

FORUM

VAN RONKENDE RECLAME TOT ROMMELIGE RAMSJ

Baukje de Koning, Selma Ruiter en Linda van Zonneveld publiceerden in december 2019 in dit blad een artikel over de Intelligence and Development Scales-2 (IDS) en het onderliggende CHC-model. Volgens Marcel Veenman zetten De Koning c.s. met veel aplomb profilering op basis van het CHC-model in de etalage. ‘Gedegen argumentatie en empirische onderbouwing voor de validiteit en bruikbaarheid van die profilering ontbreken echter. Daarmee zetten zij (school-)psychologen vooralsnog op het verkeerde been.’ De Koning c.s. reageren op deze kritiek.

Laat ik maar met de deur in huis vallen. Het artikel van De Koning, Ruiter en van Zonneveld (2019) over de IDS-2 test en

het CHC-model als state-of-the-art theorie voor het meten van intelligentie, munt uit in oppervlakkigheid. De auteurs stellen dat de subschalen van de Wechsler-tests aan kritiek onderhevig zijn

vanuit de wetenschap en diagnostische praktijk.

Nu is er niets op tegen om een kritische blik te werpen op deze wereldwijd gebruikte intelligentietest, maar het terzijde schuiven

van de Wechsler-schalen vraagt dan wel om meer uitleg dan alleen ‘a-theoretisch en daarom arbitraire metingen’ (De Koning c.s., p. 39). Wat is er dan theoretisch mis met het onderscheid tussen perfoormaal en verbaal IQ en waarom kan de diagnostische praktijk niet uit de voeten met dit onderscheid? Hoewel ik mij bij de antwoorden op deze vragen wel iets kan voorstellen (zie Watkins, 2000), blijft de grond van deze kritiek ongewis in het artikel van De Koning et al. (2019). Evenmin wordt beargumenteerd waarom het alternatieve CHC-model dan wél een brug kan slaan tussen wetenschap en de diagnostische praktijk. Ook ontbeert het artikel een bespreking van de substantiële kritiek op het CHC-model, waarin niet alleen de theoretische onderbouwing, maar ook de validiteit van tests en profielscores gebaseerd op het CHC-model worden gehekelde (onder meer de auteurs van een special issue in 2019 van *Applied Measurement in Education*, 32(3)). Sommigen betogen zelfs dat het CHC-model vooral wordt gebruikt als vehikel voor de commerciële uitbating van intelligentietests (Canivez & Youngstrom, 2019).

WAT HAD ER MOETEN WORDEN VERTELD?

Het CHC-model is niet plotsklaps uit de lucht komen vallen. Een korte geschiedschrijving van het intelligentie-onderzoek is hier onontbeerlijk, te beginnen bij Thurstone’s *primary mental abilities* als voorloper van het tweede stratum in het CHC-model versus Spearman’s ‘G’ als voorloper van

het derde stratum in het CHC-model van algemene intelligentie (Spearman, 1923; Thurstone, 1938). Thurstone en Spearman bestreden elkaar in eerste instantie te vuur en te zwaard over het bestaan van G. In 1945 moest Thurstone toch toegeven dat secundaire factoranalyse op zijn primaire factoren wel degelijk een algemene factor opleverde (Veenman, 2005). Dit debat tussen opvattingen van algemene intelligentie (G) versus intelligentie als meervoudige factoren wordt tot op de dag van vandaag gevoerd.

Cattell (1963), de eerste ‘C’ in CHC, maakte het onderscheid tussen *fluid* en *crystallized intelligence* (Gf en Gc), dat wil zeggen tussen analytisch redeneren en kennis-gebaseerde intelligentie. Cattell’s categorieën zijn terug te vinden in het tweede stratum van het CHC-model. Horn, de ‘H’ in CHC, was een aanhanger van Cattell betreffende de *fluid-crystallized*-dimensie (Horn & Cattell, 1966), maar hij wees het bestaan van G categorisch af.

Voorgaande zijn relevante pogingen geweest om structuur in de meting van intelligentie aan te brengen door middel van factoranalyse op testdata.

Uiteindelijk zette John B. Carroll (1993), de tweede C uit CHC, een kers op de taart met de publicatie van *Human cognitive abilities: A survey of factor analytic studies*. In dit boek, een turf van zo’n 800 pagina’s, veegt Carroll het factor-analytisch onderzoek van 461 studies naar intelligentie bijeen in een majeure her-analyse. Dat resulteert in zijn ‘Three Stratum Theory of Cognitive

Abilities’, zoals weergegeven in Figuur 1 (De Koning et al., 2019, p. 41). Ere wie ere toekomt! Ik vermoed dat De Koning c.s. hebben nagelaten kennis te nemen van Carroll, gezien de secundaire verwijzing naar Resing en het ontbreken van een literatuurverwijzing naar Carroll (1993) in de literatuurlijst. Onduidelijk blijft voor mij uit het artikel van De Koning et al. wat het CHC-model nu eigenlijk toevoegt aan de drie-strata theorie van Carroll, behalve de betwistbare belofte van profilering.

PROFILERING VAN COGNITIEVE VAARDIGHEDEN

Carroll (1988) stelt dat factoren uit een factoranalyse statistische abstracties zijn die niet één-op-één kunnen worden gelijkgesteld aan cognitieve processen. Een factor beschrijft uitsluitend de gemeenschappelijke variantie die subtests delen, waarbij subtests doorgaans op meer dan één factor laden (Veenman, 2005). Met een psychologische interpretatie van dergelijke psychometrische modellen dient men echter terughoudend te zijn. Carroll (1988) beschouwde factoren hooguit als benaderingen (*approximations*) van cognitieve processen, die verdere onderbouwing behoeven door experimenteel psychologisch onderzoek. In weerwil van zulk onderzoek in de laatste decennia van de vorige eeuw (bijv. Rowe, 1991; Sternberg, 1982), zijn psychometrische intelligentie en informatieverwerkingsprocessen nog altijd verschillende takken van sport (Lohman, 1989; Neiser et al., 1996).

Profilering van cognitieve vaardigheden is gebaseerd op de veronderstelling dat psychometrische factoren bestaande psychologische entiteiten representeren. Daar valt wat op af te dingen. Thurstone probeerde destijds al profilering op basis van zijn primaire factoren empirisch te onderbouwen, maar zelfs in zijn selectieve steekproef trof hij bij 90% nauwelijks profielverschillen aan (Veenman, 2005). Volgens Watkins (2000) laten profielscores van de wisc en andere intelligentietests een gebrek aan betrouwbaarheid, validiteit en diagnostische bruikbaarheid zien. Profielscores blijken nauwelijks of geen toegevoegde waarde te hebben ten opzicht van overall IQ-scores. Desondanks blijft profilering onverminderd populair onder schoolpsychologen (Watkins, 2000). In het CHC-model wordt profilering op factorniveau van het tweede stratum toegepast. Deze profielscores zouden een 'nauwkeuriger beeld' geven van algemene intelligentie en van de sterktes en zwaktes van het kind op brede cognitieve deelvaardigheden (De Koning et al., 2019, p. 42).

Nu wreekt de aanpak van De Koning c.s. zich, die uitsluitend gericht is op protagonisten van het CHC-model, zoals McGrew en Flanagan. Kritische beschouwingen over het CHC-model in het algemeen en profilering in het bijzonder worden hier node gemist. Die kritiek betreft de onverenigbaarheid van de Horn-Cattell-theorie met de drie strata-theorie van Carroll (Canivez & Youngstrom, 2019), de onvolkomenheid van tests op basis van het

CHC-model (McGill & Dombrowski, 2019; Wasserman, 2019), de gebrekkige replicatie van factorstructuren uit het CHC-model (Wasserman, 2019) en de geringe validiteit en bruikbaarheid van profilering (McGill, Dombrowski & Canivez, 2018; McGill et al., 2016; Styck, Beaujean, & Watkins, 2019; Wasserman, 2019). Met veel aplomb zetten De Koning c.s. profi-

lering op basis van het CHC-model in de etalage. Gedegen argumentatie en empirische onderbouwing voor de validiteit en bruikbaarheid van die profilering ontbreken echter. Daarmee zetten zij (school-)psychologen vooralsnog op het verkeerde been, want Carroll's drie-strata-theorie of het CHC-model lenen zich niet klakkeloos voor cognitieve profilering met



Sommigen betogen dat het CHC-model vooral wordt gebruikt als vehikel voor de commerciële uitbating van intelligentietests

diagnostische doeleinden. Dat hoor je de uitgevers van CHC-gebaseerde tests niet zeggen, want dat

verkoopt niet lekker (Canivez & Youngstrom, 2019).

DE TOETERS EN BELLEN VAN DE IDS-2

De IDS-2 beoogt meer dan alleen intelligentie te meten. Maar laten we eerst het 'functie-bereik intelligentie' onder de loep nemen, dat zes van de tien CHC-factoren bestrijkt (De Koning c.s., p. 42). Wat direct opvalt, is dat veel subtests ontbreken die doorgaans in een of andere vorm zijn terug te vinden in bestaande intelligentietests (Carroll, 1993). Geen vocabulaire als belangrijke maat voor Gc en Gf (Lohman, 1989)¹. Geen Gf -tests zoals analogieën, (cijfer)reeksen afmaken, of lineaire syllogismen. En ook geen reken-snelheid. Bovendien zijn een aantal van de overige CHC-factoren ondergebracht in de aparte categorie van schoolse vaardigheden van de IDS-2, maar die ontbreken vreemd genoeg weer in de berekening van algemene intelligentie (De Koning c.s., p. 42). Onduidelijk blijft waarom.

Verder blijken subtests van het functie-bereik intelligentie relatief vaak een beroep te doen op de bewerking van figuren (50-64%). Wellicht is dat gedaan om het cultuur-fair-gehalte van de intelligentiemeting op te krikken (De Koning c.s., pp. 46-47), maar daarmee wordt wel een bias naar visuele verwerking geïntroduceerd. Tot slot is het navrant dat bijna de helft van de subtests om een relatief passieve vorm van respons vraagt zoals naleggen, nazeggen of doorstrepen, waarmee

¹ 'Tegenstellingen noemen' (subtest 14 van de IDS-2) is niet hetzelfde als vocabulaire of woordbetekenissen (zie Lohman, 1989).

De gymnastiekoefeningen van de grove motoriek en het moeren draaien en kralenrijen van de fijne motoriek doen tamelijk archaisch en ridicuul aan

de IDS-2 dus minder een beroep op actieve productie doet.

Concluderend zijn er genoeg redenen voor twijfel over de huidige configuratie van subtests in de IDS-2 als betere afspiegeling van het intellectuele repertoire dan reeds bestaande intelligentietests. Voor de diagnostische praktijk is het aanbevelenswaardig om eerst verder onderzoek naar de validiteit van de IDS-2 af te wachten en tot nader order gebruik te blijven maken van de WISC, GIT2, SON-R en RAKIT.

De IDS-2 heeft een aparte schaal voor executief functioneren of 'zelfsturend gedrag' (De Koning c.s., p. 46). Het artikel maakt niet duidelijk op welke executieve functies de vier subtests zijn gebaseerd, namelijk werkgeheugen, planning, inhibitie, of mentale flexibiliteit (Veenman, 2017)². Ongewis blijft ook hoe executieve functies zich verhouden tot algemene intelligentie. Er

is voldoende evidentie dat de relatie tussen intelligentie, zelfsturend gedrag en leerprestaties complexer is dan nu wordt gesuggereerd (Veenman, 2017). Tot slot dient men de methodologische vraag te stellen of zelfsturend gedrag wel adequaat kan worden gemeten met een beperkte set aan tests en taakjes (Veenman & van Cleef, 2019).

Het ontwikkelingsdomein van de IDS-2 beslaat vier schalen: psycho-motoriek, sociaal-emotionele vaardigheden, werkhouding en de eerder genoemde schoolse vaardigheden. Behalve schoolvordering, staan die schalen geheel los van het CHC-model. Hoewel ik me bij de relevantie van visuele motoriek voor het functioneren op school nog iets kan voorstellen, doen de gymnastiekoefeningen van de grove motoriek en het moeren draaien en kralenrijen van de fijne motoriek tamelijk archaisch en ridicuul aan. Zo testte men meer dan een halve eeuw geleden fabrieksarbeiders. Bij sociaal-emotionele vaardigheden dient een kanttekening te worden gemaakt dat iets herkennen, aangeven of benoemen niet

hetzelfde is als het daadwerkelijk kunnen reguleren of handelen (Veenman & van Cleef, 2019). Dit is in essentie het verschil tussen declaratieve kennis ('kennen') en procedurele kennis ('kunnen'). Kennen is een vereiste, maar niet voldoende voorwaarde voor kunnen (Miller, 1990; Veenman, 2017). Het is dus maar de vraag of de drie tests daadwerkelijk sociaal-emotionele vaardigheden meten.

Tot slot is het oordeel over de werkhouding gebaseerd op de observaties van de testleider. Natuurlijk dient een testleider oog te hebben voor de testbereidheid van het kind, maar voor oordelen over consciëntieusheid zou ik eerder vertrouwen op bestaande tests zoals de DAPP of SPFT. Voor prestatie-motivatie is de PMT-K het aangewezen instrument. Wat onduidelijk blijft in het artikel van De Koning c.s. is hoe de informatie uit deze zes verschillende functie-bereiken moet worden gecombineerd en gewogen. Zelfs de beknopte case-study (De Koning c.s., pp. 47-49) biedt daarvoor nauwelijks aanknopingspunten.

De Koning c.s. (p. 46) noemen het een 'uniek' voordeel dat de IDS-2 een breed onderzoek naar verschillende functies combineert in één test. Je hoeft dan geen afzonderlijke meetinstrumenten aan te schaffen: meer waar voor de prijs van één! Dat commerciële argument heeft echter een keerzijde. De Koning c.s. (p. 43) stellen weliswaar dat niet de gehele IDS-2 hoeft te worden afgenomen bij een specifieke vraagstelling, toch kan er een tendens ontstaan om maar een zo

breed mogelijke screening toe te passen. Meer informatie betekent echter niet altijd meer betekenisvolle informatie. Brede screening draagt het risico van overdiagnostisering in zich. Het lokt een hagelschotbenadering uit, waarbij er altijd wel een vogel uit de lucht valt. Ieder kind krijgt zo zijn eigen labeltje. Overdiagnostisering kan zowel optreden bij analyses op subtestniveau en profielniveau, als bij analyses over de zes functie-bereiken.

PSYCHOMETRISCHE VALIDITEIT

In weerwil van afname bij een normeringssteekproef van 1646

kinderen (De Koning c.s., p. 43), ontbreken allerlei psychometrische gegevens over de IDS-2. Het aantal items per subtest wordt niet vermeld. Indices voor interne consistentie of betrouwbaarheid van de subtests ontbreken. Er wordt niet onderzocht of de data van de IDS-2 een replicatie van de CHC-factorstructuur opleveren, evenmin als de convergente validiteit met andere intelligentietests. Psychometrisch blijft het oorverdovend stil. Dat maakt dit artikel tot een reclamefolder voor Hogrefe Uitgevers met de IDS-2 als schamel uitverkoopje. De IDS-2-profielen op intelligentieniveau én op functie-bereik vormen

(vooralsnog) veel te dun ijs om een behandelingsplan voor interventie of therapie op te baseren.

In memoriam Jan J. Elshout
(† 19 december 2019).

OVER DE AUTEUR

Dr. Marcel V.J. Veenman is eigenaar van het Instituut voor Metacognitie Onderzoek. Ook is hij lid van de redactieraad van *De Psycholoog*. E-mail: mvjveenman@kpnmail.nl.

Literatuur

- Canivez, G.L. & Youngstrom, E.A. (2019). Challenges to the Cattell-Horn-Carroll theory: Empirical, clinical, and political implications. *Applied Measurement in Education*, 32, 232-248.
- Carroll, J. B. (1988). Cognitive abilities, factors, and processes. *Intelligence*, 12, 101-109.
- Carroll, J. B. (1993). *Human cognitive abilities. A survey of factor-analytic studies*. Cambridge: Cambridge Univ. Press.
- Cattell, R. B. (1963). Theory of fluid and crystallized intelligence: A critical experiment. *Journal of Educational Psychology*, 54, 1-22.
- Horn, J.L. & Cattell, R. B. (1966). Refinement and test of the theory of fluid and crystallized intelligence. *Journal of Educational Psychology*, 57, 253-270.
- Koning, B. de, Ruiter, S. & Zonneveld, L. van (2019). De Intelligentie and De ontwikkeling Scales-2. *De Psycholoog*, 54(12), 38-50.
- Lohman, D.F. (1989). Human intelligenc: An introduction to advances in theory and research. *Review of Educational Research*, 59, 333-373.
- McGill, R. J. & Dombrowski, S.C. (2019). Critically reflecting on the origins, evolution, and impact of the Cattell-Horn-Carroll (CHC) model. *Applied Measurement in Education*, 32, 216-231.
- McGill, R. J., Dombrowski, S.C. & Canivez, G.L. (2018). Cognitive profile analysis in school psychology: History, issues, and continued concerns. *Journal of School Psychology*, 71, 108-121.
- McGill, R. J., Styck, K. M., Palomares, R.S. & Hass, M.R. (2016). Critical issues in specific learning disability identification: What we need to know about the PSW model. *Learning Disability Quarterly*, 39, 159-170.
- Miller, G.E. (1990). The assessment of clinical skills/competence/performance. *Academic Medicine*, 65(9), S63-S67.
- Neisser, U. et al. (1996). Intelligence: Knowns and unknowns. *American Psychologist*, 51, 77-101.
- Rowe, H.A.H. (1991). *Intelligence: Reconceptualization and measurement*. Hillsdale, NJ: Erlbaum.
- Spearman, C. (1923). *The nature of intelligence and the principles of cognition*. London: MacMillan.
- Sternberg, R.J. (1982). *Handbook of human intelligence*. Cambridge: Cambridge University Press.
- Styck, K.M., Beaujean, A.A. & Watkins, M.W. (2019). Profile reliabilities of cognitive ability subscores in a referred sample. *Archives of Scientific Psychology*, 7, 119-127.
- Thurstone, L.L. (1938). *Primary mental abilities*. Chicago: University of Chicago Press.
- Veenman, M.V.J. (2005). Thurstone's scales of primary abilities. In K. Kempf-Leonard (Ed.), *Encyclopedia of social measurement*. Volume 3 (pp. 811-815). San Diego: Academic Press.
- Veenman, M.V.J. (2017). Learning to self-monitor and self-regulate. In R. Mayer & P. Alexander (Eds.), *Handbook of research on learning and instruction*, 2nd revised edition (pp. 233-257). New York: Routledge.
- Veenman, M.V.J. & Cleef, D. van (2019). Measuring metacognitive skills for mathematics: Students' self-reports vs. on-line assessment methods. *ZDM International Journal on Mathematics Education*, 51, 691-701.
- Wasserman, J. D. (2019). Deconstructing CHC. *Applied Measurement in Education*, 32, 249-268.
- Watkins, M. W. (2000). Cognitive profile analysis: A shared professional myth. *School Psychology Quarterly*, 15, 465-479.

² Executieve functies worden twee keer genoemd in de beknopte case-study (De Koning c.s., pp. 47-48), maar dat blijkt onvoldoende en bovendien inconsistente informatie te zijn.

REACTIE

STIJLLOZE SOFISMEN

De reactie van Marcel Veenman op ons artikel over de IDS-2 begint quasistijlvol, met een rollende alliteratie in de titel. Al gauw blijkt dat het stuk vol staat met kritieken die aanmatigend zijn of zelfs onjuist en geen opening bieden voor een professionele discussie.

In een tijdschrift als *De Psycholoog* is het uiten van gefundeerde kritiek zeer gewenst. Dit levert een discussie op met als doel te komen tot een hoger niveau van kennis en begrip. De wijze waarop de kritiek verkondigd wordt, is belangrijk om te komen tot een respectvolle, open discussie. Een allereerste voorwaarde is begrip van het doel van het artikel en de essentie van de boodschap. Schadelijk voor de professionele discussie is anderen de maat nemen en zelf op vele manieren uit de bocht vliegen. We zullen ons beperken tot een korte inhoudelijke reactie op enkele punten die Veenman bespreekt in zijn stuk.

In tegenstelling tot wat Veenman suggereert, pretenderen

wij in geen enkel opzicht dat we met dit artikel het CHC-model volledig en met alle mitsen en maren hebben besproken. Het doel van het artikel is het informeren van professionals over een nieuw instrument voor Nederland op het gebied van algemene intelligentie- en ontwikkelingsdiagnostiek. Wij beschouwen dit instrument als een goede aanvulling op de al bestaande intelligentietests. Dit idee hebben we enthousiast, dat zeker, maar ook met professionele distantie toegelicht. Dat wellicht hier en daar ons enthousiasme wat te veel doorklinkt, nemen we ter harte.

Veenman concludeert dat wij de Wechsler-tests opzij schuiven als a-theoretisch en daarom als arbitraire metingen van intelligentie. Veenman noemt er zelfs een paginanummer bij waar dat beschreven zou staan. Wij kunnen deze passage gelukkig niet vinden en kunnen ons ook geheel niet vinden in deze bewering. Wat wij hier wel doen, is een idee beschrij-

ven dat heeft postgevat in de literatuur. Dat idee betreft vanzelfsprekend niet specifieke tests in het bijzonder, maar testresultaten van intelligentietests in het algemeen.

Volgens Veenman was een korte geschiedschrijving over het CHC-model op zijn plaats geweest. Hij doet meteen een voorzet van wat volgens hem zeker nog verteld had moeten worden over de ontwikkeling tot wat we nu kennen als het CHC-model. Gegeven dat de IDS-2 gebaseerd is op het CHC-model is de huidige vorm van belang, en niet de geschiedenis ervan. Hadden we deze meer toegelicht, dan zou de relevantie vooral gelegen hebben in de huidige discussie die gevoerd wordt over het model (zie o.a. Gignac & Watkins, 2015; Beaujean & Benson, 2018; Schneider & McGrew, 2018; Rey-Mermet et al., 2019). Belangrijk in deze discussie is de empirische ondersteuning van het model, zoals dat ook onderliggend is aan de structuur van de IDS-2.

Veenman schrijft dat Carroll's drie-strata-theorie of het CHC-model door testontwikkelaars en -uitgevers klakkeloos geleend worden voor cognitieve profilering met diagnostische doeleinden. Veenman beweert hierbij dat gedegen argumentatie en empirische onderbouwing zouden ontbreken. Veenman doelt

hiermee op alle intelligentietests die het CHC-model gebruiken als basis voor de opbouw en inhoud van de test, waaronder alle algemene intelligentietests in Nederland: de Wechsler-schalen, de RAKIT-2 en de IDS-2. Dit is een verbazingwekkende en aanmatigende bewering, die volledig voorbijgaat aan het gegeven dat wetenschappers deze instrumenten weloverwogen hebben samengesteld op basis van de theorie met ondersteunende uitgebreide psychometrische analyses en validatiestudies. Dit geldt dus voor de Wechsler-schalen, de RAKIT-2 en de IDS-2. De theoretische achtergrond, psychometrische eigenschappen en validatiestudies van de IDS-2 staan beschreven in de handleiding behorend bij het instrument (Grob & Hagmann, 2018; Grieder & Grob, submitted), en samengevat voor professionals in *Kind & Adolescent* (Ruiter, Visser & Timmerman, 2020).

Veenman beweert zonder enige onderbouwing dat profielscores

van nauwelijks of geen waarde zijn. Dat we in de interpretatie de meeste waarde behoren te hechten aan het totaal IQ, is evident. Het in het geheel geen aandacht besteden aan de opbouw van dit totaal IQ zou echter een gemiste kans zijn in het uitvoeren van handelingsgerichte diagnostiek. Goed opgeleide gedragsdeskundigen en schoolpsychologen zijn prima in staat op een genuanceerde en verantwoorde manier de aanvullende informatie die de profielscores bieden te interpreteren. Zij weten als geen ander dat de resultaten uit intelligentieonderzoek geen zuivere maten zijn voor de latente constructen die zij pretenderen te onderzoeken. Zij weten ook als geen ander dat het IQ en onderliggende scoreprofielen nooit de enige maten zijn om een behandelingsplan voor interventie of therapie op te baseren. Daar kan Veenman nog wat van leren.

Ten slotte willen we graag dit nog kwijt. Veenman krijgt de ruimte om in *De Psycholoog* zijn punt te maken. Hij krijgt hiermee

ook de ruimte om collega-professionals, testontwikkelaars en uitgevers, de redactieraad van *De Psycholoog* (waarvan Veenman overigens zelf deel uitmaakt) en de reviewers van het artikel neer te sabelen. Uit zijn stuk blijkt dat Veenman hierbij noch gehinderd is door enige vorm van concrete praktijkgerichte kennis, noch door enige persoonlijke collegiale discussie. Veenman zet een zorgvuldig geconstrueerd en onderzocht instrument weg als een 'schamel uitverkoopje'. Volgens Veenman publiceert dit tijdschrift aldus willens en wetens 'ronkende reclame en rommelige ramsj'.

Het zal je maar gezegd worden.

OVER DE AUTEURS

Dr. Baukje de Koning-Veenstra, dr. Selma de Ruiter en drs. Linda Zonneveld zijn respectievelijk als psycholoog en onderzoeker, als gedragswetenschapper en onderzoeker, en als testleider en projectmedewerker verbonden aan De Kinderacademie in Groningen. E-mail: baukje.veenstra@dekinderacademie.com.

Goed opgeleide gedragsdeskundigen en schoolpsychologen zijn prima in staat op een genuanceerde en verantwoorde manier de aanvullende informatie die de profielscores bieden te interpreteren

Literatuur

Beaujean, A.A. & Benson, N.F. (2018). Theoretically-Consistent Cognitive Ability Test Development and Score Interpretation. *Contemporary School Psychology*, 23, 126-137.

Gignac, G.E. & Watkins, M.W. (2015). There may be nothing special about the association between working memory capacity and fluid intelligence. *Intelligence*, 52, 18-23.

Grieder, S. & Grob, A. (2019, submitted) Exploratory Factor Analyses of the Intelligence and Development Scales-2 (IDS-2). Implications for Theory and Practice.

Grob, A. & Hagmann-von Arx, P. (2018). IDS-2. *Intelligentie- en ontwikkelingsschalen voor kinderen en jongeren. Verantwoording en psychometrie. Nederlandse bewerking door Selma Ruiter, Linda Visser en Marieke Timmerman*. Amsterdam: Hogrefe Uitgevers B.V.

Rey-Mermet, A., Gade, M., Souza, A.S., Bastian, C.C. & Oberauer, K. (2019). Is executive control related to Working Memory Capacity and Fluid Intelligence? *Journal of Experimental Psychology: General*, DOI: 10.1037/xge0000593.

Ruiter, S.A.J., Visser, L. & Timmerman, M.E. (2020). Integratieve diagnostiek met de Intelligence and Development Scales-2 (IDS-2). Een intelligentie- en algemene ontwikkelingstest voor kinderen en jongeren tussen 5 en 20 jaar. *Kind & Adolescent*, 41, 4-30.

Schneider, W.J. & McGrew K., 2018. The Cattell-Horn-Carroll theory of cognitive abilities. In: Flanagan, D. & Harrison P. (Red.), *Contemporary intellectual assessment: theories, tests, and issues*. 4e druk. New York: Guilford, pp. 78-109.

ANTWOORD

VOORKOM BIAS IN MENSELIJK OORDEEL

De reactie van De Koning c.s. wordt het best getypeerd met 'aangebrand'. Tja... *if you can't stand the heat, stay out of the kitchen*. Ik laat de schimpscheuten van De Koning c.s. voor wat ze zijn. Evenals ik de tendentieuze en onjuiste weergave van mijn woorden verder laat rusten. Dus ter zake. In hoeverre reageren De Koning c.s. inhoudelijk op mijn opmerkingen bij hun eerdere artikel?

1. Onvoldoende theoretische onderbouwing van het CHC-model: geen inhoudelijke reactie, want 'die geschiedenis is niet van belang', aldus De Koning c.s. Zo geredeneerd kunnen alle psychologiebibliotheken worden opgedoekt. Leer en waardeer het gedachtengoed van onze illustere voorgangers en zie: *l'histoire se répète*. Overigens ontbreekt ook de wel relevant geachte empirische onderbouwing van het CHC-model in het artikel van De Koning c.s. Dan ben je toch echt op Carroll aangewezen.
2. Problematische profilering: geen 'onderbouwde' reactie. De Koning c.s. betogen dat gedragsdeskundigen en schoolpsychologen heus wel weten wat 'latente constructen' (sic!) zijn. Hiermee gaan De Koning c.s. een fundamentele discussie

over het nut en de validiteit van profilering uit de weg.

3. Twijfels bij de testconstructie van de IDS-2: geen inhoudelijke reactie. Natuurlijk zijn de testmakers met de beste intenties te werk gegaan, maar dat laat onverlet dat er kritisch mag én moet worden gekeken naar het resultaat. Dan volstaat een verwijzing naar de testhandleiding niet.
4. Risico van overdiagnostisering: genegeerd. Het ratjetoe aan subtests in de IDS-2 en kapitalisatie van kans op fouten van de eerste soort blijft onbesproken.
5. Gebrekkige psychometrische verantwoording: weggewuifd. De Koning c.s. denken weg te komen met een verwijzing naar de testhandleiding. Maar zolang je de IDS-2 niet hebt aangeschaft, is de handleiding niet beschikbaar op de website van Hogrefe.

Volgens De Koning c.s. moeten we maar op de blauwe (of bruine) ogen van makers en gebruikers van de test vertrouwen. In het beste geval is dat een naïef standpunt, in het slechtste geval wetenschappelijk onverantwoord. Objectieve testmeting beoogt juist bias in het menselijk oordeel tegen te gaan. Mijn jarenlange ervaring als onderwijspsycholoog (én opleider van onderwijs- en schoolpsychologen) heeft mij geleerd dat gebruikers van intelligentietests veelal onvoldoende kennis hebben van zowel de intelligentietheorie als de testleer en psychometrika. De valkuilen daarvan heb ik getracht te benoemen. Het artikel van De Koning c.s. riep serieuze vragen op die niet worden beantwoord in hun reactie.

I rest my case. De IDS-2, ik zou er (vooralsnog) mijn geld niet op zetten.

