

Flinterdunne en flagrante Flynn-effecten bij licht verstandelijk gehandicapte kinderen: aanbevelingen voor het gebruik van de WISC-III¹

Wim Pesch^a & Albert Ponsioen^b

^a Pluryn / J.P. Heije

^b OZC/Amstelmonde en Landelijk Kenniscentrum LVG

Samenvatting

De discussie rond de recent uitgebracht nieuwe WISC-intelligentietest (de WISC-III) draagt in ieder geval bij tot een relativering van een IQ-score. Zeker bij licht verstandelijk gehandicapte kinderen is dit hard nodig gezien de waarde die men onder meer in het kader van zorgtoewijzing en indicatiestelling aan deze score hecht. In dit artikel wordt een voorstel gedaan voor een verantwoord gebruik van de WISC-III bij licht verstandelijk gehandicapte kinderen.

Inleiding:

Het eerste waarmee licht verstandelijk gehandicapte (LVG) kinderen² in een psychodiagnostisch onderzoek in de regel te maken krijgen is een IQ-test. Per definitie is het intelligentieniveau³ een belangrijk diagnostisch kenmerk van een LVG-kind en het speelt in de klinische praktijk vaak een doorslaggevende rol in de besluitvorming omtrent indicatiestelling en zorgtoewijzing. Door de toenemende protocolisering, onlosmakelijk verbonden met de ontwikkeling naar vraaggestuurde zorg waarbij indicatiestelling een professie op zich aan het worden is en het subjectieve klinische oordeel door 'harde' (doch vooral papieren) diagnostische gegevens wordt overschaduwd, wordt aan een IQ-score zwaar getild. Het IQ-getal als resultaat van een

¹ Dit artikel is in een aangepaste vorm aangeboden voor publicatie in *De Psycholoog*

² Waar LVG-kinderen staat, wordt LVG-kinderen en -jongeren bedoeld.

³ 'Licht verstandelijk gehandicapte jongeren zijn personen tot circa 21 jaar die in hun ontwikkeling zijn belemmerd en die zich op grond van hun lagere intellectueel functioneren én beperkte sociale redzaamheid niet (zonder hulp) kunnen handhaven in één of meer reguliere maatschappelijke verbanden (gezin, school, werk, groep, leeftijdsgenoten, buurt). Met speciale hulp hebben zij een redelijke kans op een zekere mate van zelfstandigheid als volwassenen.' (NVGz-nota, 1995).

gerenommeerde intelligentietest wekt immers de suggestie een 'stabiel en hard' gegeven te zijn.

Het toeval wil dat vrij recent van de twee meest gebruikte intelligentietests in Nederland, de Wechsler schalen, nieuwe versies zijn uitgebracht, de WAIS-III en de WISC-III. Dit werd tijd daar de normen van de vorige versie voor volwassenen (WAIS) meer dan dertig jaar oud waren en die van de vorige kinderversie (de WISC-RN) ongeveer 20 jaar oud. Het enthousiasme over de meer eigentijdse vormgeving en inhoud van de beide tests temperde alras na een aantal publicaties in de Psycholoog met kritische kanttekeningen omtrent de betrouwbaarheid, de validiteit en de 'scheve' verdeling van de normgroepen (Tellegen, 2002a, Tellegen, 2002b). Voor de WISC-III is de ontstane situatie des te pijnlijker daar deze door het dienstencentrum van de beroepsvereniging van psychologen (het NIP) wordt uitgegeven. Dat de aankondigingen van kritische artikelen over dit onderwerp op het titelblad van de Psycholoog tot nu toe alleen de WAIS-III betreffen (tot 'weg met de nieuwe WAIS?' aan toe, de Psycholoog – jaargang 37 – nummer 9) is wellicht een indicatie voor de gevoeligheden rond dit thema. Het feit dat het NIP Dienstencentrum de WISC-III tot nader orde uit de handel heeft genomen geeft aan dat men nu toch wel ernst neemt met de situatie (NRC, "Twijfels over IQ-test", 2 april 2003).

Om nu gevolg te geven aan de oproep van Tellegen om de WISC-III (en de WAIS-III) terug te sturen naar de uitgevers (niet goed, geld terug?) en de test voorlopig niet meer te gebruiken gaat ons te ver. Wij sluiten ons aan bij de opvatting van een vermaard onderzoeker op het gebied van de verstandelijk gehandicaptenzorg (Kraijer, 1994) dat '*in principe iedere nieuwe test beter is dan een verouderde*' en adviseren een behoedzaam gebruik van de test met enige aanpassingen in het scoreprotocol voor de LVG-groep.

Over het behoedzame gebruik kunnen wij kort zijn. Het adagium 'elk nadeel heeft z'n voordeel' van een vermaard voetbalphilosoof in acht nemend (Davids, 1998), is de aan te bevelen hantering van betrouwbaarheidsintervallen een prijzenswaardige nuancering van een IQ-score die voor elke (oude en nieuwe) intelligentietest zou moeten opgaan. In weinig rapporten van onderzoeken waarin de oude WISC-RN werd gebruikt wordt deze nuancering echter aangegeven (laat staan dat er een woord wordt gerept over de verouderde normering waarop men de resultaten baseert!). In het vervolg van dit artikel gaan wij nader in op het bijzondere gebruik van de WISC-III bij LVG-kinderen. Wij baseren ons hierbij op een

reeds aangelegd databestand met 98 LVG-kinderen waarbij de WISC-III is afgenomen. Om de ervaringen met de WISC-III goed te kunnen duiden wordt eerst een samenvatting gegeven van intelligentieonderzoek bij LVG-kinderen vanaf het begin van de jaren tachtig van de vorige eeuw. De databestanden die sinds deze periode in het Orthopedagogisch Centrum Jan Pieter Heije, onderdeel Pluryn Werkenrode Groep te Oosterbeek worden aangelegd, werden voor dit doel geraadpleegd.

Intelligentieonderzoek bij LVG-kinderen:

Uit longitudinaal onderzoek is bekend dat de intelligentie nagenoeg constant blijft gedurende een mensenleven. Recent bekend geworden onderzoek van Flynn (referentie) maakt echter duidelijk dat de door een zelfde IQ-test gemeten intelligentie, met ongewijzigde normgegevens, minder constant is: in 30 jaar werd een toename van ongeveer 10 IQ-punten gevonden. Factoren als leereffecten bij test-hertesten en het 'doorsijpelen' van de door de test gemeten kennis in de maatschappij worden

verondersteld hierbij een rol te spelen. Het opnieuw normeren van een IQ-test is dus noodzakelijk, hoe langer men daarmee wacht des te groter wordt dit 'Flynn-effect'. Zo zou met WISC-RN (normgegevens uit 1986) het IQ momenteel zo'n 6 punten hoger ingeschat worden. Het is echter de vraag of dit Flynn-effect zich in dezelfde mate voordoet voor de kinderen die de uiteinden van de IQ-verdeling vertegenwoordigen. In tabel 1 worden de resultaten van LVG-kinderen getest met de aloude WISC-R (groene uitvoering) uit 1974 vergeleken met die van de in de jaren tachtig nieuw uitgebrachte WISC-RN (gele uitvoering). Daar in die periode vanaf de introductie in 1986 van de WISC-RN de normgegevens van de WISC-R minstens 14 jaar oud waren, zou men een Flynn-effect van minstens 3 punten moeten vinden. Voor het performaal IQ zien we in tabel 1 wel een Flynn-effect, voor het totaal IQ slechts de helft van het voorspelde Flynn-effect en voor het verbaal IQ is er zelfs een negatief effect.

Tabel 1. Vergelijking van de WISC-R met de WISC-RN

Meting binnen dezelfde LVG-populatie	Verbaal IQ	Performaal IQ	Totaal IQ
Met de WISC-R uit 1974	65,3 (N=103)	75,4 (N=118)	67,9 (N=103)
Met de WISC-RN uit 1986	66,7 (N=100)	71,1 (N=100)	66,2 (N=100)
Flynn-effect	- 1,4	+ 4,3	+ 1,7

Omdat het in deze (tussen-subject) analyse gaat om twee verschillende groepen kinderen is tevens een (binnen-subject) analyse gedaan op een en dezelfde groep kinderen waarbij meermalen dezelfde WISC-RN test werd afgenomen. De periode tussen twee testafnames bedroeg gemiddeld 3 jaar en 4 maanden. De resultaten van deze analyse worden samengevat in tabel 2.

Tabel 2. Vergelijking van herhaalde WISC-RN metingen

Meting tussen subjecten	Verbaal IQ	Performaal IQ	Totaal IQ
Oudste meting WISC-RN	71,0 (N=147)	75,1 (N=147)	70,3 (N=147)
Recente meting WISC-RN	68,5 (N=147)	76,8 (N=147)	69,8 (N=147)
Flynn-effect met een gemiddeld verschil van 3 jaar; 4 mnd.	- 2,5	+ 1,7	- 0,5

Tabel 2 laat in de verbale intelligentie een negatief Flynn-effect (-2,5) zien en binnen de performale intelligentie een licht positief effect (+1,7). Het totaal IQ laat nauwelijks enige verandering (-0,5) zien in ruim drie jaar. Een verdere uitwerking in tijd via een mediaansplitsing is te zien in tabel 3.

Tabel 3. Vergelijking van herhaalde WISC-RN metingen tussen (LVG-)subjecten

Beneden de mediaan: van 2 jaar en 0 maanden tot 2 jaar en 8 maanden	Verbaal IQ	Performaal IQ	Totaal IQ
Oudste meting WISC-RN	72,1 (N=73)	76,8 (N=73)	71,67 (N=73)
Recente meting WISC-RN	69,6 (N=73)	77,4 (N=73)	70,55 (N=73)
Flynn-effect	- 2,5	+ 0,6	- 1,1
Boven de mediaan: van 2 jaar en 9 maanden tot 10 jaar en 3 maanden	Verbaal IQ	Performaal IQ	Totaal IQ

Oudste meting WISC-RN	70,0 (N=74)	73,3 (N=74)	68,9 (N=74)
Recente meting WISC-RN	67,5 (N=74)	76,1 (N=74)	69,0 (N=74)
Flynn-effect	- 2,5	+ 2,8	- 0,1

Naarmate de tijd vordert, zien we in tabel 3 geen Flynn-effect in het verbaal IQ, eerder een constant negatief Flynn-effect (-2,5). Het 'verbale' vervluchtigt veelal en de gegevens boven de mediaan laten bij de verbale subtests het meeste negatieve Flynn-effect zien bij: *Begrijpen* (-1,2), *Woordenschat* (-0,7) en *Overeenkomsten* (-0,3). Het Romeinse gezegde '*verba volant*' (*het woord vervliegt*) geldt nog steeds en zeker voor LVG-kinderen. Daarentegen zien we in de handelingsintelligentie in tijd wel enig leer- of Flynn-effect (2,8). De gegevens boven de mediaan laten bij de performale subtests *Plaatjes ordenen*

(0,8), *Blokatronen* (0,6), *Doolhoven* (0,6) en *Figuur Leggen* (0,3) het meeste Flynn-effect zien. De conclusie dat als er sprake is van een Flynn-effect dit alleen voor het performale IQ geldt, is in het licht van de huidige WISC-III discussies belangrijk, maar eigenlijk niet zo verrassend. Als het bij het Flynn-effect vooral handelt om leereffecten dan zou men bij moeilijk lerende kinderen niet veel effect verwachten. Bovendien is de ervaring dat bij deze groep kinderen het leren zich vooral op het visueel-concrete vlak afspeelt. (Van Weelden, 1988; Ponsioen & Van der Molen, 2002).

Tabel 4. Vergelijking van herhaalde WISC-RN metingen tussen subjecten

WISC-RN metingen binnen 2 jaar (tijdsspanne van minimaal 9 maanden tot maximaal 23 maanden)	Verbaal IQ	Performaal IQ	Totaal IQ
Oudste meting WISC-RN	67,3 (N=12)	77,7 (N=12)	69,33 (N=12)
Recente meting WISC-RN	68,6 (N=12)	78,4 (N=12)	70,67 (N=12)
Flynn-effect	+ 1,3	+ 0,8	+ 1,3

Incidenteel, bijvoorbeeld bij onduidelijkheden tijdens een internaatsaanmelding in combinatie met een gebrekkige communicatie of rapportage, wilde het wel eens voorkomen dat een LVG-kind per ongeluk twee keer werd getest binnen de officieel voorgeschreven 'rustperiode' van 2 jaar in verband met test-hertesteeffecten. Uit deze gegevens blijkt dat we in deze korte periode wel enig leer- of Flynn-effect zien binnen het verbaal IQ en totaal IQ (zie tabel 4). Met andere woorden voor een LVG-kind zien we wel enig

leereffect (of test-hertesteeffect) binnen een korte tijdsspanne, maar dit Flynn-effect op langere termijn lijkt voor de LVG-doelgroep toch miniem.

Maar hoe anders zijn de resultaten met de nieuwe WISC-III test: het Flynn-effect bij 56 nieuwe WISC-III afnamen bedraagt 6,3 punten, ten opzichte van eerdere IQ metingen (zie tabel 5) en 6,1 punten ten opzichte van de oude WISC-RN-metingen (zie tabel 6). In beide gevallen laat het verbale IQ een fors Flynn-effect (7,5 en 8,1) zien in tegenstelling tot het performale IQ.

Tabel 5. Vergelijking van eerdere intelligentiemetingen met de WISC-III

Meting tussen subjecten met LVG-niveau	Verbaal IQ	Performaal IQ	Totaal IQ
Meest recente IQ-meting	73,2 (N=52)	76,7 (N=52)	72,4 (N=56)
De nieuwe WISC-III	65,7 (N=56)	72,6 (N=56)	66,1 (N=56)
Flynn-effect	+ 7,5	+ 4,1	+ 6,3

Tabel 6. Vergelijking van de WISC-RN met de WISC-III

Meting tussen subjecten met LVG-niveau	Verbaal IQ	Performaal IQ	Totaal IQ
Meest recente WISC-RN meting	73,4 (N=48)	76,1 (N=48)	71,9 (N=48)
De nieuwe WISC-III	65,3 (N=48)	72,4 (N=48)	65,8 (N=48)
Flynn-effect	+ 8,1	+ 3,7	+ 6,1

Waarschijnlijk heeft dit 'Flynn-effect' meer te maken met de schoonheidsfouten die bij de

ontwikkeling van de nieuwe versie gemaakt zijn (Tellegen, 2002a; Tellegen, 2002b). Zo kan men kritiek hebben op de samenstelling van de

normgroepen, vooral ten aanzien van de leeftijdscategorieën 15 en 16 jaar. Zo zou het onderwijsniveau in de gebruikte steekproeven te hoog zijn. Dit maakt dat de gevonden lagere scores bij de LVG-groep een onderschatting van het IQ kan inhouden.

Een hiermee samenhangend probleem is de gevonden factorstructuur van de WISC-III. Bij de vorige WISC-versie, de WISC-RN, werden de (toen) 12 subtests samengevat door drie factoren: een verbale factor (V), een ruimtelijke factor (R) en een numerieke factor (N; samengesteld uit de subtests *Rekenen*, *Cijferreeksen* en *Substitutie*). Een factoranalyse op een databestand met de WISC-RN-scores van meer dan 800 LVG-kinderen gaf deze zelfde oplossing in drie factoren te zien. Met de WISC-III wordt de numerieke factor echter niet meer gevonden. De derde factor bij deze test, de Verwerkingssnelheid (Vs), wordt nu gevormd door de subtest *Substitutie* en de nieuwe subtest *Symbool Vergelijken*. Deze factor is erg visueel van aard en sluit daarom aan bij de sterke kanten van het cognitieve profiel van een LVG-kind. In de tabellen 8 en 9 komt dit verschil tussen de Vs-score en de N-score (berekend uit de normgegevens van de 'oude WISC-RN') duidelijk naar voren.

De veronderstelling is dat de N-factor voor LVG-kinderen met een lager opleidingsniveau belangrijker is dan voor kinderen met een hoger opleidingsniveau, daar veel van de kennis van LVG-kinderen minder geautomatiseerd is (o.a.

rekenvaardigheden). Bij deze kinderen zal het werkgeheugen daardoor meer belast worden en minder efficiënt kunnen worden ingezet bij de taken die op de N-factor laden (*Rekenen*, *Cijferreeksen* en *Substitutie*). Deze veronderstelling is getoetst door middel van een factoranalyse op het databestand van 98 kinderen. Een exploratieve Principale Componenten Analyse (PCA) met varimax rotatie werd uitgevoerd op de scores van de kinderen waarbij alle 13 subtest waren afgenomen (N = 85). De resultaten van deze analyse zijn ondergebracht in tabel 7.

Op het databestand werden vervolgens confirmatieve factoranalyses uitgevoerd om te beoordelen welk factormodel de data het beste beschrijft: het WISC-III driefactorenmodel of het vanuit de exploratieve analyse gevonden vierfactorenmodel. In figuur 1 worden de beide modellen uitgebeeld.

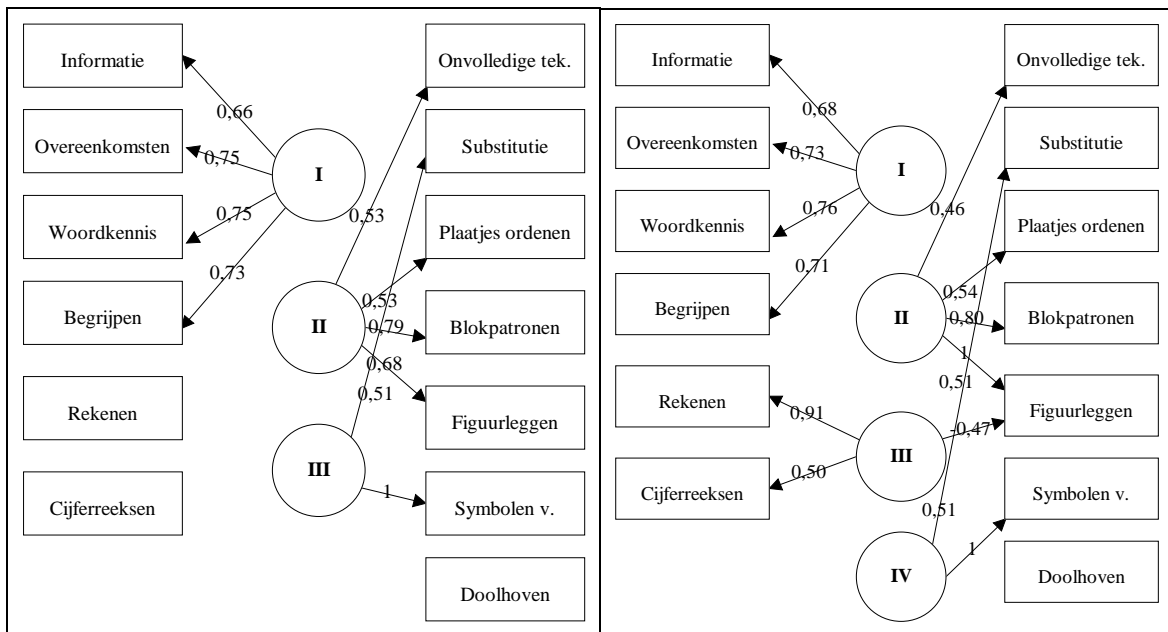
Uit de Chi²-toetsen blijkt het vierfactorenmodel de data beter samen te vatten ($\chi^2 = 67,2$, $p = 0,218$). De volgende aanvullingen zijn hierbij van belang:

- de ruimtelijke factor (2^e factor) omvat tevens de subtest *Doolhoven*
- de numerieke factor (3^e factor) bestaat uit de subtest *Rekenen*, *Cijferreeksen* (zoals verwacht), maar tevens uit de subtest *Figuurleggen*.

Tabel 7. Principale Componenten Analyse met Varimax rotatie (alleen de waarden groter dan 0,4 zijn opgenomen)

	Component			
	1	2	3	4
informatie	,705			
overeenkomsten	,810			
rekenen			,701	
woordkennis	,756			
begrijpen	,809			
cijferreeksen			,802	
onvolledige tekeningen		,749		
substitutie				,883
plaatjes ordenen		,595		
blokpatronen		,701		
figuurleggen		,731		
symbolen vergelijken				,783
doolhoven		,490		

Extraction Method: Principal Component Analysis. Rotation Method: Varimax with Kaiser Normalization (rotation converged in 6 iterations).



Figuur 1. Factormodellen met drie gecorreleerde latente factoren (links; $\chi^2 = 150,469$, $p = 0,000$, $df = 66$) en met vier gecorreleerde latente factoren (rechts; $\chi^2 = 67,166$, $p = 0,218$, $df = 59$) met de bijbehorende gestandaardiseerde regressiecoëfficiënten.

De analyses laten zien dat voor 'de' LVG-groep de drie factoren van de WISC-III de resultaten onvoldoende beschrijven. In het scoringsprotocol van de WISC-III worden bovendien de subtest *Cijferreeksen* en *Doolhoven* niet in de berekening van de factorscores meegenomen. Juist voor de LVG-groep blijken dit belangrijke subtests (respectievelijk voor de numerieke factor N en voor de ruimtelijke factor R). In samenwerking met het NIP Dienstencentrum worden vanuit het Landelijk Kenniscentrum LVG de WISC-III-gegevens van de LVG-kinderen verzameld en in de normgroepen opgenomen. Vooruitlopend op de resultaten van het nadere onderzoek naar validiteit en betrouwbaarheid van de WISC-III wordt in de onderstaande paragraaf van dit artikel een voorstel gedaan voor een aanpassing van het scoringsprotocol voor LVG-kinderen, waarmee naar ons inzicht meer recht wordt gedaan aan de cognitieve relevantie en de onderlinge samenhang van de subtests voor deze groep kinderen.

Voorstel WISC-III-gebruik bij LVG-kinderen:

Om de WISC-III met de huidige tekortkomingen toch enigszins verantwoord te kunnen gebruiken bij de LVG-populatie wordt in deze afsluitende paragraaf een aantal voorstellen gedaan met betrekking tot de berekening van de IQ-scores, het gebruik van de factorscores en het vermelden van betrouwbaarheidsintervallen.

Aanbevolen wordt om voor de berekening van het verbale IQ de subtest *Cijferreeksen* mee te nemen en voor het performale IQ de subtest *Doolhoven*. De bevinding van het NDC dat deze

subtests voor de gebruikte normgroepen statistisch weinig toevoegen, blijkt voor de LVG-groep niet op te gaan (zie de vorige paragraaf). Deze aangepaste berekeningsmethode⁴ vormt geen bezwaar indien men maar melding maakt via welke route men het IQ berekent. In tabel 8 wordt voor een 12-jarige LVG-jongere de voorgestelde berekeningsmethode ('LVG-methode')⁵ vergeleken met de standaardmethode ('NDC-methode').

Omdat er bij de LVG-groep sprake is van een 4-factorenoplossing, waarbij de derde factor voor een deel overlapt met de F3IQ-factor van de vorige WISC-RN, wordt het gebruik van deze factor bij LVG-kinderen aanbevolen. Bij deze groep kinderen kan de Verwerkingssnelheidsfactor een vertekend beeld geven als deze als 'leerfactor' wordt opgevat. Met het gebruik van de in totaal vier factorscores wordt bovendien aangesloten bij de factorstructuur van de WAIS-III. De factorscores voor de LVG-groep zijn in tabel 9 ondergebracht.

Daar de definitieve normgegevens (en de hernieuwde COTAN-beoordeling) nog op zich laten wachten, wordt aanbevolen naast de IQ-factorscores de betrouwbaarheidsintervallen aan te geven. Deze aanbeveling is van algemene betekenis en niet afhankelijk van de huidige status van de WISC-III. De waarde van een IQ-score is zeker bij LVG-

⁴ In samenspraak tussen het Landelijke Kenniscentrum LVG en het NDC overeengekomen.

⁵ De som van de nu 6 subtests voor zowel het verbale als het performale gedeelte wordt met 5/6 vermenigvuldigd om van de factorscores gebaseerd op 5 subtests gebruik te kunnen blijven maken.

kinderen betrekkelijk te noemen (Ponsioen & Van der Molen, 2002). Van de huidige discussie over de

WISC-III en de WAIS-III kan wat dit aangaat een louterende werking uitgaan.

Tabel 8. IQ-factorscores WISC-III volgens de officiële NIP-IQ-berekening en de voorgestelde LVG-berekening, mét de subtests Cijferreeksen en Doolhoven (zie de tekst)

	Officiële NDC scoring		LVG-methode (inclusief de subtests Cijferreeksen en Doolhoven)	
	IQ-niveau	95% betrouwbaarheid	IQ-niveau	95% betrouwbaarheid
Verbaal IQ	74	68-84	77	71-87
Performaal IQ	64	59-79	71	65-85
Totaal IQ	67	62-77	72	66-81
VB Verbaal begrip	70	64-81	70	64-81
PO perceptuele organisatie	63	58-79	63	58-79
VS verwerkingssnelheid	88	79-100	88	79-100
Numerieke factor (F3IQ WISC-RN)			78	<i>geen gegevens</i>

Tabel 9. WISC-III-factorscores

	Gem.	Minimum	Maximum	SD
verbaal IQ	67	56	91	8
performaal IQ	75	49	104	12
totaal IQ	68	54	92	8
verbaal begrip	68	56	95	8
performale organisatie	77	49	106	13
verwerkingssnelheid	75	55	121	14
sequentieer factor*	66	45	100	12
ruimtelijk analytische factor*	80	48	108	14

* Gebaseerd op WISC-RN normgegevens

Literatuur:

Davids, H. (1998). *Je moet schieten, anders kun je niet scoren en andere uitspraken van Johan Cruijff*. 's Gravenhage: BZZTôH.

Flynn, J.R. (1987). Massive IQ gains in 14 nations: What IQ tests really measure. *Psychological Bulletin*, 101, 171-191.

Kraijer, D.W. (1994). *Zwakzinnigheid, autisme en aan autisme verwante stoornissen: Classificatie, diagnostiek, prevalentie, specifieke problematiek, opvoeding en behandeling*. Lisse: Swets & Zeitlinger.

NVGz-nota (1995). *Vorm en kleur. Beleidsnota van de orthopedagogische centra voor licht verstandelijk gehandicapte jongeren van de Nederlandse*

Vereniging voor Gehandicaptenzorg Utrecht: NVGz.

Ponsioen, A.J.G.B. & Van der Molen, M.J. (2002). *Cognitieve vaardigheden van licht verstandelijk gehandicapte kinderen en jongeren. Een onderzoek naar mogelijkheden*. Ermelo: Landelijk Kenniscentrum LVG.

Tellegen, P. (2002a). De WISC-III NL. Een illusie armer. *De Psycholoog*, 37 (11). 607-610.

Tellegen, P. (2002b). Afname van de WAIS-III of WISC-III NL. Verantwoord of verstandig? *De Psycholoog*, 37 (12). 677-679.

Van Weelden, J. (1988). *Moeilijk lerende kinderen*. Rotterdam: Lemniscaat.