

Bedankt voor het downloaden van dit artikel. De artikelen uit de (online)tijdschriften van Uitgeverij Boom zijn auteursrechtelijk beschermd. U kunt er natuurlijk uit citeren (voorzien van een bronvermelding) maar voor reproductie in welke vorm dan ook moet toestemming aan de uitgever worden gevraagd.

# Boom

Behoudens de in of krachtens de Auteurswet van 1912 gestelde uitzonderingen mag niets uit deze uitgave worden verveelvoudigd, opgeslagen in een geautomatiseerd gegevensbestand, of openbaar gemaakt, in enige vorm of op enige wijze, hetzij elektronisch, mechanisch door fotokopieën, opnamen of enig andere manier, zonder voorafgaande schriftelijke toestemming van de uitgever.

Voor zover het maken van kopieën uit deze uitgave is toegestaan op grond van artikelen 16h t/m 16m Auteurswet 1912 jo. Besluit van 27 november 2002, Stb 575, dient men de daarvoor wettelijk verschuldigde vergoeding te voldoen aan de Stichting Reprorecht te Hoofddorp (postbus 3060, 2130 KB, [www.reprorecht.nl](http://www.reprorecht.nl)) of contact op te nemen met de uitgever voor het treffen van een rechtstreekse regeling in de zin van art. 16l, vijfde lid, Auteurswet 1912.

Voor het overnemen van gedeelte(n) uit deze uitgave in bloemlezingen, readers en andere compilatiewerken (artikel 16, Auteurswet 1912) kan men zich wenden tot de Stichting PRO (Stichting Publicatie- en Reproductierechten, postbus 3060, 2130 KB Hoofddorp, [www.cedar.nl/pro](http://www.cedar.nl/pro)).

*No part of this book may be reproduced in any way whatsoever without the written permission of the publisher.*

[info@boomamsterdam.nl](mailto:info@boomamsterdam.nl)  
[www.boomuitgeversamsterdam.nl](http://www.boomuitgeversamsterdam.nl)



# Opinie: Uniformiteit in de kwalitatieve beschrijving van scores op prestatietaken

- ▶ Marc P.H. Hendriks
- ▶ Bart A.W. Mol
- ▶ Roy P.C. Kessels

■ **Samenvatting** — Het beschrijven van kwantitatief bepaalde testscores is een essentieel aspect van het neuropsychologisch onderzoek. Het gebruik van uniforme, kwalitatieve beschrijvingen van de scores op (cognitieve) tests zorgt voor een heldere, ondubbelzinnige communicatie binnen de beroepsgroep en verbetert het maatschappelijk belang van de klinisch-neuropsychologische diagnostiek. Hoewel de noodzaak hiervan breed onderschreven wordt, was er tot nu toe geen (inter)nationaal geaccepteerde overeenstemming over de wijze waarop scores worden geclassificeerd en beschreven. In 2018 heeft een consensusbijeenkomst van de American Academy of Clinical Neuropsychology (AACN) plaatsgevonden, met twee doelen: (1) het formuleren van een voorstel voor het classificeren van testscores die normaal en niet-normaal verdeeld zijn, en (2) het definiëren van ‘stoornis’ en de toepassing bij individuele cases (Guilmette e.a., 2020). In dit artikel hebben we de conclusies en aanbevelingen getracht te vertalen in een voorstel voor de Nederlandse praktijk.

## Inleiding

Een van de maatregelen die de overheid heeft gesteld naar aanleiding van de aanpak van het coronavirus, is dat we minstens anderhalve meter afstand van elkaar moeten houden. Een kwantitatieve regel die geen onduidelijkheid geeft. Iedereen weet immers hoe groot deze afstand is (met een gering betrouwbaarheidsinterval). Verwarrend is echter dat deze kwantitatieve regel op verschillende manieren kwalitatief omschreven kan worden, afhankelijk van de context; zo zal het aantal mensen binnen de toegestane anderhalve meter bij de kapper, in een horecagelegenheid of in een park omschreven kunnen worden als ‘druk’, ‘gemiddeld’, respectievelijk ‘niet zo druk’. Bovendien geldt dat er in verschillende landen ook weer

verschillende definities worden gehanteerd als het gaat om een kwalitatief begrip als ‘voldoende veilige afstand’. Zonder deze begrippen te expliciteren of ze eenduidig te gebruiken ligt miscommunicatie op de loer.

Als het om het kwalitatief beschrijven van kwantitatieve test scores in de (neuro)psychologische diagnostiek gaat, accepteren we als beroepsgroep deze onduidelijke situatie al een groot aantal jaren (Hendriks e.a., 2014). Recent zijn de aanbevelingen van een in 2018 gehouden consensusconferentie van de American Academy of Clinical Neuropsychology (AACN) gepubliceerd (Guilmette e.a. 2020). Deze conferentie had een tweetal doelstellingen. Op de eerste plaats was het doel om tot overeenstemming te komen over uniforme, kwalitatieve beschrijvingen van resultaten van cognitieve (prestatie)tests waarvan de scores normaal of niet-normaal verdeeld zijn. Ten tweede wilde men komen tot een definitie van het begrip ‘stoornis’ en de toepassing hiervan bij casus-specifieke klinische besluitvorming in de praktijk. In dit artikel willen we de voorstellen en aanbevelingen die op basis van de consensusconferentie worden gedaan, tegen het licht houden van de huidige praktijk van de neuropsychologische diagnostiek in Nederland en een voorstel formuleren tot het gebruik van uniforme kwalitatieve beschrijvingen van testprestaties. Uiteraard is het ook onze doelstelling om hiermee de communicatie binnen de beroepsgroep van de neuropsychologische diagnostiek te verbeteren, maar ook de communicatie tussen klinisch neuropsychologen enerzijds en de verwijzers en patiënten en hun directbetrokkenen anderzijds. Zoals het voorstel van de AACN gebaseerd is op consensus, is de voorliggende vertaling naar de Nederlandse praktijk, dus (nog) niet als richtlijn bedoeld, maar als een aanzet tot het bereiken van een consensus; deze staat dan ook open voor discussie. Het is ook nadrukkelijk niet bedoeld als een instructie over de wijze waarop neuropsychologische testgegevens moeten worden geïnterpreteerd. Anderhalve meter is immers anderhalve meter, en het omschrijven van scores is wat anders dan het interpreteren ervan; dit laatste vraagt een systematisch diagnostisch besluitvormingsproces waarbij een neuropsycholoog testprestaties en informatie uit verschillende andere bronnen als de (hetero)anamnese en de gedragsobservaties integreert om de geformuleerde vraagstelling over een specifieke patiënt te beantwoorden.

In de huidige klinische praktijk is de in *De Psycholoog* gepubliceerde indeling van Resing en Blok (2002; in 2015 door Resing aangepast, waarbij de term ‘zwakzinnig’ vervangen is door ‘intellectueel beperkt’) over de beschrijving van de intelligentiescores de enige classificatie waar in de beroepsgroep overeenstemming over bestaat. Deze indeling is in het kader van de neuropsychologische diagnostiek niet goed toepasbaar. De voorge-

stelde kwalitatieve beschrijvingen als ‘licht intellectueel beperkt’ verwijzen immers naar verstandelijke of intellectuele vermogens. Vergelijkbare normscores op een willekeurige andere neuropsychologische test (zoals de Stroop Kleur-Woord Test) kunnen uiteraard niet omschreven worden als ‘licht intellectueel beperkt’. Overigens bestaat er niet alleen in de praktijk van de klinische neuropsychologie ontevredenheid over deze onduidelijkheid. Recent hebben Ruiters en collega's (2017) over deze indeling en het beschrijven van de intelligentiescores ook een voorstel tot herziening van de classificatie gedaan. Gezien het wereldwijde gebruik van de Wechslerschalen stellen zij voor de bijbehorende beschrijvingen (Wechsler, 2012) als uitgangspunt te nemen. Voor het beschrijven van de intervallen stellen zij voor kwalitatieve beschrijvingen te gebruiken als ‘laag’, ‘zeer hoog’, en ‘laag in het gemiddelde bereik’.

Voor het beschrijven van zowel scores op een intelligentietest als op andere cognitieve prestatietests hebben ook volgens ons bredere, test- en domeinonafhankelijke termen, zoals gemiddeld, hoog, zwak of zeer laag de voorkeur, omdat deze bij alle scores op prestatietaken kunnen worden toegepast. We hebben de indruk dat binnen de klinische neuropsychologie het rapporteren met behulp van dergelijke omschrijvingen vaker wordt gedaan dan het rapporteren in testscores. Dit vergroot de leesbaarheid van de rapportages, maar juist omdat geen overeenstemming is over de exacte betekenis van de omschrijvingen, is hier sprake van een schijnovereenstemming, en daardoor juist meer verwarring. Zo wordt de term ‘laaggemiddeld’ soms gebruikt voor scores die laag in het bereik liggen van  $-1$  tot  $+1$  SD, en soms voor scores die weer net onder de  $-1$  SD liggen, tot  $-1,5$  SD.

Een belangrijke voorwaarde voor de gebruikte terminologie is dat deze niet veroordelend is of kan worden gezien als een klinische vaststelling, maar dat ze uitsluitend een positie aanduidt in de normaalverdeling van de testscores. De klinische interpretatie dient in dit opzicht ook een onafhankelijk proces te zijn, zodat de psycholoog die verantwoordelijk is voor de interpretatie, bepaalt of er sprake is van een neuropsychologische beperking of stoornis. In het Consensus Conference Statement wordt terecht onderscheid gemaakt tussen classificatie van normaal-verdeelde scores en van niet-normaal verdeelde scores, bovendien worden aanbevelingen gedaan over de definitie en het hanteren van de term ‘stoornis’ (Guilmette e.a., 2020).

### Normaal-verdeelde scores

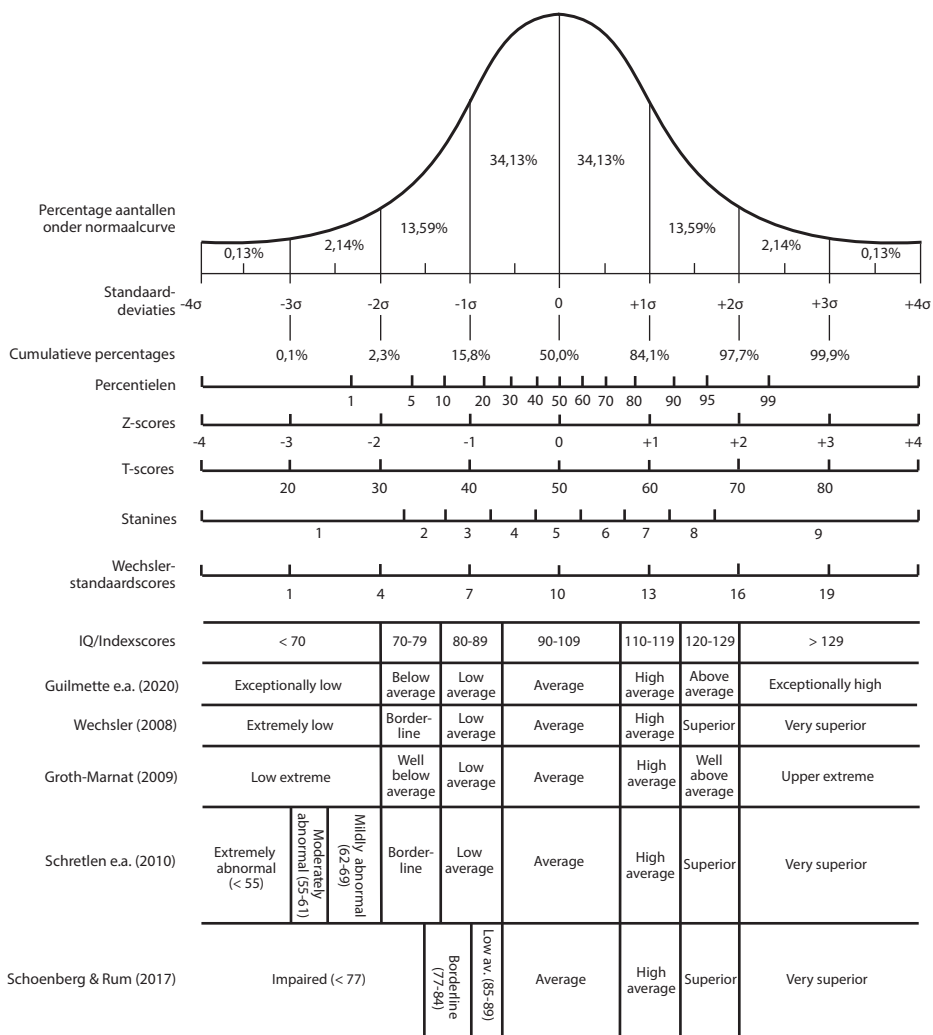
De uitgangspunten bij het formuleren van de aanbevelingen waren:

- ▶ De interpretatie van scores verschilt van de kwalitatieve beschrijving ervan.
- ▶ Er kan geen sprake zijn van een stoornis van testcores, maar van functies.
- ▶ Eenvoud van de beschrijving zal de communicatie verbeteren.
- ▶ Kwalitatieve beschrijvingen moeten gebaseerd zijn op de frequentie en gemeenschappelijkheid van prestaties, niet van pathologie.

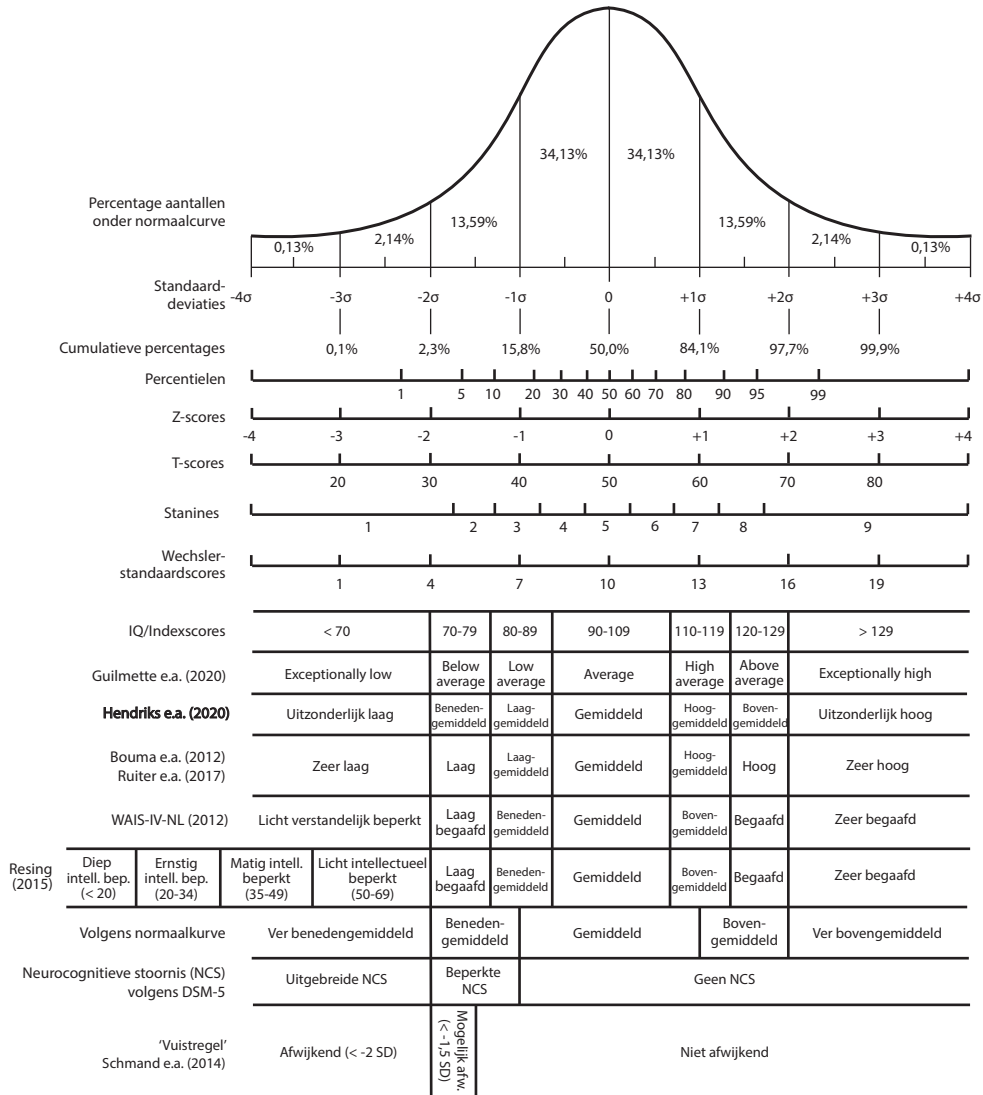
In de internationale literatuur is een beperkt aantal classificatiemodellen beschreven (Figuur 1, p. 170). De beschreven modellen verschillen in de omschrijvingen van de intervallen en ook in de mate waarin er meer of minder verfijning wordt aangebracht aan de onderkant van de classificaties in de normaalverdeling. Tegelijkertijd vertonen ze opvallend veel overlap in gehanteerde categorieën. De meest recente modellen volgen goeddeels de Wechslerindeling, waarbij scores geclassificeerd worden in zeven intervallen van ongeveer een standaarddeviatie van 0,7 (Figuur 1). Een alternatief model betreft een vijfpuntsschaal, waarbij de intervallen worden begrensd door de standaarddeviaties in de normaalverdeling (Figuur 2, p. 171). Dit model sluit echter minder goed aan bij de gepresenteerde praktijkmodellen. Het wekt dan ook geen verwondering dat er in de consensusbijeenvakkomsten relatief snel een voorkeur ontstond voor een categorisering volgens Wechsler's zevenpuntsschaal.

De beschrijving van de kwalitatieve labels leidde echter tot aanzienlijk meer debat, met name over de twee intervallen onder en boven de classificatie 'gemiddeld' (standaardscores 90-109). Na ampel beraad is overeengekomen dat de zeven intervallen van hoog naar laag worden omschreven als: 'exceptionally high' (> 98 percentiel), 'above average' (percentiel 91-97), 'high average' (percentiel 75-90), 'average' (percentiel 25-74), 'low average' (percentiel 9-24), 'below average' (percentiel 2-8), 'exceptionally low' (percentiel < 2) (Guilmette e.a., 2020).

Ondanks de hierboven beschreven onduidelijkheid in de huidige Nederlandse praktijk kunnen we vaststellen dat de verschillende kwantitatieve indelingen die in de Nederlandse literatuur beschreven zijn (zoals die van Bouma e.a., 2012 en Ruiter e.a., 2017), de meeste overeenkomst vertonen met de zevenpuntsschaal van Wechsler (Figuur 2). De indeling in zeven classificatiecategorieën sluit dus ook hier het beste aan bij de gangbare praktijk. De door Guilmette e.a. (2020) voorgestelde kwalitatieve beschrijvingen van de intervallen daarentegen zullen mogelijk in



FIGUUR 1 Verdeling van standaardscores volgens de normaalverdeling en de verschillende Noord-Amerikaanse kwalitatieve beschrijvingen



FIGUUR 2 Verdeling van standaardscores volgens de normaalverdeling en de verschillende kwalitatieve beschrijvingen binnen het Nederlandse taalgebied

Nederland wel discussie oproepen. Dit geldt vooral voor de kwalificaties 'below average' en 'above average', waarvoor na een uitvoerige discussie binnen de AACN-werkgroep uiteindelijk is gekozen. Gezien de overeenstemming met de kwantitatieve indeling van Bouma en collega's (2012) en Ruiter e.a. (2017) kan tevens gekozen worden voor hun kwalitatieve beschrijving (respectievelijk 'laag' en 'hoog'). Ook in de consensusbijeenkomst is in eerste instantie voorgesteld te kiezen voor 'low' en 'high' in plaats van 'below average' en 'above average'. Om twee redenen heeft men uiteindelijk toch niet voor 'laag' en 'hoog' gekozen. Op de eerste plaats vond men de connotatie van de termen 'low' en 'high' meer veroordelend, terwijl het vermijden hiervan een van de uitgangspunten van de conferentie was. Ten tweede beschouwde men 'laag' en 'hoog' als meer ambigu dan 'beneden- en bovengemiddeld'. Wij hebben zoals aangegeven niet de intentie hierover discussie te voorkomen, maar stellen wel voor de kwalitatieve beschrijving in de Nederlandse praktijk af te stemmen op de weloverwogen overeenstemming binnen de AACN. Dit betekent dat het de voorkeur heeft de classificatie 'below average' en 'above average' in het voorstel van Guilmette e.a. (2020) in Nederland als 'benedengemiddeld' en 'bovengemiddeld' te beschrijven (zie Figuur 2). Gevolg hiervan is wel dat het label 'benedengemiddeld' in het nieuwe voorstel een andere betekenis krijgt dan het heeft in de gangbare beschrijving van intelligentiescores in Nederland (Guilmette e.a., 2020; Wechsler, 2012, zie Figuur 2). Het risico op inconsistent gebruik van terminologie wordt echter voorkomen, als men er ook in Nederland voor kiest om, conform het consensusadvies, gebruik te maken van het voorgestelde classificatiemodel in plaats van prestatiebeschrijvingen zoals die worden voorgesteld in handleidingen van afzonderlijke tests (waaronder de WAIS-IV-NL). Dit zal de uniformiteit in de rapportage vergroten en daarmee zal de communicatie van het neuropsychologisch onderzoek helderder worden.

Een mogelijk discussiepunt waarover in Guilmette e.a. (2020) niet wordt gerapporteerd, is dat ondanks de beschreven kritiek op het gebruik van 'laag' en 'hoog', er toch wordt voorgesteld een percentiel van 2 of lager als 'uitzonderlijk laag' en percentielen van 98 en hoger als 'uitzonderlijk hoog' te beschrijven.

### *Niet-normaalverdeelde scores*

Sommige functies (zoals het kennen van het lichaamsschema of het links-rechtsonderscheid), of testprestaties (zoals een kloktekening, cognitieve screeners of prestatievaliditeitstest) zijn echter niet-normaal verdeeld, omdat ze als doel hebben een afwijkende prestatie in dichotome zin te meten (bijvoorbeeld wel of niet onderpresteren, wel of niet cognitief aan-



gedaan). In de consensusbijeenkomst zijn drie typen tests tegen het licht gehouden: tests met een sterk scheve verdeling, tests die de aanwezigheid of afwezigheid van pathognomische kenmerken onderzoeken, en prestatievaliditeitstests. Bij deze taken en tests wordt het gebruik van percentielen geadviseerd in plaats van standaardcores en wordt aangeraden om hogere scores te beschrijven als vallend 'binnen normale grenzen' (Figuur 1). Bij taken die de validiteit van testprestaties moeten vaststellen, zoals de Amsterdams Korte Termijn Geheugentest (Schmand e.a., 1999), wordt een beschrijving volgens een driepuntsschaal geadviseerd; scores vallen binnen het 'valide', 'onbepaald valide', of 'invalide bereik'. Deze adviezen zijn naar onze mening – voor zover niet al gehanteerd – gemakkelijk te incorporeren in de Nederlandse praktijk.

#### *Definiëring en hantering van het begrip 'stoornis'*

Bij de beoordeling van afwijkende of 'gestoorde' (*impaired* in het Engels) testprestaties worden (inter)nationaal uiteenlopende criteria gebruikt. Zo worden er verschillende afkappunten gehanteerd, variërend van 1, 1,5 of 2 standaarddeviaties lager (en hoger) dan het gemiddelde (Guilmette e.a., 2020; Schmand e.a., 2014). Guilmette e.a. (2020) waarschuwen hierbij voor een te concreet en te zeer aan de testscore gebonden klinisch oordeel als onderdeel van de interpretatie; een enkele testscore heeft immers een zeer geringe klinische betekenis (in de zin van zowel de betrouwbaarheid als de validiteit), als deze niet in het totale scoreprofiel en in de context van de patiënt wordt geïnterpreteerd. Zij omschrijven dit als een 'onacceptabele methode om tot klinische conclusies te komen' (p. 440). Er wordt dan ook terecht een fundamenteel onderscheid gemaakt tussen het beschrijven van testresultaten en de interpretatie ervan; scores zijn immers niet gestoord, maar cognitieve functies. Waar (de beschrijving van) scores nog gevat kunnen worden in een uniform classificatiesysteem, vraagt de vaststelling van eventuele achteruitgang of van stoornissen in (cognitief) functioneren om een professioneel, individueel hypothese toetsend diagnostisch besluitvormingsproces. Testprestaties worden daarbij niet alleen afgezet tegen normgroepen (interpersoonlijke vergelijking), maar krijgen juist betekenis tegen de achtergrond van contextuele factoren zoals gerapporteerde en geobserveerde klachten en problemen en hun ontstaansgeschiedenis, het patroon van testprestaties op verschillende tests, verwachtingen van prestaties op grond van (schattingen van) het premorbide functioneren en demografische factoren (intrapersoonlijke vergelijking), in combinatie met theoretische en empirische kennis over hersen-gedragrelaties. Tegen deze achtergrond beschrijft het consensusdocument het begrip stoornis (*impairment*) als:

Neuropsychological impairment is abnormal neurocognitive or neurobehavioral capacity. Impairment may result from loss of previously acquired skill or result from atypical development, may be transient or fixed across time, and can have variable impact on functional capacity and disability. Test scores, per se, do not define impairment. A combination of factors, including test scores that deviate from expectations, and other findings related to functional capacity, identify neuropsychological impairment. (Guilmette e.a., 2020, pp. 449-450)

Vervolgens expliciteren zij verschillende factoren die hierbij in ogen-schouw genomen moeten worden, waaronder gangbare intra-individuele variabiliteit in testprestaties en het vóórkomen van lage testscores in de gewone bevolking. Een criterium voor het beoordelen van testprestaties als ‘lager dan verwacht’ wordt hierbij niet beschreven. De praktijk waarin hiervoor – veelal impliciet – uiteenlopende afkappunten worden gehanteerd, blijft daarmee in stand. In dit verband zouden wij ook stelling willen nemen tegen het gebruik van de beschrijvingen (*cognitief tekort* en (*cognitieve*) *beperking* voor neuropsychologische testprestaties los van de context (dus enkel gerelateerd aan de standaardscores. Een ‘cognitief tekort’ wordt dan gezien als een prestatie van meer dan 1 standaarddeviatie onder het normgemiddelde, maar niet meer dan 2 SD); hier is door Guilmette e.a. (Figuur 2) nu juist de kwalificatie ‘benedengemiddeld’ of ‘laag-gemiddeld’ voor voorzien. De term *tekort* dient volgens ons enkel gebruikt te worden in het kader van een profielanalyse (dus met inachtneming van de context), bijvoorbeeld met als doel tot een sterkte-zwakteanalyse te komen. Een zwakte in een dergelijk profiel zou als een *tekort* gezien kunnen worden (hoewel het gebruik van een dergelijke term ook een waardeoordeel suggereert, denk aan de werkwoorden *tekortkomen* of *tekortschieten*). Bovendien kent de term (*cognitief*) *tekort* ook geen equivalent in de internationale (Engelstalige) literatuur. Van een beperking daarentegen is pas sprake wanneer een cognitieve *stoornis* leidt tot problemen in het dagelijks leven op het niveau van dagelijkse activiteiten (conform het ICF-model, World Health Organization, 2001). Dit is nooit enkel en alleen op basis van afkapwaarden op cognitieve tests vast te stellen, maar vereist altijd informatie over het dagelijks leven van de patiënt. De term *cognitieve beperking* is derhalve een contradictio in terminis.

Deze kritische kanttekening in acht nemend, sluiten de door de AACN-werkgroep opgestelde aanbevelingen grotendeels aan bij en zijn ze toepasbaar in Nederland.

## Tot slot

Dit Consensus Conference Statement biedt een kans om internationaal en ook in Nederland te komen tot een meer eenduidige wijze van rapporteren en interpreteren van cognitieve testprestaties (maar ook tot niet-cognitieve functie- en vaardigheidstests, of zelfs tests naar het persoonlijkheidsfunctioneren), op grond van 'expert-gebaseerde richtlijnen' of 'best practice'. Het wijst opnieuw op een bekend probleem in onze beroepsuitoefening dat klaarblijkelijk in de beroepsgroep niet eenvoudig was op te lossen, terwijl zowel daarbinnen als daarbuiten de behoefte aan uniformiteit en een heldere communicatie groot is. We nemen dan ook aan dat eenieder het belang van een dergelijk eenduidig systeem volmondig onderschrijft. We hebben geconcludeerd dat de meeste aanbevelingen goed aansluiten bij de Nederlandse praktijk, of tamelijk eenvoudig te incorporeren zijn. Wij kunnen behalve een aantal kleine verschillen in de labeling van de classificaties geen onoverkomelijke knelpunten signaleren. Vanuit het onmiskenbaar overstijgende belang van een (inter)nationaal uniform gehanteerd classificatiesysteem, pleiten wij er – in navolging van onze Amerikaanse collegae – dan ook voor om het beschreven classificatiesysteem ook in Nederland te gebruiken in plaats van classificaties uit verschillende testhandleidingen. Tevens wordt aanbevolen om in (een bijlage van de) rapportages een tabel of figuur op te nemen om toe te lichten welke betekenis de gehanteerde kwalitatieve beschrijvingen hebben. Zo kunnen we aan zee, in het park, in de supermarkt, of bij de kapper 'voldoende' interpersoonlijke afstand houden in coronatijd, in de wetenschap dat hieronder in Nederland iets anders wordt verstaan dan in Australië (anderhalve, respectievelijk twee meter afstand).

Tot slot willen wij benadrukken dat het in Figuur 1 gepresenteerde voorstel niet als richtlijn is bedoeld, maar als startpunt voor een discussie over een dergelijke indeling, met als doel consensus te bereiken. De auteurs zijn hiervoor uiteraard bereikbaar.

**Marc P.H. Hendriks** Radboud Universiteit, Nijmegen; Academisch Centrum voor Epileptologie, Kempenhaeghe, Heeze, e-mail: m.hendriks@donders.ru.nl.

**Bart A.W. Mol** Radboud Universiteit, Nijmegen; klinisch neuropsycholoog/klinisch psycholoog aan de vrijgevestigde praktijk Cerno, Nijmegen en aan Tactus Verslavingszorg, Zutphen.

**Roy P.C. Kessels** Radboud Universiteit, Nijmegen; klinisch neuropsycholoog aan het Radboudumc, Nijmegen en aan Vincent van Gogh voor Geestelijke Gezondheidszorg, Venray.

## Literatuur

- Bouma, A., Mulder, J., Lindeboom, J. & Schmand, B. (2012). *Handboek neuropsychologische diagnostiek* (2e dr.). Amsterdam: Pearson Assessment and Information B.V.
- Groth-Marnat, G. (2009). *Handbook of psychological assessment* (5th ed.). Hoboken, New Jersey: John Wiley & Sons, Inc.
- Guilmette, T.J., Sweet, J.J., Hebben, N., Koltai, D., Mahone, E.M., Spiegler, B.J., Stucky, K., Westerveld, M. & Conference participants (2020). American Academy of Clinical Neuropsychology consensus conference statement on uniform labeling of performance test scores. *The Clinical Neuropsychologist*, 34, 437-453.
- Hendriks, M., Kessels, R., Gorissen, M., Schmand, B. & Duits, A. (2014). *Neuropsychologische diagnostiek: De klinische praktijk*. Amsterdam: Uitgeverij Boom.
- Resing, W.C.M. (2015). Praktische aspecten van intelligentiemeting: Afname, scoring en interpretatie. In: W.C.M. Resing (red.), *Handboek intelligentietheorie en testgebruik* (pp. 41-62). Amsterdam: Pearson.
- Resing, W.C.M. & Blok, J.B. (2002). De classificatie van intelligentiescores: Voorstel voor een eenduidig systeem. *De Psycholoog*, 37, 244-249.
- Ruiter, S.A.J., Hurks, P.P.M. & Timmerman, M.E. (2017). IQ-score is dringend aan modernisering toe: Naar een nieuwe interpretatie en classificatie van de geschatte intelligentie. *Kind en Adolescent Praktijk*, 16(1), 16-23.
- Schmand, B., Hendriks, M., Kessels, R., Gorissen, M. & Duits, A. (2014). Interpretatie. In: M. Hendriks, R. Kessels, M. Gorissen, B. Schmand & A. Duits (red.), *Neuropsychologische diagnostiek: De klinische praktijk* (pp. 209-232). Amsterdam: Uitgeverij Boom.
- Schmand, B., De Sterke, S. & Lindeboom, J. (1999). *Amsterdamse Korte Termijn Geheugentest, AKTG*. Lisse: Swets & Zeitlinger.
- Schoenberg, M.R. & Rum, R.S. (2017). Towards reporting standards for neuropsychological study results: A proposal to minimize communication errors with standardized qualitative descriptors for normalized test scores. *Clinical Neurology and Neurosurgery*, 162, 72-79. doi:10.1016/j.clineuro.2017.07.010.
- Schretlen, D.J., Testa, S.M. & Pearlson, G.D. (2010). *Calibrated Neuropsychological Normative System professional manual*. Lutz, FL: Psychological Assessment Resources.
- Wechsler, D. (2012). *WAIS-IV-NL: Wechsler Adult Intelligence Scale – Fourth Edition –Nederlandstalige bewerking: Technische handleiding en Afname en scoringshandleiding*. Amsterdam: Pearson Assessment and Information.
- World Health Organization (2001). *The International Classification of Functioning, Disability and Health: ICF*. Genève: WHO.