla TONI-4, ponieważ pozycje testowe we wszystkich wydaniach są praktycznie identyczne. W TONI-2 do każdej wersji testu dodano pięć pozycji, a w TONI-3 z każdej wersji testu usunięto po 10 pozycji. W TONI-4 przywrócono 10 usuniętych pozycji z TONI-2 oraz dodano 5 nowych pozycji, tak że ogólna liczba w każdej z wersji testu wynosiła 60. Innymi słowy, ponad 85% pozycji we wszystkich czterech wydaniach testu pozostało niezmiennych.

W niniejszej publikacji przedstawiamy zestawienie badań z lat 1982–2009 za tekstem oryginalnym (Johnsen, Brown, Sherbenou, 2010), dodatkowo rozszerzając je o krótki przegląd wybranych badań z lat 2010–2020.



W kolejnej części omawiamy recenzje dotyczące poprzednich trzech wydań testu: TONI (1982), TONI-2 (1990), TONI-3 (1997).

Recenzje TONI

Pierwsze wydanie *Test of Nonverbal Intelligence* – TONI (Brown, Sherbenou, Johnsen, 1982) opracowano dla osób w wieku od 5;0 do 85;11 lat. Dwie równoważne wersje testu, obejmujące po 50 pozycji, znormalizowano w grupie 1929 osób badanych zamieszkających w 28 stanach. Test nie był ograniczony czasowo, a jego wykonanie zajmowało około 15 minut. Diagnosta pokazywał instrukcję, a osoba badana wskazywała odpowiedzi, co eliminowało potrzebę słuchania, mówienia, czytania czy pisanie w prezentowaniu lub odpowiadaniu na materiał testowy. Do wyjaśnienia osobie badanej zasad rozwiązywania testu zastosowano sześć pozycji próbnych. Każda pozycja testowa składała się z zestawu abstrakcyjnych figur, w których brakowało jednej lub więcej figur. Zadanie polegało na zidentyfikowaniu zależności pomiędzy figurami i wyborem prawidłowej odpowiedzi spośród czterech lub sześciu opcji.

Umiejętnością poznawczą, którą od początku badał TONI, było rozwiązywanie problemów – silny i ogólny komponent inteligencji, stosunkowo łatwy w pomiarze z wykorzystaniem abstrakcyjnych (figuralnych) treści, umożliwiających zredukowanie kulturowego

i językowego obciążenia testu. TONI nie został zaprojektowany w celu zastąpienia szerokich baterii testów inteligencji, ale raczej w celu zapewnienia alternatywnej metody oceny, wówczas gdy umiejętności poznawcze, językowe oraz motoryczne osoby badanej uniemożliwiały zastosowanie tradycyjnych testów.

W swoich analizach recenzenci testu byli ogólnie konstruktywni i prezentowali pozytywne nastawienie do narzędzia:

TONI ma wiele pozytywnych cech, jest m.in. obiektywny, łatwy w przeprowadzeniu i można go wykonać we względnie krótkim czasie. (Roberts, 1990, s. 100)

TONI będzie szczególnie pomocny w badaniach przesiewowych funkcjonowania intelektualnego uczniów w przypadkach, w których użycie narzędzi werbalnych będzie niewskazane lub niemożliwe. (Salvia, Ysseldyke, 1988, s. 197)

Wydaje się być obiecującym narzędziem pomiaru niewerbalnej umiejętności rozwiązywania problemów. (McLoughlin, Lewis, 1986, s. 167)

Formuła TONI oparta na rozwiązywaniu problemów zredukowała kulturowe obciążenie, o które łatwo w większości standaryzowanych testów inteligencji. Dodatkowo, niewerbalna treść oraz możliwość wskazywania odpowiedzi uczyniły TONI o wiele bardziej obiektywną miarą umiejętności intelektualnych dla mniejszości etnicznych oraz osób niepełnosprawnych ruchowo niż wiele innych obciążonych werbalnie miar. (Harrington, 1985, s. 791)

Wygłąda na to, że test standaryzowano na starannie dobranej grupie, reprezentatywnej dla danej populacji [...]. Wielkość korelacji dowodzi równoczesnej trafności TONI w badanej próbie. (Clark, 1985, s. 1581)

Opracowanie TONI pod względem psychometrycznym wydaje się zadowalające. Trafność fasadowa spełnia zamierzony niewerbalny cel. Pozycje są dobrze zeskalowane pod względem trudności i psychologicznej złożoności. (Mayo, 1985, s. 1582–1583)

Recenzenci rekomendowali potencjalne ulepszenia, takie jak wyjaśnienie instrukcji do przeprowadzania, dokładniejsze przeanalizowanie koncepcji teoretycznych, na których opiera się test, rozszerzenie zarówno zasad bazowych, jak i sufitu, zwiększenie liczebności próby normalizacyjnej oraz zwiększenie liczby badań nad trafnością i rzetelnością.

Statystyczne właściwości TONI opiniowano przy użyciu obiektywnych kryteriów opracowanych w *A Consumer's Guide to Tests in Print* (Hammill, Brown, Bryant, 1989), a ogólna ocena brzmiała: „rekomendowany”. Trafność testu otrzymała ocenę „wysoce rekomendowany”. Cechy normalizacyjne i spójność wewnętrzna zostały ocenione jako „rekomendowane”. Jednak stabilność oceniono jako „nierekomendowana”, z powodu zastosowanej wówczas metodologii.

Recenzje TONI-2

Te wcześniejsze uwagi krytyczne były przekonujące i większość z nich została uwzględniona podczas rewizji testu w 1990 roku. Szczególną uwagę poświęcono wielkości i składowi próby normalizacyjnej, badaniom potwierdzającym stabilność testu, jego trafności w odniesieniu do populacji o specjalnych potrzebach, a także strukturze czynnikowej testu. Zakres testu został rozszerzony, ale próby rozszerzenia zasad bazowych nie zakończyły się sukcesem prawdopodobnie dlatego, że konstrukt rozwiązywania problemu mierzony w sposób stosowany w TONI nie jest po prostu rozwojowo stabilny w przypadku dzieci poniżej 5. roku życia. Z powodu podobieństwa pomiędzy TONI i TONI-2 oryginalna próba normalizacyjna została zachowana i wzmocniona poprzez dodanie 835 nowych badanych. Ta nowa próba (N = 2764) odpowiadała cechom demograficznym populacji USA w momencie publikacji. Do górnego zakresu każdej wersji TONI-2 dodano po pięć pozycji. Pozycje trudniejsze były odpowiedzią na krytykę, że utalentowani i zdolni badani mogą nie być w stanie uzyskać prawdziwego wyniku w TONI, odnoszącego się do ich rzeczywistych zdolności intelektualnych.

Stabilność w TONI oceniono przy użyciu bezpośredniej metody test-retest w porównaniu dwóch alternatywnych wersji, co jest tak

naprawdę miarą spójności wewnętrznej, a nie stabilności. Kwestia ta została sprostowana w TONI-2. Rzetelność TONI-2 została ponadto zbadana w populacji osób o specjalnych potrzebach, które nie mówiły po angielsku lub miały ograniczoną biegłość w posługiwaniu się językiem angielskim oraz u których zdiagnozowano dysleksję, trudności w nauce, uraz głowy oraz ponadprzeciętne zdolności intelektualne. Badanie to uzupełniło poprzednie badanie rzetelności testu uwzględniające osoby z niepełnosprawnością intelektualną, trudnościami w nauce oraz uszkodzeniem słuchu.

Oryginalny podręcznik diagnosty TONI cytował osiem badań nad trafnością – trzy dotyczące populacji ogólnej oraz pięć odnoszących się do różnorodnych populacji o specjalnych potrzebach. W TONI-2 odnotowano 20 badań, w tym zarówno pracę autorek, jak i niezależne badania opublikowane w recenzowanych czasopiśmie naukowych. Zaprezentowano istotne dane dotyczące dzieci, które nie mówiły po angielsku lub miały ograniczoną znajomość języka angielskiego, a w szczególności dzieci, których pierwszym językiem był hiszpański lub dialekt rdzennych mieszkańców Ameryki. Dodano wiele danych korelacyjnych oraz z analizy regresji, a także analizę czynnikową, która nie wygenerowała silnych podczynników, co podtrzymało hipotezę, że TONI jest testem jednoczynnikowym. Uzyskano dane wskazujące na precyzję TONI-2 w różnicowaniu grup osób badanych o szczególnych cechach intelektualnych i językowych oraz ich związek z tak istotnymi zmiennymi, jak wiek chronologiczny, osiągnięcia naukowe, predyspozycje intelektualne i inteligencja ogólna.

Tak jak swój poprzednik, TONI-2 został poddany obiektywnej ocenie w *A Consumer's Guide to Tests in Print* (Hammill, Brown, Bryant, 1992). Ogólna ocena dla TONI-2 nadal brzmiała „rekomendowany”. Stabilność, najniżej oceniany komponent TONI, zyskał lepszą ocenę – „rekomendowana” w TONI-2. Podobnie było z cechami normalizacyjnymi. Dowody empiryczne trafności oraz spójności wewnętrznej otrzymały najwyższą ocenę „wysoce rekomendowane”.

McLoughlin i Lewis (1994) określili TONI-2, tak jak jego poprzednią wersję, „obietującym narzędziem pomiaru niewerbalnych umiejętności rozwiązywania problemów” (s. 198). Murphy

(1992) przyznał, że twórcy TONI-2 „zajęli się dobrze [wcześniej] krytykowanymi kwestiami” i polecił podręcznik testowy jako „jasny i kompletny” (s. 970). Powołując się na ulepszenia techniczne i dotyczące formuły, które zostały wprowadzone w ramach rewizji, Aiken (1996) uznał TONI-2 za „jeden z najpopularniejszych testów inteligencji używanych w szkołach publicznych” (s. 204). Murphy (1992), który przyznał, że TONI-2 odnosi się do większości swojej krytyki, dostrzegł tylko jedną „prawdziwą uwagę krytyczną” dotyczącą procedury przeprowadzania testu w małej grupie, która została wyeliminowana w TONI-3.

Natomiast Watson (1992) nie zgadzał się ze swoimi kolegami i orzekł, że „większość podniesionej krytyki TONI była nadal aktualna w przypadku TONI-2”. Uważał, że „procedury standaryzacyjne powinny zostać poprawione, zwłaszcza w zakresie nawiązywania relacji, wyjaśniania celu i zastosowania testu osobom badanym oraz instrukcji bezsłownych” (s. 970). Zakwestionował potrzebę „kolejnego testu inteligencji”, jakim był TONI-2 (s. 972). Sodowsky, Gonzalez i Kuo-Jackson (1998) kwestionowali „brak stronniczości w zakresie języka angielskiego” TONI-2, opierając się na istotnej korelacji testu „z podtestem Język w baterii osiągnąć SRA” (.74 dla wersji A i .78 dla wersji B, s. 259). Dodali, że „TONI-2 prawdopodobnie nie jest lepszą wersją poprzedniej wersji testu” (s. 260).

Biorąc tę krytykę do serca, twórcy podręcznika TONI-2 wyjaśnili filozoficzne i teoretyczne podstawy TONI oraz wskazali, że formuła TONI wraz z jego solidnymi właściwościami psychometrycznymi, wskazując, że to naprawdę wyjątkowy test spośród głównych testów inteligencji. Kamphaus (1993) uznał, że „podręcznik pozostawia trochę do życzenia pod względem szczegółów” (s. 333), ale nie wyjaśnił, jakie „szczegóły” miał na myśli.

Recenzje TONI-3

TONI-3 w bardziej klarowny sposób określał zasady oceny inteligencji niewerbalnej i dokładnie omawiał filozoficzne i teoretyczne podstawy testu dla ogólnej oceny inteligencji i umiejętności. Zarówno jakość, jak i liczba skumulowanych badań nad rzetelnością i trafnością wzrosła, dając coraz bogatszy obraz wartości testu, jego

użyteczności i praktyczności. Pozycje, w tym ich tradycyjne wskaźniki oraz nowe miary stronniczości, zostały poddane ponownej analizie ze szczególnym uwzględnieniem cech psychometrycznych. Wyeliminowano 10 pozycji z każdej wersji testu, co nie wpłynęło na ogólną techniczną jakość TONI. Obie wersje testu w TONI-3, zawierające po 45 pozycji, można było szybciej ukończyć, a nie ucierpiały na tym rzetelność i trafność.

Dane normalizacyjne zebrane do TONI-3 dały doskonałe wyniki w zakresie reprezentatywności demograficznej. Bardziej zaawansowane technologie umożliwiły zastosowanie lepszych metod badania stronniczości pod względem rasy, pochodzenia etnicznego oraz płci. Przedstawiono badania dokumentujące rzetelność testu w przypadku dorosłych, a narzędzia komputerowe, niedostępne w 1982 roku, umożliwiły zaprojektowanie odświeżonych rysunków do testu.

Zmiany te zostały dobrze przyjęte przez krytyków. Atlas (2001) w swojej recenzji do *Mental Measurements Yearbook* Instytutu Buros uważał za „niezwykłe, że test charakteryzuje się dobrze skonstruowanymi normami, a jednocześnie stawia tak minimalne wymagania wobec osoby badanej, że prawie wszystkie populacje, z wyjątkiem niewidomych, mogą być oceniane pod kątem inteligencji postrzeganej jako zdolność do rozumowania abstrakcyjnego [...] TONI-3 zapewnia właśnie taką skalę” (s. 1259). Inni recenzenci zgadzali się, że test spełniał użyteczną rolę w przesiewowym badaniu umiejętności myślenia niewerbalnego (McLoughlin, Lewis, 2008; Sattler, 2001), szczególnie w przypadku osób nie posługujących się jęz. angielskim, osób z ograniczeniem ruchowym (DeThorne, Schaefer, 2004) oraz z uszkodzeniem słuchu czy deficytami neurologicznymi (Friedman i in., 2003). Pierangelo i Giuliani (1998) dodali, że „test jest szczególnie odpowiedni dla osób z licznymi niepełnosprawnościami” (s. 65).

Ten podręcznik to „zbiór informacji historycznych, przeglądów badań, refleksji na temat ograniczeń, opisanych w sposób dostępny dla studentów i praktyków po kursie z zakresu psychometrii” (Atlas, 2001, s.1260). Podręcznik prezentuje „kompleksowy przegląd testów inteligencji i skal niewerbalnych oraz obejmuje kilka

interesujących zagadnień, jak na przykład zastosowanie norm dla populacji specjalnych” (Strauss, Sherman, Spreen, 2006, s. 271).

Kilku krytyków zwraca uwagę na komputerowo stworzone rysunki do testu (Baron, 2004; Reynolds, Fletcher-Janzen, 2004). Reynolds i Fletcher-Janzen (2004) stwierdzili, że „rysunki w TONI-3 zostały znacznie ulepszone” (s. 943).

Recenzenci określili dane normalizacyjne z TONI-3 jako „imponujące” (Kroese, 2003, s. 620) oraz „kompleksowe, inkluzywne i dobrze skonstruowane” (Strauss i in., 2006, s. 269). Wskazywali także na znaczącą liczbę zgromadzonych badań wspierających rzetelność i trafność TONI-3 (DeMauro, 2001; Gregory, 2000; Hoge, 1999; Kroese, 2003; Reynolds, Fletcher-Janzen, 2004; Salvia, Ysseldyke, Bolt, 2007; Strauss i in., 2006). Dwudziestoletnie dane „dokumentują związek testu z innymi miarami inteligencji, ze skalami osiągnięć i osobowości, jego skuteczność w prawidłowym różnicowaniu grup oraz jego strukturę czynnikową [...]. [Ponadto] zbadano potencjalną stronniczość elementów testowych i stwierdzono, że jest ona nieistotna” (Reynolds, Fletcher-Janzen, 2004, s. 943).

Zidentyfikowano także obszary wymagające poprawy. Atlas (2001) miał na przykład obawy, że osoby szczególnie uzdolnione i utalentowane oraz pełnosprawne mogą uważać niewerbalne instrukcje za „niepotrzebnie uduziwione” (s. 1259). Kwestie rzetelności zostały podniesione przez kilku recenzentów; zwłaszcza rzetelność test-retest w starszych grupach wiekowych (Athanasίου, 2000), rzetelność alternatywnych wersji (Sattler, 2001; Strauss i in., 2006) oraz spójność wewnętrzną dla niektórych zakresów wiekowych (Flanagan, Harrison, 2005). Strauss i in. (2006) rekomendowali także zbadanie roli efektów praktycznych.

Kilku recenzentów zaleciło dodatkowe badania nad trafnością u dzieci wykazujących specyficzne zaburzenia językowe (Kroese, 2003), osób utalentowanych (Sattler, 2001) oraz osób z podgrup zróżnicowanych pod względem rasowym, etnicznym i kulturowym (Strauss i in., 2006). Salvia i in. (2007) uznali, że dowody potwierdzające trafność testu są nadmiernie uzależnione od jego związku z zewnętrznymi środkami kryterialnymi. Kilku

recenzentów zasugerowało również, że tabele mogłyby zostać poprawione pod kątem przydatności klinicznej poprzez raportowanie średnich i odchyłeń standardowych w odniesieniu do równoległych badań nad trafnością (Sattler, 2001; Strauss i in., 2006). Recenzenci chcieli zobaczyć dane odnoszące się do związków pomiędzy TONI-3 a nowszymi wersjami skal Wechslera oraz poprzednich wersji TONI (Strauss i in., 2006). Ostatnia obawa dotyczyła potencjalnego efektu podłogi w przypadku dzieci z autyzmem (Handleman, Delmolino, 2005) lub w przypadku niepełnosprawności intelektualnej (Sattler, 2001; Strauss i in., 2006).

2

Rzetelność =

Istnieją trzy ważne źródła błędów w testach, takich jak TONI: błędy losowe treści, błędy losowe czasu oraz błędy wynikające z różnic pomiędzy osobami przyznającymi punkty. Badania takie określa się zazwyczaj mianem współczynników rzetelności. Współczynniki rzetelności wykorzystane do tego celu muszą osiągnąć wartość co najmniej .80, a pożądane współczynniki to .90 lub powyżej (Aiken, Groth-Marnet, 2006; Salvia i in., 2007). Kolejną część opracowania zajmuje się testem TONI w odniesieniu do powyższych trzech źródeł błędu.

Błąd losowy treści

Błędy losowe treści odnoszą się do stopnia jednorodności pomiędzy pozycjami testu. Celem testu jest pomiar pewnych cech, zdolności lub treści. Błąd testu jest niewielki, jeśli elementy testowe są ze sobą powiązane w jakimś zakresie przy pomiarze tych cech. Gdy dana pozycja nie jest powiązana z innymi pozycjami w teście, prawdopodobnie mierzy ona inną treść niż pozostałe pozycje, co zwiększa błąd losowy treści.

Błąd losowy treści wpływa na spójność wewnętrzną testu. Do oszacowania spójności wewnętrznej testu stosuje się kilka miar statystycznych, różniących się pod względem wspólnego współczynnika korelacji. *Współczynnik alfa* (Cronbach, 1951) wskazuje

stopnień, do którego pozycje testu korelują ze sobą i jest jedną z bardziej rygorystycznych procedur wykorzystywanych do oceny spójności wewnętrznej narzędzia. *Rzetelność alternatywnych wersji* (kiedy jedna wersja wykonana jest bezpośrednio po innej) to kolejna procedura wykorzystywana do oceny spójności wewnętrznej. W tej części opisano badania spójności wewnętrznej TONI wśród ogólnej populacji oraz populacji o specjalnych potrzebach przy użyciu współczynnika alfa oraz metody alternatywnych wersji.

Próby populacji ogólnej

Trzy badania sprawdzały spójność wewnętrzną TONI przy wykorzystaniu prób złożonych z przekroju osób z populacji ogólnej. Każde z tych badań wykorzystywało współczynnik alfa oraz metodę alternatywnych wersji (bezpośrednie przeprowadzanie) do oceny spójności wewnętrznej przy użyciu prób normalizacyjnych z poprzednich wydań TONI. Dane te podsumowano w tabeli 1.

Średnie współczynniki alfa zawarte w tabeli 1 są najwierniejszymi ocenami rzetelności testu, a ich wielkości (.80 i .90) pokazują, że obie wersje TONI mogą być używane z pełnym przekonaniem. Współczynniki dla wersji A i wersji B są praktycznie identyczne, co jest kolejnym dowodem na to, że obie wersje są równoważne.

TABELA 1

Badania spójności wewnętrznej (pominięto dziesiętne)

Przedział wiekowy (lata)	Badanie 1 L. Brown i in.(1982)				Badanie 2 L. Brown i in. (1990)				Badanie 3 L. Brown i in. (1997)			
	Współczynnik alfa n = 400		Alternatywne wersje (bezpośrednie przeprowadzenie) N = 1888		Współczynnik alfa n = 900		Alternatywne wersje (bezpośrednie przeprowadzenie) N = 2119		Współczynnik alfa n = 3451		Alternatywne wersje (bezpośrednie przeprowadzenie) N = 3451	
	Wersja A	Wersja B	r _{AB}	Wersja A	Wersja B	r _{AB}	Wersja A	Wersja B	Wersja A	Wersja B	r _{AB}	r _{AB}
6	78	78	82	81	85	86	89	91	89	91	85	85
7			82	95	94	84	93	93	93	93	79	79
8	-	-		94	92	85	90	91	90	91	79	79
9	-	-	78	93	94	81	89	90	89	90	85	85
10	-	-		93	96	85	89	89	89	89	79	79
11	-	-		95	96	82	91	92	91	92	85	85
12	-	-	85	97	97	85	93	93	93	93	86	86
13	88	90		95	95	84	93	93	93	93	86	86
14				95	97	81	93	94	93	94	79	79

TABELA 1
Badania spójności wewnętrznej (pominięto dziesiątne)

Przedział wiekowy (lata)	Badanie 1 L. Brown i in.(1982)			Badanie 2 L. Brown i in. (1990)			Badanie 3 L. Brown i in. (1997)		
	Współczynnik alfa n = 400	Alternatywne wersje (bezpośrednie przeprowadzenie) N = 1888		Współczynnik alfa n = 900	Alternatywne wersje (bezpośrednie przeprowadzenie) N = 2119		Współczynnik alfa n = 3451	Alternatywne wersje (bezpośrednie przeprowadzenie) N = 3451	
	Wersja A	Wersja B	r _{AB}	Wersja A	Wersja B	r _{AB}	Wersja A	Wersja B	r _{AB}
15	-	-		96	97	90	93	93	76
16	-	-	88	96	97	86	93	93	83
17	-	-		96	96	87	92	92	74
18	-	-		97	97	90	91	91	75
19	-	-	-	96	97	92	94	95	90
20-29	-	-	84	98	98	85	94	95	90
30-49	91	92	80	98	99	82	94	94	88
50-85	91	91	95	97	98	95	95	95	89
Średnia	88	89	85	96	96	86	92	93	84

Uwaga: r_{AB} – korelacja pomiędzy wersją A i wersją B.



Próby populacji o specjalnych potrzebach

Testy nie są jednakowo rzetelne w przypadku wszystkich podgrup. Ponieważ TONI używany jest zazwyczaj w przypadku diagnoz klinicznych, wiedza, czy test jest rzetelny w przypadku określonych populacji o specjalnych potrzebach, jest niezwykle istotna. Na przestrzeni lat zebrano pokaźną liczbę takich badań dla TONI (patrz podsumowanie w tabeli 2). Dla udogodnienia, badania podzielono na trzy kategorie: (1) podgrupa pochodzenie etniczne/płeć (2) podgrupa hiszpańskojęzyczna oraz (3) grupy specjalne. Uśrednione współczynniki rzetelności dla tych podgrup znajdują się w prawej kolumnie tabeli. Z wyjątkiem podgrupy osób z autyzmem, te uśrednione współczynniki mają ciągle dużą wielkość, a aż 12 z 16 współczynników wynosi .90 lub więcej, podobnie jak w przypadku tych wymienionych w tabeli 1 dla populacji ogólnej. W kwestii podgrupy z autyzmem, Edelson, Schubert, Edelson (1998) stwierdzili, że współczynnik rzetelności .60 jest do zaakceptowania, zważywszy na fakt, że zachowanie osoby z autyzmem może znacząco się różnić z dnia na dzień.

TABELA 2

*Spójność wewnętrzna w przypadku wybranych podgrup
(pominięto dziesiętne)*

Podgrupa	Badanie	N	Współczynnik alfa		Alternatywne wersje (bezpośrednie przeprowadzenie) r_{AB}	Średnia
			Wersja A	Wersja B		
Etniczna/płeć						
Latynosi	L. Brown i in. (1997)	386	95	95	–	95
Afro- amerykanie	L. Brown i in. (1997)	444	94	94	–	94
Kobiety	L. Brown i in. (1997)	1738	95	95	–	95
Mężczyźni	L. Brown i in. (1997)	1713	96	96	–	96
Hiszpańskojęzyczna						
Meksyk	Garcia (1988)	1652	88	86	–	87
Meksyk	Garcia (1988)	1652	–	–	90	90
Chile	Prado (1989)	118	91	–	–	91
Grupy specjalne						
Autyzm	Edelson i in. (1998)	76	–	–	60	60
Głuchota	L. Brown i in. (1997)	46	–	–	87	87
Głuchota	L. Brown i in. (1997)	76	95	95	–	95
Trudności w nauce	L. Brown i in. (1982)	27	–	–	92	92
Trudności w nauce	L. Brown i in. (1997)	287	94	94	–	94
Niepełnospr. intelekt.	L. Brown i in. (1982)	10	–	–	81	81
Uzdolnienia	L. Brown i in. (1990)	49	92	–	–	92
Uzdolnienia	L. Brown i in. (1990)	65	91	–	–	91
Uzdolnienia	L. Brown i in. (1997)	134	92	92	–	92

Uwaga: r_{AB} – współczynnik korelacji pomiędzy wersją A i wersją B.

Błąd losowy czasu

Stabilność oceniana jest poprzez badania błędu losowego czasu, który zajmują się stopniem, do jakiego wynik testu pozostaje stabilny na przestrzeni czasu. Celem jest stworzenie stabilnych narzędzi, których wyniki mogą służyć za podstawę decyzji diagnostycznych. Jeśli test nie jest stabilny pomiędzy jednym okresem testowania a drugim, błąd losowy będzie dość duży, a współczynniki stabilności testu będą niskie.

Istnieją dwa główne sposoby badania stabilności. Pierwszą jest odroczonego czasu metoda *test-retest*, w której test jest przeprowadzany w tej samej grupie uczniów dwukrotnie w odstępie określonego czasu, a dwa zestawy uzyskanych wyników są skorelowane. Druga – *alternatywne wersje rzetelności* – korzysta z równoważnych alternatywnych wersji testu. Upływ czasu pomiędzy testowaniem może się wahać od kilku dni do (rzadko) kilku lat. Wysokie współczynniki korelacji test-retest lub alternatywnych wersji (odroczonego czasu przeprowadzanie) odzwierciedlają stabilne wyniki testu w okresie interwencji. Wyniki badań stabilności z wykorzystaniem TONI-2 i TONI-3 podsumowano w tabeli 3.

McGhee i Lieberman (1990) wykorzystali prostą formułę test-retest, przeprowadzając wersję A i wersję B w 3-tygodniowym odstępie u 50 losowo wybranych pierwszo- i drugoklasistów. Rzetelność test-retest dla wersji A wyniosła .89, a dla wersji B .83.

L. Brown, Sherbenou i Johnsen (1990) badali stabilność TONI-2, wykorzystując metodę odroczonego czasu wersji alternatywnych. Wersję A przeprowadzono u 39 osobach, a wersję B 1 tydzień później. Szacowana stabilność wyniosła .85.

L. Brown, Sherbenou i Johnsen (1997) badali stabilność TONI-3 przy wykorzystaniu metody test-retest. Liczba respondentów wynosiła 170 i obejmowała trzy grupy wiekowe: 13, 15 oraz 19–40 lat. Obie wersje testu przeprowadzono w odstępie 1 tygodnia. Uzyskane współczynniki test-retest zawarte w tabeli 3 wynoszą od .89 do .94 i wskazują na bardzo niewielki błąd pomiaru.

Podsumowując, otrzymane współczynniki stabilności TONI-2 i TONI-3 jasno dowodzą wysokiego stopnia stabilności obu wersji testu.

TABELA 3
Współczynniki stabilności (pominięto dziesiątne)

Badanie	N	Wiek/klasa badanego	Metoda		Współczynnik rzetelności		
			Test–retest	Alternatywne wersje (odroczone przeprowadzanie)	r_{AA}	r_{BB}	r_{AB}
McGhee Lieberman (1990)	50	Klasy 1, 2	3 tygodnie	–	89	83	–
L. Brown i in.(1990)	39	8–16 lat	–	1 tydzień	–	–	85
L. Brown i in.(1997)	43	13 lat	1 tydzień	–	92	94	–
L. Brown i in.(1997)	87	15 lat	1 tydzień	–	92	93	–
L. Brown i in.(1997)	40	19–40 lat	1 tydzień	–	90	89	–
Średni współczynnik					92	92	85

Uwaga: r_{AA} – współczynnik korelacji pomiędzy wersją A w czasie 1 i wersją A w czasie 2; r_{BB} – współczynnik korelacji pomiędzy wersją B w czasie 1 i wersją B w czasie 2; r_{AB} – współczynnik korelacji pomiędzy wersją A i wersją B.

Błąd wynikający z różnic pomiędzy osobami przyznającymi punkty

Trzeci rodzaj rzetelności dotyczy błędu testu spowodowanego różnicami w punktowaniu przez diagnostów. W testach takich jak TONI, niewymagających oceny diagnosty do ustalenia poprawności bądź niepoprawności odpowiedzi, zróżnicowanie w punktacji można przypisać niewłaściwemu zastosowaniu zasad startu i przerywania testu, błędom w dodawaniu oraz/lub błędom w zapisie. Błąd osoby oceniającej można zmniejszyć dzięki jasnym procedurom przeprowadzania, szczegółowym wytycznym dotyczącym punktacji oraz możliwości przeciwiczenia przyznawania punktów. Twórcy testu powinni jednak przedstawić liczbę błędów, którą można przypisać różnicy pomiędzy osobami przyznającymi punkty. Zgodnie z zaleceniami Anastasi i Urbiny (1997) dwóch pracowników departamentu badań PRO-ED niezależnie przyznało punkty w 25 parach kompletnych protokołów TONI-3, losowo wybranych z próby normalizacyjnej (L. Brown i in., 1997). Próba miała zakres wiekowy 8–18 lat. Powiązany wskaźnik zgodności wyniósł .99 dla każdej wersji, co dało przekonujący dowód na to, że prawdopodobieństwo błędu osoby przyznającej punkty został praktycznie wyeliminowany.

Podsumowanie wyników rzetelności

Ogólną rzetelność TONI podsumowano w tabeli 4, w której pokazano wyniki wszystkich zbadanych źródeł błędów testów (błąd losowy treści, błąd losowy czasu, błąd różnicy pomiędzy osobami przyznającymi punkty).

TABELA 4

Podsumowanie rzetelności TONI w odniesieniu do czterech źródeł błędów testu (pominięto dziesiętne)

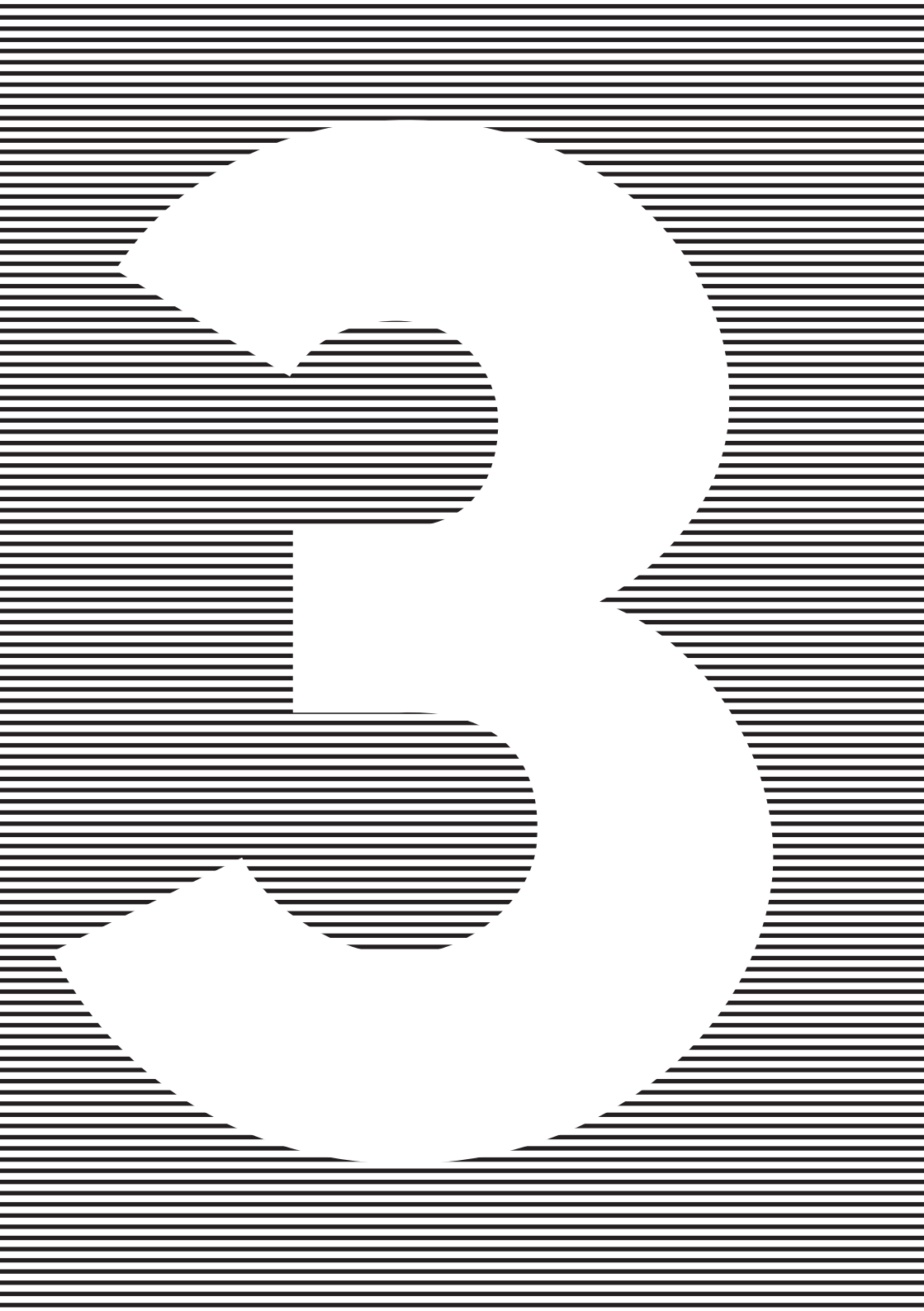
Współczynniki rzetelności

Wydanie	Współczynnik alfa		Alternatywne wersje (bezpośrednie przeprowadzanie)		Alternatywne wersje (odroczone przeprowadzanie)		Test-retest		Różnice pomiędzy osobami przysługującymi punkty	
	Wersja A	Wersja B	r_{AB}	r_{AB}	r_{AB}	r_{AB}	Wersja A	Wersja B	Wersja A	Wersja B
TONI (L. Brown i in., 1982)	88	89	85	–	–	–	–	–	–	–
TONI-2 (L. Brown i in., 1990)	96	96	86	85	–	–	–	–	–	–
TONI-3 (L. Brown i in., 1997)	92	93	84	–	92	92	92	92	99	99
Średni wynik	93	94	85	85	92	92	92	92	99	99
Źródła błędów testu ^a	Losowy treści, niejednorodność treści		Losowy treści	Losowy czasu, losowy treści	Losowy czasu, losowy treści	Losowy czasu	Różnice w przyznaniu punktów			

Uwaga: r_{AB} – współczynnik korelacji pomiędzy wersją A i wersją B.

^a Te źródła różnic w błędach pochodzą z *Psychological Testing* (7th ed., p. 101), A. Anastasi i S. Urbina, 1997, Upper Saddle River, NJ: Prentice Hall.

Współczynniki alfa i współczynniki wersji alternatywnych (natychmiastowe przeprowadzanie) wymienione w tabeli są uśrednionymi współczynnikami podanymi w tabeli 1. Wersje alternatywne (odroczone przeprowadzanie) oraz współczynniki test–retest pochodzą z uśrednionych współczynników zawartych w tabeli 3. Różnice wynikające z błędów osób punktujących omówiono w poprzedniej części. Rzetelność jest ciągle wysoka (.80 i .90), co dowodzi rzetelności w różnych wydaniach TONI.



Trafność testu musi zostać zbadana wielokrotnie i w różnorodnych okolicznościach, aby mogła dać jakikolwiek stopień pewności. Określanie trafności testu jest procesem ciągłym. W tej części podsumowano ponad 25 lat badań nad trafnością TONI. Badania podzielono na dwie główne kategorie: badania nad siłą prognostyczną testu oraz badania nad jego konstruktami teoretycznymi.

Trafność kryterialna/prognostyczna

Według Anastasi i Urbiny (1997) „Procedury trafności prognostycznej wskazują na skuteczność testu w przewidywaniu poziomu wykonania określonych czynności przez osobą badaną” (s. 118). Wykonanie testu może być bezpośrednią lub pośrednią miarą tego, co test ma na celu prognozować. Test taki jak TONI, który ma na celu pomiar zdolności umysłowych, powinien być dobrze skorelowany z innymi testami, które również mierzą zdolności umysłowe, a jako niewerbalna miara inteligencji ogólnej, powinien być jeszcze bardziej skorelowany z innymi testami mierzącymi inteligencję, wykorzystującymi wersję niewerbalną.

Ze względów praktycznych, cztery wydania TONI to zasadniczo ten sam test. Z wyjątkiem opcjonalnych instrukcji werbalnych dodanych w TONI-4, instrukcje do przeprowadzenia oraz formuła odpowiedzi są zasadniczo niezmienione oraz, jak zauważono

wcześniej, wszystkie, oprócz pięciu pozycji TONI-4, były obecne w poprzednich wydaniach testu. Dlatego badania nad trafnością poprzednich wersji testu można ekstrapolować w TONI-4.

Trafność narasta w miarę upływu czasu i nabiera coraz większego znaczenia w miarę rozszerzania się zakresu badań, dając szerszy i pełniejszy obraz testu. Ważna jest sama objętość dostępnych badań. Gdy badacze w różnych okresach i w różnych miejscach niezależnie od siebie uzyskują określone wyniki, wzmacnia to wypływające z nich wnioski. Gromadzenie dostępnych badań na przestrzeni czasu daje duże możliwości i szeroki zakres działań, ponieważ różni badacze koncentrują się na poszczególnych aspektach trafności testu, potwierdzając, klarując oraz poszerzając nasze rozumienie użyteczności teoretycznej, a także praktycznej przydatności testu. W tabeli 5 podsumowano ponad 25 lat badań nad trafnością korelacyjną. Testy kryterialne wymienione są w lewej kolumnie; numery badań w drugiej kolumnie odnoszą się do konkretnych badań, w których pierwotnie zgłoszono współczynniki (charakterystyka badań – patrz Załącznik A). Rodzaj inteligencji mierzonej testem kryterialnym znajduje się po prawej stronie tabeli.

Hopkins (2002) określa współczynniki pomiędzy 0,00 i 0,09 jako „bardzo niskie,” współczynniki pomiędzy 0,10 a 0,29 jako „niskie,” współczynniki pomiędzy 0,30 a 0,49 jako „umiarkowane,” współczynniki pomiędzy 0,50 a 0,69 jako „wysokie,” a współczynniki pomiędzy 0,70 a 0,99 jako „bardzo wysokie.” Uśrednione współczynniki podano na dole tabeli 5.

TABELA 5

*Korelacja TONI z miarami inteligencji ogólnej, werbalnej i niewerbalnej
(pominięto dziesiętne)*

Miara kryterium	° numer badania	Miara inteligencji					
		Ogólne		Werbalne		Niewerbalne	
		Wersja A	Wersja B	Wersja A	Wersja B	Wersja A	Wersja B
BVA	34	–	–	45	–	–	–
CELF-R Bierny	21	–	–	68	–	–	–
CELF-R Bierny	25	–	–	45	–	–	–
CELF-R Czynny	21	–	–	64	–	–	–
CELF-R Czynny	21	–	–	58	–	–	–
CELF-R Ogólne	21	–	–	67	–	–	–
CELF-R Ogólne	25	–	–	58	–	–	–
CPM Ogólne	38	–	–	–	–	77	–
CTONI Obrazkowe niewerb. IQ	27	–	–	–	–	43	–
CTONI Obrazkowe niewerb. IQ	28	–	–	–	–	74	72
CTONI Geometr. niewerb. IQ	27	–	–	–	–	84	–
CTONI Geometr. niewerb. IQ	28	–	–	–	–	64	64
CTONI Ogólne niewerb. IQ	27	–	–	–	–	82	–
CTONI Ogólne niewerb. IQ	28	–	–	–	–	76	74
DTLA-A Ogólna inteligencja	12	44					
DTVI Ogólna	26	–	–	–	–	61	–
DTVI Ogólna	22	–	–	–	–	62	–
DTVI i Wzory z Klocków WISC	33	–	–	–	–	74	–

TABELA 5 cd.

*Korelacja TONI z miarami inteligencji ogólnej, werbalnej i niewerbalnej
(pominięto dziesiętne)*

Miara kryterium	a numer badania	Miara inteligencji					
		Ogólne		Werbalne		Niewerbalne	
		Wersja A	Wersja B	Wersja A	Wersja B	Wersja A	Wersja B
GAS	46	63					
KABC							
Przetwarzanie umysłowe	12	53	–	–	–	–	–
KABC							
Przetwarzanie umysłowe	19	53	–	–	–	–	–
KABC Przetw. sekwencyjne	12	52	–	–	–	–	–
KABC Przetw. sekwencyjne	19	26	–	–	–	–	–
KABC Przetw.	12	46	–	–	–	–	–
KABC Przetw.	19	58	–	–	–	–	–
KABC Skala niewerbalna	12	–	–	–	–	47	–
KABC Skala niewerbalna	19	–	–	–	–	49	–
LIPS Ogólne	3	–	–	–	–	89	83
LIPS-R Brief	40	–	–	–	–	32	–
MGF	36	44	–	–	–	–	–
MGF – Głusi	37	41	–	–	–	–	–
OLMAT Ogólny	1	–	–	86	87	–	–
PPVT-III	43	–	–	29	–	–	–
QT Ogólna	11	–	67	–	–	–	–
QT Ogólna	8	–	66	–	–	–	–
RDLS	42	–	–	83	–	–	–

TABELA 5 cd.

*Korelacja TONI z miarami inteligencji ogólnej, werbalnej i niewerbalnej
(pominięto dziesiętne)*

Miara kryterium	° numer badania	Miara inteligencji					
		Ogólne		Werbalne		Niewerbalne	
		Wersja A	Wersja B	Wersja A	Wersja B	Wersja A	Wersja B
SAGES Powiązany program D	14	52	54	–	–	–	–
SAGES Powiązany program D	10	52	–	–	–	–	–
SAGES Rozumowanie	10	–	–	–	–	50	–
SAGES Inf. nabyte w szkole	10	–	–	47	–	–	–
SAT Ilościowe	40	–	–	37	–	–	–
SAT Werbalne	40	–	–	25	–	–	–
SAT Ogólne	40	41	–	–	–	–	–
SATA Ogólne predyspozycje	12	60	–	–	–	–	–
SATA Rozumowanie niewerb.	12	–	–	–	–	53	–
SATA Rozumowanie werb.	12	–	–	43	–	–	–
SATA Rozumowanie jakościowe	12	–	–	–	–	64	–
SIT-R Ogólne	9	–	–	27	–	–	–
SIT-R Ogólne	25	–	–	52	–	–	–
SIT-R Ogólne	6	–	–	71	–	–	–
SPM Ogólne	5	–	–	–	–	92	92
TOLD-I:2 Język mówiony	13	–	–	55	–	–	–
TOLD-I:2 Język mówiony	10	–	–	49	–	–	–

TABELA 5 cd.

*Korelacja TONI z miarami inteligencji ogólnej, werbalnej i niewerbalnej
(pominięto dziesiętne)*

Miara kryterium	a numer badania	Miara inteligencji					
		Ogólne		Werbalne		Niewerbalne	
		Wersja A	Wersja B	Wersja A	Wersja B	Wersja A	Wersja B
TOLD-P:2 Ogólne	10	–	–	24	–	–	–
UNIT Podskala rozumowania logicznego	41	–	–	–	–	56	–
WAIS-R Wersja diadyczna Pełna skala	12	48	–	–	–	–	–
WAIS-R Skala werbalna IQ	31	–	–	57	51	–	–
WAIS-R Skala werbalna IQ	20	41	–	–	–	–	–
WAIS-R Skala werbalna IQ	24	–	–	42	–	–	–
WAIS-R Skala wykonaniowa IQ	31	–	–	–	–	75	76
WAIS-R Skala wykonaniowa IQ	20	–	–	44	–	–	–
WAIS-R Skala wykonaniowa IQ	24	–	–	–	–	64	–
WAIS-R Pełna skala IQ	31	73	71	–	–	–	–
WAIS-R Pełna skala IQ	20	50	–	–	–	–	–
WAIS-III Zadania z matrycami	39	–	–	–	–	48	–
WAIS-III Werbalne IQ	40	–	–	18	–	–	–

TABELA 5 cd.
*Korelacja TONI z miarami inteligencji ogólnej, werbalnej i niewerbalnej
 (pominięto dziesiętne)*

		Miara inteligencji					
		Ogólne		Werbalne		Niewerbalne	
Miara kryterium	° numer badania	Wersja A	Wersja B	Wersja A	Wersja B	Wersja A	Wersja B
WISC-R Wersja diadyczna Pełna skala	12	30	–	–	–	–	–
WISC-R Pełna skala	2	62	92	–	–	–	–
WISC-R Pełna skala	4	76	87	–	–	–	–
WISC-R Pełna skala	14	71	75	–	–	–	–
WISC-R Pełna skala	7	44	–	–	–	–	–
WISC-R Pełna skala	11	66	–	–	–	–	–
WISC-R Pełna skala	16	51	–	–	–	–	–
WISC-R Pełna skala	17	87	–	–	–	–	–
WISC-R Pełna skala	19	27	–	–	–	–	–
WISC-R Pełna skala	8	68	–	–	–	–	–
WISC-R Pełna skala	23	59	–	–	–	–	–
WISC-R Pełna skala	18	48	–	–	–	–	–
WISC-R Pełna skala	20	50	–	–	–	–	–
WISC-R Pełna skala	6	72	–	–	–	–	–

TABELA 5 cd.

*Korelacja TONI z miarami inteligencji ogólnej, werbalnej i niewerbalnej
(pominięto dziesiętne)*

Miara kryterium	a numer badania	Miara inteligencji					
		Ogólne		Werbalne		Niewerbalne	
		Wersja A	Wersja B	Wersja A	Wersja B	Wersja A	Wersja B
WISC-R Wykonaniowa	2	–	–	–	–	62	87
WISC-R Wykonaniowa	4	–	–	–	–	89	80
WISC-R Wykonaniowa	14	–	–	–	–	60	60
WISC-R Wykonaniowa	7	–	–	–	–	44	–
WISC-R Wykonaniowa	11	–	–	–	–	63	–
WISC-R Wykonaniowa	16	–	–	–	–	52	–
WISC-R Wykonaniowa	17	–	–	–	–	81	–
WISC-R Wykonaniowa	19	–	–	–	–	30	–
WISC-R Wykonaniowa	8	–	–	–	–	68	–
WISC-R Wykonaniowa	23	–	–	–	–	58	–
WISC-R Wykonaniowa	18	–	–	–	–	42	–
WISC-R Wykonaniowa	20	–	–	–	–	44	–
WISC-R Wykonaniowa	6	–	–	–	–	77	–
WISC-R Werbalna	2	–	–	69	70	–	–
WISC-R Werbalna	4	–	–	46	63	–	–

TABELA 5 cd.

*Korelacja TONI z miarami inteligencji ogólnej, werbalnej i niewerbalnej
(pominięto dziesiętne)*

Miara kryterium	° numer badania	Miara inteligencji					
		Ogólne		Werbalne		Niewerbalne	
		Wersja A	Wersja B	Wersja A	Wersja B	Wersja A	Wersja B
WISC-R Werbalna	7	–	–	34	–	–	–
WISC-R Werbalna	16	–	–	38	–	–	–
WISC-R Werbalna	17	–	–	61	–	–	–
WISC-R Werbalna	19	–	–	32	–	–	–
WISC-R Werbalna	8	–	–	57	–	–	–
WISC-R Werbalna	23	–	–	58	–	–	–
WISC-R Werbalna	18	–	–	68	–	–	–
WISC-R Werbalna	20	–	–	41	–	–	–
WISC-R Werbalna	6	–	–	61	–	–	–
WISC III Org. percepcyjna	34	–	–	–	–	40	–
WISC III Org. percepcyjna	29	–	–	–	–	67	–
WISC III Szybkość przetw.	29	–	–	–	–	57	–
WISC III Skala werbalna IQ	30	–	–	59	53	–	–
WISC III Skala werbalna IQ	44	–	–	59	–	–	–
WISC III Skala werbalna IQ	35	–	–	88	–	–	–
WISC III Skala wykon. IQ	30	–	–	–	–	56	58
WISC III Skala wykon. IQ	44	–	–	–	–	48	–
WISC III Skala wykon. IQ	32	–	–	–	–	67	–

TABELA 5 cd.

*Korelacja TONI z miarami inteligencji ogólnej, werbalnej i niewerbalnej
(pominięto dziesiętne)*

Miara kryterium	a numer badania	Miara inteligencji					
		Ogólne		Werbalne		Niewerbalne	
		Wersja A	Wersja B	Wersja A	Wersja B	Wersja A	Wersja B
WISC III Skala wykon. IQ	35	–	–	–	–	89	–
WISC III Pełna skala IQ	30	63	63	–	–	–	–
WISC III Pełna skala IQ	44	59	–	–	–	–	–
WISC III Pełna skala IQ	35	93	–	–	–	–	–
WISC IV Pamięć robocza	45	–	–	–	50	–	–
WISC IV Pamięć robocza – edukacja specjalna	45	–	–	–	44	–	–
WISC IV Pamięć robocza – uzdolnienia	45	–	–	–	59	–	–
WISC IV Rozumowanie percepc.	45	–	–	–	54	–	–
WISC IV Rozumowanie percepc. edukacja specjalna	45	–	–	–	55	–	–
WISC IV Rozumowanie percepc. uzdolnienie	45	–	–	–	32	–	–
WISC IV Szybkość przełw.	45	–	–	–	22	–	–

TABELA 5 cd.

*Korelacja TONI z miarami inteligencji ogólnej, werbalnej i niewerbalnej
(pominięto dziesiętne)*

Miara kryterium	° numer badania	Miara inteligencji					
		Ogólne		Werbalne		Niewerbalne	
		Wersja A	Wersja B	Wersja A	Wersja B	Wersja A	Wersja B
WISC IV Szybkość przetw. – edukacja specjalna	45	–	–	–	33	–	–
WISC IV Szybkość przetw. – uzdolnienie	45	–	–	–	21	–	–
WISC IV Rozumienie werb.	45	–	–	–	53	–	–
WISC IV Rozumienie werb. – edukacja specjalna	45	–	–	–	58	–	–
WISC IV Rozumienie werb. – uzdolnienie	45	–	–	–	50	–	–
WISC IV Pełna skala IQ	45	70	–	–	–	–	–
WISC IV Pełna skala IQ – edukacja specjalna	45	67	–	–	–	–	–
WISC IV Pełna skala IQ – uzdolnienie	45	54	–	–	–	–	–
WJPEB Werbalne	15	80	–	–	–	–	–
WJPEB Rozumowanie	15	79	–	–	–	–	–
WJPEB Szybkość przetwarzania	15	–	–	–	–	72	–
WJPEB Pamięć	15	–	–	80	–	–	–

TABELA 5 cd.

**Korelacja TONI z miarami inteligencji ogólnej, werbalnej i niewerbalnej
(pominięto dziesiątę)**

Miara kryterium	Miara inteligencji						
		Ogólne		Werbalne		Niewerbalne	
	^a numer badania	Wersja A	Wersja B	Wersja A	Wersja B	Wersja A	Wersja B
WJ III Myślenie wzrokowo-przestrzenne	34	–	–	–	–	37	–
WRAML	33	–	–	–	–	39	–
Średnie w badaniach	–	59	75	55	53	65	77
Średnie w formatach		62		54		71	
Wielkość ^b		Wysoka		Wysoka		Bardzo wysoka	

Uwaga: BVA – *Bilingual Verbal Ability Tests* [Dwujęzyczne testy zdolności werbalnych] (Muñoz-Sandoval, Cummins, Alvarado, Ruef, 1998); CELF-R – *Clinical Evaluation of Language Fundamentals-Revised* [Kliniczna ocena podstaw językowych – wyd. poprawione] (Semel, Wig, Secord, 1987); CPM – *Colored Progressive Matrices* [Test kolorowych matryc Ravena] (Raven, 1965); CTONI – *Comprehensive Test of Nonverbal Intelligence* [Kompleksowy test inteligencji niewerbalnej] (Hammill, Pearson, Wiederholt, 1996); DTLA-A – *Detroit Tests of Learning Aptitude-Adult* [Badania predyspozycji do nauki dla dorosłych] (Hammill, Bryant, 1991); DTVI – *Developmental Test of Visual-Motor Integration* [Test integracji wzrokowo-ruchowej] (Beery, 1989); GAS – *Global Assessment Scale* [Skala oceny ogólnej] (Endicott, Spitzer, Fleiss, Cohen, 1976); KABC – *Kaufman Assessment Battery for Children* [Bateria diagnozy dzieci Kaufmanów] (Kaufman, Kaufman, 1983); LIPS – *Leiter International Performance Scale* [Międzynarodowa wykonaniowa skala Leitera] (Leiter, 1948); LIPS-R Brief – *Leiter International Performance Scale-Revised* [Międzynarodowa wykonaniowa skala Leitera – wyd. poprawione] (Roid, Miller, 1997); MGF – *Modified Gottschaldt Figures* [Zmodyfikowane figury Gottschaldta] (Thurstone, 1944); OLMAT – *Otis-Lennon Mental Ability Test* [Badanie zdolności umysłowych Otisa i Lennona] (Otis, Lennon, 1970); PPVT-III – *Peabody Picture Vocabulary Test-Third Edition* [Obrazkowy test słownikowy Peabody – wyd. III] (Dunn, Dunn, 1997); QT – *Quick Test* [Szybki test] (Ammons, Ammons, 1962); RDLS – *Reynell Developmental Language Scales* [Skala rozwoju językowego Reynell] (Department of Health and Education, Hong Kong, 1987); SAGES – *Screening Assessment for Gifted Elementary Students* [Przesiewowa diagnoza dla uzdolnionych uczniów szkoły podstawowej] (Johnsen, Corn, 1987); SAT – *Scholastic Aptitude Test* [Badanie predyspozycji szkolnych] (Educational Testing Service, 2002); SATA – *Scholastic Abilities Test for Adults* [Badanie predyspozycji szkolnych dla dorosłych] (Bryant, Patton, Dunn, 1991); SIT-R – *Slosson Intelligence Test-Revised* [Test inteligencji Slossona – wyd. popra-

wione] (Slosson, 1991); SPM – *Standard Progressive Matrices* [Test matryc Ravena] (Raven, 1938); TOLD-I:2 – *Test of Language Development–Intermediate: Second Edition* [Badanie rozwoju języka – średniozaawansowane: wyd. drugie] (Newcomer, Hammill, 1988a); TOLD-P:2 – *Test of Language Development– Primary: Second Edition* [Badanie rozwoju języka – podstawowe: wyd. drugie] (Newcomer, Hammill, 1988b); UNIT – *Universal Nonverbal Intelligence Test* [Uniwersalny test inteligencji niewerbalnej] (Bracken, McCallum, 1996); WAIS-R – *Wechsler Adult Intelligence Scale–Revised* [Skala inteligencji dla dorosłych Wechslera – wyd. poprawione] (Wechsler, 1981); WAIS III – *Wechsler Adult Intelligence Scale–Third Edition* [Skala inteligencji Wechslera dla dorosłych – wyd. trzecie] (Wechsler, 1997); WISC-R – *Wechsler Intelligence Scale for Children–Revised* [Skala inteligencji Wechslera dla dzieci – wyd. poprawione] (Wechsler, 1974); WISC-III – *Wechsler Intelligence Scale for Children–Third Edition* [Skala inteligencji Wechslera dla dzieci– wyd. trzecie] (Wechsler, 1991); SPED – *special education* [edukacja specjalna]; WISC-IV – *Wechsler Intelligence Scale for Children–Fourth Edition* [Skala inteligencji dla dzieci Wechslera – wyd. czwarte] (Wechsler, 2003); WJPEB – *Woodcock-Johnson Psycho-Educational Battery* [Psychoedukacyjna bateria Woodcocka-Johnson] (Woodcock, Johnson, 1977); WJ III – *Woodcock-Johnson III Tests of Cognitive Abilities* [Badanie zdolności poznawczych Woodcocka-Johnson III] (Woodcock, McGrew, Mather, 2001); WRAML – *Wide Range Assessment of Memory and Learning* [Szerokoza-kresowa diagnoza pamięci i nauki] (Sheslow, Adams, 1990).

^a Opisy badań według numerów – patrz Załącznik A.

^b Na podstawie kryteriów Hopkinsa (2002) dla interpretacji współczynników korelacji.

Te uśrednione współczynniki mają zakres od wyników wysokich do bardzo wysokich, co tłumaczy silne związki pomiędzy TONI a miarami kryterialnymi. Silnym dowodem jest także podobieństwo wyników badań w szerokich przedziałach wiekowych i w różnych populacjach, obejmujących zazwyczaj osoby z autyzmem, trudnościami w uczeniu się, dysleksją, ubytkiem słuchu, urazem głowy, lekką niepełnosprawnością, niepełnosprawnością intelektualną i wysokimi zdolnościami (uzdolnieniami) intelektualnymi.

Trafność prognostyczna TONI została zbadana przez wielu badaczy stosujących analizy wielokrotnej korelacji i regresji. Badacze tacy jak Vance, Hankins, Reynolds (1988), Selbst (1997), Mackinson (1996), Bostantjopoulou, Kiosseoglou, Katsarou, Alevriadou (2001), Gonzalez (1994) opracowali dane pokazujące trafność TONI. W większości badań stwierdzono, że za pomocą TONI można prognozować wyniki w innych testach inteligencji. Vance i in. (1988) wykorzystali procedurę wielokrotnej korelacji w celu określenia użyteczności TONI i *The Quick Test* (Ammons, Ammons, 1962) w prognozowaniu Pełnej skali IQ ze Skali Inteligencji

Wechslera dla Dzieci – Wersja Zrewidowana (*Wechsler Intelligence Scale for Children-Revised*; Wechsler, 1974). Doszli do wniosków, że obie miary są równie dobrymi predyktorami Pełnej Skali IQ WISC-R. Selbst (1997) także używał analizy wielokrotnej regresji z ilorazem TONI-2 jako zmienną kryterialną oraz 11 wynikami podtestów Skali Inteligencji Wechslera dla Dzieci – Edycja Trzecia (*Wechsler Intelligence Scale for Children-Third Edition*; WISC-III; Wechsler, 1991) jako zmiennymi predykcyjnymi. Wyniki także wskazały na istotną wielokrotną korelację na poziomie .71. Z 11 zmiennych predykcyjnych włączonych do równania, cztery podtesty miały istotną wagę w równaniu regresji. Były to Układanki, Wiadomości, Kodowanie oraz Powtarzanie Cyfr. Bostantjopoulou i in. (2001) wykorzystali regresję wielokrotną do zbadania zdolności TONI-2 do prognozowania wyników pacjentów z chorobą Parkinsona i wyników kontrolnych w teście Kolorowych Matryc Ravena. Ponownie znaleziono istotny związek ($p < .001$) zarówno w przypadku pacjentów ($R^2 = .53$), jak i grupy kontrolnej ($R^2 = .59$).

W swoich badaniach Gonzalez (1994) wykorzystwała krok po kroku analizę wielokrotnej regresji liniowej w celu zbadania możliwości TONI do prognozowania rozwoju poznawczego dzieci latynoskiego pochodzenia. Wykazała, że wyniki TONI prognozowały wyniki zarówno werbalnych, jak i niewerbalnych zadań kategoryzacji pojęciowej. Jedno z badań sprawdzało możliwość wykorzystania TONI do prognozowania wyników w testach osiągnięć. W tym badaniu Mackinson, Leigh, Blennerhassett i Anthony (1997) wykorzystali procedurę wielokrotnej korelacji w celu określenia użyteczności TONI-2 oraz WISC-III w prognozowaniu osiągnięć. Odnotowali, że oba testy prognozowały wyniki składowe *Stanford Achievement Test-Eighth Edition* (Harcourt Educational Measurement, 1989), z najistotniejszym związkiem dla ortografii ($r = .71$), rozwiązywania problemów matematycznych ($r = .66$) oraz rozumowania ($r = .62$). We wszystkich tych badaniach TONI był istotnym predyktorem inteligencji, rozwoju poznawczego oraz osiągnięć.

Trafność konstruktu

„Trafność konstruktu testu oznacza zakres, do jakiego można wykorzystać test w celu zmierzenia konstruktu teoretycznego lub cechy” (Anastasi, Urbina, 1997, s. 126). Stopień, w jakim można zidentyfikować podstawowe cechy testu, i zakres, w jakim cechy te odzwierciedlają model teoretyczny, na którym opiera się test, mówi o jego trafności. W celu wykazania tego rodzaju trafności zastosowano trójetapową procedurę. Po pierwsze, zidentyfikowano konstrukty, które mają odpowiadać za wyniki testu. Po drugie, zaproponowano hipotezy dotyczące wykonania testu i zidentyfikowanych konstruktyw. Po trzecie, hipotezy te zostały zweryfikowane lub obalone metodami logicznymi lub empirycznymi. W tej części omówiono pięć możliwych do przetestowania propozycji dotyczących głównych konstruktyw leżących u podstaw TONI.

1. Mierzona inteligencja jest zgodna z wzorem rozwojowym do okresu późnego wieku nastoletniego, stabilizuje się w późnym wieku średnim, a następnie stopniowo spada. Wyniki surowe TONI powinny być zgodne z tym samym schematem chronologicznym.
2. TONI ma na celu pomiar inteligencji ogólnej. Wyniki TONI powinny silnie korelować z testami osiągnięć szkolnych.
3. TONI ma na celu pomiar inteligencji ogólnej. Jego wyniki powinny różnicować grupy osób uznawanych za intelektualnie przeciętne, ponadprzeciętne i poniżej przeciętnej.
4. TONI stworzono do pomiaru rozumowania abstrakcyjnego i rozwiązywania problemów, teoretycznie uznawanych za ogólne składniki inteligencji, podobne do czynnika g Spearmana. Zadania TONI powinny ułożyć się w jeden duży czynnik.
5. TONI zakładał redukcję użycia języka i eliminację obciążenia kulturowego podczas przeprowadzania testu. Zadania TONI powinny wykazywać minimalną stronniczość.

Dane na poparcie tych punktów można znaleźć poniżej.

Związek z wiekiem

Związek wieku z wynikami surowymi TONI badano sześciokrotnie. Trzy z tych badań przeprowadzili autorzy (L. Brown i in., 1982, 1990, 1997). Średnie i odchylenia standardowe dla TONI-3 przedstawione w tabeli 6 pokazują, że TONI-3 jest silnie powiązany z wiekiem osoby badanej w wieku 6;0 do 17;11 ($r = .63$ dla wersji A, $r = .60$ dla wersji B). Powyżej 17. roku życia wyniki surowe stabilizują się aż do 60. roku życia i następnie zaczynają spadać. Wzór ten jest charakterystyczny dla testów inteligencji wykonywanych przez dorosłych (np. Wechsler, 1958). Podobny wzór odnotowano w tabelach norm dla TONI i TONI-2.

Prado, Gatica i Rojas (1993) uzyskali podobne wyniki badań dla ich demograficznie reprezentatywnej próby 328 chilijskich uczniów w wieku od 6 do 15;11 lat. Średnie wyniki surowe miały zakres od 9,93 dla 6-latków do 28,85 dla 15-latków. Porównanie uszeregowania i zakresów punktacji Scheffego w jednorocznych przedziałach dało w efekcie cztery jednorodne grupy: (a) 6-latków, (b) 7–8-latków (c) 9–10-latków oraz (d) 11–15-latków. W analizie wariancji dotyczącej wpływu wieku, płci i statusu społeczno-ekonomicznego na wyniki badań, jedynym znaczącym efektem interakcji był wiek. Barrett (2000) zauważył także silny związek z wiekiem ($r = .35$) na podstawie wykonania TONI-3 przez 112 jamajskich uczniów szkoły podstawowej.

Związek z osiągnięciami szkolnymi

Wyniki testu zdolności intelektualnych uznaje się za dobre predyktory osiągnięć szkolnych. Dlatego spodziewać się można umiarkowanych do wysokich pozytywnych relacji pomiędzy wynikami TONI i wynikami ze skal znormalizowanych mierzących osiągnięcia szkolne.

TABELA 6

*Średnie wyników surowych oraz odchylenia standardowe dla
TONI-3 w 20 przedziałach wiekowych*

Przedział wiekowy (lata)	Wersja A			Wersja B	
	N	M	SD	M	SD
6	154	9	6	9	7
7	159	14	6	14	7
8	180	17	6	17	6
9	138	18	6	18	6
10	204	21	6	21	6
11	164	22	7	22	8
12	110	24	8	24	8
13	137	24	8	23	8
14	189	25	8	25	8
15	201	28	8	27	8
16	170	29	9	29	9
17	180	30	7	30	7
18	132	30	8	30	8
19–29	277	29	10	29	10
30–39	252	31	10	31	10
40–49	330	30	10	30	10
50–59	202	30	8	30	7
60–69	127	24	9	24	10
70–79	91	20	9	20	10
80–89	54	18	12	16	12

Jedenaście tego typu badań przedstawiono w tabeli 7. Charakterystyka prób wykorzystanych w tych badaniach znajduje się w Załączniku B. W tabeli 7 przedstawiono dane dotyczące czterech kategorii osiągnięć szkolnych: osiągnięcia ogólne, język, matematyka i inne (takie jak nauki społeczne, nauki ścisłe i umiejętności wyszukiwania). Jeśli badacz uzyskał współczynniki zarówno dla wersji A, jak i dla wersji B, uśredniano je na potrzeby tej tabeli. Następnie uśredniono współczynniki dla każdej z czterech kategorii, a współczynniki te zostały przedstawione na dole tabeli 7. Wszystkie wyniki są istotne i mają umiarkowaną lub wysoką wielkość, co jest wzmocnione przez szeroki zakres wiekowy osób badanych oraz podobieństwo wyników w populacjach osób, które są w normie intelektualnej, głuchych, niedosłyszących, z dysleksją lub innymi trudnościami w uczeniu się.

Różnice pomiędzy grupami

Test inteligencji powinien różnicować wyniki różnych grup osób, określanych jako utalentowane, w normie intelektualnej i z niepełnosprawnościami.

TABELA 7

Korelacja TONI z miarami osiągnięć szkolnych (pominięto dziesiętne)

Miara kryterium	Numer badania ^a	Ogólne osiągnięcia	Język ^b	Matematyka	Inne ^c
DAB-2 Rozumienie tekstu pisanego	9	–	59	–	–
DATA-2 Ogólne osiągnięcia	6	57	–	–	–
DATA-2 Czytanie	6	–	52	–	–
DATA-2 Matematyka	6	–	–	46	–
DATA-2 Pisanie	6	–	40	–	–
DATA-2 Czytanie ze zrozumieniem	6	–	48	–	–
DATA-2 Obliczenia matematyczne	6	–	–	48	–
DATA-2 Rozumowanie matematyczne	6	–	–	41	–
DATA-2 Nauki przyrodnicze	6	–	–	–	45
DATA-2 Nauki społeczne	6	–	–	–	33
DATA-2 Umiejętność wyszukiwania	6	–	–	–	36
DATA-2 Identyfikacja słów	6	–	51	–	–
DATA-2 Ortografia	6	–	36	–	–
DATA-2 Kompozycja pisemna	6	–	33	–	–
KTEA-II Czytanie złożone	10	–	33	–	–
KTEA-II Matematyka złożona	10	–	–	43	–
KTEA-II Bateria ogólna	10	38	–	–	–

TABELA 7 cd.

Korelacja TONI z miarami osiągnięć szkolnych (pominięto dziesiętne)

Miara kryterium	Numer badania ^a	Ogólne osiągnięcia	Język ^b	Matematyka	Inne ^c
ITBS Język	1	–	84	–	–
ITBS Język	6	–	81	–	–
ITBS Czytanie ze zrozumieniem	1	–	69	–	–
ITBS Czytanie ze zrozumieniem	2	–	55	–	–
ITBS Czytanie ze zrozumieniem	6	–	70	–	–
ITBS Matematyka	1	–	–	69	–
ITBS Matematyka	2	–	–	65	–
ITBS Matematyka	6	–	–	59	–
SAT Czytanie ze zrozumieniem	3	–	74	–	–
SAT Pojęcia matematyczne	3	–	–	77	–
SAT/DHH Czytanie ze zrozumieniem	8	–	49	–	–
SAT/DHH Język	8	–	50	–	–
SAT/DHH Pojęcia liczbowe	8	–	–	66	–
SAT/DHH Stosowanie matematyki	8	–	–	62	–
SAT/DHH Obliczenia matematyczne	8	–	–	58	–
SAT/DHH Ortografia	8	–	71	–	–
SATA Ogólne	7	56	–	–	–
SATA Matematyka	7	–	–	36	–
SATA Czytanie	7	–	44	–	–
SATA Pisanie	7	–	57	–	–

TABELA 7 cd.

Korelacja TONI z miarami osiągnięć szkolnych (pominięto dziesiętne)

Miara kryterium	Numer badania ^a	Ogólne osiągnięcia	Język ^b	Matematyka	Inne ^c
SATA Przesiewowe	7	60	–	–	–
SATA Obliczenia matematyczne	7	–	–	39	–
SATA Słownictwo z czytania	7	–	64	–	–
SATA Zastosowania matematyki	7	–	–	33	–
SATA Interpunkcja	7	–	58	–	–
SATA Kompozycja pisemna	7	–	54	–	–
SATA Czytanie ze zrozumieniem	7	–	26	–	–
SRA Złożone	11	80	–	–	–
SRA Czytanie	1	–	75	–	–
SRA Matematyka	1	–	–	78	–
SRA Język	1	–	76	–	–
TORC-R Słownictwo ogólne	3	–	41	–	–
TORC-R Podobieństwa syntaktyczne	3	–	33	–	–
TORC-R Czytanie akapitów	3	–	36	–	–
TORC-R Kolejność zdań	3	–	40	–	–
TORC-R Czytanie ze zrozumieniem	3	–	46	–	–
TORC-R Czytanie ze zrozumieniem	5	–	27	–	–
WJPEB-R Czytanie wielowymiarowo	9	–	72	–	–

TABELA 7 cd.

Korelacja TONI z miarami osiągnięć szkolnych (pominięto dziesiętne)

Miara kryterium	Numer badania ^a	Ogólne osiągnięcia	Język ^b	Matematyka	Inne ^c
WJPEB-R Matematyka wielowymiarowo	9	–	–	75	–
WJPEB-R Wiedza wielowymiarowo	9	–	–	–	56
WRMT-R Atak słów	11	–	37	–	–
WRMT-R Rozumienie fragmentu	11	–	32	–	–
WRAT Czytanie	4	–	33	–	–
WRAT Ortografia	4	–	33	–	–
WRAT Arytmetyka	4	–	–	38	–
WRAT-R Czytanie	6	–	28	–	–
WRAT-R Ortografia	6	–	30	–	–
WRAT-R Arytmetyka	6	–	–	34	–
Średni współczynnik		60	52	56	43
Wielkość ^d		Wysoka	Wysoka	Wysoka	Umiar- kowana

Uwaga: DAB-2 – *Diagnostic Achievement Battery–Second Edition* [Diagnostyczna bateria osiągnięć – wyd. drugie] (Newcomer, 1990); DATA-2 – *Diagnostic Achievement Test for Adolescents–Second Edition* [Diagnostyczny test osiągnięć dla młodzieży – wyd. drugie] (Newcomer, Bryant, 1993); KTEA-II – *Kaufman Test of Educational Achievement–Second Edition* [Test osiągnięć edukacyjnych Kaufmanów – wyd. drugie] (Kaufman, Kaufman, 2004); ITBS – *Iowa Tests of Basic Skills* [Badanie podstawowych umiejętności Iowa] (Hieronymus, Lindquist, Hoover, 1982); SAT – *Stanford Achievement Test* [Test osiągnięć Stanford] (Madden, Gardner, Rudman, Karlsen, Merwin, 1973); SAT/DHH – *Stanford Achievement Test–Eighth Edition adapted for deaf or hard of hearing* [Test osiągnięć Stanford – wyd. ósme, adaptacja dla głuchych i niedosłyszących] (Holt, Traxler, Allen, 1992); SATA – *Scholastic Abilities Test for Adults* [Badanie predyspozycji szkolnych dla dorosłych] (Bryant, Patton, Dunn, 1991); SRA – *SRA Achievement Series* [Seria osiągnięć SRA] (Naslund, Thorpe, Lefever, 1978); TORC-R – *Test of Reading Comprehension–Revised* [Badanie czytania ze zrozumieniem – wyd. poprawione] (V.L. Brown, Hammill, Wiederholt, 1986); WJPEB-R – *Woodcock-Johnson Psycho-Educational Battery–Revised* [Psychoedukacyjna bateria Woodcock-Johnson – wyd. poprawione] (Woodcock,

Johnson, 1989); WRMT-R – *Woodcock Reading Mastery Tests–Revised* [Badanie zaawansowania czytania Woodcocka – wyd. poprawione] (Woodcock, 1998); WRAT – *Wide Range Achievement Test* [Szerokozakresowy test osiągnięć] (J.F. Jastak, Jastak, 1965); WRAT-R – *Wide Range Achievement Test–Revised* [Szerokozakresowy test osiągnięć – wyd. poprawione] (S.R. Jastak, Wilkinson, 1984).

^a Opisy badań według numerów – patrz Załącznik B.

^b Język obejmuje pisanie, ortografię i czytanie.

^c Inne zawiera miary, takie jak nauki społeczne, przyrodnicze oraz umiejętności wyszukiwania.

^d Na podstawie kryteriów Hopkinsa (2002) dla interpretacji współczynników korelacji.

Można na przykład postawić hipotezę, że osoby, u których zdiagnozowano niepełnosprawność intelektualną, poradzą sobie bardzo słabo w TONI, natomiast osoby uważane za intelektualnie uzdolnione poradzą sobie z tym testem wyjątkowo dobrze.

Średnie wyniki dla wielu podgrup osób w normie i o wyjątkowym statusie podano w tabeli 8; charakterystyka prób wykorzystanych w tych badaniach została opisana w Załączniku C. Tabelę 8 podzielono na trzy części: badania podgrup zróżnicowanych pod względem płci/etnicznym, badania podgrup o wyjątkowym statusie i innych podgrup. Badania podgrup poza USA zarówno z populacji w normie, jak i tych o specjalnych potrzebach są dostępne, ale nie zostały uwzględnione w tej tabeli. Badania te przeprowadzono w państwach, takich jak: Australia, Kanada, Chile, Chiny, Grecja, Indie, Włochy, Jamajka, Korea i Tajwan; opis tych badań można znaleźć w Załączniku C.

Zgodnie z oczekiwaniami, średnie wyniki głównych podgrup mieszczą się w zakresie wyników przeciętnych. Wyniki w podgrupach, które uznaje się za mające trudności poznawcze (np. autyzm, niepełnosprawność), są znacznie poniżej wyników podgrup płciowych/etnicznych, podczas gdy podgrupy o wyjątkowym statusie, o których wiadomo, że nie mają znaczących poznawczych trudności (np. trudności w uczeniu się, ADHD, trudności emocjonalne), mieszczą się w statystycznej normie lub w zakresie przeciętnym (tzn. 90 do 110). Średnie wyniki w tabeli wspierają trafność teoretyczną TONI. Zgodnie z oczekiwaniami, podgrupa osób z niepełnosprawnością intelektualną miała wyniki charakterystyczne



dla takich osób (tj. około dwa odchylenia standardowe poniżej średniej), a średnia dla podgrupy utalentowanej była o około jedno odchylenie standardowe powyżej średniej. Chociaż można było się spodziewać, że iloraz uzdolnionej próby wyniesie około 5 lub więcej punktów, możliwe jest, że jest on nieco obniżony przez kwestie związane z nazewnictwem: w wielu okręgach szkolnych – a cała populacja uzdolniona jest w wieku szkolnym – nomenklatura uzdolnionych jest w rzeczywistości skrótem od „uzdolnionych i utalentowanych”, co oznacza, że tak określani uczniowie mogą być uzdolnieni intelektualnie lub mieć określone predyspozycje. Ci ostatni uczniowie mogą mieć unikalne zdolności, ale nie mieć wysokich możliwości intelektualnych. Jeśli tacy uczniowie zostaną włączeni do podgrupy osób uzdolnionych, czego nie możemy ustalić, ich obecność może obniżyć wyższe wyniki uczniów, którzy są naprawdę uzdolnieni intelektualnie.

TABELA 8

Średnie wyniki TONI dla różnych podgrup

Podgrupa główna	Numer badania	Wersja A	Wersja B	Średnia podgrupami
Kobiety	20	101	–	
Kobiety	1	99	100	99
Kobiety	46	97	–	
Mężczyźni	20	99	100	
Mężczyźni	41	–	104	101
Mężczyźni	1	99	–	
Afroamerykanie	20	94	95	95
Afroamerykanie	1	96	–	
Dwujęzyczni	25	97	–	
Dwujęzyczni, równoważna znajomość obu j.	25	100	–	
Dwujęzyczni, przeważający angielski	25	99	–	
Dwujęzyczni, przeważający hiszpański	25	92	–	97
Angielski jako drugi język	20	93	–	
Ograniczona płynność w języku angielskim	3	102	–	
Latynoski i latynoamerykański	20	96	95	
Latynoski i latynoamerykański	30	98	–	96
Podgrupa o statusie wyjątkowym				
ADHD	20	98	98	98
Autyzm/całościowe zaburzenia rozwojowe	24	89	–	
Autyzm/całościowe zaburzenia rozwojowe	21	96	–	
Autyzm/całościowe zaburzenia rozwojowe	37	95	–	

TABELA 8 cd.

Średnie wyniki TONI dla różnych podgrup

Podgrupa główna	Numer badania	Wersja A	Wersja B	Średnia podgrupami
Autyzm/całościowe zaburzenia rozwojowe	28	105	–	
Autyzm/całościowe zaburzenia rozwojowe	23	90	–	
Autyzm/całościowe zaburzenia rozwojowe	36	87	–	92
Autyzm/całościowe zaburzenia rozwojowe	45	100	–	
Autyzm/całościowe zaburzenia rozwojowe: werbalne	24	90	–	
Autyzm/całościowe zaburzenia rozwojowe: niewerbalne	24	79	–	
Autyzm/całościowe zaburzenia rozwojowe: problemy z zachowaniem	22	84	98	
Głuchota	20	98	–	
Głuchota/niedosłuch	19	90	–	
Głęboki niedosłuch	19	94	–	
Ciężki ubytek słuchu	19	88	–	96
Średni ubytek słuchu	19	86	–	
Niedosłuch	16	98	–	
Implant ślimakowy	31	97	–	
Aparat słuchowy	31	111	–	
Opóźniony rozwój	18	69	68	69
Niepełnosprawność intelektualna	2	66	71	
Niepełnosprawność intelektualna	20	71	–	
Dysleksja	6	93	99	98
Dysleksja	20	100	101	

TABELA 8 cd.

Średnie wyniki TONI dla różnych podgrup

Podgrupa główna	Numer badania	Wersja A	Wersja B	Średnia podgrupami
Choroba afektywna dwubiegunowa	32	100	–	
Depresja	33	101	–	98
Zaburzenia emocjonalne	20	93	–	
Uzdolnienia	20	111	109	
Uzdolnienia – szkoła średnia	8	119	–	113
Uzdolnienia – szkoła podstawowa	8	113	–	
Uraz głowy	7	90	–	90
Trudności w nauce	20	95	95	
Trudności w nauce	12	93	–	
Trudności w nauce	43	99	–	93
Trudności w nauce	3	83	–	
Trudności w nauce	13	87	–	
Trudności w nauce czytania	20	96	95	
Inna podgrupa				
Przestępstwo z użyciem przemocy	44	–	95	
Przestępstwo bez użycia przemocy	44	–	100	
Przestępstwo bez użycia przemocy	34	89	–	
Przemoc domowa	34	89	–	
Przemoc ogólna	34	88	–	93
Gwałt	29	93	–	
Kazirodztwo	29	90	–	
Molestowanie dzieci	29	95	–	
Zażywanie narkotyków	40	97	–	

TABELA 8 cd.

Średnie wyniki TONI dla różnych podgrup

Podgrupa główna	Numer badania	Wersja A	Wersja B	Średnia podgrupami
Młodzież w niekorzystnej sytuacji	15	85	–	
Skierowanie w przypadku trudności w nauce	38	96	–	
Skierowanie w przypadku trudności w nauce	42	91	–	
Skierowanie w przypadku pomocy psychologicznej	4	89	–	
Skierowanie w przypadku problemów w klasie	17	90	–	92
Skierowanie w przypadku trudności w nauce	11	96	–	
Skierowanie do poradni dla dzieci	9	91	–	
Skierowanie do poradni dla dzieci	10	92	–	
Choroba Parkinsona	26	91	–	91
Edukacja specjalna/Natywni Amerykanie	41	85	–	85

Dowody z analizy czynnikowej

Parmar (1988) przeanalizowała strukturę czynnikową wyników TONI-2 otrzymanych z wykorzystaniem wersji A i B (cechy charakterystyczne próby – patrz Załącznik C, Badanie 5). Wykazała jedynie niewielkie różnice w ładunkach pozycji na podstawie wykonania testu przez hinduskich badanych.

Brown i in. przeanalizowały strukturę czynnikową wyników obu wersji TONI-3, opierając się na wykonaniu próby normatywnej testu. Zastosowano rotacje ortogonalne i ukośne. Uzasadnienie dla czynników rotacyjnych można znaleźć głównie u Thurstone'a (1947) i Cattella (1978), którzy twierdzili, że procedura ta upraszcza strukturę czynnikową, co uściśla interpretację oraz zwiększa powtarzalność w przypadku różnych prób danych.

Varimax (Kaiser, 1958) jest najpowszechniej stosowaną metodą rotacji, dającą rozwiązanie, które upraszcza interpretację, ponieważ każda zmienna (w tym przypadku każda badana pozycja) ma tendencję do powiązania tylko z jednym lub co najwyżej z niewielką liczbą czynników. W pierwszej analizie czynnikowej głównej składowej badania przeprowadzonego przez Brown i in. (1997) wyłonił się pojedynczy silny czynnik ogólny odpowiadający za dwie trzecie wariancji. Ładunki stanowiły 67% w wersji A i 68% w wersji B. Zdaje się to potwierdzać wcześniejsze wyniki Parmar (1988) dotyczące pojedynczego dominującego czynnika w TONI-2.

Thurstone (1947) rekomendował rotacje ukośne, rozluźniające nieco ograniczenia analizy ortogonalnej, a tym samym oferujące jeszcze większe uproszczenia w interpretacji. Szybką i prostą dostępną metodą jest rotacja promax (Cohen, Cohen, Aiken, West, 2002). Brown i in. (1997) użyły w celu określenia macierzy docelowej wcześniejszej metody rotacji varimax, którą następnie metoda rotacji promax próbowała dopasować do prostego modelu lub struktury. Wyniki skośnych rotacji prawie zawsze interpretuje się, patrząc na korelacje pomiędzy zrotowaną osią a pierwotnymi zmiennymi, interpretowanymi jako ładunki. W próbie TONI-3 ładunki zaprezentowane w tabelach 9 i 10 pokazują, że pierwszy wymiar ponownie wyodrębnia pojedynczy główny lub ogólny

czynnik z dwoma mniejszymi czynnikami odpowiadającymi łącznie za 97% i 99% wariancji, kolejno w wersji A i B.

Analiza stronniczości pozycji testowych

Termin *stronniczość* używany jest na określenie wielu odmiennych, czasem niekonwencjonalnych, a często sprzecznych sposobów. W opracowywaniu narzędzi diagnozy stosuje się go bardzo konkretnie, jako statystyczny deskryptor błędu w prognozowaniu przynależności do grupy na podstawie wyników testu. „Ocena stronniczości jest czysto obiektywną, empiryczną, statystyczną i ilościową kwestią całkowicie niezależną od subiektywnych ocen wartości i kwestii etycznych dotyczących uczciwości testów i ich zastosowania. Stronniczość psychometryczna jest zbiorem statystycznych cech charakterystycznych dla danego testu oraz dwóch lub więcej określonych podpopulacji” (Jensen, 1980, s. 375).

L. Brown i in. (1997) wykazały brak stronniczości pozycji testowych TONI-3 poprzez (a) opis treści testu pod względem potencjalnej stronniczości, (b) uwzględnienie docelowych grup demograficznych w próbie normalizacyjnej proporcjonalnie do ich występowania w danych ze spisu powszechnego w USA, (c) dostarczanie oddzielnych informacji na temat rzetelności i trafności dla grup docelowych oraz (d) wykazanie, że pozycje testowe są odpowiednie dla grup specjalnych, podobnie jak dla populacji ogólnej. Te procedury oraz ich wpływ opisano poniżej.

Testy są stronnicze, jeśli ich zawartość i zadania są nieodpowiednie dla danej grupy w takim zakresie, który niekorzystnie wpływa na wykonanie testu. Testy wykorzystujące treści i zadania niewerbalne są postrzegane jako z natury mniej stronnicze niż testy wykorzystujące język mówiony i pisany (Gonzales, 1982; Holtzman i Wilkinson, 1991; Hoy, Gregg, 1994), ponieważ diagności muszą przyjąć mniej założeń kulturowych i językowych dotyczących badanych osób. Treści i zadania niewerbalnych testów inteligencji są więc odpowiednie dla większej liczby osób. Zalety testów niewerbalnych są oczywiste w przypadku ich wykorzystywania w badaniu osób niesłyszących, posługującym się innymi językami niż dominujący lub mających zaburzenia językowe. Podobnie testy, które zakładają wskazywanie odpowiedzi, mają mniej negatywnych

skutków dla osób z poważnymi zaburzeniami w zakresie motoryki małej. Treść TONI jest niewerbalna i wymaga jedynie wskazywania odpowiedzi, przez co test ten może być mniej stronniczy niż w przypadku testów zawierających treści werbalne, wymagających skomplikowanych odpowiedzi motorycznych, takich jak mówienie, pisanie, rysowanie lub manipulowanie przedmiotami, szczególnie w przypadku osób, które nie są biegłe w stosowaniu danego języka lub są językowo czy też motorycznie niepełnosprawne.

Testy niewerbalne nie są jednak całkowicie pozbawione stronniczości kulturowej (Hoy, Gregg, 1994). Na przykład, wiele testów niewerbalnych wykorzystuje obrazy zamiast języka, ale te obrazy powiązane są z kulturą, która może być lub nie być znana poszczególnym podgrupom osób. Każdy materiał wzrokowy sprawia, że test jest nieodpowiedni do stosowania w przypadku osób, które mają problemy ze wzrokiem lub percepcją. Nawet sama koncepcja „testowania” jest obciążona kulturowo i potencjalnie stronnicza.

TABELA 9

Rotacyjne czynniki promax TONI-3 wersja A

Pozycja testowa	Czynnik		
	1	2	3
A38	.76	.38	.10
A36	.73	.39	.11
A30	.76	.45	.12
A41	.74	.42	.11
A42	.72	.40	.11
A40	.75	.46	.11
A37	.70	.50	.13
A39	.49	.22	.06
A33	.68	.57	.12
A27	.63	.52	.12
A21	.70	.67	.16
A34	.69	.67	.14
A29	.65	.60	.15
A35	.67	.64	.13
A44	.38	.14	.05
A28	.59	.60	.17
A25	.56	.54	.12
A45	.24	.09	.03
A43	.24	.09	.03
A16	.26	.71	.49
A19	.41	.71	.32
A15	.31	.68	.43
A22	.51	.70	.24
A23	.32	.66	.39
A26	.55	.70	.22
A18	.30	.64	.38

TABELA 9 cd.

Rotacyjne czynniki promax TONI-3 wersja A

Pozycja testowa	Czynnik		
	1	2	3
A17	.19	.66	.57
A12	.17	.63	.53
A32	.42	.64	.28
A24	.64	.71	.19
A31	.61	.68	.18
A14	.18	.61	.56
A1	.12	.24	.66
A8	.09	.38	.70
A2	.11	.21	.59
A5	.10	.27	.62
A3	.10	.21	.58
A11	.13	.41	.65
A4	.13	.25	.54
A13	.14	.42	.63
A7	.13	.48	.65
A20	.13	.46	.60
A6	.14	.41	.57
A10	.16	.54	.63
A9	.15	.60	.64
% wariancji wyjaśnionej przez każdy czynnik ignorujący inne czynniki	48	59	35

Anastasi i Urbina (1997) słusznie zauważyli, że „żadne badanie nie może być powszechnie stosowane lub równie *sprawiedliwe* w przypadku wszystkich kultur” (s. 342). Dlatego narzędzia takie jak TONI należy raczej określać jako mniej obciążone kulturowo niż jako całkowicie neutralne kulturowo.

Stronniczość testu można zminimalizować poprzez podkreślenie różnorodności kulturowej w próbie normalizacyjnej oraz poprzez włączenie do grupy normalizacyjnej osób o szerokim zakresie zdolności i niepełnosprawności (Anastasi, Urbina, 1997; R.J. Cohen, Swerdlik, 2002; Crocker, Algina, 2008; Salvia i in., 2007). Grupa normalizacyjna TONI jest wyraźnie reprezentatywna ze względu na płeć, rasę, pochodzenie etniczne, poziom wykształcenia, dochody rodziny, region geograficzny oraz obecność grup o wyjątkowym statusie. Grupy te są reprezentowane w tym samym procencie, co grupy zgłoszone w najnowszym spisie powszechnym w USA i zostały stratyfikowane według wieku.

Testy uważa się za bezstronne, jeśli charakteryzują się podobnymi właściwościami statystycznymi dla wszystkich grup (np. Anastasi, Urbina, 1997; R.J. Cohen, Swerdlik, 2002; Crocker, Algina, 2008; Jensen, 1980; Kline, 2005; Wilson, 2005). Oznacza to, że jeśli na wyniki testu nie wpływa żadna forma stronniczości, to względna kolejność pozycji nie będzie się różnić w zależności od grupy. W praktyce oznacza to, że współczynniki rzetelności i trafności będą w przybliżeniu takie same dla wszystkich podgrup w ramach próby normalizacyjnej. Jeśli jednak nie są, to „jasne jest, że wyniki testu nie będą jednakowymi [rzetelnymi lub trafnymi] miarami dla obu grup” (Jensen, 1980, s. 430). Z wyjątkiem autyzmu, współczynniki alfa dla wybranych podgrup (przedstawione w tabeli 2) są jednakowo wysokie dla wszystkich grup, co sugeruje niską stronniczość testu. Biorąc pod uwagę duże zróżnicowanie w zachowaniu 76 dzieci z zaburzeniami ze spektrum autyzmu na przestrzeni kolejnych dni biorących udział w badaniu, Edleson, Shubert i Edelson (1998) uznali, że współczynnik rzetelności .60 jest wystarczający do zaakceptowania. Wiele badań wykorzystujących różnorodne podgrupy etniczne, rasowe i umiejętności (patrz tabela 8) dowiodło, że TONI jest trafnym narzędziem zarówno w przypadku tych podgrup, jak i w przypadku próby normalizacyjnej jako całości.

Stronniczość narzędzia została także ograniczona w TONI poprzez zastosowany sposób wyboru pozycji testowych, włączający klasyczne techniki analizy pozycji, analizy cech krzywych danej pozycji (Camilli, Shepard, 1994) oraz zastosowanie wartości Delta (Jensen, 1980) dla każdej pozycji TONI w powiązanych podgrupach płciowych, rasowych, etnicznych i umiejętności.

TABELA 10

Rotacyjne czynniki promax TONI-3 wersja B

Pozycja testowa	Czynnik		
	1	2	3
B9	.74	.39	.13
B5	.76	.42	.12
B3	.75	.43	.14
B10	.71	.38	.13
B8	.73	.43	.15
B4	.70	.42	.13
B6	.71	.43	.14
B12	.73	.49	.13
B1	.60	.32	.12
B11	.72	.62	.14
B2	.50	.25	.10
B7	.50	.29	.14
B22	.44	.75	.34
B14	.46	.75	.35
B33	.35	.70	.42
B13	.50	.70	.29
B19	.35	.68	.39
B30	.29	.71	.52
B31	.41	.67	.34

TABELA 10 cd.

Rotacyjne czynniki promax TONI-3 wersja B

Pozycja testowa	Czynnik		
	1	2	3
B23	.28	.67	.51
B25	.42	.61	.25
B18	.52	.65	.26
B20	.61	.66	.22
B24	.24	.62	.50
B21	.29	.56	.33
B15	.67	.66	.18
B17	.56	.62	.22
B16	.46	.54	.17
B41	.13	.44	.78
B43	.12	.41	.74
B38	.13	.42	.73
B39	.11	.38	.68
B40	.09	.28	.56
B27	.16	.55	.71
B34	.15	.53	.70
B37	.08	.20	.50
B36	.15	.54	.69
B26	.16	.57	.69
B29	.15	.54	.67
B28	.18	.60	.69
B45	.06	.16	.43
B32	.17	.56	.65
B42	.06	.13	.39
B35	.17	.59	.64

TABELA 10 c.d.

Rotacyjne czynniki promax TONI-3 wersja B

Pozycja testowa	Czynnik		
	1	2	3
B44	.02	.05	.23
% wariancji wyjaśnionej przez każdy czynnik ignorujący inne czynniki	44	60	43

W tabelach 11 (wersja A) i 12 (wersja B) przedstawiono standardowe wskaźniki stronniczości związane z wykonaniem każdej pozycji TONI przez pięć podgrup. W tabeli 13 przedstawiono liczbę istotnych wskaźników stronniczości w tych grupach, a w tabeli 14 podsumowanie korelacji pomiędzy wartościami Delta dla tych samych pięciu par podgrup. Bardzo nieliczne standardowe wskaźniki stronniczości wynosiły 2.00 lub miały większą wartość, a wszystkie korelacje Delta były bardzo duże, co empirycznie wskazuje, że pozycje testowe TONI-3 nie są szczególnie stronnicze w odniesieniu do grup zróżnicowanych pod względem płci, rasy, etnicznym lub wyjątkowych pod względem intelektualnym i umiejętności uczenia się

Podsumowując, TONI wydaje się być ważną miarą inteligencji, której diagności mogą używać z dużą pewnością, zwłaszcza w przypadku badania osób, wobec których inne testy inteligencji mogą być stronnicze (pod względem kulturowym, językowym lub innym). Zachęcamy specjalistów do dalszego badania testu z wykorzystaniem różnych prób badanych, procedur statystycznych i związanych z nimi miar oraz do dzielenia się wynikami w celu włączenia ich do kolejnych wydań Podręcznika TONI lub tego kompendium. Ciągłe gromadzenie danych poszerzy naszą wiedzę i będzie stanowić wskazówkę w przyszłej rewizji testu, tak jak miało to miejsce w trakcie opracowywania TONI-4.

TABELA 11

Standardowe wskaźniki stronniczości związane z wykonaniem przez pięć dychotomicznych grup TONI-3 wersja A

Wersja A	Grupy dychotomiczne				Z trudno- ściami w nauce/ Inni
	Mężczyźni/ Kobiety	Afroamery- kanie/Inni	Latynosi/ Inni	Uzdolnieni/ Inni	
A1	0.89	0.68	0.33	0.56	0.80
A2	0.36	0.98	0.99	0.32	0.33
A3	1.02	1.26	0.45	0.45	1.34
A4	0.47	0.67	0.26	1.57	1.09
A5	0.91	0.10	0.65	1.26	1.10
A6	0.39	0.13	1.18	0.35	0.39
A7	0.95	0.93	0.19	0.67	0.82
A8	0.34	0.57	0.54	0.21	0.44
A9	1.00	0.07	0.63	0.38	0.23
A10	0.02	0.68	0.64	0.94	1.76
A11	0.52	0.51	0.49	0.79	0.55
A12	0.15	0.80	0.40	0.40	0.40
A13	0.13	0.75	0.46	0.65	1.03
A14	0.70	0.63	2.25	1.19	0.30
A15	0.42	0.77	0.35	0.91	0.92
A16	0.79	0.44	0.12	0.61	0.57
A17	0.61	0.41	0.91	0.40	0.19
A18	0.33	0.69	2.11	0.08	1.29
A19	1.06	0.07	1.11	0.12	0.03
A20	0.21	0.17	0.11	1.22	1.96
A21	2.36	1.14	0.91	1.49	0.18
A22	1.16	1.88	1.82	1.36	0.85
A23	0.20	0.24	0.85	0.09	1.10

TABELA 11 cd.

Standardowe wskaźniki stroniczości związane z wykonaniem przez pięć dychotomicznych grup TONI-3 wersja A

Wersja A	Grupy dychotomiczne					Z trudno- ściami w nauce/ Inni
	Mężczyźni/ Kobiety	Afroamery- kanie/Inni	Latynosi/ Inni	Uzdolnieni/ Inni		
A24	0.19	0.38	0.34	0.95		0.24
A25	0.60	0.64	1.19	1.52		0.66
A26	1.68	0.68	0.76	3.17		0.38
A27	0.55	0.57	2.15	1.91		0.59
A28	1.49	0.22	0.22	1.67		0.95
A29	1.80	0.74	1.38	2.56		1.25
A30	0.03	1.59	0.65	1.00		1.32
A31	0.02	0.78	1.22	1.20		1.04
A32	1.15	0.39	1.86	0.81		0.45
A33	1.21	1.58	0.88	1.86		1.40
A34	0.48	1.03	0.62	2.07		2.56
A35	1.45	0.73	1.60	1.34		0.91
A36	0.51	0.49	0.13	0.49		3.19
A37	1.18	0.95	0.12	0.71		2.34
A38	1.86	2.35	1.67	1.42		0.25
A39	0.62	1.98	1.14	0.21		0.01
A40	0.51	2.34	0.59	1.35		1.22
A41	1.63	1.50	1.17	1.41		0.01
A42	1.28	1.84	0.02	0.82		1.58
A43	0.41	2.07	0.35	0.37		0.29
A44	2.19	1.23	1.02	0.33		0.24
A45	0.16	0.93	0.77	0.16		0.55

TABELA 12

Standardowe wskaźniki stronniczości związane z wykonaniem przez
pięć dychotomicznych grup TONI-3 wersja B

Wersja B	Grupy dychotomiczne				Z trudno- ściami w nauce/ Inni
	Mężczyźni/ Kobiety	Afroamery- kanie/Inni	Latynosi/ Inni	Uzdolnieni/ Inni	
B1	0.14	0.08	1.31	0.30	0.82
B2	0.31	1.44	0.21	0.21	0.02
B3	0.60	1.88	0.10	1.60	0.50
B4	0.35	0.49	0.08	0.33	0.72
B5	0.11	0.61	0.71	1.19	1.49
B6	0.81	1.16	0.16	0.26	0.31
B7	0.38	0.74	1.47	0.06	0.43
B8	0.18	0.74	1.52	1.93	0.07
B9	0.52	0.67	0.02	1.33	0.35
B10	0.59	1.40	1.06	0.14	0.45
B11	0.29	1.14	1.04	2.06	0.04
B12	0.07	1.77	0.31	0.63	0.34
B13	1.29	2.34	1.15	0.55	0.91
B14	1.48	0.79	1.45	1.50	1.05
B15	0.44	0.61	1.32	0.37	0.53
B16	1.04	1.45	0.42	0.46	0.46
B17	0.09	1.25	0.79	1.30	0.89
B18	0.40	0.49	1.30	1.41	0.36
B19	0.59	0.78	1.12	1.17	0.74
B20	0.39	0.79	1.48	0.05	2.30
B21	0.34	1.15	3.47	0.76	0.83
B22	1.40	0.59	0.19	0.30	0.46
B23	1.48	0.70	1.49	1.05	1.20

TABELA 12 cd.

Standardowe wskaźniki stroniczości związane z wykonaniem przez pięć dychotomicznych grup TONI-3 wersja B

Wersja B	Grupy dychotomiczne				
	Mężczyźni/ Kobiety	Afroamery- kanie/Inni	Latynosi/ Inni	Uzdolnieni/ Inni	Z trudno- ściami w nauce/ Inni
B24	1.36	0.47	0.74	2.25	0.53
B25	0.37	2.16	0.57	1.30	0.17
B26	0.55	1.62	1.36	1.22	1.58
B27	0.74	1.45	0.39	0.62	1.28
B28	0.19	1.96	1.16	1.18	1.96
B29	1.00	0.73	1.00	1.02	1.72
B30	0.77	0.39	0.21	0.65	0.08
B31	1.36	0.95	1.63	0.25	2.34
B32	0.76	0.55	0.88	2.59	0.17
B33	1.13	1.18	2.35	1.80	3.62
B34	0.56	0.50	0.46	0.17	2.17
B35	0.37	1.14	0.10	1.40	2.33
B36	0.37	1.76	0.58	1.89	1.63
B37	1.90	1.05	0.31	0.12	0.27
B38	0.85	2.11	0.78	0.06	0.10
B39	0.44	1.86	1.66	0.38	1.37
B40	1.38	1.29	0.87	1.62	0.27
B41	3.30	1.03	0.29	0.26	0.69
B42	1.18	1.29	0.63	1.65	0.65
B43	0.20	1.25	2.04	1.23	1.55
B44	0.96	0.30	0.76	0.13	0.18
B45	1.80	1.75	1.42	1.42	0.09

TABELA 13

Liczba istotnych wskaźników stronniczości wykrytych w pięciu grupach dychotomicznych dla TONI-3

Grupy dychotomiczne	Wersja A	Wersja B
Mężczyźni/Kobiety	2	1
Afroamerykanie/Inni	3	3
Latynosi/Inni	2	4
Uzdolnieni/Inni	3	2
Z trudnościami w nauce/ Inni	6	2

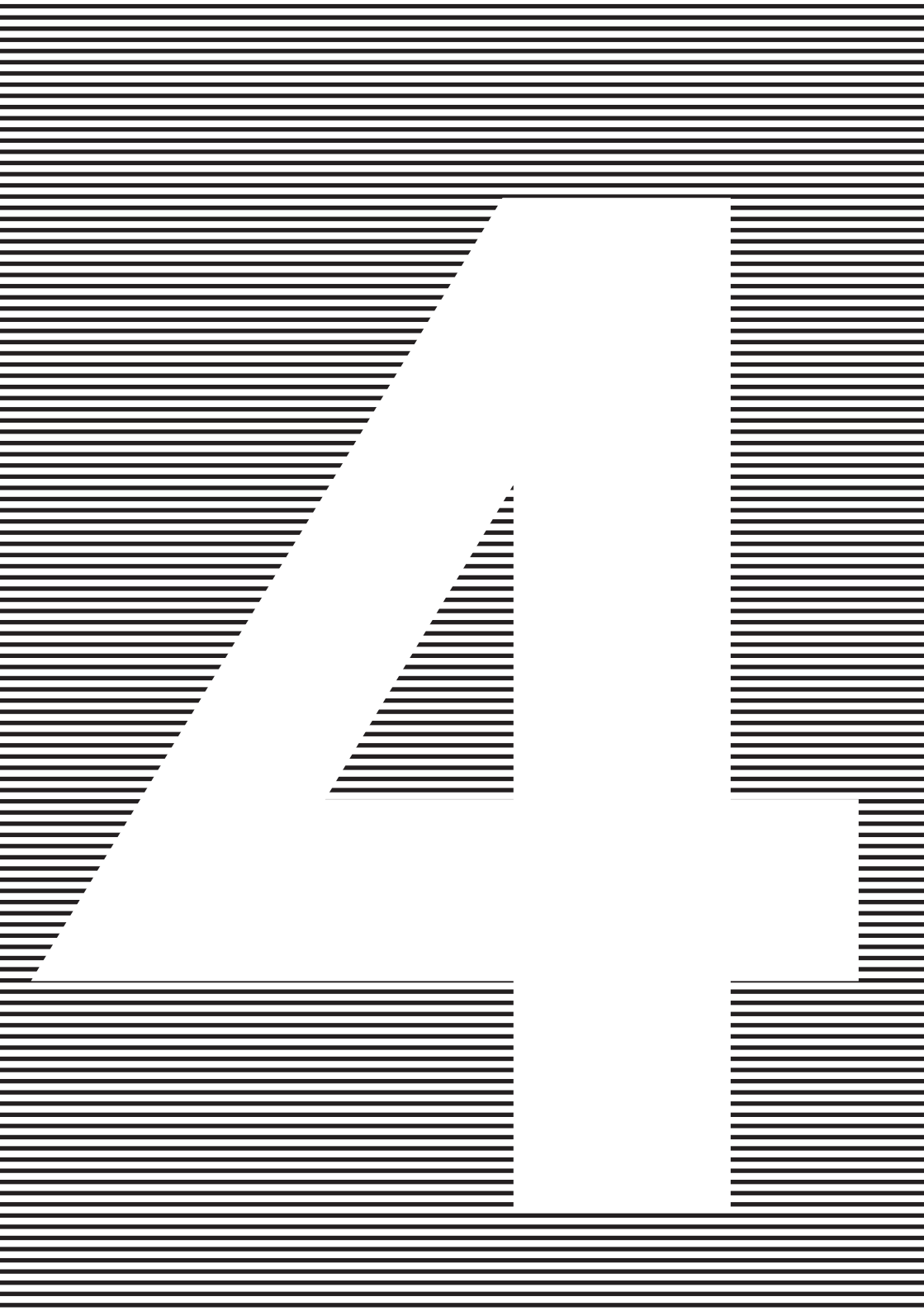
TABELA 14

Korelacja pomiędzy wartościami Delta związanymi z siedmioma podgrupami próby normalizacyjnej TONI-3 (pominięto dziesiątne)

Grupy dychotomiczne	Wersja A	Wersja B
Mężczyźni/Kobiety	99	99
Afroamerykanie	99	98
ESL/Inni	98	98
Latynosi/Inni	98	99
Głuchota/Niedosłuch	99	99
Uzdolnieni/Inni	98	99
Z trudnościami w nauce/ Inni	99	99

Uwaga: ESL – Język angielski jako drugi język





Nowe amerykańskie wydanie TONI-4



Przy tworzeniu TONI-4 autorzy dokładnie analizowali uwagi i komentarze recenzentów, wyniki wszystkich poprzednich badań przeprowadzonych z wykorzystaniem TONI oraz inne sugestie użytkowników testu. Zostały one połączone z własnymi pomysłami autorów na to, w jaki sposób zwiększyć użyteczność testu. Test został ulepszony zwłaszcza w następujących zakresach:

1. Zebrano nowe dane normalizacyjne w celu zaprezentowania zmian demograficznych odzwierciedlonych w *Statystycznym Abstrakcie USA* (U.S. Census Bureau, 2009). Uwzględniono pochodzenie etniczne, rasę, płeć, region geograficzny, wykształcenie rodziców oraz status społeczno-ekonomiczny, a dane normalizacyjne zostały stratyfikowane według wieku, płci, regionu geograficznego, pochodzenia etnicznego i rasy.
2. Zmniejszono stroniczość związaną z płcią, rasą, pochodzeniem etnicznym i niepełnosprawnością.
3. Rzetelność i trafność testu zostały wykazane w poszczególnych grupach zróżnicowanych pod względem płci, niepełnosprawności, etnicznym/rasowym, jak również dla populacji ogólnej.
4. Wskazano współczynniki rzetelności dla całej próby normalizacyjnej oraz dla poszczególnych podgrup pod względem zdolności poznawczych, rasy, pochodzenia etnicznego i płci.
5. Dodano badania nad trafnością ze szczególnym uwzględnieniem tych podgrup.

- 
6. Podano średnie i odchylenia standardowe dla wszystkich badań nad trafnością kryterialną.
 7. Pozycje testowe zostały poddane ponownej ocenie zarówno za pomocą analiz w celu uwzględnienia pozycji dobrych, jak i analiz funkcjonowania pozycji różnicujących w celu wyeliminowania pozycji stronicznych.
 8. Dodano pozycje odnoszące się do niższego poziomu funkcjonowania intelektualnego w celu poprawy efektu podłogi dla podgrupy osób niepełnosprawnych intelektualnie.
 9. Zbadano adekwatność techniczną pozycji z każdej z poprzednich edycji TONI; zachowano lub przywrócono pozycje adekwatne technicznie.
 10. Podano korelacje między wersjami we wszystkich grupach wiekowych i korelacje między TONI-4 i TONI-3.
 11. Podano instrukcje zarówno werbalne, jak i niewerbalne, a także dane równoważne dla osób, które nie mają zaburzeń językowych ani motorycznych.
 12. Zaoferowano diagnostom pracującym z osobami badanymi niemówiącymi w języku angielskim instrukcje w głównych językach obcych.
 13. Pozycje zostały uporządkowane i posortowane w taki sposób, aby zrównoważyć wersję A i wersję B pod względem poziomu trudności i pomiaru umiejętności rozwiązywania problemów.

Od czasu przeglądu badań i recenzji wykonanego przez autorki testu za lata 1982–2009 (Johnsen, Brown, Sherbenou, 2010) pojawił się szereg badań z wykorzystaniem trzeciej oraz czwartej wersji TONI. Część badaczy koncentrowała się na wykazaniu związku wyników TONI-4 z innymi uznanymi skalami inteligencji, szczególnie ciekawe wydają się te poświęcone grupom specjalnym. Samakkaew (2019) w badaniach dzieci wykorzystanych seksualnie dowodzi obecność wysokiej korelacji ($r = 0,763$) wyników TONI-4 ze skalą WISC-III (z istotnością statystyczną na poziomie 0,01). Badania wykazały ponadto wysokie korelacje TONI-4 z Indekssem Organizacji Percepcyjnej i Indekssem Szybkości Przetwarzania pochodzącymi z tajskiej wersji WISC-III (odpowiednio $r = 0,749$ i $0,758$; z istotnością statystyczną na poziomie 0,01).

TONI był również wykorzystywany na szeroką skalę w badaniach wielu innych grup specjalnych, na przykład TONI-3 w badaniach dzieci z zaburzeniami ze spektrum autyzmu i innymi zaburzeniami rozwojowymi (Edelson 2005; Gillespie-Lynch i in., 2020), dzieci niesłyszących i niedosłyszących (Cassiano Barbosa, Lukasova, Mecca, Macedo, 2013), TONI-4 w badaniach dzieci z trudnościami w uczeniu się (Adlof i in., 2017), a także pacjentów ze schizofrenią (Chen i in., 2021), gdzie okazał się być użytecznym narzędziem do oceny funkcjonowania intelektualnego osób badanych.

Ponadto prowadzono badania dotyczące równoważności testu lub walidacji krzyżowej między innymi w Chile, Meksyku czy Indiach (Johnsen, 2017). Pomimo tego, że wersja TONI-4 nie wymaga używania języka i zawiera treści neutralne kulturowo, umożliwiając jego stosowanie bez względu na kraj zamieszkania czy też język, podkreślano, że amerykańskie normy mogą nie być odpowiednie do stosowania wobec populacji innych krajów czy kultur, stąd też chcąc zapewnić najwyższą rzetelność i trafność narzędzia, powinno się przeprowadzać jego pełną normalizację w różnych populacjach osób badanych.

Zalety TONI-4

Aktualnie, na podstawie dotychczasowych recenzji oraz wprowadzonych ulepszeń w obecnej wersji narzędzia (Handleman and Delmolino 2005; McLoughlin and Lewis, 2005; Reynolds and Fletcher-Janzen 2004; Salvia i in. 2007; Strauss i in. 2006), wskazuje się następujące jego zalety:

- test jasno określa zasady diagnozowania inteligencji niewerbalnej oraz wskazuje na teoretyczne i filozoficzne podstawy badania inteligencji ogólnej oraz predyspozycji intelektualnych;
- jest użytecznym narzędziem służącym do szybkiej oceny zdolności rozumowania niewerbalnego u dzieci i dorosłych;
- pozycje testowe konsekwentnie sprawdzają umiejętność rozwiązywania problemów niewerbalnych na materiale figuralnym i zostały opracowane przy zastosowaniu zarówno klasycznych technik analizy pozycji testowych, jak i zaawansowanych metod statystycznych;
- jakość i liczba badań dotyczących rzetelności i trafności narzędzia obrazuje szeroki zakres jego przydatności. W wyczerpujący

sposób autorzy dokumentują związek testu z innymi miarami inteligencji i osiągnięć, jego skuteczność w różnicowaniu poszczególnych grup oraz strukturę czynnikową;

- TONI-4 ma dwie wersje (A i B), co daje możliwość sprawdzenia efektywności programów terapeutycznych i wprowadzonych interwencji;
- zawiera nowe, dobrze skonstruowane i kompleksowo opracowane normy, reprezentatywne dla populacji USA. Dane normalizacyjne zostały ponadto przygotowane w podziale na wiek, płeć, obszar geograficzny, pochodzenie etniczne oraz rasę;
- test jest nieograniczony czasowo, umożliwiając uwzględnienie różnic indywidualnych, w szczególności dla osób, które mogą potrzebować więcej czasu na wykonanie zadań. Włączenie pozycji treningowych umożliwia badającemu ocenę adekwatności narzędzia dla danej osoby badanej;
- w przypadku osób biegle posługujących się językiem angielskim można przeprowadzić test, korzystając z ustnych instrukcji, natomiast w przypadku osób, które nie rozumieją lub nie mówią po angielsku dostępne są instrukcje niewerbalne;
- test jest szczególnie przydatny w diagnostyce inteligencji osób z niepełnosprawnością sprzężoną, można nim zbadać prawie wszystkie grupy specjalne, z wyjątkiem osób niewidomych.
- podręcznik jest napisany bardzo jasnym i przejrzystym językiem, jest przydatny zarówno dla studentów przeszkolonych w zakresie psychometrii, jak i praktyków;
- nowsze rysunki, generowane komputerowo są znacznie ulepszone, wyraźne i przejrzyste.

Ograniczenia TONI-4

Ze względu na systematyczne uwzględnianie krytycznych uwag pojawiających się przy tworzeniu nowych wersji TONI, w TONI-4 wskazano jedynie kilka ograniczeń:

- TONI-4 nie mierzy szerokiego wachlarza umiejętności intelektualnych i umożliwia ocenę jedynie umiejętności rozwiązywania problemów w wersji niewerbalnej, na materiale figuralnym;
- pomimo że badania trafności były przeprowadzone w grupach specjalnych, nie stworzono oddzielnych dla nich norm;
- zaproponowano więcej badań dotyczących relacji pomiędzy dwiema wersjami testu, różnymi edycjami narzędzia, a także

efektów „podłogi” w badaniu dzieci z zaburzeniami ze spektrum autyzmu oraz niepełnosprawnością umysłową. Jak wspomniano powyżej, TONI-4 nie mierzy szerokiego wachlarza umiejętności intelektualnych, ale jest to wada jedynie wtedy, gdy test nie jest właściwie używany. Zachęcamy wszystkich stosujących TONI-4 do samodzielnego opracowywania specjalistycznych norm dla grup specjalnych oraz kontynuowania badań z wykorzystaniem TONI-4 w celu określenia jego przydatności wobec różnych grup osób badanych.



Bibliografía

Adlof, S.M., Scoggins, J., Brazendale, A., Babb, S., & Petscher, Y. (2017). Identifying Children at Risk for Language Impairment or Dyslexia With Group-Administered Measures. *Journal of speech, language, and hearing research: JSLHR*, 60(12), 3507-3522. doi:10.1044/2017_JSLHR-L-16-0473

Brown, L., Sherbenou, R.J., & Johnsen, S.K. (1995). *Test de inteligencia no verbal: Apreciación de la habilidad cognitiva sin influencia del lenguaje*. Madrid, España: TEA Ediciones, S.A.

Brown, L., Sherbenou, R., & Johnsen, S. (2014). *Test of Nonverbal Intelligence-4. [ZEKA TESTİ UYGULAYICI YETİŞTİRME SERTİFİKA KURSU]*. Konya, Turkey.

Cassiano Barbosa, A.C., Lukasova, K., Meca, T.P., & Macedo, E.C. (2013). Intelligence assessment of deaf students with TONI 3. Evaluación de la inteligencia de los alumnos sordos con TONI 3. *Psico-USF*, 18(2), 183-191. doi:10.1590/s1413-82712013000200002

Chen, K.-W., Lee, Y.-C., Yu, T.-Y., Cheng, L.-J., Chao, C.-Y., & Hsieh, C.-L. (2021). Test-retest reliability and convergent validity of the Test of Nonverbal Intelligence-Fourth Edition in Patients with Schizophrenia. *BMC Psychiatry*, 21. doi:10.1186/s12888-021-03041-4

Edelson, M.G. (2005). A car goes in the garage like a can of peas goes in the refrigerator: Do deficits in real-world knowledge affect the assessment of intelligence in individuals with autism? *Focus on Autism and Other Developmental Disabilities*, 20, 2-9.

Gillespie-Lynch, K., Hotez, E., Zajic, M., Riccio, A., DeNigris, D., Kofner, B., Luca, K. (2020). Comparing the writing skills of autistic and non-autistic university students: A collaboration with autistic university students. *Autism*, 24(7), 1898-1912. doi:10.1177/1362361320929453

Handleman, J.S., & Delmolino, L.M. (2005). Assessment of children with autism. In D.B. Zager (Ed.), *Autism spectrum disorders:*



Identification, education, and treatment (3rd ed., pp. 269–294). London: Routledge.

Johnsen, S.K. (2017). Test of Nonverbal Intelligence: A Language-Free Measure of Cognitive Ability. In R.S. McCallum (Ed.), *Handbook of Nonverbal Assessment* (pp. 185–206). Cham: Springer International Publishing.

McLoughlin, J.A., & Lewis, R.B. (2005). *Assessing special students (6th ed.)*. Columbus, OH: Merrill.

Reynolds, C.R., & Fletcher-Janzen, E. (Eds.). (2004). *Concise encyclopedia of special education: A reference for the education of the handicapped and other exceptional children and adults (2nd ed.)*. Hoboken, NJ: Wiley.

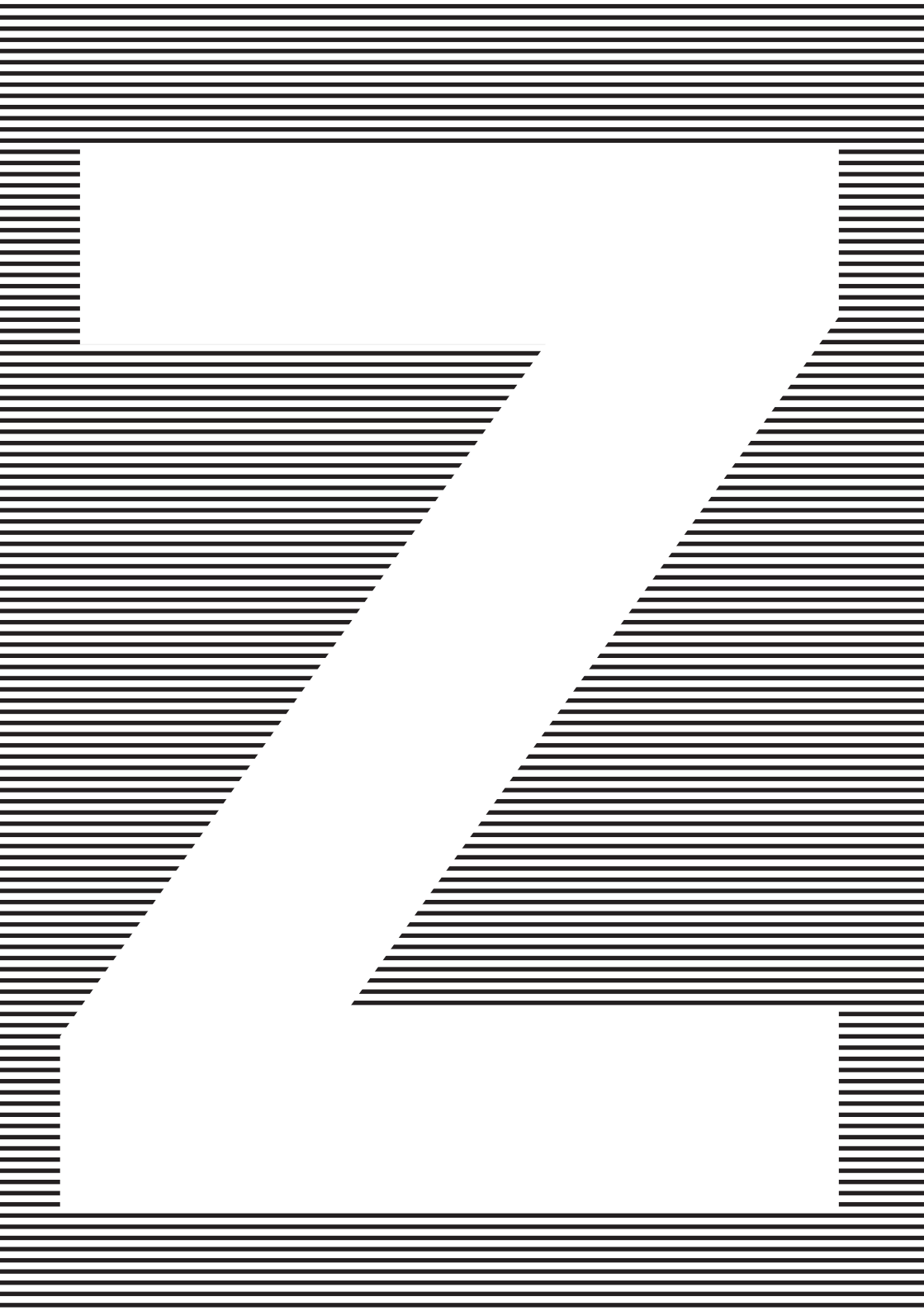
Salvia, J., Ysseldyke, J.E., & Bolt, S. (2007). *Assessment in special and inclusive education (10th ed.)*. Boston: Houghton Mifflin.

Strauss, E., Sherman, E.M.S., & Spreen, O. (2006). *A compendium of neuropsychological tests: Administration, norms, and commentary (3rd ed.)*. Oxford: Oxford University Press.

Wu-Tien, W., Chung-chien, T., Jyh-fen, H.J.W., Hsin-tai, L., & Ching-chih, K. (1995). *The chinese version of the test of nonverbal intelligence (TONI-2) and its related studies*. Taipei, Taiwan: Taiwan Normal University.

Wanitchaya, S. (2019) Relationships between Thai Version Intelligence WISC-III Test Scores and TONI-4 in Sexually Abused Children. *Royal Thai Air Force Medical Gazette*, 65(1), 20–27.





Załącznik A

Charakterystyka badań korelujących TONI z umiejętnościami

Nr badań	Źródło	N	Opis próby
1	L. Brown, Sherbenou, Johnsen (1982)	25	- zakres wiekowy = 7.1–10.3 lat ($M = 8.1$ lat) - mężczyźni = 13, kobiety = 12 - Biali = 10, Afroamerykanie = 3, inni = 1 - rozwój typowy
2	L. Brown i in. (1982)	16	- zakres wiekowy = 10.7–13.5 lat ($M = 12.1$ lat) - mężczyźni = 12, kobiety = 4 - Biali = 10, Afroamerykanie = 6 - trudności w uczeniu się
3	L. Brown i in. (1982)	30	- zakres wiekowy = 14.8–17.2 lat ($M = 16.1$ lat) - mężczyźni = 26, kobiety = 4 - Biali = 28, Afroamerykanie = 2 - głusi
4	L. Brown i in. (1982)	10	- średni wiek = 11.5 lat - mężczyźni = 7, kobiety = 3 - Biali = 4, Afroamerykanie = 6 - niepełnosprawność intelektualna objęta edukacją
5	L. Brown i in. (1982)	16	- zakres wiekowy = 16.1–18.8 lat ($M = 17.9$ lat) - mężczyźni = 12, kobiety = 4 - Biali = 15, Afroamerykanie = 1 - głusi
6	Bond, Kennon (1982)	71	- zakres wiekowy = 7–18 lat - mężczyźni = 43, kobiety = 28 - Biali = 52, Afroamerykanie = 19 - Memphis, TN

Nr badań	Źródło	N	Opis próby
7	Haddad (1986)	66	• średni wiek = 9.4 lat ($SD = 1.8$) • mężczyźni = 51, kobiety = 15 • Biali = 41, Afroamerykanie = 25 • trudności w uczeniu się
8	Vance, Hankins, Brown (1986)	51	• zakres wiekowy = 7.2–15.5 lat • Biali = 31, Afroamerykanie = 20 • rozwój typowy • Northern Delaware
9	Nasca (1988)	221	• sześciolatki przedszkolne • intelektualnie uzdolnieni
10	Salas (1988)	61	• średni wiek = 10.4 lat • mężczyźni = 48, kobiety = 13 • Biali = 53, Afroamerykanie = 8 • dysleksja
11	Vance, Hankins, Reynolds (1988)	89	• zakres wiekowy = 6.3–16.8 lat ($M = 11.1$ lat) • mężczyźni = 61, kobiety = 28 • Biali = 52, Afroamerykanie = 37 • umiarkowanie niepełnosprawni
12	Brown, Sherbenou, Johnsen (1990)	94	• zakres wiekowy = 6.0–53.0 lat ($M = 23.9$ lat) • mężczyźni = 39, kobiety = 55 • Biali = 91, inni = 3 • rozwój typowy
13	Brown i in. (1990)	139	• zakres wiekowy = 6.0–17.8 lat ($M = 8.4$ lat) • mężczyźni = 65, kobiety = 74 • Biali = 134, Afroamerykanie = 1, inni = 4 • intelektualnie uzdolnieni
14	Brown i in. (1990)	40	• głównie Biali • trudności w uczeniu się
15	Brown i in. (1990)	103	• zakres wiekowy = 10–83 lat ($M = 30.8$ lat) • Biali = 97, Afroamerykanie = 2, inni = 3 • urazy głowy
16	Kowall, Watson, Madak	30	• zakres wiekowy = 6.4–11.6 lat ($M = 9.1$ lat) • mężczyźni = 20, kobiety = 10 • skierowani do poradni dla dzieci • przedmieścia • status socjoekonomiczny – klasa średnia • Kanadyjczycy

Nr badań	Źródło	N	Opis próby
17	Kowall i in (1990)	22	- zakres wiekowy = 6.5–11.8 lat ($M = 8.5$ lat) - mężczyźni = 14, kobiety = 8 - skierowani do poradni dla dzieci - przedmieścia - status socjoekonomiczny – klasa średnia - Kanadyjczycy
18	Whorton, Morgan (1990)	46	- średni wiek = 11.4 lat - mężczyźni = 28, kobiety = 18 - Biali = 29, Rdzenni Amerykanie = 7 - skierowani na badanie psychologiczne
19	Lassiter, Bardos (1992)	36	- zakres wiekowy = 7–12 lat ($M = 9.8$ lat, $SD = 1.5$) - mężczyźni = 28, kobiety = 8 - zaburzenia uczenia - południowe stany USA
20	Coleman, Scribner, Johnsen, Evans (1993)	37	- zakres wiekowy = 16–21 lat ($M = 17.9$ lat) - mężczyźni = 26, kobiety = 11 - Amerykanie meksykańskiego pochodzenia, ograniczona znajomość języka angielskiego = 15 - zaburzenia uczenia - Teksas
21	Powell, Germani (1993)	53	- zakres wiekowy = 8–12 lat - mężczyźni = 40, kobiety = 13 - osoby posługujące się wyłącznie ojczystym językiem angielskim - zdiagnozowane zaburzenia komunikacji
22	Cash (1994)	205	- zakres wiekowy = 5.7–18.4 lat ($M = 10.2$ lat) - mężczyźni = 107, kobiety = 98 - zaburzenia słuchu = 205 - Szkoła dla Głuchych w Utah
23	D'Amato, Lidiak, Lassiter (1994)	74	- zakres wiekowy = 6–15 lat ($M = 10.2$ lat, $SD = 2.5$) - mężczyźni = 59, kobiety = 15 - skierowani z trudnościami szkolnymi - w większości niski i średni status socjoekonomiczny - południowe i środkowozachodnie stany USA

Nr badań	Źródło	N	Opis próby
24	Auerbach (1995)	65	- zakres wiekowy = 22–75 lat ($M = 39.7$ lat) - mężczyźni = 29, kobiety = 36 - Biali = 58, Afroamerykanie = 6, Hiszpanie = 1 - mieszkańcy domów „nadzorowanych” = 62, mieszkający niezależnie = 3 - niepełnosprawność intelektualna lub ciężka niepełnosprawność fizyczna
25	Perez, Slate, Neeley, McDaniel, Baggs, Layton (1995)	40	- średni wiek = 10.9 lat - mężczyźni = 29, kobiety = 11 - Biali = 40 - Skierowani do poradni zaburzeń mowy i słuchu - środkowopółnocne stany USA
26	Frey (1996)	311	- zakres wiekowy = 5.8–17.11 lat ($M = 14.6$ lat, $SD = 2.4$) - mężczyźni = 155, kobiety = 156 - w większości niski status socjoekonomiczny, - śródmieścia, Afroamerykanie - większość zaniedbujących lub przestępczych środowisk z orzeczeniami sądu rodzinnego - podopieczni ośrodka Św. Elżbiety w Waszyngtonie, DC
27	Hammill, Pearson, Wiederholt (1996)	43	- młodszy wiek szkolny - różne zaburzenia uczenia się
28	Hammill i in. (1996)	550	- zakres wiekowy = 19–50 lat - struktura demograficzna wg. spisu powszechnego z 1990
29	Mackinson (1996)	27	- zakres wiekowy = 6.2–14.9 lat ($M = 11.3$ lat) - Biali = 10, Afroamerykanie = 15, Latynosi = 1, inni = 1 - głęboki ubytek słuchu = 13, znaczny ubytek słuchu = 8, umiarkowany ubytek słuchu = 6 - głuche dzieci słyszących rodziców = 19, głuche dzieci głuchych rodziców = 8 - ciężki stan = 15, lżejszy stan = 12 - dzienna szkoła dla głuchych - duże metropolie w stanach wschodniego wybrzeża USA
30	Brown, Sherbenou, Johnsen (1997)	34	- zakres wiekowy = 7–17 lat - mężczyźni = 27, kobiety = 7 - Biali = 29, mniejszości = 2 - zaburzenia uczenia

Nr badań	Źródło	N	Opis próby
31	Brown i in. (1997)	19	- zakres wiekowy = 16–19 lat - mężczyźni = 14, kobiety = 5 - Biali = 17, mniejszości = 2 - zaburzenia uczenia
32	Mackinson, Leigh, Blennerhassett, Anthony (1997)	27	- zakres wiekowy = 6.2–14.9 lat ($M = 11.3$ lat) - mężczyźni = 15, kobiety = 12 - Biali = 10, Afroamerykanie = 15, Latynosi = 1, inni = 1 - głusi i z zaburzeniami słuchu - dzienna szkoła dla głuchych
33	Brinson (1999)	60	- Afroamerykanie z anemią sierpowatą i udarem mózgu w wywiadzie = 20 (zakres wiekowy 7.4–16.9 lat, $M = 11.7$; mężczyźni = 13, kobiety = 7) - Afroamerykanie z anemią sierpowatą = 20 (zakres wiekowy = 7.1–16.9 lat, $M = 12.2$; mężczyźni = 11, kobiety = 9; 2 w edukacji specjalnej) - Afroamerykanie = 20 (zakres wiekowy = 7.2–16.3 lat; $M = 11.6$; mężczyźni = 5, kobiety = 15; 2 w edukacji specjalnej) - Klinika Pediatrii Uniwersytetu na Florydzie i w Birmingham, AL
34	Alvarado (1999)	90	- klasy 1 i 2 - mężczyźni = 45, kobiety = 45 - dwujęzyczni - południowy Teksas
35	Barrett (2000)	217	- zakres wiekowy = 6–17 lat - mężczyźni = 101, kobiety = 116 - szkoły niepubliczne = 101, szkoły publiczne = 116 - Jamajczycy
36	Daisley (2000)	58	- zakres wiekowy = 18–36 lat ($M = 21.1$) - mężczyźni = 25, kobiety = 33 - dwa duże wschodnie amerykańskie uniwersytety
37	Daisley (2000)	50	- zakres wiekowy = 18–28 lat ($M = 22$) - mężczyźni = 22, kobiety = 28 - głuchota prelingwalna - dwa duże wschodnie amerykańskie uniwersytety

Nr badań	Źródło	N	Opis próby
38	Bostantjopoulou, Kiosseoglou, Katsarou, Alevriadou (2001)	122	- dorośli z depresją, z Grecji = 75 (średni wiek = 68.4 [SD = 4.08]; mężczyźni = 44, kobiety = 31; 7.2 lat edukacji [SD = 3.76]) - dorośli bez depresji, z Grecji = 47 (średni wiek = 67.8 lat [SD = 4.41]; mężczyźni = 28; kobiety = 19; 6.2 lat edukacji [SD = 3.66])
39	Sedore (2002)	40	- zakres wiekowy = 18–23 lat ($M = 19.7$, $SD = 1.4$) - mężczyźni = 20, kobiety = 20 - Biali = 40 - uczniowie college'u - wschodnie Kentucky
40	Hopper (2002)	46	- zakres wiekowy = 18.0–20.10 lat - mężczyźni = 13, kobiety = 33 - Biali = 46 - uczniowie college'u - północno-zachodnie stany USA
41	Edelson (2005)	35	- zakres wiekowy = 4–18 lat ($M = 10.12$, $SD = 3.8$) - mężczyźni = 23, kobiety = 12 - autyzm - Amerykanie = 29, Włosi = 6
42	Chan, Cheung, Leung, Cheung, Cheung (2005)	46	- zakres wiekowy = 5–6 lat - mężczyźni = 31, kobiety = 15 - typowy rozwój = 27, autyzm = 19 - Chińczycy posługujący się językiem ojczystym - Hong Kong
43	Strattman, Hodson (2005)	75	- zakres wiekowy = 7.2–9.2 lat ($M = 7.10$) - mężczyźni = 30, kobiety = 45 - Biali = 47, Afroamerykanie = 21, Amerykanie pochodzenia azjatyckiego = 7 - typowy rozwój - klasa II - środkowozachodnie metropolitalne szkoły publiczne USA
44	Essix, Sapp, Abbott, Kohler (2006)	75	- zakres wiekowy = 7.9–16.5 lat ($M = 10.7$, $SD = 2.45$) - mężczyźni = 41, kobiety = 34 - Afroamerykanie = 75 - skierowani ze specyficznymi trudnościami w uczeniu się

Nr badań	Źródło	N	Opis próby
45	Johnson (2007)	90	- zakres wiekowy = 6.0–16.1 lat ($M = 11.09$, $SD = 2.68$) - mężczyźni = 46, kobiety = 44 - Rdzeni amerykańanie (Sioux) - Standardowa edukacja = 30 (mężczyźni = 15, kobiety = 15) - edukacja specjalna = 30 (mężczyźni = 21, kobiety = 9) - uzdolnieni = 30 (mężczyźni = 10, kobiety = 20) - północnowschodnia Południowa Dakota
46	Karp (2007)	56	- zakres wiekowy = 22–56 lat ($M = 41.6$) - mężczyźni = 29, kobiety = 27 - Biali = 4, Afroamerykanie = 47, Latynosi = 5 - światlica środowiskowa = 30, ośrodek dziennej terapii = 26 - Nowy Jork

—

Załącznik B

Charakterystyka badań korelujących TONI do osiągnięć edukacyjnych

Nr badań	Źródło	N	Opis próby
1	Brown, Sherbenou, Johnsen (1982)	112	- zakres wiekowy = 7.0–12.6 lat ($M = 9.5$ lat) - mężczyźni = 58, kobiety = 54 - Biali = 90, Afroamerykanie = 21, inni = 19 - rozwój typowy
2	Brown i in. (1982)	16	- zakres wiekowy = 10.7–13.5 lat ($M = 12.1$ lat) - mężczyźni = 12, kobiety = 4 - Biali = 10, Afroamerykanie = 6 - trudności w uczeniu się
3	Brown i in. (1982)	16	- zakres wiekowy = 16.1–18.8 lat ($M = 17.9$ lat) - mężczyźni = 12, kobiety = 4 - Biali = 15, Afroamerykanie = 1 - głusi
4	Haddad (1986)	66	- średni wiek = 9.4 lat ($SD = 1.8$) - mężczyźni = 51, kobiety = 15 - Biali = 41, Afroamerykanie = 25 - trudności w uczeniu się
5	Salas (1988)	61	- średni wiek = 10.4 lat - mężczyźni = 48, kobiety = 13 - Biali = 53, Afroamerykanie = 8 - dysleksja
6	L. Brown, Sherbenou, Johnsen (1990)	94	- zakres wiekowy = 6.0–53.0 lat ($M = 23.9$ lat) - mężczyźni = 39, kobiety = 55 - Biali = 91, inni = 3 - rozwój typowy

Nr badań	Źródło	N	Opis próby
7	L. Brown i in. (1990)	40	• głównie Biali • trudności w uczeniu się
8	Mackinson (1996)	27	• zakres wiekowy = 6.2–14.9 lat ($M = 11.3$ lat) • Biali = 10, Afroamerykanie = 15, Latynosi = 1, inni = 1 • głęboki ubytek słuchu = 13, znaczny ubytek słuchu = 8, umiarkowany ubytek słuchu = 6 • głuche dzieci słyszących rodziców = 19, głuche dzieci głuchych rodziców = 8 • ciężki stan = 15, lżejszy stan = 12 • dzienna szkoła dla głuchych • duże miasta w stanach wschodniego wybrzeża USA
9	L. Brown, Sherbenou, Johnsen (1997)	20	• wiek = 8–16 lat • mężczyźni = 16, kobiety = 4 • Biali = 18, mniejszości = 2 • zaburzenia uczenia
10	Essix, Sapp, Abbott, Kohler (2006)	75	• zakres wiekowy = 7.9–16.5 lat ($M = 10.7$, $SD = 2.4$) • mężczyźni = 41, kobiety = 34 • Afroamerykanie = 75 • skierowani z podejrzeniem specyficznych trudności w uczeniu się
11	Ouellette (2006)	60	• zakres wiekowy = 9.5–10.5 ($M = 9.10$, $SD = 3.5$ miesięcy) • mężczyźni = 22, kobiety = 38 • klasa IV • rozwój typowy

Załącznik C

Charakterystyka demograficzna grup badanych

Nr badań	Źródło	N	Opis próby
1	Bond, Kennon (1982)	71	- zakres wiekowy = 7–18 lat - mężczyźni = 43, kobiety = 28 - Biali = 52, Afroamerykanie = 19 - Memphis, TN
2	L. Brown, Sherbenou, Johnsen (1982)	10	- zakres wiekowy = 8.1–13.7 lat ($M = 11.4$ lat) - mężczyźni = 7, kobiety = 3 - Biali = 4, Afroamerykanie = 6 - niepełnosprawność intelektualna
3	Zavala, Mims (1983)	20	- zakres wiekowy = 7.3–12.3 lat ($M = 10.6$ lat) - mężczyźni = 16, kobiety = 4 - zaburzenia uczenia = 10 - ograniczona znajomość języka angielskiego - zastosowano instrukcję w j. hiszpańskim i j. angielskim - klasy I–IV szkoły podstawowej - przedmieścia Południowej Kalifornii
4	Vance, Hankins, Brown (1986)	89	- zakres wiekowy = 6.4–16.10 lat - mężczyźni = 61, kobiety = 28 - Biali = 52, Afroamerykanie = 37 - skierowani na badania psychologiczne - klasy I–X - miejskie i wiejskie szkoły - Maryland i Delaware

Nr badań	Źródło	N	Opis próby
5	Parmar (1988)	91	- zakres wiekowy = 7–9 lat ($M = 7.9$ lat) - rozwój typowy = 73, niepełnosprawność intelektualna = 18 - prywatne szkoły - osoby posługujące się językiem indyjskim - mieszkańcy Indii
6	Salas (1988)	61	- średni wiek = 10.4 lat - mężczyźni = 48, kobiety = 13 - Biali = 53, Afroamerykanie = 8 - Dysleksja - publiczne szkoły w centralnym Teksasie
7	Brown, Sherbenou, Johnsen (1990)	103	- zakres wiekowy = 10.83 lat ($M = 30.8$ lat) - uraz głowy - Pocatello, Idaho
8	Brown i in. (1990)	114	- średni wiek = 14.8 lat - wybitnie uzdolnieni = 49 - Plano Independent School District (ISD), Teksas - średni wiek = 6.2 - uzdolnieni = 65 - Plano ISD i Eanes ISD, Teksas
9	Kowall, Watson, Madak (1990)	30	- zakres wiekowy = 6.4–11.6 lat ($M = 9.1$ lat) - mężczyźni = 20, kobiety = 10 - skierowani do poradni dla dzieci - przedmieścia - status socjoekonomiczny – klasa średnia - Kanadyjczycy
10	Kowall i in (1990)	22	- zakres wiekowy = 6.5–11.8 lat ($M = 8.5$ lat) - mężczyźni = 14, kobiety = 8 - rdzenni Kanadyjczycy - skierowani do poradni dla dzieci - miasto - status socjoekonomiczny – klasa niższa - Kanada
11	Whorton, Morgan (1990)	17	- średni wiek = 11.4 lat - skierowani z powodu trudności w uczeniu się - rdzenni Amerykanie
12	Lassiter, Bardos (1992)	36	- zakres wiekowy = 7–12 lat ($M = 9.8$ lat, $SD = 1.5$) - mężczyźni = 28, kobiety = 8 - zaburzenia uczenia - południowe stany USA

Nr badań	Źródło	N	Opis próby
13	Coleman, Scribner, Johnsen, Evans (1993)	37	- zakres wiekowy = 16–21 lat ($M = 17.9$ lat) - mężczyźni = 26, kobiety = 11 - zaburzenia uczenia - Amerykanie meksykańskiego pochodzenia = 37 - ograniczona znajomość języka angielskiego = 15 - środkowy Teksas
14	Prado, Gatica, Rojas (1993)	328	- zakres wiekowy = 6.0–15.11 lat - mieszkańcy Valdivia, Chile
15	Sam (1993)	100	- zakres wiekowy = 16–21 lat - mężczyźni = 49, kobiety = 51 - Biali = 24, Afroamerykanie = 32, Azjaci = 13, Hiszpanie = 25, inni = 6 - młodzież w niekorzystnej sytuacji korzystająca z programów wsparcia - Północna Kalifornia
16	Cash (1994)	205	- zakres wiekowy = 5.7–18.4 lat ($M = 10.2$ lat) - mężczyźni = 107, kobiety = 98 - zaburzenia słuchu = 205 - Szkoła dla Głuchych w Utah
17	D'Amato, Lidiak, Lassiter (1994)	74	- zakres wiekowy = 6–15 lat ($M = 10.2$ lat, $SD = 2.5$) - mężczyźni = 59, kobiety = 15 - skierowani z trudnościami szkolnymi - w większości niski i średni status socjoekonomiczny - południowe i środkowo-zachodnie stany USA
18	Auerbach (1995)	65	- zakres wiekowy = 22–75 lat ($M = 39.7$ lat) - mężczyźni = 29, kobiety = 36 - Biali = 58, Afroamerykanie = 6, Hiszpanie = 1 - mieszkańcy domów nadzorowanych = 62, mieszkający niezależnie = 3 - niepełnosprawność intelektualna lub ciężka niepełnosprawność fizyczna
19	Mackinson (1996)	27	- zakres wiekowy = 6.2–14.9 lat ($M = 11.3$ lat) - Biali = 10, Afroamerykanie = 15, Latynosi = 1, inni = 1 - głęboki ubytek słuchu = 13, znaczny ubytek słuchu = 8, umiarkowany ubytek słuchu = 6 - dzienna szkoła dla głuchych - duże metropolie w stanach wschodniego wybrzeża USA

Nr badań	Źródło	N	Opis próby
20	Brown, Sherbenou, Johnsen (1997)	3451	- zakres wiekowy = 5.0-89.0 lat - Afroamerykanie = 444 - ADHD = 131 - głusi = 76 - dysleksja = 51 - zaburzenia emocjonalne = 55 - język angielski jako drugi język = 90 - mężczyźni = 1713, kobiety = 1738 - uzdolnieni = 134 - Hiszpanie = 386 - zaburzenia uczenia = 287 - niepełnosprawność intelektualna = 27 - Biali = 29, mniejszości = 2 - zaburzenia uczenia - problemy z czytaniem = 31 - próba odzwierciedlająca populację wg. danych ze spisu powszechnego
21	Edelson, Edelson, Jung (1998)	39	- zakres wiekowy = 5-18 lat (M = 9.59) - mężczyźni = 37, kobiety = 2 - werbalni = 17, komunikatywni = 9, kilka słów = 2 - autyzm - głównie Biali
22	Edelson i in. (1998)	39	- zakres wiekowy = 6-38 lat (M = 12.77) - mężczyźni = 33, kobiety = 6 - autyzm i problemy behawioralne - werbalni = 23, komunikatywni = 8, kilka słów = 3 - głównie Biali
23	Edelson i in. (1998)	39	- zakres wiekowy = 3-15 lat (M = 9.39) - mężczyźni = 37, kobiety = 2 - werbalni = 7, komunikatywni = 23, kilka słów = 9 - autyzm - Tajwańczycy = 39
24	Edelson, Schubert, Edelson (1998)	258	- zakres wiekowy = 4-41 lat (M = 12.06, SD = 7.08) - mężczyźni = 193, kobiety = 65 - werbalni = 222, niewerbalni = 36 - autyzm - głównie Biali

Nr badań	Źródło	N	Opis próby
25	Alvarado (1999)	90	• klasy 2 i 3 • mężczyźni = 45, kobiety = 45 • dwujęzyczni, dominujący j. angielski = 26, dwujęzyczni, zrównoważony = 34, dwujęzyczni, dominujący j. hiszpański = 30 • edukacja dwujęzyczna = 48, edukacja regularna = 37, uzdolnieni i utalentowani = 4, edukacja specjalna = 4 • Południowy Teksas
26	Alveriadou, Katsarou, Bostantjopoulou, Kiosseoglou, Mentenopoulos (1999)	37	• średni wiek = 62.1 lat (SD = 6.8) • mężczyźni = 23, kobiety = 14 • Choroba Parkinsona bez demencji • Grecy
27	Barrett (2000)	217	• zakres wiekowy = 6–17 lat • mężczyźni = 101, kobiety = 116 • szkoły niepubliczne = 101, szkoły publiczne = 116 • Jamajczycy
28	Kern, Cauller, Dodd (2000)	14	• zakres wiekowy = 5–10 lat • autyzm i zaburzenia rozwojowe • Północny Teksas
29	Valliant, Gauthier, Pottier, Kosmyrna (2000)	54	• wiek = 18 lat albo starsi • mężczyźni = 54 • przestępstwa powszechne = 20, gwałt = 14, kazirodztwo = 9, molestowanie dziecka = 11 • rekrutowani z więzienia okręgowego Sudbury, Ontario, Kanada
30	Aridi (2002)	100	• zakres wiekowy = 18–65 lat (M = 36.48, SD = 11.25) • mężczyźni = 28, kobiety = 72 • Łatynosi • Washington, DC, obszar metropolitalny
31	Tharpe, Ashmead, Rothpletz (2002)	28	• zakres wiekowy = 8–14 lat • mężczyźni = 13, kobiety = 15 • implanty ślimakowe = 9, aparaty słuchowe = 9, bez aparatuowania = 10 • rekrutowani z kliniki słuchu

Nr badań	Źródło	N	Opis próby
32	Lagace, Kutcher, Robertson (2003)	44	• średni wiek = 19.4 (SD = 2.9) • mężczyźni = 17, kobiety = 27 • średnia oceny = 11.2 (SD = 1.1) • zaburzenie dwubiegunowe typu I
33	Lagace, i in. (2003)	30	• średni wiek = 18.5 (SD = 2.8) • mężczyźni = 9, kobiety = 21 • średnia oceny = 10.3 (SD = 1.6) • ciężka depresja
34	Valliant, deWit, Bowes (2004)	103	• średni wiek = 33 lata dla przestępców • mężczyźni = 103 • napaść domowa = 46, napaść ogólna = 23, przestępstwo bez przemocy = 19, uczelnia = 15 • rekrutowani z więzienia okręgowego i lokalnego uniwersytetu • Sudbury, Ontario, Kanada
35	Chan, Cheung, Leung, Cheung, Cheung (2005)	27	• zakres wiekowy = 5–6 lat (M = 5.74, SD = 0.3) • mężczyźni = 15, kobiety = 12 • populacja ogólna • Chińczycy posługujący się językiem ojczystym • Hong Kong
36	Chan, i in. (2005)	19	• zakres wiekowy = 5–6 lat (M = 5.77, SD = 0.59) • mężczyźni = 16, kobiety = 3 • autyzm • Chińczycy posługujący się językiem ojczystym • Hong Kong
37	Edelson (2005)	35	• zakres wiekowy = 4–18 lat (M = 10.12, SD = 3.8) • mężczyźni = 23, kobiety = 12 • Amerykanie = 29, Włosi = 6 • autyzm
38	Lyall (2005)	118	• zakres wiekowy = 6.0–16.11 lat • Afroamerykanie = 40, Latynosi = 78 • Osoby posługujące się j. angielskim • miejskie szkoły o niskich dochodach w New Jersey
39	Michelsen (2005)	18	• zakres wiekowy = 14.18 lat • uczniowie języka angielskiego jako drugiego języka • mężczyźni z Korei Południowej

Nr badań	Źródło	N	Opis próby
40	Yip, Lee (2005)	200	- zakres wiekowy = 18.45 lat ($M = 28$, $SD = 5.7$) - rekrutowani z lokalnych barów i klubów tanecznych - użytkownicy ekstazy = 100 - Hong Kong
41	Johnson (2007)	90	- zakres wiekowy = 6.0–16.1 lat ($M = 11.09$, $SD = 2.68$) - mężczyźni = 46, kobiety = 44 - Rdzeni Amerykanie (Sioux) - Standardowa edukacja = 30 (mężczyźni = 15, kobiety = 15) - edukacja specjalna = 30 (mężczyźni = 21, kobiety = 9) - uzdolnieni = 30 (mężczyźni = 10, kobiety = 20) - południowa Południowa Dakota
42	Essix, Sapp, Abbott, Kohler (2006)	75	- zakres wiekowy = 7.9–16.5 lat ($M = 10.7$) - mężczyźni = 41, kobiety = 34 - Afroamerykanie = 75 - skierowani z powodu trudności w uczeniu się
43	Richardson, Harris, Gerken (2006)	32	- zakres wiekowy = 18.23 lat - trudności w uczeniu się = 16, bez trudności w uczeniu się = 19 - zajęcia na studiach licencjackich z psychologii - Uniwersytet Arizony
44	Greenfield, Valliant (2007)	59	- wiek = 18 lat albo starsi - mężczyźni = 59 - przemoc = 20, gwałt = 14, bez przemocy = 19, uniwersytet = 20 - rekrutowani z więzienia okręgowego = 39 - rekrutowani z uniwersytetu = 20 - Sudbury, Ontario, Kanada
45	Lewis, Murdoch, Woodyatt (2007)	20	- zakres wiekowy = 9–17 lat ($M = 11$, $SD = 2$) - mężczyźni = 16, kobiety = 4 - zaburzenia ze spektrum autyzmu - Queensland, Australia
46	Perera i in. (2007)	98	- średni wiek = 30.2 ($SD = 5.2$) - kobiety = 98 - Biali = 40, Azjaci = 34, Afroamerykanie = 13, inni = 11 - Zamieszkali w odległości 2 mil od World Trade Center w miesiąc po 11 września

—