



De toepasbaarheid van meetinstrumenten bij migrantenkinderen en hun ouders

Lenette Rietveld

Nederlands Jeugdinstituut
Infolijn t (030) 230 65 64
e infojeugd@nji.nl
i www.nji.nl

September 2010

Meetinstrumenten zijn doorgaans gebaseerd op theoretische uitgangspunten van westerse oorsprong en ook de betrouwbaarheid en validiteit is vaak alleen onderzocht bij autochtonen. De tests zijn doorgaans niet genormeerd voor migrantenkinderen, zodat de score van het individuele kind alleen vergeleken kan worden met de scores van een autochtone groep. Vergelijkend onderzoek naar scores van migranten en autochtonen ontbreekt vaak of is bij een beperkt aantal instrumenten alleen uitgevoerd voor de grootste migrantengroepen, namelijk Turken en Marokkanen. Onderzoek naar het gebruik van instrumenten bij migranten is met name verricht naar intelligentietesten in het onderwijs of naar de voorspellende waarde van capaciteitstests op het functioneren in een werkomgeving. Over de toepasbaarheid van tests voor sociaal-emotioneel functioneren en persoonlijkheid bij migrantenkinderen is veel minder bekend.

In dit document wordt ingegaan op:

- de vergelijkbaarheid van testcores bij autochtonen en migranten;
- etnocentrisme in een test;
- de vraag of een test onafhankelijk kan of moet zijn van culturele invloeden;
- de gevolgen van verkeerde diagnostiek bij migrantenkinderen;
- mogelijke criteria voor instrumentkeuze;
- aandachtspunten voor de afname en interpretatie van instrumenten;
- meetinstrumenten die onderzocht zijn bij migrantenkinderen.

Dit document is onderdeel van het dossier culturele diversiteit in opgroeien en opvoeden. Het gehele dossier is in te zien op www.nji.nl/diversiteit.

Bias of vertekening

Bij het testen van migrantenkinderen en hun ouders kan er niet zomaar van worden uitgegaan dat de bewezen betrouwbaarheid, validiteit en normering van testen ook voor deze groep geldt. Hiervoor zijn diverse redenen. De belangrijkste reden is bias, ook wel vertekening of partijdigheid genoemd. Scores van tests hebben daardoor een verschillende betekenis voor diverse groepen. Bleichrodt & Van den Berg (2004) onderscheiden verschillende soorten bias:

- **Itembias:** de verschillen in itemscores tussen autochtonen en migranten worden niet veroorzaakt door het kenmerk dat het item beoogt te meten. Een voorbeeld van itempartijdigheid is een verhaalsom om rekenvaardigheid te meten. Bij migranten kan dit item meer taalvaardigheid dan rekenvaardigheid meten. Items kunnen ook voor verschillende culturen verschillende betekenissen hebben. Dit is bijvoorbeeld het geval bij het onderscheiden en waarderen van gedrag van kinderen in de Child Behavior Checklist (CBCL). Het CBCL item 'vermijd anderen aan



te kijken' is in de westerse cultuur klinisch significant, maar kan bijvoorbeeld in een andere cultuur juist een gewenst opvoedingsdoel zijn. Welk gedrag te jong is voor de leeftijd, afwijkend of ongewenst kan verschillen tussen culturen. Het CBCL item 'Doet te jong voor zijn/haar leeftijd' is dus gevoelig voor culturele verschillen.

- Testbias : de betekenis van de test verschilt voor bepaalde groepen. Bij intelligentietests is het de vraag of die inderdaad intelligentie meten of eerder taalvaardigheid of bekendheid met de Nederlandse cultuur. Factoren die de betekenis van een test kunnen beïnvloeden zijn taalvaardigheid, bijvoorbeeld gezegdes en dubbele ontkenningen, sociale wenselijkheid of instemmingstendencie, bijvoorbeeld eerder 'ja' dan 'nee' zeggen op een vraag en de mate van bekendheid met tests, bijvoorbeeld met gehanteerde begrippen of testtactieken (gokken als je het antwoord niet weet).
- Predictiebias of differentiële predictie: de voorspellende waarde van de test voor de praktijk is verschillend voor autochtonen en migranten. Een rekentest op de basisschool (voorspeller) is bijvoorbeeld bedoeld om het wiskundecijfer (criterium) in het voortgezet onderwijs te voorspellen. Op basis van onderzoek onder autochtone kinderen is bijvoorbeeld bepaald dat een score van 10 op een rekentest samenhangt met een gemiddeld wiskundecijfer tussen de 4 en 5. Stel dat migrantenkinderen bij een score van 10 op de rekentest gemiddeld een hoger wiskundecijfer halen dan autochtone kinderen, nl. tussen de 6 en 7. De test levert dan een ondervoorspelling (een te negatieve voorspelling) op voor migrantenkinderen als wordt uitgegaan van het verband voor autochtone kinderen tussen testscore en wiskundecijfer.
- Criteriumbias: ook wat voorspeld moet worden, het criterium, kan partijdig zijn. Als bijvoorbeeld in bovenstaand voorbeeld de betekenis van het wiskundecijfer anders is voor autochtone en migrantenkinderen, dan is het niet mogelijk om hiermee de vergelijkbaarheid van de rekentest voor autochtone en migrantenkinderen te evalueren. Het is dan immers niet duidelijk of een verschil in de voorspellende waarde van de rekentest komt door de partijdigheid van het criterium.

Etnocentrisme

Behalve dat migranten door testen benadeeld of verkeerd beoordeeld kunnen worden, kan er ook sprake zijn van etnocentrisme in de test. Dit is het geval wanneer in de items formuleringen gebruikt worden waarin de eigen cultuur als universeel of superieur wordt afgeschilderd. Dit hoeft niet direct bias te veroorzaken, maar bij de testconstructie wordt er dan geen rekening mee gehouden dat de respondent kan behoren tot een migrantengroep. Een voorbeeld is een item uit de Nederlandse Persoonlijkheds-Vragenlijst (NPV): "Volgens mij komen de beste bruggenbouwers uit ons land".

Cultuurvrije test?

Het ontwikkelen van een volledig cultuurvrije test is onmogelijk. Het meten van intelligentie en persoonlijkheid is immers gebaseerd op ideeën en constructen van de werkelijkheid en dus altijd cultuurgebonden. Het gebruik van cultuurafhankelijke instrumenten kan zinvol zijn als de geteste persoon moet kunnen functioneren in een bepaalde cultuurgebonden situatie. Een test om de haalbaarheid van een bepaalde opleiding te bepalen meet bijvoorbeeld ook de Nederlandse taalvaardigheid. Daarom is het belangrijk om vooraf te bepalen of het bij de vraagstelling om de persoon zelf gaat of om het functioneren in een (werk- of leer)omgeving (Van den Berg, 1999).

Gevolgen van verkeerde diagnoses door tests

Intelligentietests spelen een grote rol in de indicatiestelling voor het speciaal onderwijs. Uitslagen van persoonlijkheidstests zijn van invloed op de diagnose en de aanpak van gedragsproblemen en wegen onder andere mee bij de indicaties voor uithuisplaatsing en geestelijke gezondheidszorg (ggz). Wanneer tests niet zijn afgestemd op migrantenkinderen, dan kan dit tot gevolg hebben dat zij door een verkeerde diagnose hulp krijgen die niet aansluit of schadelijk is. Daarnaast kunnen zij ten onrechte wel of geen indicatie krijgen voor speciaal onderwijs, ggz of orthopedagogische tehuizen (Kouratovsky, 2002).

Criteria voor het kiezen van instrumenten

De Commissie Testaangelegenheden Nederland (COTAN) van het NIP (de beroepsvereniging van en voor psychologen) beoordeelt instrumenten aan de hand van de volgende criteria:

- Uitgangspunten van testconstructie;
- Kwaliteit van het testmateriaal;
- Kwaliteit van de handleiding;
- Normen;
- Betrouwbaarheid;
- Begripsvaliditeit;
- Criteriumvaliditeit.

Bij de beoordeling houdt de COTAN ook rekening met de toepasbaarheid voor migranten. Bij de kwaliteit van het testmateriaal wordt bekeken of items vrij zijn van racistische, etnocentrische, seksistische en voor bepaalde bevolkingsgroepen kwetsende inhoud. De aanwezigheid van gegevens over mogelijke verschillen tussen subgroepen, bijvoorbeeld migranten en autochtonen, telt mee bij de beoordeling van het criterium normen. Bij de beoordeling van de begripsvaliditeit wordt gelet op de aanwezigheid van gegevens over mogelijke itembias bij verschillende groepen (Evers e.a., 2009). Hierbij worden items opgespoord die voor verschillende groepen een verschillende betekenis hebben.

Andere criteria voor de keuze van instrumenten zijn: 1) de specifieke vraagstelling en 2) de situatie van degene die getest wordt. De taalvaardigheid en het opleidingsniveau bepalen bijvoorbeeld de keuze van het instrument. Wanneer een kind nog maar kort in Nederland verblijft, is het bijvoorbeeld niet zinvol om een verbale test af te nemen.

Bij het bepalen van de mogelijke geschiktheid van een instrument voor migranten noemen deskundigen onderstaande kenmerken van instrumenten ((Van de Vijver e.a., 2001).

- in de instructie en de items is het taalgebruik duidelijk, eenduidig en eenvoudig;
- in de normgroepen is een representatief aantal migrantenkinderen opgenomen;
- de normgroepen zijn onderscheiden naar herkomst en verblijfsduur;
- er is zowel een vergelijking mogelijk met de scores van de migrantengroep waaruit het kind afkomstig is als met de scores van de gehele normeringsgroep;
- bij het meten van intelligentie wordt een test gekozen die zowel intelligentie in aanleg als aangeleerde schoolse vaardigheden meet;
- er is ook een versie van het instrument beschikbaar in de moedertaal.

Aandachtspunten voor testgebruik

De professional moet niet alleen een gedegen afweging maken tussen mogelijke meetinstrumenten, maar heeft ook interculturele competenties nodig om op een adequate manier te kunnen communiceren en effectief te kunnen werken met migranten, zoals aandacht voor de eigen waarden en normen en kennis van verschillende culturele achtergronden (Knipscheer & Kleber, 2004). Deskundigen geven onderstaande adviezen voor testgebruik bij migrantengroepen (Van de Vijver e.a., 2001) (Van den Berg, 1999):

- Informatie verzamelen over het kind en zijn (groot)ouders, over het land en streek van herkomst, de etniciteit, de taalkennis, het opleidingsniveau, de verblijfsduur, de generatiestatus en de integratiegeschiedenis.
- Nagaan of de persoon de instructies en items voldoende begrijpt en of hij bekend is met testafnames, bijvoorbeeld door een taaltest af te nemen of hem voorbeeldopgaven te laten maken. Geef indien nodig een uitgebreide instructie;
- Inzicht verkrijgen in de kennis van de moedertaal en bepalen of een test in een andere taal dan het Nederlands zinvol en beschikbaar is. De voorkeur gaat dan uit naar een vragenlijst die in het land van herkomst is vertaald en/of bewerkt, omdat het simpelweg vertalen van Nederlandse uitdrukkingen als ‘in de put zitten’ niet voldoende is;
- Wanneer ouders de vragenlijst thuis invullen, is het zinvol om hen naast een Nederlandse vragenlijst ook een vragenlijst en instructie in de eigen taal te sturen. Vraag vervolgens na of de ouder de vragenlijst alleen heeft ingevuld of met hulp van anderen;
- Scores van het individuele kind indien mogelijk zowel vergelijken met de eigen groep als met de gehele normeringgroep;
- beoordelaars en psychologen afkomstig uit migrantengroepen ook laten kijken naar de resultaten;
- In de rapportage vermelden in hoeverre de test bruikbaar is in relatie tot de culturele achtergrond en de taalkennis van de onderzochte.

Voorbeelden van meetinstrumenten onderzocht bij migrantenkinderen

Kouratovsky (2002) bespreekt tests die veel gebruikt worden in de jeugdzorg en geestelijke gezondheidszorg. Voor onderstaande tests voor cognitieve ontwikkeling is de toepasbaarheid bij Turkse en Marokkaanse jeugdigen onderzocht en heeft de COTAN de betrouwbaarheid en de validiteit als voldoende tot goed beoordeeld. De SON-R 2,5 tot 7 jaar is tevens onderzocht bij Surinaamse jeugdigen, de RAKIT en de MCT-M bij Surinaamse en Antilliaanse jeugdigen.

- Snijders-Oomen Niet-Verbale intelligentietest (SON) voor 2,5 tot 7 jaar en 5,5 tot 17 jaar;
- De Revisie Amsterdamse Kinder Intelligentie Test (RAKIT) voor 4 tot en met 11 jaar;
- Leertest voor Etnische Minderheden (LEM) voor 5- tot en met 7-jarigen;
- Multiculturele Capaciteiten Test - Middelbaar niveau (MCT-M) voor 15 jaar en ouder.

Als het kind afkomstig is uit een groep waarvoor de toepasbaarheid van de test niet is onderzocht, dan adviseert Kouratovsky de LEM en de MCT-M. Deze tests zijn speciaal ontwikkeld voor migrantenkinderen. De SON is bruikbaar voor migrantenkinderen, taal- en spraakgestoorden en geremde of niet-gemotiveerde kinderen.

Voor het meten van het sociaal emotioneel functioneren en persoonlijkheidskenmerken zijn de volgende tests volgens Kouratovsky goed toepasbaar:

- [Child Behavior Checklist \(CBCL\)](#) voor 2- tot 3-jarigen en 4- tot 18-jarigen;
- [Familie Relatie Test \(FRT\)](#) voor 4- tot en met 11-jarigen;
- De Thematic Apperception Test (TAT) voor 4 jaar en ouder.

De COTAN heeft de betrouwbaarheid en validiteit van de CBCL als voldoende tot goed en die van de FRT als onvoldoende beoordeeld. De CBCL en de FRT zijn volgens Kouratovsky wel goed toepasbaar, omdat deze de wisselwerking tussen kind en omgeving in kaart brengen. De TAT is een projectieve test waarmee de persoonlijkheid wordt bepaald door de manier waarop de jeugdige het testmateriaal waarneemt en interpreteert. Deze test is bruikbaar gebleken bij diverse migrantengroepen. De TAT staat niet in de COTAN.

Bleichrodt & Van de Vijver (2001) onderzochten de geschiktheid voor migrantenkinderen van de meest gebruikte intelligentietesten. Evenals Kouratovsky noemen Bleichrodt & Van de Vijver de RAKIT en de SON-R versies als voorbeeld van testen die onderzocht zijn bij migrantenkinderen. De LEM en de LIR (leerpotentieeltest voor Inductief Redeneren) zijn geschikt om het leerpotentieel te meten. De LIR is ontwikkeld bij twee subtests van de RAKIT. Ook beschrijven zij een niet-intelligentietest, namelijk de SCHOBL-R. Dit is een Schoolgedrag Beoordelingslijst uit 1993 die leerkrachten invullen om het schoolgedrag van 4- tot 11-jarigen te beoordelen. Ondanks de verouderde normgegevens wordt de SCHOBL-R nog steeds ingezet bij de indicatiestelling voor het speciaal onderwijs. De SCHOBL-R is onderzocht bij autochtone en migrantenkinderen. De conclusie was dat de betekenis van de SCHOBL-R scores nauwelijks varieerde voor de verschillende etnische groepen. De COTAN heeft de normen, de betrouwbaarheid en de validiteit van de SCHOBL-R als voldoende tot goed beoordeeld.

Het [Van WiechenOnderzoek \(VWO\)](#) en de [Strengths and Difficulties Questionnaire \(SDQ\)](#) zijn instrumenten die binnen de jeugdgezondheidszorg veel gebruikt worden en zijn onderzocht op hun toepasbaarheid bij migrantenkinderen. Het VWO wordt als landelijke standaard gebruikt voor ontwikkelingsonderzoek bij 0- tot 4-jarigen. In het TNO-rapport van Boere-Boonekamp e.a. (2009) staan drie onderzoeken beschreven naar de toepasbaarheid van het VWO bij migrantenkinderen. De onderzoekers concluderen dat:

- de D-score (een samenvattende maat) ook bij migrantenkinderen tot 2 jaar een bruikbare maat is voor het meten van de ontwikkeling;
- er geen structureel verschil is tussen migranten en autochtone kinderen in het verloop van de ontwikkeling per ontwikkelingsveld van het Van Wiechenonderzoek, gemeten van 1 tot 45 maanden;
- de predictieve validiteit voor het opsporen van kinderen met een globale ontwikkelingsachterstand gelijk is voor migranten- en autochtone kinderen.
- het percentage ontbrekende gegevens bij migrantenkinderen significant hoger ligt vergeleken met autochtone kinderen. Bij migrantenkinderen wordt het VWO dus minder goed ingevuld.

De [Strengths and Difficulties Questionnaire \(SDQ\)](#) is een korte vragenlijst die de psychische problematiek en vaardigheden bij kinderen van 3 tot en met 16 jaar meet. De SDQ kent verschillende versies, namelijk een voor leerkrachten, een voor ouders en een versie voor de jeugdige zelf. De ouderversie van de SDQ wordt binnen de jeugdgezondheidszorg veelvuldig gebruikt. Ouders vullen de vragenlijst zelfstandig in. Rensink (2008) onderzocht de bruikbaarheid en de doelmatigheid van de ouderversie van de SDQ bij migrantenouders in de jeugdgezondheidszorg door onder andere interviews af te nemen onder migrantenouders (n=16) en jgz-medewerkers (n=7). Zij concludeerde dat zowel de bruikbaarheid als de doelmatigheid van de SDQ te wensen over laat.

In het onderzoek van Zwirs (2006) vulden leerkrachten de SDQ in voor 6- tot 10-jarigen (n=2185). De leerkrachten rapporteerden meer probleemgedrag voor Marokkaanse jongens, minder probleemgedrag voor Turkse jongens en evenveel probleemgedrag voor Surinaamse jongens, in vergelijking met Nederlandse jongens. Het probleemgedrag van meisjes bleek niet te verschillen tussen de etnische groepen. Volgens de best-estimate procedure (diagnose door een panel van deskundigen) afgenomen bij een subgroep van 270 kinderen verschilde de totale prevalentie van psychiatrische stoornissen en gedragsstoornissen echter niet voor de verschillende etnische groepen. Zwirs onderzocht ook de screeningsvaliditeit van de externaliserende probleemschalen van de leerkrachtversie van de SDQ (n=254, 6- tot 10-jarigen). Zij concludeerde dat het mogelijk was om met vier items uit deze schalen zowel de autochtone als de allochtone kinderen met een hoge kans op een gedragsstoornis te onderscheiden van de kinderen met een lage kans op een gedragsstoornis.

De kindvolgsystemen [Kijk!](#), [Kind Observatie registratie](#), [Peuter- en kleutervolgsysteem](#) en [Ze laten het je zien..ze laten je het horen](#) zijn meetinstrumenten die binnen de Voor- en Vroegschoolse Educatie (VVE)¹ veel worden gebruikt bij migrantenkinderen, maar niet zijn onderzocht op hun toepasbaarheid bij deze doelgroep (Ince & Van den Berg, 2010).

Bronnen

- Berg, H. van den & Leest, F. van. (1999). *Praktisch testgebruik bij migranten. Wanneer zinnig, nuttig of waardeloos. De Psycholoog*, 34: 256-260.
- Bleichrodt, N. & Vijver, F. van de. (2001). *Diagnostiek bij migranten. Mogelijkheden en beperkingen van psychologische tests*. Lisse: Swets & Zeitlinger B.V.
- Bleichrodt, N. & Berg, H. van den. (2004). *Handleiding Multiculturele Capaciteiten Test Middelbaar niveau*. Amsterdam: NOA.
- Bochhah, N. e.a. (2005). *Richtlijnen gebruik diagnostische instrumenten bij etnische minderheden*. Rotterdam: Art. 1 en NIP.
- Boere-Boonekamp, M. e.a. (2009). *Onderbouwing van de validiteit van het ontwikkelingsonderzoek bij kinderen van 0 tot en met 4 jaar: het Van Wiechenonderzoek*. Leiden: TNO.
- Evers, A. e.a. (2009). *COTAN Beoordelingssysteem voor de kwaliteit van tests*. Amsterdam: NIP.
- Kouratovsky, V. (2002). *Wat is er aan de hand met Jamila? Transculturele diagnostiek in de jeugdzorg*. Utrecht, FORUM.

¹ Voor- en vroegschoolse educatie (VVE) is het stimuleren van de ontwikkeling van kinderen met kans op een taalachterstand in de periode voordat zij naar school gaan en tijdens de eerste paar schooljaren.

- Ince, D. & Berg, G. van den. (2010). *Overzichtsstudie interventies voor Migrantenjeugd: ontwikkelingsstimulering, preventie en vroeghulp*. Utrecht: Nederlands Jeugdinstituut.
- Knipscheer, J.W. & Kleber, R.J. (2004). *Een interculturele entree in Altrecht. Inhoud en resultaat van een interculturele werkwijze bij kortdurende behandeling*. Utrecht: Universiteit Utrecht, Klinische Psychologie.
- Rensink, L. (2008). *Strengths and Difficulties Questionnaire. De bruikbaarheid en doelmatigheid van het vroegsignaleringsinstrument SDQ voor het tijdig signaleren van psychosociale problemen bij allochtone kinderen*. Universiteit Utrecht.
- Vijver, F. van de (red.) (2001). *Deskundigen over het testen van etnische minderheden*. Rotterdam: Art. 1 en NIP.
- Zwirs, B. (2006). *Externalizing disorders among children of different ethnic origin in the Netherlands*. Proefschrift. ISBN: 90-393-4255-5