

RECENSIE 'HANDBOEK INTELLIGENTIETHEORIE EN TESTGEBRUIK' (2^E EDITIE, 2021)

Handboek intelligentietheorie en testgebruik, 2^e editie (2021)

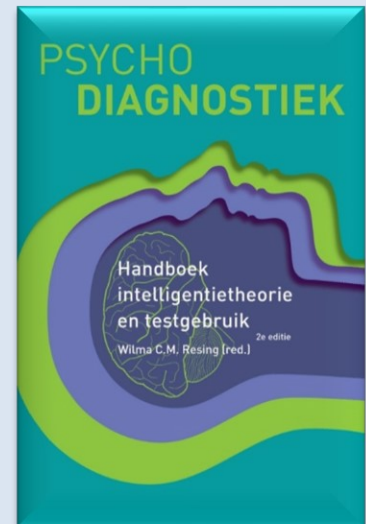
Auteur: Wilma C.M. Resing (red.)

Prijs paperback: €46,50 (ISBN: 9789043037792)

Prijs e-book: €35,50 (ISBN: 9789043039581)

Uitgever: [Pearson](#) Benelux bv, Amsterdam

In 2015 verscheen de 1^e druk van het 'Handboek intelligentietheorie en testgebruik', onder redactie van prof. dr. Wilma Resing, o.a. bekend als auteur van de RAKIT(-2) en voormalig COTAN-lid. In maart 2021 verscheen de 2^e (herziene) editie van dit boek, waarin verouderde tests zijn verwijderd, nieuwe zijn toegevoegd en nog veel meer is veranderd. Ik vind het een prettig en *up to date* boek, wat mij betreft hoort het thuis in de collectie van iedere instelling/praktijk, die regelmatig intelligentieonderzoek verricht. Klik op de afbeelding of [hier](#) voor een link naar de site van uitgever Pearson, waar u de details vindt. Het boek is overal te bestellen, het e-book is via [Studystore](#) verkrijgbaar.



Ik zal in deze recensie vooral aandacht besteden aan het WISC-V hoofdstuk (hoofdstuk 11, geschreven door Marc Hendriks, Selma Ruiters, Mark Schittekatte en Petra Hurks), nu dit voor de meeste bezoekers van mijn website en lezers van mijn Nieuwsbrief over intelligentiediagnostiek bij K&J het meest relevant is. De auteurs van dit hoofdstuk zijn allen bekende namen op het gebied van intelligentiediagnostiek. Los van de kanttekeningen die hierna volgen en soms ook over het hele boek gaan, vind ik het een prettig geschreven en helder hoofdstuk, met veel verstandige en genuanceerde opmerkingen. Ik zal een aantal punten bespreken, die ik van belang acht voor collega's die psychodiagnostisch onderzoek doen, WISC-V gebruikers in het bijzonder. Verder zal ik soms wat 'uitstapjes' maken, die niet zozeer over dit boek gaan, maar waarbij een punt uit dit boek een aanleiding was om de lezers van mijn Nieuwsbrief wat meer achtergrondinformatie te willen geven of te stimuleren tot reflectie.

Blijvende waarde van Totaal IQ bij een disharmonisch profiel

Zo ben ik bijvoorbeeld erg gecharmeerd van de aandacht voor de waarde van het TIQ bij disharmonische profielen. Nog steeds zijn er veel vakgenoten die denken dat je bij een disharmonisch profiel (nog even los van hoe je dat tegenwoordig definieert) niks meer met het TIQ *mag* of *kan*; soms denkt men zelfs dat het dan ook onvermeld zou moeten blijven (incidenteel tref ik zelfs rapportages aan, waarbij een index-IQ is weggelaten als het paarsgewijze verschil tussen de subtests van die index significant en/of als betekenisvol wordt beschouwd: nooit doen!). In eerdere Nieuwsbrieven heb ik hier al regelmatig aandacht aan besteed en ook toonaangevende internationale publicaties geciteerd, met als boodschap dat het TIQ, ongeacht de opbouw, de meest betrouwbare en valide maat blijft voor cognitief functioneren en je dit 'kind' nooit met het 'badwater' moet weggooien. Dat is overigens niet specifiek voor de WISC-V, ook in publicaties van 10-20 jaar geleden kwam dit al vrij consequent naar voren m.b.t. onder meer de WISC-III. In 2007 schreef ik bijv. een [WISC-III artikel](#) waarin al deze tekst te lezen was:

Blz. 2: 'Voor psychodiagnosten die er om soortgelijke redenen voor kiezen om het TIQ weg te laten bij een disharmonisch profiel op schaalniveau, is het goed om te weten dat internationale WISC-experts (Weiss e.a., 2006) met klem adviseren altijd het TIQ te vermelden omdat dit zelfs bij een disharmonische opbouw de beste samenvatting blijft van algeheel cognitief functioneren en nog steeds de best

correlerende maat is met schoolprestaties ('among the highest known relationships between any two variables in the field of psychology', Weiss e.a., 2006, p.93). Ook de steekproef bevat veel kinderen met disharmonische profielen.'

Het handboek waarover deze recensie gaat, zegt hierover in het WISC-V hoofdstuk o.a. het volgende:

Blz. 274: 'Overigens is het zinvol een kanttekening te plaatsen bij het verlenen van betekenis aan onderlinge verschillen tussen klinische populaties. Het is gebleken dat een discrepantie van 23 punten of meer tussen 2 indexscores¹ in de WISC-V relatief vaak voorkomt: namelijk bij 52.5% (Amerikaanse klinische groepen) tot 56.6% (Amerikaanse normgroep) van de kinderen (Wahlstrom et al., 2018). Wanneer gesteld zou worden dat het Totaal IQ niet valide is door een disharmonische opbouw van het intelligentieprofiel, dan zou dit dus het geval zijn bij meer dan de helft van de kinderen. Onderzoek laat echter zien dat de constructvaliditeit en predictieve validiteit van het Totaal IQ niet nadelig worden beïnvloed door discrepanties tussen indexscores (Daniel, 2007). Dit betekent dat, zelfs wanneer een intelligentieprofiel disharmonisch is, het Totaal IQ voorspellende waarde heeft. Daarom wordt in de technische handleiding geadviseerd om altijd het Totaal IQ te betrekken bij de interpretatie van het intelligentieprofiel. Uiteraard kan men nooit volstaan met slechts een beschrijving van het Totaal IQ en zijn er aanvullende gegevens nodig. Er zijn geen Nederlandse gegevens gepubliceerd over het percentage disharmonische profielen in de normeringssteekproef.'

En op blz. 280: 'In de handleiding van de WISC-V wordt nadrukkelijk vermeld dat ook indien een scoreprofiel sterk disharmonisch is, het Totaal IQ valide en betrouwbaar is. (...) Een disharmonisch profiel kan reden zijn om het totale intelligentieniveau als een minder goede samenvatting te zien van het algemeen intelligentieniveau, maar het blijft een valide en betrouwbare score, die in dit geval aangeeft wat de sterke en zwakke schakels zijn in het functioneren van een kind.'

In het WISC-V hoofdstuk wordt vermeld dat het TIQ uit 7 subtests bestaat en de primaire indexscores uit 10 (=5x2) subtests bestaan. Wat ik mis, is dat dit als aanvullend argument wordt meegenomen om het TIQ serieus te blijven nemen bij disharmonische profielen. Immers, een disharmonisch profiel kan dus (mede) veroorzaakt worden door subtests die op geen enkele manier invloed hebben op het TIQ. Het zou dan raar zijn als je die disharmonie wel van invloed laat zijn op het bepalen van de waarde van het TIQ.

Indexen WISC-V dekken een cognitief domein onvolledig

Ook vind ik op blz. 265 mooi omschreven dat de indexinhoud weliswaar zorgvuldig gekozen is en onderbouwd wordt door validiteits- en betrouwbaarheidsonderzoek, maar dat het niet de pretentie is om op basis van die 2 subtests een heel achterliggend domein aan vaardigheden te representeren/'dekken'. De prestatie op een index biedt dus onvoldoende onderbouwing voor stellige diagnostische conclusies:

'Een lage score op de index Visueel Ruimtelijk betekent niet dat men de conclusie kan trekken dat het cognitief domein visueel-ruimtelijke waarneming van het kind geheel beperkt is. Visueel-ruimtelijke waarneming omvat meer dan de subtest scores op Blokpatronen en Figuur Samenstellen. Het is overigens ook niet de bedoeling om op basis van het scoreprofiel (zogenoemde profielanalyse) van de WISC-V-NL een cognitief domein in zijn volledigheid te onderzoeken. Het primaire doel is om informatie te geven over het functioneren van een kind op verschillende brede cognitieve vaardigheidsdomei-

¹ Verderop in deze recensie (onderaan bij punt 7) wordt toegelicht dat het hier om *primaire* indexscores gaat, een relevante toevoeging die hier ontbreekt in de verder vrijwel letterlijk vertaalde passage uit de Engelstalige bron. Verder spreekt die bron van over een 'discrepancy of greater than 23 points between two primary index scores' en dat is niet hetzelfde als een 'discrepantie van 23 punten of meer tussen 2 indexscores'. Geen idee wie gelijk heeft, maar >23 is niet hetzelfde als ≥23.

nen en om vervolgens een uitspraak te kunnen doen over het algemene intelligentieniveau en de opbouw daarvan.’

Waak er dus voor om op basis van een index Verbaal Begrip stellige uitspraken te doen over de talige vaardigheden of op basis van de index Fluid Redeneren te concluderen dat iemand goed is in het oplossen van problemen (dat vind ik nóg zorgelijker dan op basis van de VBI iets zeggen over verbale capaciteiten). Dat gaat echt veel te ver op basis van 2 subtests die zich vooral, zoals ik het zie, richten op een enkel aspect van probleemoplossende vaardigheden, namelijk non-verbaal logisch redeneren (en ook dat is weer te kort door de bocht 😊).

Classificaties IQ-scores

Wat valt verder op in dit boek? Jarenlang was in Nederland de IQ-classificatie van [Resing & Blok \(2002\)](#) leidend, er was niet of nauwelijks [discussie](#), er waren ook geen algemeen geaccepteerde alternatieven. In de eerste druk van dit handboek in 2015 werd die indeling (minimaal) aangepast; de term ‘*verstandelijk* beperkt’ werd vervangen door ‘*intellectuele* beperking’, maar in de toelichting erop werd weer geschreven dat voor de ene term ook de andere gebruikt/gelezen kan worden. Verder waren ‘aan de onderkant’ de termen licht/matig/ernstig *zwakzinnig* vervangen door lichte/matige/ernstige/diepe of zeer ernstige *intellectuele beperking*.

In de 2021-editie van Resing is de beschrijvende (‘normatieve’) indeling vervangen door een statistische, die lijkt op de indeling van Ruiter e.a. (2017). Ruiter e.a. pasten in 2018 overigens de classificatietabel in hun 2017-artikel aan: de termen ‘*hoog/laag in het gemiddelde gebied*’ werden vervangen door ‘*hooggemiddeld/laaggemiddeld*’. Als ik hier spreek over de 2017-indeling of ‘Ruiter e.a. (2017)’, dan bedoel ik daarmee de meest recente versie daarvan (dus incl. de 2018-rectificatie). Er zijn 2 verschillen tussen de indelingen van Resing (2021) en Ruiter e.a. (2017): Resing (2021) noemt IQ’s ≥ 130 ‘*uitzonderlijk hoog*’, waar Ruiter e.a. (2017) spreken van ‘*zeer hoog*’ (onderaan de casus op blz. 283 in het WISC-V hoofdstuk valt op dat Ruiter daar toch de Resing-term ‘*uitzonderlijk*’ hanteert, een benaming die afwijkt van de eigen indeling en wat op de WISC-V uitdraai van Q-global verschijnt, dit lijkt een foutje).

Aan de onderkant van de beide classificatie-indelingen ($IQ < 70$) zitten de andere verschillen. Om te beginnen zit hier weer eenzelfde verschil tussen ‘*uitzonderlijk laag*’ en ‘*zeer laag*’. En verder houdt Resing (2021) voor een nadere differentiatie bij IQ’s < 70 dan dezelfde indeling aan als in de 2015-editie van haar boek, waarin de eerder genoemde gradaties van een ‘*intellectuele beperking*’ worden onderscheiden.

In de [IQ Scorehulp 1.19](#) vindt u onder werkblad ‘Naamgeving scores’ de genoemde indelingen volledig weergegeven (zie hieronder).

IQ-scores (Resing, 2015)	
>130	zeer begaafd
121-130	begaafd
111-120	bovengemiddeld
90-110	gemiddeld
80-89	benedengemiddeld
70-79	laag begaafd / moeilijk lerend
50-69	lichte intellectuele beperking
35-49	matige intellectuele beperking
20-34	ernstige intellectuele beperking
<20	diepe/zeer ernstige intellectuele beperking

IQ-scores (Ruiter, Hurks & Timmerman, 2017) / Resing (2021)	
≥ 130	zeer hoog / uitzonderlijk hoog
120-129	hoog
110-119	hooggemiddeld
90-109	gemiddeld
80-89	laaggemiddeld
70-79	laag
<70	zeer laag / uitzonderlijk laag <i>Resing (2021) houdt voor nadere differentiatie <70 de 2015-indeling aan (hiernaast)</i>

Mijn voorkeur heeft nog steeds de beschrijvende/normatieve classificatie (Resing (2015)). Ik vind 'moeilijk lerend' informatiever dan het neutralere en ook algemenere 'laag'. Bovendien vind ik de grenswaarden van Resing (2015) logischer; 90-110 is gemiddeld, dat is het puntgemiddelde 100 ± 10 . In de nieuwste indelingen is dat 90-109 geworden (waarom naar beneden 100-10 en naar boven dan 100+9, het puntgemiddelde moet toch precies in het midden liggen?). Deze 'asymmetrie' werkt ook eenzijdig door in alle categorieën erboven (111-120 is 110-119 geworden, 121-130 is 120-129 geworden, enz.).

Ik ben het eens met de trend en adviezen van de laatste jaren om *testscores* te beschrijven in plaats van *personen*. Dat kan wat mij betreft ook prima door te zeggen dat iemand op een bepaald niveau heeft *gepresteerd* of evt. *functioneert* (maar dat is een aanname op basis van een testprestatie), zonder dat je nu meteen zegt dat iemand ook een intellectuele beperking *heeft* of zelfs verstandelijk beperkt, moeilijk lerend of (zeer) begaafd *is* (voor het vaststellen van een verstandelijke beperking volstaat een intelligentietest ook niet en is ook informatie over het adaptieve functioneren nodig). Soms is het overigens confronterend maar ook wel 'hoe het is' en kan zo'n term veel meer dan de term '(zeer) laag' laten doordringen dat er echt iets aan de hand is en iemand een beperking heeft...

Mijn advies is om in ieder geval één van deze indelingen aan te houden en niet weer een eigen indeling te gebruiken of er een eigen variatie op aan te brengen. De indelingen van Resing (2002/2015/2021) en die van Ruiter e.a. (2017) zijn op dit moment in Nederland de enige indelingen die goed onderbouwd zijn en waarover consensus is.

En als we het *echt* niet meer weten en *out of the box* willen denken, dan verwijst ik graag naar de handleiding van mijn nieuwe combimagnetron. 😊

Vermogensniveau	Percentage
HOOG	100 %
MIDDELHOOG	67 %
GEMIDDELD	50 %
MIDDELLAAG	33 %
ONTDOOIEN	20 %
LAAG	11 %

Kwetsbaarheid materiaal

Terug naar het handboek: ik vind het een prima boek en kan het aanraden voor geïnteresseerden in dit onderwerp. Helaas heeft het geen harde kaft, wat een gemis is voor een boek dat een aantal jaren mee moet gaan om tussendoor vaak te kunnen raadplegen. De huidige kaft trekt zelfs bij voorzichtig gebruik al direct krom. Wat mij betreft zou Pearson meer aandacht mogen besteden aan hun materiaal, zodat het geschikt is voor langdurig gebruik. Het is natuurlijk een beetje muggenzifterij. Ik maak er een punt van, omdat ditzelfde bezwaar ook aan de orde is t.a.v. de stimulusboeken en handleidingen van de Wechsler tests, die eigenlijk niet gemaakt zijn voor langdurig intensief gebruik. Laat staan voor gebruik bij kinderen. Hierin kan men echt een voorbeeld nemen aan het materiaal van de KIQT+, waarbij je je geen zorgen hoeft te maken over vouwen, scheuren of onherstelbare schade door vocht of viezigheid. We kennen allemaal de Wechsler-stimulusboekjes die uit hun ringband lopen of de kleine multomap-handleidingen waarbij de geperforeerde pagina's bij de gaatjes zijn doorgescheurd. En probeer maar eens een grijs van herkenning te onderdrukken als we het hebben over staat waarin uw WISC-III ringbandhandleiding de laatste jaren verkeerde... 😊

Een 'concurrent' (het bekende rode *Handboek psychodiagnostiek voor de hulpverlening aan kinderen en adolescenten* van Tak, Bosch, Begeer & Albrecht (2014), ook wel vanuit eerdere drukken bekend als het handboek van 'Kievit') heeft een stevige harde kaft en ziet er bij ons na een aantal jaar gebruik nog prima uit. Van dit handboek is overigens een geheel herziene nieuwe editie in de maak, die

zal vermoedelijk in 2023 verschijnen en naast een uitgebreid hoofdstuk over intelligentiediagnostiek breder van opzet zijn.

Voldoende kritisch op de besproken instrumenten?

In de hoofdstukken van dit handboek onder redactie van Resing worden vaak per besproken test de sterke en zwakke kanten weergegeven in een zogenaamde 'box'. Wat daarbij opvalt, is dat de kolom met zwakke kanten altijd korter is, het hoofdstuk over de WNV (door Resing & Bleichrodt) uitgezonderd. T.a.v. de WNV ontbreekt wat mij betreft overigens nog de 'zwakte' dat dit instrument slechts uit 4 subtests bestaat (als je de non-verbale taken van de WISC-V afneemt, heb je al 8 informatieve subtests). Ook van andere instrumenten mis ik in die boxen nog wel meer zwakke kanten of beperkingen van de besproken tests.

De meest bekende zwakte, die in het veld nog steeds als een groot gemis wordt beschouwd, is het ontbreken van een adequate onderwijstabel m.b.t. bijv. de WISC-V en IDS-2, waarin per niveau van het voortgezet onderwijs voor alle IQ-maten de gemiddelde prestaties (met sd's) worden vermeld van de kinderen in de steekproef. In eerdere Nieuwsbrieven ([2019-1](#) en [2018-1](#)) is hieraan al uitgebreid aandacht besteed. Het ontbreken van een voldoende gedifferentieerde onderwijstabel maakt dat we de WISC-V minder goed kunnen gebruiken bij schoolvraagstukken en het adviseren daarover. Natuurlijk weten we allemaal dat er veel meer factoren met schoolsucces (op welk niveau dan ook) samenhangen, maar we doen juist vaak een IQ-test bij een vastgelopen leerling, omdat we willen weten of iemand het in cognitieve zin in principe aan zou moeten kunnen. Als dat al niet toereikend is (en we achten de meting betrouwbaar), dan hoeven we het over de andere factoren bij wijze van spreken niet eens meer te hebben. Goed om te weten dat er wat betreft het kritiekpunt van de onderwijstabel visieverschillen bestaan. Het is een bewuste keuze van de testuitgever geweest deze tabel niet nader te differentiëren (of bijv. vragen te beantwoorden over de opbouw van de gecombineerde havo/vwo-tabel), omdat men daarin risico's van onoordeelkundig gebruik ziet en denkt dat mensen er dan teveel aan gaan ophangen of er onterechte adviezen/conclusies op gaan baseren. De beperkt gedifferentieerde WISC-V onderwijstabel heeft wat mij betreft ook al gevolgen voor de beschreven casus, waarmee het WISC-V hoofdstuk afsluit. De conclusie is (vrij vertaald) dat deze 15-jarige jongen met zijn inzet in cognitieve zin wel havo zou moeten aankunnen, terwijl we eigenlijk onvoldoende zicht hebben op welke WISC-V testprestatie eigenlijk gemiddeld is voor havoleerlingen. We weten al jaren dat het ene (T)IQ het andere niet is en dat een gemiddelde van een bepaald onderwijsniveau op de ene test veel hoger kan liggen dan op de andere, dus we hebben echt de WISC-V-NL statistieken nodig om op basis van deze test hierover dan ook iets zinnigs te kunnen zeggen.

Bij deze casus had ik ook graag wat meer testresultaten gezien dan enkel het TIQ (waarvan alleen het 90%-betrouwbaarheidsinterval vermeld wordt), hoewel ik begrijp dat vooral het TIQ gebruikt wordt bij het meedenken over het schoolniveau, nu deze maat de meest betrouwbare/valide voorspeller vormt. In een WISC-V hoofdstuk in een handboek had ik graag ook subtestscores en andere (primaire) IQ-scores gezien en een casus die zich meer leent om nog iets meer de 'diepte' in te gaan. Maar ook daarover is discussie mogelijk; kun je het de WISC-V nog wel de diepte in of moet je het meer zien als een instrument dat screent in de breedte?

Er worden dus wel beperkingen van de tests besproken in dit handboek, maar ècht kritische geluiden over de beschreven test mis ik een beetje (ik denk dan bijv. ook aan het feit dat de SON-R en RAKIT-2 qua visuele items nog steeds deels lijken op hun gedateerde jaren '80 voorgangers en niet bepaald '21st century' zijn).

Belangen?

Testauteurs schrijven hoofdstukken over hun eigen tests

Ik heb van meerdere collega's vragen ontvangen over de objectiviteit en onafhankelijkheid van dit handboek, nu bij veel besproken tests de auteurs van diezelfde tests (of anderszins gelieerden, zoals

een 'product developer' van de IQ-testuitgever Pearson) ook auteur zijn van de hoofdstukken over hun eigen instrumenten. Dat maakt dat men minder snel verwacht dat er heel kritisch over een test en de toepasbaarheid daarvan geschreven wordt. Dat is helemaal niet gek, ik ben ook enthousiast over mijn eigen IQ Scorehulp en kan het gebruik ervan aanbevelen. In het WISC-V hoofdstuk valt bijv. op dat er een aantal categorieën van de COTAN-beoordeling zijn overgenomen, maar dat de 'onvoldoende' op criteriumvaliditeit onbenoemd blijft, terwijl de begripsvaliditeit wel genoemd wordt. Dat komt over als een selectieve weergave van de beoordeling van de psychometrische eigenschappen.

Ik snap de vragen van collega's en heb hierboven ook enkele opvallende en ontbrekende kritiekpunten genoemd. En tegelijkertijd vind ik het ook zo logisch dat mensen met expertise op dit vlak vaak ook de mensen zijn die op dat vlak dus iets gedaan hebben, zoals het publiceren over hun instrument. Wil je mensen die er verstand van hebben en erover willen schrijven, dan ontkom je er niet aan dat daar ook mensen tussen zitten die deze deskundigheid ook gebruikt hebben bij testontwikkeling. Of omgekeerd, die juist door betrokkenheid bij testontwikkeling hun deskundigheid verkregen hebben. De auteurs geven wat mij betreft geen beeld over wat je allemaal met hun tests zou kunnen, waarvan ik denk 'Dat is echt onverantwoord'. Ik vind het WISC-V hoofdstuk op dat vlak heel realistisch, er wordt niet meer uit de WISC-V gehaald dan er mijns inziens inzit.

COTAN beoordeelt instrumenten van eigen commissieleden

Vanuit dezelfde kritische houding zouden overigens net zo goed vragen gesteld kunnen worden over het feit dat in de COTAN ook testauteurs zitten, die hun instrumenten door deze zelfde commissie laten beoordelen. Uiteraard zijn daar procedures voor en zijn testauteurs niet betrokken bij de beoordeling van hun eigen instrumenten. Het moet wat mij betreft dus bewust gehanteerd worden, maar je ontkomt er niet aan dat mensen verschillende petten kunnen hebben.

Bij het hoofdstuk 'Over de auteurs' wordt ook netjes en transparant vermeld welke banden auteurs van het handboek met bepaalde tests en commissies hebben (gehad). Ik vind de hoofdstukken ook niet geschreven als 'reclamebrochures', het is een genuanceerd geheel, waarbij ook beperkingen van de instrumenten genoemd worden (ook al mag het van mij dus soms nog wat kritischer).

Handboek over IQ-tests vanuit IQ-testuitgever Pearson

Soortgelijke kritische overwegingen hoor ik terug m.b.t. keuze van de uitgever van dit boek (testuitgever Pearson). Ik kan me goed voorstellen dat je wat eerder voor een bepaalde uitgever kiest als je daar al mee samenwerkt, omdat zij al de uitgever van je test zijn. En omgekeerd: een testuitgever van IQ-tests zal eerder interesse hebben over dit onderwerp ook een boek uit te willen geven en vraagt dan vanzelfsprekend de eigen deskundige testauteurs hiervoor. Het voordeel van een dergelijk handboek publiceren bij Pearson als uitgever van IQ-tests is bovendien dat deze uitgever de doelgroep van het boek al kent en bedient met tests.

Evenwel hoor ik collega's zich afvragen of het nou handig is om een dergelijk onafhankelijk handboek te laten uitgeven door de testuitgever van de WPPSI, WISC, WAIS, WNV, Raven's Progressive Matrices en RAKIT-2. Ik heb bij het lezen van dit handboek echter geen bijmaak ervaren, waarbij ik het idee kreeg van een subjectieve beschrijving of kritische(re) uitlatingen over tests van andere uitgevers t.o.v. de eigen testcollectie. Het is wat mij betreft dus een terechte vraag maar ook een ongegrond bezwaar t.a.v. dit boek.

Nu er in Nederland een aantal uitgevers van psychologische vakliteratuur zijn en slechts enkele actief op het vlak van intelligentietests, had men er evenwel (om iedere schijn van belangenverstremming te voorkomen) veiligheidshalve voor kunnen kiezen om een 'neutrale' uitgever te nemen, of i.i.g. niet de uitgever die nu meer dan de helft van alle instrumenten uitgeeft, waaraan dit boek hoofdstukken wijdt, en in Nederland geldt als *leading* op het vlak van uitgave van intelligentietests.

COTAN Documentatie ondergebracht bij een testuitgever (Boom)

Voortbordurend op dit thema, zijn net zo goed kritische vragen te stellen over het feit dat het abonnement op de COTAN-database ([COTAN Documentatie](#)) is ondergebracht bij een andere testuitgever:

Boom. De meeste uitgevers zijn niet betrokken bij psychodiagnostisch testmateriaal, dus waarom heeft het NIP dit nu juist ondergebracht bij een *testuitgever*, die ook te beoordelen instrumenten uitgeeft en een direct commercieel belang heeft bij die beoordeling? Want laat er geen misverstand over bestaan; het COTAN-oordeel kan duidelijke consequenties hebben voor de testverkoop. Uitgevers vrezen dit oordeel of vermijden soms het laten beoordelen van hun tests uit angst voor de effecten op de verkoop bij een negatief oordeel. Er zullen ongetwijfeld afspraken gemaakt zijn om de onafhankelijkheid te waarborgen en ik ken concrete voorbeelden van negatief beoordeelde Boom-tests, waarbij ik geen enkele indruk heb van ook maar een schijn van afhankelijkheid of invloed van Boom op de beoordeling (evenmin heb ik aanwijzingen dat tests of auteurs van Boom ook maar enigszins een voorkeursbehandeling zouden krijgen), maar ik zou zelf als beroepsvereniging (NIP) eerder gekozen hebben voor een uitgever zonder belangen, zodat je niet de 'keuringsrapporten' van vlees onderbrengt bij een slager.

Errata en suggesties ter verbetering

Voor zover mij bekend, zijn er geen errata verschenen bij dit boek. Pearson is hierin altijd erg terughoudend, ik zie het ook zelden rondom tests verschijnen als er fouten in handleidingen of normtabellen gemeld worden. Als ik zelf punten doorgeef, dan hoor ik doorgaans terug dat het in een volgende druk wordt aangepast, maar dat gebeurt dan 'in stilte op de achtergrond' en zonder communicatie naar de gebruikers. Ik ken enkele uitzonderingen, bijv. eerdere WISC-V errata, deze kwamen mijns inziens echter na druk vanuit het werkveld tot stand.

Ik heb in augustus 2021 een van de auteurs van onder meer het WISC-V hoofdstuk geïnformeerd over de meeste punten uit deze recensie (bij het schrijven ontdekte ik nog enkele nieuwe punten), naar ik begrijp is het doorgegeven aan de uitgever.

Omdat een handboek als dit niet alleen als naslagwerk wordt gebruikt maar ook wordt gebruikt voor collega's die in opleiding zijn, vind ik het van belang hieronder nog enkele punten te noemen, zodat mensen die dit handboek bezitten het desgewenst in hun versie kunnen aanpassen of zich realiseren welke nuances ik bij bepaalde passages plaats.

1. Blz. 259: Het WISC-V hoofdstuk heeft op de eerste pagina een kopje met 'Testkenmerken', waar de betrouwbaarheid, begripsvaliditeit en normering worden benoemd met kwalificaties van voldoende tot goed. Hoewel het er niet letterlijk bijstaat, verwijst dit naar de COTAN-beoordeling. Zoals eerder genoemd, valt het op dat de enige 'onvoldoende' binnen deze beoordeling (de criteriumvaliditeit) onvermeld en ook elders in het hoofdstuk onbesproken blijft, terwijl de voldoende begripsvaliditeit wel genoemd wordt.
2. Blz. 259: Bij de testkenmerken van de WISC-V staat 'continue normering'. Goed dat de lezer zich realiseert dat dit niet betekent dat de WISC-V software (Q-global of Q-interactive) met *dagnormen* werkt, zoals dat bijv. wel bij de RAKIT-2 en SON-R aan de orde is. Helemaal in het begin van de WISC-V-NL hoorde ik van collega's terug dat er kleine verschillen waren tussen digitale WISC-V scoring en scoring met de handleidingstabellen, die met leeftijdscohorten à 4 maanden werken (bij dagnormen zou je anders per leeftijdsjaar 365 tabellen moeten publiceren). Op een gegeven moment verdwenen die scoreverschillen, omdat ook de digitale scoring overstapte op de cohorttabellen per leeftijdsgroep. Bij navraag bevestigde de testuitgever dit, maar er is bij mijn weten nooit proactief of op de online productpagina van de WISC-V breed over gecommuniceerd naar gebruikers. Naar ik begrijp doelt deze vermelding in het handboek op gebruik van de continue normeringsmethode, die bij de WISC-V is gebruikt en moet dit niet opgevat worden als digitale scoring a.d.h.v. dagnormen.
3. Blz. 262: Hier staat ten onrechte beschreven dat subtest Plaatjesreeksen over *ruimtelijk* werkgeheugen gaat, terwijl het om *visueel* werkgeheugen gaat (zonder ruimtelijke component).

4. Blz. 263: In het hele hoofdstuk worden de termen ‘aanvullende’ en ‘secundaire’ subtests of indexen door elkaar gebruikt. Helemaal prima en doe ik waarschijnlijk ook in mijn Nieuwsbrieven. Het is vooral goed dat de gebruiker zich realiseert dat het om precies hetzelfde gaat en beide termen naar dezelfde subtests/indexen verwijzen (d.w.z. alles wat niet onder de primaire subtests/indexen valt).
5. Blz. 265-268: Ik vind het onwenselijk dat testitems zijn gepubliceerd, ter illustratie bij de subtestbeschrijvingen. Dit zijn geen demonstratie-items, maar 4 vragen van verbale subtests die meetellen in de scoring, en daarnaast ook nog een enkele cijferreeks. Nu het alleen een illustratief doel dient, had ik liever gezien dat er fictieve items opgenomen waren, nu dit boek vrij verkrijgbaar is en we testitems willen beschermen. Ook wordt er een testinstructie-citaat uit de handleiding gegeven, dat in werkelijkheid helemaal niet zo in de testhandleiding staat. Op blz. 266 wordt bij de subtest Woordenschat gesuggereerd dat een testvraag letterlijk als volgt zou zijn: ‘*Wat betekent het woord “...”?*’ (de puntjes zijn mijn aanpassing om niet ook in deze recensie een testitem weg te geven), terwijl in de testhandleiding deze vraag (wederom zijn de puntjes mijn aanpassing) geformuleerd is als: ‘*.... Wat is een ...?*’
6. Blz. 267: Wat mij betreft had bij de beschreven meetpretentie van subtest Matrix Redeneren nog toegevoegd mogen worden dat het om *non-verbaal* (en) *logisch* redeneervermogen gaat.
7. Blz. 274: Hier staat een prachtige tekst over de behouden waarde van het TIQ bij disharmonische profielen: ‘*Het is gebleken dat een discrepantie van 23 punten of meer tussen 2 indexscores in de WISC-V relatief vaak voorkomt: namelijk bij 52.5% (Amerikaanse klinische groepen) tot 56.6% (Amerikaanse normgroep) van de kinderen (Wahlstrom et al., 2018).*’

Ik ben even gaan kijken in de geciteerde bron, waarin ik op blz. 268 van dat Engelstalige boek onderstaande tekst vond. Ik wilde vooral nakijken of de statistieken alleen betrekking hadden op 2 *primaire* indexscores en dat bleek het geval (ik heb het even vet gemaakt in het citaat hieronder) en zou mijns inziens in het hoofdstuk nog aangevuld mogen worden.

Verder heb ik hieronder nog wat meer geciteerd uit de bron, omdat er nog wat aanvullende onderbouwing (o.a. met de WPPSI-IV) wordt beschreven om nog duidelijker te maken dat het TIQ (Full Scale IQ; FSIQ) van waarde blijft bij een disharmonisch profiel. Ik neem het hier op als relevante aanvullende informatie voor u als psychodiagnost, het is niet bedoeld als verbeterpunt voor het handboek:

*‘Research indicates that the construct and predictive validity of the FSIQ are not affected by an index score discrepancy (Daniel, 2007). It is quite typical to have a discrepancy of greater than 23 points between two **primary** index scores: For the WISC-V, 56.6% of the normative sample (i.e., 1,246 of the 2,200 children) had such a discrepancy, and 52.5% of the special group study cases (261 of the 497 children from these groups) had such a discrepancy (Kaufman et al., 2016). Similar frequencies were found among the WPPSI-IV normative and special group study samples (Raiford & Coalson, 2014). Given the vast evidence in support of the predictive validity of g and the FSIQ (Daniel, 2007; Deary & Johnson, 2010; Deary, Strand, Smith, & Fernandes, 2007; Johnson, Deary, & Iacono, 2009; S. B. Kaufman et al., 2012), it is counterintuitive to assume that for almost 60% of children, the FSIQ is not valid due to such discrepancies. Identifying index score discrepancies can reveal rich clinical information about a child, which is why comprehensive assessment of these diverse cognitive domains is important. Furthermore, more specific domains of intellectual ability do not show the same broad degree of predictive validity as does g (Hartmann, Larsen, & Nyborg, 2009; Kotz, Watkins, & McDermott, 2008; Reeve & Charles, 2008). The FSIQ consistently provides essential, clinically rich information when practitioners are attempting to understand the expression of intelligent behavior in real-world settings (Jacobson, Delis, Hamilton, Bondi, & Salmon, 2004).’*

8. Blz. 281: In box 11.2, met een hele scheve verhouding tussen sterke en zwakke punten van de WISC-V, mis ik persoonlijk nog dat iedere primaire indexscore slechts is opgebouwd uit 2 subtests, waardoor het betreffende domein vooral screenend in kaart gebracht wordt, maar minder 'body' heeft of 'gedekt'/vertegenwoordigd wordt, dan wanneer het uit meer subtests zou bestaan (vgl. het WISC-III VIQ, dat uit 5 subtests bestond, bijna net zoveel als nu het WISC-V TIQ, wat in de WISC-III nog uit 10 subtests was opgebouwd). Elders (blz. 265) staat in dit hoofdstuk overigens heel mooi en genuanceerd beschreven dat het ook niet de pretentie is om het hele domein met de betreffende indexvulling te onderzoeken en dat je ook in conclusies erover voorzichtig moet zijn. Een citaat daarvan trof u al eerder in deze recensie aan. Ik zou het dan in deze box 11.2 nog laten terugkomen, zodat gebruikers zich bewust blijven van deze beperking. De WISC-V screent in de breedte, maar is minder verdiepend dan bijv. de WISC-III, die een beperkter aantal domeinen grondiger dekte.
- Zoals eerder gezegd en ook benoemd als punt waarover discussie is, mis ik bij de zwakke punten t.a.v. de WISC-V in deze box de onvoldoende gedifferentieerde onderwijstabel.
9. Blz. 283: Onder de grafiek wordt de term 'uitzonderlijk laag/hoog' gebruikt, wat past binnen de indeling van Resing (2021). Maar de indeling van Ruiter e.a. (2017), auteur van dit hoofdstuk en ook in lijn met wat er op de Q-global uitdraai staat, noemt het 'zeer laag/hoog'. Die wisselende termen geven ruis, hoewel er niet bijstaat welke indeling dit betreft. Sowieso moet de '>' bij 130 een '≥' zijn (conform de besproken classificatie-indelingen), want 130 valt nu precies tussen 2 categorieën en hoort nu nergens bij.
10. Verder zijn er nog wat eindredactionele details. Uitgebreider dan ik het hier benoem, heb ik e.e.a. in 2021 doorgegeven aan een van de auteurs:
- Er zijn inconsistenties tussen schrijfwijzen in de testhandleiding/Q-global en dit handboek. Denk hierbij bijv. aan hoofdlettergebruik in de namen van subtests en indexen of het gebruik van de Engelstalige of Nederlandstalige term (Fluid Redeneren vs. Fluïde r/Redeneren). Ik vind het jammer dat het boek niet in lijn is met de handleidingen/Q-global waar het wel netjes consistent wordt weergegeven.
 - Blz. 269: Onderaan wordt gesproken over de WISC-IV-NL, maar die is nooit verschenen. Nederland sloeg de WISC-IV over en ging na de WISC-III-NL gelijk over op de WISC-V-NL. Hier had 'WISC-IV' moeten staan, verwijzend naar de internationale edities.
 - Blz. 272: Onderaan wordt de WISC-III gerefereerd met jaartal 2003, elders in dit hoofdstuk wordt consequent 2005 aangehouden, dat zal hier ook 2005 moeten zijn. Voor de gene die van achtergrondinfo houdt: de WISC-III-NL verscheen in 2002, een coproductie tussen het NIP en Boom. Inhoudelijk was er veel te doen over vooral de normering, er volgde een erratum in oktober 2003 en in april 2005 volgde een herziene handleiding, waarin ook de normen waren aangepast. Er had vooral een statistische correctie plaatsgevonden omdat de normen erg streng waren ten gevolge van de relatieve oververtegenwoordiging van hoger opgeleide kinderen in de WISC-III steekproef. De 2002-editie van de WISC-III-NL had o.a. een 'onvoldoende' COTAN-beoordeling van de normen, de 2005-beoordeling was op dit punt verbeterd naar een 'voldoende'.
 - Blz. 277: Bovenaan wordt Amerikaans equivalentieonderzoek van Daniel & Wahlstrom (2018) besproken. Er staat dat het de WISC-IV betreft, maar moet de WISC-V zijn (zo staat het ook in de literatuurlijst en zowel het jaartal als de context wijzen ook in die richting).
 - Blz. 286: de WPPSI-III staat per ongeluk geschreven als WPSSI-III.
 - Ook vond ik enkele minimale foutjes in de literatuurlijst (bij de referentie van Raiford e.a. (2016) mist bijv. de voorlaatste auteur (James A. Holdnack) en staan op die plek drie puntjes ('...').

Alles samenvattend, vind ik dit een prima handboek, waarmee we weer een aantal jaar vooruit kunnen. Een aanrader voor mensen die betrokken zijn bij intelligentiediagnostiek! En... mocht u een foutje in deze recensie aantreffen, laat het me gerust weten. 😊 Ik pas het graag aan.

Utrecht, 15-10-2022

Yaron Kaldenbach

GZ-psycholoog

www.apollopraktijk.nl

ykaldenbach@hotmail.com