



FACULTEIT PSYCHOLOGIE EN
PEDAGOGISCHE WETENSCHAPPEN

Academiejaar 2015 - 2016

Tweedekansexamenperiode

Culturele Equivalentie en Vertekening in de Cognitieve Vaardigheidstest volgens het CHC-model (CoVaT-CHC).

Masterproef II neergelegd tot het behalen van de graad van

Master of Science in de Psychologie, afstudeerrichting Bedrijfspsychologie en Personeelsbeleid

Promotor: Prof. Johnny Fontaine

Begeleider: Julie Rosseel

01103034

Ester van Agtmaal

Dankwoord

Als laatste stap in mijn 5-jarige opleiding en het behalen van mijn diploma, dien ik deze masterproef in. Het onderwerp van de masterproef heeft zo een interesse in mij opgewekt, waardoor ik beslist heb om volgend jaar schoolpsychologie bij te studeren. Graag zou ik enkele mensen willen bedanken die het schrijven van deze masterproef mee mogelijk hebben gemaakt.

Eerst en vooral zou ik graag mijn promotor, professor Johnny Fontaine, willen bedanken voor de begeleiding en tips die hij mij gaf, die cruciaal geweest zijn bij het schrijven van deze masterproef.

Vervolgens wil ik ook graag dr. Mark Schittekatte bedanken, waar ik altijd bij terecht kon voor theoretische en praktische vragen, ondanks dat hij niet mijn officiële begeleider was.

Ook wil ik Walter Magez en Marlies Tierens bedanken voor de training in de afname van de CoVaT-CHC en omdat zij altijd beschikbaar waren indien ik praktische en theoretische vragen had met betrekking tot het model.

Uiteraard gaat mijn dank ook uit naar de scholen en kinderen die hebben deelgenomen aan deze studie. De afnames waren een zeer leuke ervaring voor mij.

Tenslotte zou ik graag mijn vrienden en familie willen bedanken voor het voorzien van de nodige ontspanning en de continue steun tijdens stressvolle momenten.

Abstract

Deze masterproef onderzocht culturele equivalentie en vertekening in de cognitieve vaardigheidstest volgens het CHC-model. Een verkorte versie van de test werd afgenomen van 224 kinderen uit het 4^{de}, 5^{de} en 6^{de} leerjaar in 3 scholen. Oorspronkelijk werden de kinderen onderverdeeld in een Vlaamse groep, een Europese en een niet-Europese groep. Omdat analyses geen significante scoreverschillen aantoonde voor de Europese groep en de niet-Europese groep werden de 2 groepen samengenomen tot 1 groep. De resulterende allochtone groep bestond uit 77 kinderen, de Vlaamse groep bestond uit 147 kinderen. Er werden significante verschillen gevonden voor de subtesten van vloeiende intelligentie, gekristalliseerde intelligentie en verwerkingssnelheid waarbij de allochtone groep gemiddeld lager scoorde dan de Vlaamse groep. Aan de hand van logistische regressie werd itemvertekening nagegaan. Bij de verbale subtests van gekristalliseerde intelligentie en kortetermijngeheugen werden vertekende items gedetecteerd, waarbij de vertekening de allochtone groep soms benadeelde en soms bevoordeelde. Het verband tussen acculturatie-oriëntatie en de 4 schaalscores was enkel positief significant voor gekristalliseerde intelligentie waarbij kinderen die het meest aangepast waren aan de gastcultuur hogere scores behaalden. Factoranalyse resulteerde in 1 factor voor de gehele groep en de Vlaamse groep, met positieve factorladingen. Voor de allochtone groep werden 2 factoren teruggevonden, met zowel negatieve als positieve factorladingen.

Inhoudstafel

Literatuurstudie.....	2
Het Begrip 'Intelligentie'	2
Algemene intelligentie 'g' volgens Spearman.....	2
Intelligentie volgens Thurstone.....	3
Intelligentie volgens het CHC-model.....	3
IQ en Maatschappelijke Relevantie.....	5
CoVaT-CHC Model.....	5
Ontstaan en ontwikkeling CoVaT-CHC.....	5
Eigenschappen CoVaT-CHC-basisversie.....	6
Modules en subtesten.....	6
Culturele Scoreverschillen voor de CoVaT-CHC.....	7
Equivalentie en Vertekening	8
Equivalentie.....	9
Functionele equivalentie.....	9
Structurele equivalentie.....	9
Metrische equivalentie.....	9
Volledige-score equivalentie.....	10
Vertekening.....	10
Constructvertekening.....	11
Methodevertekening.....	11
Itemvertekening.....	13
Intelligentie en SES.....	15
Intelligentie en Acculturatie.....	16
Onderzoeksvraag en Hypotheses.....	18
Methode.....	20
Steekproef.....	20
Opzet.....	20
Meetinstrumenten.....	21
CoVaT-CHC.....	21
Acculturatie vragenlijst.....	21

Demografische variabelen.	22
Schoolresultaten.	22
Procedure.....	22
Resultaten	23
Interne Consistentie.....	23
Multiple regressie analyse.	23
Item vertekening.....	26
Puntreeksen.	27
Tegenstellingen.	27
Geheugen A.....	29
Aparte Foutenanalyse Tegenstellingen.....	30
Effecten van Achtergrondvariabelen	32
Schoolresultaten “Nederlands” en “Wiskunde”	32
SES.	33
Geslacht.....	33
Leeftijd.....	33
Effect van Acculturatiestrategie.....	33
Achtergrondvariabelen en acculturatie-strategie.....	34
Equivalentie.....	35
Discussie	36
Bespreking resultaten	36
Culturele scoreverschillen.....	36
Itemvertekening.....	38
SES.	39
Acculturatie-oriëntatie.....	39
Equivalentie.....	40
Beperkingen	40
Toekomstig onderzoek.....	41
Conclusie	41
Referenties	42

Het begrip ‘intelligentie’ is al jaren voer voor discussie voor zowel wetenschappers als leken. Vooral het debat over al dan niet reële culturele verschillen in intelligentie ligt maatschappelijk gevoelig. Er is namelijk een robuuste bevinding dat, op bijna elke cognitieve vaardigheidstest, minderheidsgroepen gemiddeld lager scoren dan meerderheidsgroepen. Zo toonde ook de CoVaT-CHC, een Vlaamse intelligentietest, significant lagere scores aan voor de niet-Europese kinderen voor vloeiende intelligentie ($p < .01$), gekristalliseerde intelligentie ($p < .01$), kortetermijngeheugen ($p < .01$) en visuele informatieverwerking ($p < .01$) in vergelijking met de kinderen met Nederlands als moedertaal. Voor de testen van verwerkingssnelheid werden er geen verschillen gevonden. Niet-Nederlandstalige Europese kinderen verschilden significant van de Nederlandstalige groep voor vloeiende intelligentie ($p < .05$) en gekristalliseerde intelligentie ($p < .01$). Tussen de niet-Nederlandstalige Europese en de niet-Europese kinderen werden geen significante verschillen gevonden (Bos, Magez, Tierens, Van Parijs & Decaluwé, 2014).

Deze masterproef onderzoekt in welke mate de scoreverschillen voor de CoVaT-CHC te wijten zijn aan reële verschillen in cognitieve vaardigheden, dan wel moeten worden toegeschreven aan culturele vertekening in het meetinstrument.

Eerlijke diagnostiek is het uiteindelijke doel van deze masterproef. Meer bepaald wordt onderzocht of de CoVaT-CHC in zijn huidige vorm gebruikt kan worden voor etnische minderheden. Eerst wordt de itemvertekening nagegaan aan de hand van Differential Item Functioning via logistische regressie. Verwacht wordt dat de verbale subtests meer vertekende items zullen bevatten dan de performale subtests. Bij kinderen waarvan minstens 1 ouder geboren is in een niet-Europees land werd onderzocht of er een verband bestaat tussen de acculturatie-oriëntatie van het kind en de testresultaten. Verwacht wordt dat kinderen die zich meer richten op de gastcultuur hogere testresultaten zullen behalen waarbij het effect het grootste zal zijn op de verbale subtests, vergeleken met kinderen die meer gericht zijn op de thuiscultuur. Als laatste werd een indicatie van equivalentie nagegaan aan de hand van factoranalyse.

Literatuurstudie

Het Begrip 'Intelligentie'

De vraag of intelligentie één algemene capaciteit is of uit meerdere capaciteiten bestaat, wordt door verschillende wetenschappers anders beantwoord. In wat volgt wordt de visie op intelligentie volgens Spearman, Thurstone en het CHC-model weergegeven.

Algemene intelligentie 'g' volgens Spearman.

Charles Spearman (1904) legde de basis voor de theorie over algemene intelligentie. Deze theorie stelt dat de prestatie op cognitieve vaardigheidstesten beïnvloed wordt door één brede mentale capaciteit, de "g-factor". Spearman merkte op dat hoewel mensen kunnen uitblinken in bepaalde domeinen, er een tendens is dat mensen die het goed doen in het ene gebied de neiging hebben om het ook goed te doen in andere gebieden. Dit zou te wijten zijn aan de g-factor. Hij geloofde dat alle resultaten op intelligentietesten beïnvloed werden door de onderliggende g-factor. Bijgevolg worden interindividuele verschillen op cognitieve vaardigheidstesten grotendeels verklaard door interindividuele verschillen in de g-factor. Naast de g-factor erkent Spearman ook het bestaan van "s-factoren", die specifiek zijn voor het oplossen van één bepaald type taak. Voorbeelden van s-factoren zijn een verbale factor en een mathematische factor.

De robuuste bevinding dat subtesten van een IQ-batterij altijd positief gecorreleerd zijn, "positive manifold", wordt vaak gezien als evidentie voor het bestaan van de g-factor. De correlaties tussen scores op een wiskundige test of verbale test zouden te wijten zijn aan eenzelfde factor die een deel van de variatie verklaart, de g-factor. Echter zijn er ook alternatieve verklaringen voor de positive manifold. Volgens het model van Van Der Maas et al. (2006) wordt de positive manifold verklaard door mutualisme. De positieve intercorrelaties zouden puur ontstaan door positieve interacties tussen cognitieve processen tijdens de ontwikkeling. Zo zou dankzij de ontwikkeling van verbaal redeneren, ook het logisch redeneren verbeteren. Positive manifold is dus een robuust verschijnsel, maar er geen zijn sluitende verklaringen voor.

Intelligentie volgens Thurstone.

Thurstone (1938) maakte bezwaar tegen het idee van één algemene intelligentiefactor. Hij ontkent het bestaan van de *g*-factor en stelt dat er zeven specifieke intelligentie factoren zijn: “verbaal begrip”, “woordvlotheid”, “rekenkundig inzicht”, “ruimtelijk visueel inzicht”, “associatief geheugen”, “waarnemingssnelheid” en “logisch redeneren”. Thurstone beweerde dat deze factoren statistisch onafhankelijk van elkaar zijn; goed zijn in verbaal begrip voerspelt niet of men goed is in alle andere factoren.

Verscheidene onderzoeken hebben aangetoond dat de *g*-factor meer variantie verklaart bij groepen die laag scoren op intelligentietesten. Bij groepen die hoog scoren zijn het vooral de specifieke factoren die de variantie verklaren (Detterman, 1991; Detterman & Daniel, 1989; Abad et al., 2003; Colom et al., 2002; Deary et al., 1996; Hunt, 1995; Legree, Pfifer & Grafton, 1996).

Intelligentie volgens het CHC-model.

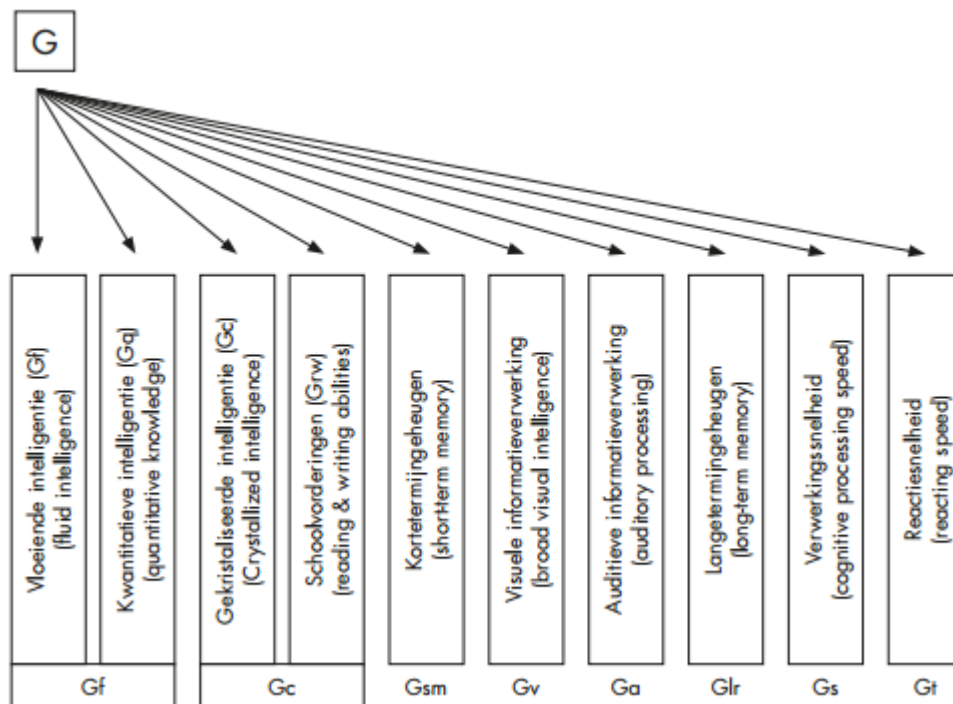
Een consensusmodel dat een compromis biedt voor de bevindingen van Spearman en Thurstone is het CHC-model. Volgens de internationale literatuur is het CHC-model de meest uitgewerkte, empirisch en psychometrisch onderbouwde intelligentietheorie (Neisser et al., 1996; Gottfredson, 1998; Keith & Reynolds, 2010; Flanagan & Harrison, 2012). Het model ziet intelligentie als meerledig en is hiërarchisch opgebouwd. De hiërarchie situeert zich op 3 strata.

Op stratum III bevindt zich het breedste en meest algemene niveau van intelligentie, de ‘G’ factor. De ‘G’ factor is betrokken bij complexe hogere-orde cognitieve processen (Gustaffson & Undheim, 1996; Jensen, 1998; McGrew & Woodcock, 2001) en is de beste voorspeller van algemene schoolprestaties en professioneel succes (McDermott et al., 1990, in McGrew & Flanagan, 1998).

Op stratum II staan de brede cognitieve vaardigheden (BCV), die elk een eigen inhoud hebben. Ze kunnen gezien worden als eerder basale en stabiele cognitieve eigenschappen in een persoon, die het gedrag sturen en beïnvloeden op verschillende domeinen (Carroll, 1993). De belangrijkste BCV zijn vloeiende intelligentie (*Gf*) en gekristalliseerde intelligentie (*Gc*) (Catell, 1971, 1987; Horn, 1985). Vloeiende intelligentie omvat inductieve en deductieve redeneervaardigheden, die beïnvloed worden door biologische

en neurologische factoren. Het wordt omschreven als de vaardigheid om met nieuwe problemen op een efficiënte manier om te gaan. Gekristalliseerde intelligentie bestaat voornamelijk uit verworven kennis. Het is een maat voor de accumulatie van kennis en vaardigheden gedurende de levensloop en is het resultaat van de invloed van acculturatie, normen, tradities en omgeving waaronder scholing op intelligentie (Cattell, 1957, 1971). De andere 8 BCV zijn “Kwantitatieve Kennis” (Gq), “Schoolvorderingen” (Grw), “kortetermijngeheugen” (Gsm), “Visuele informatieverwerking” (Gv), “Auditieve informatieverwerking” (Ga), “Langetermijngeheugen” (Glr), “Verwerkingsnelheid” (Gs) en “Reactiesnelheid” (Grw) (Carroll, 1993). De brede cognitieve vaardigheden correleren positief met ‘G’, maar verschillen op vlak van de sterkte van de lading. Zo laadt vloeiende intelligentie hoger op G dan reactiesnelheid. Zie figuur 1 voor een schematische voorstelling van de sterkte van de ladingen; hoe dichterbij G gelegen, hoe sterker de lading.

Elk van de brede cognitieve vaardigheden omvatten een reeks nauwe cognitieve vaardigheden (NCV) op stratum I. Zij vertegenwoordigen grotere specialisaties van vaardigheden die de effecten van ervaring en leren weerspiegelen in vaak zeer specifieke manieren (Carroll, 1993, p. 634). Het is op dit niveau dat intelligentietesten, zoals de CoVaT-CHC aansluiten (Magez & Stinissen, 2010).



Figuur 1. Schematische voorstelling van het CHC-model (Magez & De Cleen, 2007)

IQ en Maatschappelijke Relevantie

Het verband tussen scores op een intelligentietest (IQ¹) en succes in het leven is meermaals onderzocht geweest. Hunt (2011) concludeert dat er sprake is van probabilistische verbanden tussen testcores en een variëteit aan maatschappelijke uitkomsten. In het algemeen kan gesteld worden dat hoe hoger de testscore is, hoe groter de kans dat men succesvol is in verscheidene maatschappelijk domeinen. Deze domeinen omvatten leerprestaties, on-the-job performance, inkomen, burgerlijke staat, gezondheid en levensduur (Neisser et al., 1996; Gottfredson, 1997, 2007; Herrnstein & Murray, 1994). De meest valide voorspeller van toekomstig succes is algemene intelligentie (Hunter & Hunter, 1984; Ree & Earles, 1992). Zo is de gemiddelde predictieve waarde van een sterk *g*-geladen cognitieve vaardigheidstest gelijk aan $r = 0.51$ (Schmidt & Hunter, 1998). Een interessante bevinding is dat hoe hoger een test *g*-geladen is, hoe groter het gemiddelde scoreverschil is tussen de meerderheidsgroep en de minderheidsgroep (Hough, Oswald, & Ployhart, 2001). Deze bevinding werd ook gevonden bij het eerste vergelijkend onderzoek van de CoVaT-CHC waarbij voor Gf en Gc de grootste culturele verschillen werden gevonden (Bos et al., 2014).

CoVat-CHC Model

Ontstaan en ontwikkeling CoVaT-CHC.

De CoVaT-CHC Basisversie (Magez, Tierens, Van Huynegem, Van Parijs, Decaluwé, & Bos, 2015) is de enige cognitieve vaardigheidstest in Vlaanderen die gebaseerd is op het Cattell-Horn-Carroll model. De test werd ontwikkeld op vraag van VCLB De Wissel-Antwerpen en VCLB Voor- en Noorderkempen. Er was nood aan een nieuw instrument om de cognitieve mogelijkheden van leerlingen uit het buitengewoon Lager Onderwijs (BuLO) in beeld te brengen in de overgangsbegeleiding van het BuLO naar het Buitengewoon Secundair Onderwijs (BuSO). De reeds bestaande intelligentietesten differentieerden te weinig. Daarnaast wou het Vlaams Agentschap voor Personen met een Handicap (VAPH) een schatting van het IQ van Vlaamse BuLO-

¹ Een intelligentiequotiënt of IQ duidt de relatieve score aan van een persoon op een intelligentietest vergeleken met zijn/haar leeftijdsgroep (Hunt, 2011).

leerlingen. De voorwaarde waar het model moest aan voldoen, was dat het gebaseerd moest zijn op het CHC-model. Zo werd de eerste editie van de test in 2011-2012 ontwikkeld door Walter Magez in samenwerking met Thomas More Antwerpen, voor leerlingen van 12 en 13 jaar uit het gewoon en buitengewoon onderwijs. In 2012-2013 werd er door Thomas More Antwerpen, samen met de Centra voor leerlingenbegeleiding (CLB) uit de provincie Antwerpen, gewerkt aan de normering voor leerlingen van 10 tot 14 jaar, zowel voor het gewoon als buitengewoon onderwijs (Magez & Tierens, 2015).

Eigenschappen CoVaT-CHC-basisversie.

De CoVaT-CHC-basisversie (Magez et al., 2015) is geschikt om de cognitieve vaardigheden van kinderen vanaf 9;6 jaar tot 13;11 jaar in kaart te brengen en kan zowel individueel als groepsgewijs afgenomen worden. De test onderzoekt vijf van de brede cognitieve vaardigheden van het CHC-model aan de hand van de modules Gf (vloeiende intelligentie), Gc (gekristalliseerde intelligentie), Gv (visuele informatieverwerking) en Gsm/Gs (kortetermijngeheugen/verwerkingssnelheid). Elke module bestaat uit minimum twee subtests. Tijdens een afname worden de instructies visueel ondersteund met behulp van een diapresentatie. Na de diapresentatie volgt er een oefenmodule, pas als de instructies begrepen zijn, wordt er overgegaan op de eigenlijke test. Er zijn geen instap-, afbreek- of terugkeerregels.

Modules en subtesten.

De eerste module in de test is vloeiende intelligentie (Gf) en bestaat uit de subtests “puntreeksen” en “figuurreeksen”. De tweede module is gekristalliseerde intelligentie (Gc). De subtests ervan zijn “schiftingen” en “tegenstellingen”. Visuele informatieverwerking (Gv) is de derde module en is de vaardigheid om gebruik te maken van visuele beelden om problemen op te lossen. Subtests zijn “gedraaide figuren” en “dozen plooiën”. De laatste module omvat kortetermijngeheugen (Gsm) met de subtests “geheugen A” en “geheugen B”, en verwerkingssnelheid (Gs) met als subtest “geheime code”. Gsm is de vaardigheid om gedurende korte tijd informatie vast te houden, te verwerken en te gebruiken. Gs is de vaardigheid om cognitieve taken vloeiend en automatisch uit te voeren. Dit zijn taken die bijna iedereen juist heeft indien voldoende tijd gegeven wordt (Magez et al., 2015). In deze test wordt er gebruik gemaakt van een

tijdslimiet bij de afname. In tabel 1 wordt een overzicht en omschrijving van de subtests gegeven.

Tabel 1

Overzicht en interne structuur van de CoVaT-CHC Basisversie (Magez et al.,2015)

BCV/MODULE	SUBTEST	BESCHRIJVING
Gf (Vloeiende intelligentie)	Puntreeksen	De puntreeksen tweemaal verder zetten op basis van de gevonden regel voor de opbouw van de reeks.
	Figuurreeksen	De tekenreeks driemaal verder zetten op basis van de te ontdekken regel, geldend voor de opbouw van de reeks.
Gc (Gekristalliseerde intelligentie)	Schiftingen	Uit een reeks van vijf woorden zoeken welk woord niet of het minst past in de reeks.
	Tegenstellingen	Uit de geboden alternatieven het woord zoeken dat (het meest) tegengesteld is aan het gegeven woord.
Gv (Visuele informatieverwerking)	Gedraaide Figuren	Snel "zien" hoe een gegeven tweedimensionale figuur kan draaien in het (platte) vlak.
	Dozen Plooien	Zien hoe een open geplooide doos (tweedimensionaal) kan teruggeplooid worden tot een driedimensionaal "lichaam".
Gsm/Gs (Korte Termijngeheugen/ Verwerkingsnelheid)	Geheugen A	Gedurende een vastgelegde tijd woorden van buiten leren/memoriseren, om onmiddellijk erna zoveel mogelijk van deze woorden uit het geheugen neer te schrijven/schriftelijk te reproduceren.
	Geheime Code	Gedurende een bepaalde tijd zo snel en zo juist mogelijk schriftelijk uitvoeren van een eenvoudige, zich herhalende opdracht.
	Geheugen B	Gedurende een vastgelegde tijd dezelfde woorden als bij Geheugen A van buiten leren/memoriseren, om onmiddellijk erna zoveel mogelijk van deze woorden op een voorgestructureerde wijze te reproduceren.

Culturele Scoreverschillen voor de CoVaT-CHC

Voor het normeringsonderzoek in het schooljaar 2012-2013 werd de CoVaT-CHC afgenomen bij kinderen met het Nederlands als moedertaal (n =1064), met een niet-Nederlandse Europese moedertaal (n =109) en bij kinderen met een niet-Europese moedertaal (n =262). Analyses toonden significant lagere scores aan voor de niet-Europese kinderen voor vloeiende intelligentie ($p < .01$), gekristalliseerde intelligentie (p

<.01), kortetermijngeheugen ($p <.01$) en visuele informatieverwerking ($p <.01$) in vergelijking met de kinderen met Nederlands als moedertaal. Voor de testen van verwerkingssnelheid werden er geen verschillen gevonden. Europese kinderen verschilden significant van de Nederlandstalige groep voor vloeiende intelligentie ($p <.05$) en gekristalliseerde intelligentie ($p <.01$) maar niet voor de andere BCV. Tussen de Europese en de niet-Europese kinderen werden geen significante verschillen gevonden (Bos, Magez, Tierens, Van Parijs, & Decaluwé, 2014).

De observatie dat minderheidsgroepen relatief lager scoren dan meerderheidsgroepen wordt meermaals teruggevonden. Bij onderzoek in de Verenigde Staten is er de robuuste bevinding dat Afro-Amerikanen ongeveer 1 standaarddeviatie (15 punten op een IQ-schaal) lager scoren dan blanke Amerikanen voor een variëteit aan cognitieve testen. Hispanics² bevinden zich tussenin (Hunter & Hunter, 1984; Sackett & Wilk, 1994). Bij globale studies met de Raven Progressive Matrices, een niet-verbale intelligentietest, werden telkens dezelfde resultaten gevonden: Blanken en Aziaten scoren het hoogst, Zwarten het laagst. Andere etnische groepen scoren tussenin (Rushton, Cvorovic, & Bons, 2007; Rushton & Skuy, 2000; Rushton, Skuy, & Fridjohn, 2003; Lynn, Backhoff, & Contreras, 2005).

Equivalentie en Vertekening

De vraag rijst of deze crossculturele verschillen in testcores het resultaat zijn van datgene wat de test beoogt te meten, namelijk cognitieve vaardigheden. Dit is niet het geval indien de test of een item vertekend is. Pas wanneer er sprake is van volledige-score equivalentie, kan men valide vergelijkingen maken. Een instrument is equivalent wanneer het hetzelfde kenmerk meet in verschillende contexten (Horn & McArdle, 1992). De concepten equivalentie en bias verwijzen niet naar de intrinsieke karakteristieken van een testinstrument, maar naar de karakteristieken van crossculturele vergelijkingen van die test. (Van de Vijver & Tanzer, 1997). In wat volgt worden equivalentie en vertekening apart besproken.

² term die gebruikt wordt om te verwijzen naar immigranten uit Spaans of Portugees sprekende landen die in de VS wonen, synoniem voor latino's (Hunt, 2011).

Equivalentie

Er zijn 4 niveaus van equivalentie, namelijk functionele, structurele, metrische en volledige-score equivalentie. De niveaus zijn hiërarchisch gerangschikt. Zo moet er eerst voldaan zijn aan functionele, structurele en metrische equivalentie vooraleer er sprake kan zijn van volledige-score equivalentie (Fontaine, 2005).

Functionele equivalentie.

Functionele equivalentie houdt in dat dezelfde theoretische variabele die men beoogt te meten (bv. numerieke kennis), het testgedrag van de verschillende culturele groepen verklaart. Functionele equivalentie kan worden nagegaan door de empirische studie van het nomologische netwerk in alle culturele groepen (Fontaine, 2005). Zo beoogt men bijvoorbeeld convergente relaties te vinden tussen testcores en schoolresultaten van leerlingen.

Structurele equivalentie.

Structurele equivalentie, een tweede vorm van equivalentie, biedt antwoord op de vraag of dezelfde geobserveerde variabelen gebruikt kunnen worden als operationalisatie van de latente variabele voor alle culturele groepen. Dit houdt in dat wat er gevraagd wordt in het testinstrument, relevant en representatief moet zijn voor elke groep. Verder mag er geen sprake zijn van triviale relaties tussen de geobserveerde en latente variabelen. Er worden echter geen restricties gelegd op het gewicht noch het intercept (Fontaine, 2005).

Metrische equivalentie.

Metrische equivalentie is de derde vorm van equivalentie. Bij het nagaan van metrische equivalentie, onderzoekt men of het mogelijk is om op basis van de latente variabele, kwantitatieve vergelijkingen tussen de culturele groepen te maken. Meer bepaald moeten de factorladingen gelijk zijn over de groepen heen. Zo zal een item dat hoog laadt op de latente factor in de autochtone groep, ook hoog moet laden in de allochtone groep. Er moet benadrukt worden dat op het niveau van metrische

equivalentie, enkel relatieve vergelijkingen mogelijk zijn tussen culturele groepen. Men mag echter geen gemiddelden vergelijken (Fontaine, 2005).

Volledige-score equivalentie.

Bij volledige equivalentie kan men scores direct vergelijken tussen culturele groepen. Dit niveau van equivalentie vereist dat zowel het intercept als het gewicht, de factorladingen, gelijk zijn voor elke groep. Dit houdt in dat wanneer twee personen dezelfde positie hebben op de latente trek (vb. Gf) exact dezelfde geobserveerde score verwacht wordt, onafhankelijk van de culturele groep waar toe men behoort (Fontaine, 2005). Enkel en alleen op het niveau van volledige score-equivalentie mogen de scores op de CoVaT-CHC direct worden vergeleken. Eenzelfde niveau van intelligentie komt dan gemiddeld genomen op eenzelfde manier tot uiting in het geobserveerde gedrag.

Vertekening

Bias of vertekening treedt op wanneer scoreverschillen over indicatoren van een bepaald construct (bv. percentage van kinderen die weten dat een dozijn gelijk is aan twaalf) niet overeenkomen met de verschillen in de onderliggende latente trek (bv. numerieke kennis). Inferenties op basis van vertekende scores zijn niet valide (Van de Vijver & Tanzer, 1997). Volgens Choenni en van der Zwan (1986) zal de score op een intelligentietest voor allochtone kinderen altijd een zekere vertekening van hun capaciteiten inhouden, aangezien volledige cultuurvrije intelligentietests niet kunnen bestaan. Het effect van vertekening is het sterkste wanneer de test aangeboden wordt in een taal die de kinderen niet goed beheersen (Extra & Verhoeven, 1985). Zo bleek dat de verbale subtest 'Vocabulary' van de GATB zo sterk vertekend was dat het de score op de test naar omlaag haalt met 0,92 SD, wat leidde tot een onderschatting van g met 1,8 IQ punten (te Nijenhuis & Van der Flier, 2003). Performantie subtesten, zoals verwerkingssnelheid, zijn het minst onderhevig aan testvertekening en zouden dus een betere inschatting geven van de cognitieve vaardigheden van allochtonen (Sierens et al., 2006; Schittekatte & Hurks, 2015; Maes, Smeets, & Schittekatte, 2015). De term "vertekening" is de verzamelnaam geworden voor storende factoren in crossculturele

vergelijkingen van testcores. Volgende taxonomie onderscheidt 3 vormen van vertekening: constructvertekening, methodevertekening en itemvertekening (Van de Vijver & Leung, 1997a, 1997b; Van de Vijver & Poortinga, 1997).

Constructvertekening.

Constructvertekening ontstaat wanneer het gemeten construct niet identiek is voor alle culturen of wanneer gedragingen die kenmerkend zijn voor het construct niet hetzelfde zijn voor alle culturen (Van de Vijver & Tanzer, 2004). Zo kan de definitie van het construct 'intelligentie' een onvolledige overlap vertonen tussen culturen. Bij het merendeel van testbatterijen die algemene intelligentie meten, wordt de nadruk vooral gelegd op redeneren, vergaren van kennis, geheugen en niet zozeer op de sociale aspecten van intelligentie. Echter is er empirische evidentie dat die sociale aspecten van intelligentie meer prominent zijn in niet-Westerse culturen (Super, 1983).

Constructvertekening kan ook het gevolg zijn van domein onderrepresentatie (Fontaine, 2005). Een voorbeeld hiervan kan gevonden worden in onderzoek naar "filial piety", gedragingen geassocieerd met het zijn van een goede zoon of dochter (Ho, 1996). Vergeleken met Westerse samenlevingen, hebben kinderen in de Chinese samenleving meer en ander soort verplichtingen jegens hun ouders (Ho, 1996). Het begrip "filial piety" is veel uitgebreider in China. Daarom is een vergelijking over culturen omtrent "filial piety" enorm vatbaar voor constructvertekening. Een Westers meetinstrument zou niet alle aspecten van "filial piety" bevatten die relevant zijn in de Chinese context. Omgekeerd zou een Chinese test items bevatten die irrelevant zijn voor mensen uit de Westerse cultuur.

Methodevertekening.

Methodevertekening verwijst naar problemen die voortvloeien uit de manier waarop een studie is uitgevoerd (Meiring, Van de Vijver, Rothmann & Barrick, 2005). Het gaat niet om de methode effecten op zich, maar om de differentiële impact van de methode effecten (Fontaine, 2005). Als de methodevariantie op eenzelfde manier tot uiting komt in de verschillende culturen, dan kan men de resultaten nog steeds vergelijken

indien men controleert voor de methodevariantie. Maar indien het methode effect in de ene culturele groep groter is dan in de andere, kan men niet meer rechtstreeks vergelijken. Van de Vijver (2002) onderscheidt 3 vormen van methodevertekening: steekproefvertekening, instrumentvertekening en afnamevertekening.

Steekproefvertekening.

De eerste vorm is 'steekproefvertekening of sample bias. Deze vorm van vertekening ontstaat wanneer leden van de steekproef systematisch van elkaar verschillen met betrekking tot andere variabelen dan de doelvariabele, die de test ongewenst beïnvloeden. Zo worden er bijvoorbeeld vaak verschillen gevonden in de scholingsgraad tussen culturele groepen, die als confounder kunnen optreden bij het meten van intelligentie. Verschil in motivatie tussen groepen kan ook een bron zijn van steekproefvertekening. Zo zal men minder gemotiveerd zijn indien men al vaker heeft deelgenomen aan een psychologische test dan indien men nog nooit heeft deelgenomen en de testsituatie dus nieuw is (Van de Vijver & Tanzer, 2004).

Instrumentvertekening.

De tweede vorm van methodevertekening is vertekening in het instrument, die veroorzaakt kan worden op 3 manieren. Ten eerste kan de specificiteit van de iteminhoud aan de basis liggen van instrumentvertekening. Wanneer testsubjecten bekend zijn met het stimulusmateriaal kan dit resulteren in een onbedoelde gemakkelijkerheid, wat tot een overschatting zal leiden. Analoog zal gebrek aan vertrouwdheid een onbedoelde moeilijkerheid met zich meebrengen, met een onderschatting als gevolg (Sarnacki, 1979.) Een voorbeeld hiervan werd gevonden door Deregowski en Serpell (1971). In hun onderzoek moesten Schotse en Zambiaanse kinderen miniatuurmodellen van voertuigen en dieren (conditie 1) ofwel foto's van de miniatuurmodellen sorteren (conditie 2). Hoewel er geen crossculturele verschillen werden gevonden bij het sorteren van de miniatuurmodellen, behaalden de Zambiaanse kinderen lagere scores dan de Schotse wanneer men foto's moest sorteren.

De specificiteit van het responseformaat kan ook zorgen voor instrumentvertekening. In onderzoek van Serpell (1979) moesten Schotse en Zambiaanse kinderen patronen reproduceren via vier verschillende manieren: met potlood, met

plasticine, via handpositie en met ijzerdraad. De Schotse kinderen behaalden de hoogste score in de conditie met het potlood. De Zambiaanse kinderen scoorden beter dan de Schotse kinderen wanneer het patroon via ijzerdraad gereproduceerd moest worden. Bij reproductie via handstand en plasticine werden er geen verschillen vastgesteld. Beide groepen waren even weinig vertrouwd met deze twee methoden.

De laatste vorm van instrumentvertekening wordt veroorzaakt door verschillen in antwoordstijlen. Specifieke antwoordformaten hebben mogelijks een differentiële impact op de antwoordstijlen, bijvoorbeeld sociale wenselijkheid. Zo hebben Hispanics de neiging om meer de uiteinden van een schaal aan te duiden dan Anglo-Amerikanen, maar enkel en alleen indien een 5-punten schaal wordt gebruikt. Wanneer een 10-punten schaal wordt gebruikt valt dit effect weg (Hui & Triandis, 1989).

Afnamevertekening.

De derde vorm van methodevertekening is afnamevertekening. Zo kunnen er gemakkelijk communicatieproblemen ontstaan tussen de interviewer en de geïnterviewde, vooral wanneer ze een verschillende moedertaal en culturele achtergrond hebben (Gass & Varone, 1991).

Itemvertekening.

Een laatste vorm van vertekening is itemvertekening, ook wel Differential Item Functioning (DIF) genoemd (Berk, 1982; Holland & Wainer, 1993). In tegenstelling tot construct- en methodevertekening, heeft itemvertekening betrekking op vertekening op itemniveau. Bij itemvertekening is de psychologische betekenis van het item niet identiek hetzelfde voor verschillende culturen, bijvoorbeeld door een slechte vertaling (Helms-Lorenz & Van de Vijver, 1995).

Daarnaast kan er ook ruisvariantie ontstaan door het ongewild introduceren van bepaalde processen of trekken. Bijvoorbeeld het woord 'dozijn' als item in een numerieke vaardigheidstest. Men meet dan eerder taalvaardigheid in plaats van rekenvaardigheid. (Fontaine, 2005). Zo is het mogelijk dat items in de geheugentest van de CoVaT-CHC voor een bepaalde culturele groep eerder peilen naar woordkennis dan naar de kwaliteit van het geheugen. Het vergelijken van testcores tussen verschillende groepen is in dat soort situatie niet valide (Van de Vijver & Tanzer, 2004).

Culturele normen en waarden kunnen ook ongewild geïntroduceerd worden in testen, waarbij kennis van de dominante gastcultuur noodzakelijk is om de items correct op te lossen. Zo werd aangetoond dat het item “Wat is spek” van de WISC-R zeer moeilijk is om te beantwoorden voor Turkse en Marokkaanse leerlingen. Waarschijnlijk is dit te wijten aan het feit dat varkensvlees taboe is in de Islamitische cultuur (Van de Vijver, Willemse, & Van de Rijt, 1993). In dit geval meet de test de mate van assimilatie aan de dominante gastcultuur (Helms-Lorenz & Van de Vijver, 1995). Dit probleem is echter niet beperkt tot de WISC-R. Voor de meest gebruikte psychologische testen in Nederland werd geconcludeerd dat alle testen, vaak impliciete, verwijzingen naar de Nederlandse waarden en normen bevatten (Hofstee, 1990; Hofstee et al., 1990).

Psychometrisch gezien wordt er een onderscheid gemaakt tussen twee soorten van itemvertekening. Wanneer de itemvertekening enkel het intercept beïnvloedt, spreekt men van uniforme item bias. In dit geval heeft de ene culturele groep systematisch een hogere kans om het item correct op te lossen, onafhankelijk van het niveau van de onderliggende latente trek, vaak geoperationaliseerd als de somscore voor de subtest (Shepard, Camilli, & Averill, 1981). Bij uniforme bias zijn indirecte vergelijkingen tussen groepen wel mogelijk, aangezien het gewicht hetzelfde blijft. Zo zal eenzelfde verandering in de latente variabele nog steeds leiden tot eenzelfde verwachte verandering in de geobserveerde score. Indien echter het gewicht (en mogelijk ook het intercept) verschillen tussen groepen, spreekt men van niet-uniforme bias. De mate van vertekening voor een respondent in een bepaalde groep is dan wel afhankelijk van zijn/haar positie op de latente variabele. De vertekening is dan niet uniform voor alle posities op de latente variabele, wat zelfs een relatieve vergelijking onmogelijk maakt (Fontaine, 2005). Zo zullen personen uit verschillende groepen met eenzelfde positie voor de latente variabele, niet dezelfde verwachte score op het item behalen (Shepard et al., 1981). In deze masterproef wordt Differential Item Functioning onderzocht via logistische regressie, waarmee men zowel uniforme als niet-uniforme bias kan nagaan bij items die dichotoom gescoord zijn (Mellenbergh, 1982; Van de Vijver & Leung, 1997a, 1997b).

Niet elke vorm van vertekening heeft een even sterk nadelig effect op de testcores. Bij een vergelijking van de construct-, methode- en itemvertekening van intelligentietests bij autochtone en allochtone basisschoolleerlingen in Nederland, was het vooral methodebias die de vergelijkbaarheid van testcores nadelig beïnvloedde. Meer

bepaald de vertrouwdheid met de taal en cultuur achter de test had de grootste effecten (Helms-Lorenz & Van de Vijver, 2007). Deze bevinding werd eerder gevonden door ander Nederlands onderzoek waaruit blijkt dat de belangrijkste bron van vertekening in tests een gebrek aan kennis van de Nederlandse taal en cultuur is (Van de Vijver & Bleichrodt, 2001).

Intelligentie en SES

Sociaaleconomische status (SES) verwijst naar de positie van mensen in de sociale stratificatie met het daaraan verbonden aanzien en prestige (Verweij, 2010). De meest gebruikte indicatoren voor SES zijn opleidingsniveau, beroepsstatus en hoogte van het inkomen van de volwassen leden van een gezin (Winkleby et al., 1992; Van Berkel-van Schaik & Tax, 1990; Jeynes, 2002). Onderzoeken in industriële landen hebben aangetoond dat er belangrijke verbanden bestaan tussen SES en IQ, met correlaties tussen 0,25 en 0,50 (Eckland, 1971; Jensen, 1969; Mascie-Taylor, 1984). Deze verbanden komen al tot uiting in de kindertijd. Zo hebben kinderen met een lage SES, reeds bij de start in de kleuterschool, slechtere wiskundige vaardigheden dan kinderen met een hogere SES (Verachtert, Van Damme, Onghena, & Ghesquière, 2009). Ook in de lagere en secundaire school behalen kinderen met een lage SES minder hoge scores op algemene testen (Frederickson & Petrides, 2008; Johnson, McGue, & Iacono, 2007). Daarnaast wordt er algemeen verondersteld dat ouders met een hoge SES meer geneigd zijn om hun kinderen aan te moedigen om deel te nemen aan probleemoplossende activiteiten dan ouders met een lage SES. Mogelijk wordt dus een deel van de culturele scoreverschillen op de CoVaT-CHC verklaard door verschillen in SES. Echter kan SES nooit de enige verklaring zijn voor deze verschillen. Hunt (2011) geeft hiervoor drie redenen. De eerste reden is statistisch: SES-metingen verklaren geen 100 % van de variantie in groepsverschillen op cognitieve vaardigheidstesten. Bovendien kan SES niet verklaren waarom er verschillen gevonden worden op cultuurarme tests, zoals de Raven Progressive Matrices. Ten tweede is dat SES wederzijds gecorreleerd is met intelligentie. Het blijft moeilijk uit te maken of het effect van ouderlijke SES op de intelligentie van kinderen biologisch al dan niet cultureel bepaald is. De laatste reden waarom SES geen volledige verklaring biedt voor intelligentieverschillen is dat SES een statistische abstractie is.

Intelligentie en Acculturatie

Aangezien kennis van de dominante gastcultuur, of meer bepaald een gebrek eraan, een belangrijke bron van vertekening kan zijn, is het belangrijk om bij de interpretatie van de testgegevens rekening te houden met acculturatie. Acculturatie slaat op de processen die zich afspelen bij langdurig cultuurcontact (Redfield, Linton & Herskovits, 1936). Deze contacten kunnen veranderingen teweegbrengen op groepsniveau en het individuele niveau, zoals veranderingen in eetpatronen, politieke voorkeur en taalgebruik (Berry, 1992). Psychologische acculturatie (Graves, 1967) verwijst naar gedragsmatige veranderingen op individueel niveau, bijvoorbeeld veranderingen in waarden en attitudes (Berry, 1992). Er zijn meerdere onderzoeken gevoerd naar het verband tussen acculturatie en intelligentie. Het classificatiesysteem van Berry (1980; Berry & Sam, 1997), is het meest toonaangevende model om acculturatie-oriëntaties te conceptualiseren (Berry, Kim, Power, Young & Bujaki, 1989). Dit model bevat twee onafhankelijke dimensies van acculturatie: hoe een immigrant omgaat met de cultuur van afkomst (cultuurbehoud) en hoe men omgaat met de cultuur van het land waarin men zich gevestigd heeft, de gastcultuur (culturele aanpassing). Een combinatie van deze twee dimensies resulteert in vier acculturatie-oriëntaties: integratie, assimilatie, separatie en marginalisatie. Bij integratie heeft men positieve attitudes ten aanzien van de thuiscultuur en de gastcultuur. Er is dan zowel sprake van het behoud van de eigen culturele identiteit als adaptatie aan de gastcultuur. Assimilatie treedt op wanneer er enkel sprake is van adaptatie naar de gastcultuur en er geen behoud is van de eigen cultuur. Het omgekeerde van assimilatie is separatie, waarbij men blijft vasthouden aan de eigen cultuur en men zich niet aanpast aan de gastcultuur. Bij marginalisatie tenslotte, verwerpt men beide culturen. Gebruik makend van Berry's classificatiesysteem onderzochten Van De Vijver, Helms-Lorenz en Feltzer (1999) de effecten van acculturatie-oriëntaties op cognitieve prestatie bij 118 Nederlandse migrantkinderen. Cognitieve prestatie werd gemeten aan de hand van Snijders-Oomen Nonverbal Intelligence Test, schoolrapporten en een reactiesnelheidstest op de computer. Er werd gevonden dat zowel oudere kinderen als kinderen die een voorkeur hadden voor de integratie-oriëntatie, hogere scores behaalden op de cognitieve prestatie metingen.

Ook in België werd het effect van acculturatie op een intelligentietest onderzocht. Meer bepaald werd de relatie onderzocht tussen de acculturatie-oriëntaties van allochtone

kinderen en de prestaties op de WISC-III test in Vlaanderen (Fontaine, Schittekatte, De Clercq, & Groenvynck, 2007). De acculturatie-oriëntatie werd nagegaan aan de hand van Arends-Tóth's acculturatieschaal (2003), waarbij de items ofwel vragen naar cultuurbehoud ofwel naar culturele aanpassing. Er werden twee studies uitgevoerd, waarbij in beide studies de kinderen van Marokkaanse of Turkse herkomst waren. In de eerste studie werden kinderen uit het lager onderwijs getest. Men vond een gemiddeld totaal IQ van 90.19, een gemiddeld verbaal IQ van 89.70 en een gemiddeld perfoormaal IQ van 92.90. Deze scores lagen significant lager dan het Vlaamse gemiddelde van 100 op de WISC-III. In een tweede studie werden kinderen uit het secundair onderwijs tussen 13 en 16 jaar getest. Ook in deze groep scoorden de allochtone kinderen beduidend lager dan de Vlaamse kinderen. De grootste verschillen werden, net zoals in de eerste studie, gevonden voor het verbaal IQ. In beide studies was er een substantiële samenhang tussen de IQ-scores binnen de allochtone groep en acculturatie. Kinderen met hogere IQ-scores zijn meer gericht op de gastcultuur dan de thuiscultuur. Het omgekeerde geldt voor kinderen met lagere IQ-scores, die meer gericht zijn op de thuiscultuur dan de gastcultuur. Dit effect was het meest uitgesproken voor de verbale subtests maar trad ook op bij de performale subtests. De richting van het causale verband tussen IQ en acculturatie kan echter niet bepaald worden aan de hand van dit onderzoek. Leidt een hoger IQ tot meer aanpassing aan de gastcultuur, leidt meer aanpassing aan de gastcultuur tot een hoger IQ of is er sprake van een wederzijdse beïnvloeding tussen aanpassing en IQ? Deze vraag is nog niet beantwoord. Acculturatie-oriëntaties spelen een belangrijke rol in het onderwijs. Zo zal een migrant die van plan is om in de toekomst terug te keren naar het land van herkomst, waarschijnlijk minder inspanningen leveren in het leren van de taal en gewoonten, noch zal men de kinderen motiveren om dit te doen. Omgekeerd zal men meer inspanningen leveren indien de immigratie permanent is (Lorenz & Van de Vijver, 1995).

Onderzoeksvraag en Hypotheses

De centrale onderzoeksvraag in deze masterproef is of de CoVaT-CHC in zijn huidige vorm gebruikt kan worden voor etnische minderheden. Meer bepaald wordt er onderzocht in welke mate de scoreverschillen voor vloeiende intelligentie, gekristalliseerde intelligentie, kortetermijngeheugen en verwerkingssnelheid van de CoVaT-CHC te wijten zijn aan reële verschillen in cognitieve vaardigheden, dan wel moeten worden toegeschreven aan culturele vertekening in het meetinstrument.

Hough, Oswald en Ployhart (2001) toonden aan dat hoe hoger een test *g*-geladen is, hoe groter het gemiddelde scoreverschil is tussen de meerderheidsgroep en de minderheidsgroep. Analoog met deze bevinding en de resultaten van Bos et al. (2014) in verband met culturele scoreverschillen op de CoVaT-CHC worden volgende hypothesen gesteld:

H1: De niet-Europese kinderen scoren significant lager dan de Nederlandstalige kinderen op de subtests van vloeiende intelligentie, gekristalliseerde intelligentie kortetermijngeheugen.

H2: Europese kinderen scoren significant lager dan de Nederlandstalige groep op de subtests van vloeiende intelligentie en gekristalliseerde intelligentie.

H3: Tussen de Europese en niet-Europese kinderen worden geen significante verschillen gevonden.

H4: De grootste significante gemiddelde scoreverschillen tussen de meerderheidsgroep en minderheidsgroep worden gevonden voor de subtests van vloeiende intelligentie en gekristalliseerde intelligentie.

Niet elke subtest is even gevoelig aan itemvertekening. Aangezien de belangrijkste bron van vertekening in tests een gebrek aan kennis van de Nederlandse taal en cultuur is (Van de Vijver & Bleichrodt, 2001) wordt de volgende hypothese gesteld:

H5: De verbale subtests “Tegenstellingen” en “Geheugen A” zullen vertekende items bevatten die in het nadeel spelen voor de allochtone groep.

Er werd aangetoond dat er belangrijke verbanden bestaan tussen SES en IQ, met correlaties gaande van 0,25 tot 0,50 (Eckland, 1971; Jensen, 1969; Mascie-Taylor, 1984). Kinderen met een lage ouderlijke SES behalen minder hoge scores op algemene testen in het lager en secundair onderwijs (Frederickson & Petrides, 2008; Johnson, McGue, & Iacono, 2007). Uit deze bevindingen volgt de 6de hypothese:

H6: Er is een significante positieve correlatie tussen SES en de somscores op de vier subtesten.

Deze masterproef onderzoekt ook het verband tussen acculturatiestrategieën en scores op de CoVaT-CHC. In vergelijkend onderzoek naar de WISC-III resultaten was er een substantiële samenhang tussen de IQ-scores binnen de allochtone groep en acculturatie. Kinderen met hogere IQ-scores zijn meer gericht op de gastcultuur dan de thuiscultuur. Het omgekeerde geldt voor kinderen met lagere IQ-scores, die meer gericht zijn op de thuiscultuur dan de gastcultuur. Dit effect was het meest uitgesproken voor de verbale subtests maar trad ook op bij de performale subtests (Fontaine, Schittekatte, De Clercq, & Groenvynck, 2007). Baserend op deze bevindingen wordt de laatste hypothese gesteld:

H7: Allochtone kinderen die meer gericht zijn op de gastcultuur (culturele adaptatie) zullen hogere scores behalen dan allochtone kinderen die meer gericht zijn op de thuiscultuur (cultuurbehoud). Dit effect zal het grootste zijn voor de verbale subtests “Tegenstellingen” en “Geheugen A”.

Methode

Steekproef

Vier subtesten van de CoVaT-CHC werden afgenomen bij 231 leerlingen uit het 4de, 5de en 6de leerjaar bij 2 gemeentelijke basisscholen in Asse en 1 katholieke basisschool in Wommel, Vlaams-Brabant. Resultaten van 7 leerlingen werden verwijderd uit de dataset wegens ongeldigheid. Bijgevolg bestaat de resterende steekproefgrootte uit 224 leerlingen, met 106 jongens en 118 meisjes. De leeftijd varieert tussen 8 en 12 jaar, ($M = 10.22$; $SD = 0.92$). Voor de analyses werden de kinderen oorspronkelijk onderverdeeld in drie groepen op basis van de thuistaal. De Vlaamse groep bestaat uit kinderen wiens thuistaal Vlaams is en geen buitenlandse origine hebben ($n = 147$). Kinderen die een andere Europese-taal als thuistaal hebben of naast het Vlaams nog een andere Europese taal spreken, behoren tot de Europese groep ($n = 55$). De laatste groep, de Niet-Europese groep, bestaat uit kinderen die een niet-Europese taal spreken ($n = 22$). In elk van de drie groepen hebben 9 kinderen gedubbeld. Van de 224 kinderen hebben 212 de Belgische nationaliteit, 2 de dubbele Belgische-Nederlandse, 1 de dubbele Belgische-Mexicaanse, 3 de Armeense, 1 de Franse, 1 de Italiaanse, 1 de Marokkaanse, 2 de Moldavische en 1 de Roemeense nationaliteit.

Een acculturatie-vragenlijst werd voorgelegd aan 42 kinderen waarvan één of beide ouders in een Niet-Europees land zijn geboren. 31 kinderen scoorden het hoogst voor cultureel behoud. 9 kinderen hadden voor culturele aanpassing de hoogste score en 2 kinderen scoorden gelijk voor zowel behoud als aanpassing.

Er is een enorme variatie in zowel de thuistaal van de kinderen als in geboorteland van de ouders.

Opzet

Deze studie was een samenwerking tussen de faculteit Psychologie en Pedagogische Wetenschappen Universiteit Gent en Thomas More Antwerpen. De directies van de deelnemende scholen werden op de hoogte gebracht via mail of telefoon en hadden vervolgens een persoonlijk gesprek met de thesisstudente. Er werd gebruik gemaakt van een active informed consent. Enkel de kinderen die ouderlijke toestemming

hadden, namen deel aan dit onderzoek. Data verzameling werd gedaan door de thesisstudente en bestond uit de afname van de CoVaT-CHC-Basisversie 2015, biografische vragenlijsten en acculturatie vragenlijsten. De thesisstudente kreeg op voorhand een training in de afname van de CoVaT-CHC door Walter Magez en Marlies Tierens. Oorspronkelijk waren de testafnames gepland in oktober maar wegens het later uitkomen van de test, de verplichte stage van de thesisstudent en het zoeken naar een nieuwe school wegens een school die niet meer wou meewerken aan het onderzoek, vonden de afnames plaats in februari, maart en april.

Meetinstrumenten

CoVaT-CHC.

Oorspronkelijk was het de bedoeling om de hele de CoVaT-CHC-Basisversie 2015 (Magez et al.,) af te nemen. Maar aangezien een volledige testafname 2u30 min duurt en de scholen hier niet toe bereid waren, werd in overleg met Walter Magez beslist om slechts vier subtesten van vier brede cognitieve vaardigheden af te nemen, respectievelijk de subtests “Puntreeksen” (Gf), “Tegenstellingen” (Gc), “Geheugen A” (Gsm) en “Geheime Code” (Gs). Bij de subtest “Puntreeksen”, een niet-talige subtest, dient de puntreeks tweemaal te worden voortgezet op basis van de gevonden regel voor de opbouw van de reeks. Bij de subtest “Tegenstellingen” moet uit de geboden alternatieven het meest tegenovergestelde woord van het gegeven woord omcirkeld worden. Bij “Geheugen A” diende gedurende 4 minuten, 30 woorden uit het hoofd geleerd te worden om onmiddellijk daarna zoveel mogelijk worden schriftelijk te reproduceren. Bij “Geheime Code” tenslotte moet men zo snel mogelijk een schriftelijke opdracht uitvoeren die bestaat uit de juiste letter onder het juiste symbool te zetten.

Acculturatie vragenlijst.

Om de acculturatie-oriëntaties in kaart te brengen van kinderen wiens één of beide ouders in een Niet-Europees land geboren zijn, werd na de testafname een acculturatieschaal voorgelegd (Arends-Tóth, 2003). Deze schaal telt 52 vragen die ofwel peilen naar cultuurbehoud ofwel naar culturele aanpassing.

Demografische variabelen.

Om de achtergrondvariabelen van de kinderen in kaart te brengen, werd op voorhand samen met de informed consent, een biografische vragenlijst ingevuld. De variabelen die bevroegd werden waren geslacht, leeftijd, geboorteland, nationaliteit, thuistaal, geboorteland, geboorteland ouders, geboorteland grootouders, opleiding ouders, en beroep van de ouders. Verder werden ook de gezinssamenstelling en de gezinsgrootte bevroegd, maar deze werden niet onderzocht in deze masterproef. De sociaaleconomische status werd voor de analyses geoperationaliseerd als de opleiding van de moeder, waarbij een lager of secundair diploma gezien wordt als een lage SES en een diploma uit het hoger onderwijs als een hoge SES.

Schoolresultaten.

Om functionele equivalentie aan de hand van het nomologische netwerk na te gaan, werden de punten van Nederlands en Wiskunde voor het 1^{ste} semester opgevraagd aan de scholen, mits toestemming van de ouders. Slechts 1 school heeft deze punten doorgegeven, de andere scholen waren hier niet toe bereid.

Procedure

In februari, maart en april vonden de afnames van de CoVaT-CHC plaats. Deze gebeurde in de klaslokalen zelf en werden in groep afgenomen. Elke opgave werd eerst mondeling aan de hand van een diapresentatie uitgelegd. Daarna werden een aantal oefeningen individueel ingevuld en werden de oplossingen in groep besproken. Zodra alle deelnemers de opdracht begrepen hadden, werd gestart met de desbetreffende subtest. Aan de kinderen werd meegedeeld dat de test niet voor punten was en dus niet te vergelijken is met een toets op school. Dit om faalangst te beperken. De afnames van de subtesten duurden in totaal 50 minuten. Daarna werden de acculturatie vragenlijsten ingevuld, dit nam ongeveer 10 minuten in beslag.

Resultaten

Interne Consistentie

Om de interne consistentie voor de items van de vier schalen na te gaan werd een Cronbach's alfa berekend. Eerst voor de gehele groep van kinderen, daarna voor de allochtone en Vlaamse groep apart. In Tabel 3 staan de waarden voor Cronbach's alfa vermeld. Bij waarden tussen .60 en .80 spreekt van men een redelijk betrouwbare schaal. Indien alfa $>.80$ is de schaal zeer betrouwbaar (Cronbach, 1951). We kunnen concluderen dat over het algemeen de schalen voor beide groepen vrij betrouwbaar zijn, aangezien verschillen tussen de Vlaamse groep en de allochtone groep miniem zijn.

Tabel 3

Interne consistentie (Cronbach's alfa)

	Cronbach's alfa		
	Gehele groep	Vlaamse groep	Allochtonen groep
PR	.80	.79	.80
TG	.77	.76	.73
GH	.78	.77	.78
GC	.91	.91	.86

Multiple regressie analyse.

Om scoreverschillen tussen de drie groepen na te gaan werd gebruik gemaakt van MANOVA in SPSS. De Bonferonni-correctie werd toegepast om te corrigeren voor inflatie van Type I fouten. Uit de analyses die terug te vinden zijn in tabel 3 blijkt dat de niet-Europese groep significant van de Vlaamse groep verschilt voor de subtests "Tegenstellingen" ($M_{\text{verschil}} = -3.35, p < .01$) en "Geheugen A" ($M_{\text{verschil}} = -2.95, p < .05$), waarbij telkens de niet- Europese groep gemiddeld lager scoort dan de Vlaamse groep. Er wordt een marginaal significant verschil gevonden voor "Puntreeksen" ($M_{\text{verschil}} = -1.49, p = .058$), waarbij opnieuw de niet- Europese groep gemiddeld gezien lager scoort

dan de Vlaamse groep. Voor de subtest van verwerkingssnelheid werden er geen significante verschillen gevonden ($p = .40$).

Hypothese 1 stelt dat de niet-Europese groep significant lager scoort dan de Vlaamse groep voor de subtests van vloeiende intelligentie, gekristalliseerde intelligentie en kortetermijngeheugen. Aangezien uit tabel 4 blijkt dat de scoreverschillen voor de subtest van vloeiende intelligentie slechts marginaal significant zijn, wordt hypothese 1 gedeeltelijk bevestigd.

Verder blijkt uit de analyse dat dat de Europese groep significant lager dan de Vlaamse groep scoort op de subtests Puntreeksen ($M_{\text{verschil}} = -1.07, p < .05$) en Tegenstellingen ($M_{\text{verschil}} = -2.59, p < .001$). Voor de subtests van Geheugen ($p = .20$) en Verwerkingssnelheid ($p = .89$) worden geen significante verschillen gevonden.

Tussen de Europese en de niet-Europese groep werden voor de vier subtests geen significante verschillen gevonden. Uit een tussengroeps-ANOVA bleek bovendien dat de effecten verwaarloosbaar klein zijn met respectievelijk $\eta_p^2 = .004$ voor “Puntreeksen”, $\eta_p^2 = .007$ voor “Tegenstellingen”, $\eta_p^2 = .018$ voor “Geheugen A” en $\eta_p^2 = .008$ voor “Geheime Code”. Aangezien de verschillen tussen de Europese groep en de niet-Europese groep niet systematisch, niet significant en verwaarloosbaar zijn werd er beslist om deze twee groepen samen te nemen en er één groep van te maken. In totaal zijn er nu twee groepen, een Vlaamse groep ($n = 147$) en een allochtone groep ($n = 77$). Er werd een MANOVA uitgevoerd om de testcores te vergelijken tussen de twee nieuwe groepen. Deze analyse toonde aan dat de allochtone groep significant lager scoort voor “Puntreeksen” ($F(1, 222) = 9.41, p = .002, \eta^2 = 0.04$), “Tegenstellingen” ($F(1, 222) = 23.28, p < .001, \eta^2 = 0.09$), en “Geheugen A” ($F(1, 222) = 7.11, p = .008, \eta^2 = 0.03$). Voor de subtest van verwerkingssnelheid werden er geen significante verschillen gevonden ($F(1, 222) = 2.39, p = .124$). In Tabel 5 worden de gemiddelden en standaarddeviaties weergegeven.

Tabel 4

Verschillen in gemiddeldes voor de vier subtests tussen de Niet-Europese, Europese en Vlaamse groep.

Subtest	Groep		M _{verschil}	SE
Puntreeksen	Niet-EU	EU	-.42	.70
		Vlaams	-1.49	.63
	EU	Vlaams	-1.07*	.44
Tegenstellingen	Niet-EU	EU	-.76	1.26
		Vlaams	-3.35*	1.14
	EU	Vlaams	-2.59*	.79
Geheugen A	Niet-EU	EU	-1.50	1.26
		Vlaams	-2.95*	1.14
	EU	Vlaams	-1.45	.79
Geheime Code	Niet-EU	EU	4.67	6.56
		Vlaams	8.98	5.94
	EU	Vlaams	4.31	4.11

Noot. *SE* = standaard error

* $p < .05$.

Tabel 5

Gemiddelde totaalscores en standaarddeviaties voor de allochtone groep en de Vlaamse groep.

Groep		PR	TEG	GEH	GC
Allochtonen	<i>M</i>	8.57	15.82	12.75	103.20
	<i>SD</i>	2.84	4.06	5.02	23.46
Vlamingen	<i>M</i>	9.76	18.62	14.63	97.56
	<i>SD</i>	2.71	4.17	5.01	27.16
Totaal	<i>M</i>	9.35	17.66	13.99	99.50
	<i>SD</i>	2.81	4.34	5.08	26.04

Noot. *M* = gemiddelde; *SD* = standaarddeviatie, PR = subtest “puntreeksen”; TEG = subtest “tegenstellingen”; GEH = subtest “geheugen A”; GC = subtest “geheime code”.

Item vertekening

Item vertekening werd aan de hand van logistische regressie nagegaan voor de subtests van vloeiende intelligentie, gekristalliseerde intelligentie en kortetermijngeheugen. Voor verwerkingssnelheid kon itemvertekening niet worden nagegaan aangezien de subtest ervan bestaat uit periodes en niet uit aparte items. De variabele “groep” werd omgezet in een dummy variabele waarbij de Vlaamse groep werd gecodeerd als 0 en de allochtone groep als 1. De regressiecoëfficiënt *b* geeft dan aan of de allochtonen (categorie 1) gemiddeld hoger of lager scoren dan de Vlamingen (categorie 0). Voor elke item dat vertekening vertoont, wordt de Wald-statistiek ($df = 1$) en bijhorende *p*-waarde, de logistische regressiecoëfficiënt *b*, en de Nagelkerke’s pseudo- R^2 weergegeven om de grootte van het effect na te gaan. Het effect is verwaarloosbaar wanneer $R^2 < 0.035$, gemiddeld wanneer $0.035 < R^2 < 0.070$ en groot wanneer $R^2 > 0.070$ (Jodoin & Gierl, 2001). Model 1 (groep als predictor) geeft uniforme bias weer, model 2 (interactie tussen groep en somscore als predictor) geeft niet-uniforme bias weer.

Puntreeksen.

Voor de subtest “puntreeksen”, bestaande uit 15 items, werden geen vertekende items gevonden.

Tegenstellingen.

Voor de subtest “tegenstellingen” werden uit een totaal van 35 items, 5 items geïdentificeerd die vertekening vertonen. In Tabel 6 wordt een overzicht gegeven van de 5 vertekende items.

Tabel 6

Overzicht van vertekende items in de subtest “tegenstellingen” met behulp van logistische regressieanalyse.

Item	Model	<i>b</i>	<i>R</i>²	Wald	DF
Kachel	1	4.38	0.42 ³	3.69	1
	2	-.31*	0.42 ³	4.64*	1
Wild	1	-8.42*	0.27 ³	4.36*	1
	2	0.39	0.01 ¹	3.82	1
Gift	1	-4.61*	0.19 ³	4.05*	1
	2	0.28*	0.22 ³	5.18*	1
Aanwerven	1	5.15*	0.14 ³	6.63*	1
	2	-0.24*	0.18 ³	5.30*	1
Schuw	1	7.31**	0.20 ³	8.57**	1
	2	-0.44**	0.002 ¹	8.59**	1

Noot. Waarden voor *R*² aangeduid met 1, 2 of 3 wijzen op respectievelijk een klein, gemiddelde of grote effectgrootte (Jodoin & Gierl, 2001).

* *p* < .05. ** *p* < .01

Het item “kachel” vertoont niet-uniforme bias ($b = -.31, p < .05$) waarbij het verband tussen de itemscore en de schaalscore minder sterk is in de allochtone groep dan in de Vlaamse groep. Mogelijks kan dit item de Vlaamse groep een klein voordeel geven aan de onderkant van de schaalscore terwijl er een groot voordeel wordt gegeven aan de bovenkant van de schaalscore. Het effect is zeer groot ($R^2 = 0.42$)

Voor het item “wild” werd uniforme bias vastgesteld ($b = - 8.42, p < .05$), waarbij de allochtone groep benadeeld is; gegeven dezelfde schaalscore, scoort de allochtone groep lager dan de Vlaamse groep op het item. Ook dit effect is zeer groot ($R^2 = .27$)

Uniforme bias werd ook gedetecteerd voor het item “gift” ($b = - 4.61, p < .05$) waarbij de allochtone groep lager scoren dan de Vlaamse groep gegeven dezelfde schaalscore. Naast uniforme bias vertoonde het item “gift” ook niet-uniforme bias ($b = 0.28, p < .05$) waarbij het verband tussen de itemscore en de schaalscore sterker is in de allochtone groep dan in de Vlaamse groep. Beide effecten zijn groot met $R^2 = 0.19$ voor uniforme bias en $R^2 = 0.22$ voor niet-uniforme bias.

Voor het item “aanwerven” werd zowel uniforme ($b = 5.15, p < .05$) als niet-uniforme bias gedetecteerd ($b = -0.24, p < .05$). Uniforme bias speelde hier in het voordeel van de allochtone groep; gegeven dezelfde schaalscore scoort de allochtone groep hoger op het item dan de Vlaamse groep). Voor niet-uniforme bias was het verband tussen de itemscore en schaalscore minder sterk in de allochtone groep dan in de Vlaamse groep. Ook bij dit item kan men spreken van grote effecten met $R^2 = 0.14$ voor uniforme bias en $R^2 = 0.18$ voor niet-uniforme bias.

Net zoals bij het item “aanwerven” werd voor het item “schuw” zowel uniforme ($b = 7.31, p < .01$) als niet-uniforme bias gedetecteerd ($b = - 0.44, p < .05$) waarbij de allochtone groep hoger scoort op het item dan de Vlaamse groep. Het verband tussen de itemscore en schaalscore is minder sterk in de allochtone groep dan in de Vlaamse groep. Voor uniforme bias was het effect groot ($R^2 = 0.20$), voor de niet-uniforme bias ging het om een klein effect ($R^2 < 0.01$)

Geheugen A.

Voor de subtest “Geheugen A” werden 4 vertekende items gedetecteerd uit een totaal van 30 items. Voor alle items die vertekening vertoonden, ging het om een sterk effect ($R^2 > 0.07$)

Tabel 7

Overzicht van vertekende items in de subtest “Geheugen A” met behulp van logistische regressieanalyse.

Item	Model	<i>b</i>	R^2	Wald	DF
Waterpas	1	-2.44	0.22 ³	3.70	1
	2	-0.22*	0.24 ³	4.53*	1
Roos	1	2.72*	0.18 ³	6.53*	1
	2	-0.15	0.20 ³	3.21	1
Kano	1	-3.54*	0.31 ³	5.22*	1
	2	0.25*	0.34 ³	4.81*	1
Zeppelin	1	-2.54*	0.25 ³	0.05*	1
	2	0.196	0.27 ³	0.05	1

Noot. Waarden voor R^2 aangeduid met 1, 2 of 3 wijzen op respectievelijk een klein, gemiddelde of grote effectgrootte (Jodoin & Gierl, 2001).

* $p < .05$. ** $p < .01$

Het item “waterpas” vertoonde niet-uniforme bias ($b = -0.22$, $p < .05$) waarbij het verband tussen de itemscore en schaalscore is minder sterk in de allochtone groep dan in de Vlaamse groep.

Bij het item “roos” werd uniforme bias gedetecteerd ($b = 2.72$, $p < .05$); de allochtone groep scoort hoger op het item dan de Vlaamse groep gegeven dezelfde schaalscore.

Het item “kano” vertoonde uniforme ($b = -3.54, p < .05$) en niet-uniforme bias ($b = 0.25, p < .05$). Gegeven dezelfde schaalscore scoort de allochtone groep lager op het item in vergelijking met de Vlaamse groep. Het verband tussen de itemscore en de schaalscore is sterker in de allochtone groep dan in de Vlaamse groep.

Voor het item “zeppelin” werd uniforme bias gevonden ($b = -2.54, p < .05$) waarbij de allochtone groep een lagere score op het item behaald dan de Vlaamse groep gegeven dezelfde schaalscore.

Aparte Foutenanalyse Tegenstellingen

Voor de subtest “Tegenstellingen” werd een foutenanalyse uitgevoerd aan de hand van een Chi-kwadraat test en crosstabs vergelijking. Eerst werd nagegaan of er significante verschillen zijn binnen een groep voor gekozen foute items (within-group comparison). Daarna werden de allochtone groep en de autochtone groep vergeleken betreffende de gekozen foute items (between-groups comparison)

Er werden culturele significante verschillen gevonden voor de items “nat” en “wild”. Het item “nat” werd fout beantwoord door 5 autochtonen en 3 allochtonen. Binnen de autochtone groep hadden 5 kinderen “warm” aangeduid. Binnen de allochtone groep had 1 kind “warm” en 2 kinderen “fris” aangeduid. Dit verschil was significant ($X^2(1) = 4.444, p < .05$). De between-groups comparison was ook significant: allochtonen kiezen meer voor “fris”, autochtonen kiezen meer voor “warm” ($X^2(1) = 4.444, p < .05$).

Het item “wild” werd fout beantwoord door 102 autochtonen en 69 allochtonen. Binnen de autochtone groep kozen 27 kinderen voor “woest”, 40 kinderen voor “rustig” en 35 kinderen voor “kalm”. Binnen de allochtone groep kozen 6 kinderen voor “woest”, 40 kinderen voor “rustig” en 32 kinderen voor “kalm”. Voor beide groepen verschilden “woest” significant van “rustig” en “kalm” waarbij “woest” het minst gekozen werd. ($X^2(2) = 8.590, p < .05$). Between-groups comparison toonde een significant verschil aan voor “woest”, waarbij allochtonen het minst voor “woest” kiezen dan autochtonen. ($X^2(2) = 8.590, p < .05$). Voor “rustig” en “kalm” werden geen significante verschillen gevonden. In tabel 8 worden de within-groups en between-groups vergelijkingen weergegeven.

Tabel 8

Within-group comparison “nat”

			Nat	
			Warm	Fris
Groep	Alloctonen	Aantal	1 _a	2 _b
	Vlamingen	Aantal	5 _a	0 _b

Noot: elke subscript letter duidt een subset van categorieën voor het item “nat” aan waarvan de kolomverhoudingen niet significant van elkaar verschillen op het 5% significante niveau.

Tabel 9

Between-group comparison “nat”

			Groep	
			Alloctonen	Vlamingen
Nat	Warm	Aantal	1 _a	5 _b
	Fris	Aantal	2 _a	0 _b

Noot: elke subscript letter duidt een subset van categorieën voor het item “nat” aan waarvan de kolomverhoudingen niet significant van elkaar verschillen op het 5% significante niveau.

Tabel 10

Within-group comparison “wild”

			wild		
			woest	rustig	kalm
Groep	Allochtonen	Aantal	6 ^a	31 ^b	32 ^b
	Vlamingen	Aantal	27 ^a	40 ^b	35 ^b

Noot: elke subscript letter duidt een subset van categorieën voor het item “wild” aan waarvan de kolomverhoudingen niet significant van elkaar verschillen op het 5% significante niveau.

Tabel 11

Between-group comparison “wild”

			Groep	
			Allochtonen	Vlamingen
Wild	Woest	Aantal	6 _a	27 _b
	Rustig	Aantal	31 _a	40 _a
	Kalm	Aantal	32 _a	35 _a

Noot: elke subscript letter duidt een subset van categorieën voor het item “wild” aan waarvan de kolomverhoudingen niet significant van elkaar verschillen op het 5% significante niveau.

Effecten van Achtergrondvariabelen

Schoolresultaten “Nederlands” en “Wiskunde”.

In totaal werden van 91 kinderen uit 1 school de resultaten verkregen voor “Nederlands” en “Wiskunde” in het eerste semester. Er werd een zwakke, maar positieve significante correlatie gevonden tussen punten voor Wiskunde en de somscore op de subtest “puntreeksen”; hoe hoger de punten voor Wiskunde, hoe hoger de score op de subtest “puntreeksen”: $r = 0.293$, $p < .01$. Voor de punten voor Nederlands en somscore op de subtest “tegenstellingen” werd een redelijk sterk positieve significant verband gevonden ($r = 0.420$, $p < .01$). De correlatie tussen punten voor Nederlands en de somscore van de subtest “geheugen A” was niet significant ($p = .636$).

SES.

Sociaaleconomische status werd geoperationaliseerd als opleiding van de moeder. In totaal hebben 66 kinderen, waarvan 33 uit de allochtone groep en 33 uit de Vlaamse groep een lage SES. 158 kinderen waarvan 44 uit de allochtone groep en 114 uit de Vlaamse groep een hoge SES. Er werd gevonden dat SES significant positief gecorreleerd is met de schaalscore op vloeiende intelligentie ($r = 0.16, p = .02$) en gekristalliseerde intelligentie ($r = .27, p < .01$). Voor de schaalscore van kortetermijngeheugen en verwerkingssnelheid werden geen significante correlaties gevonden met respectievelijk $p = .08$ en $p = .15$.

Geslacht.

In totaal bestaat de steekproef uit 106 jongens en 118 meisjes. Er werd een significant positief verband gevonden tussen geslacht en de somscore van verwerkingssnelheid ($r = .228, p = .001$) waarbij de meisjes ($M = 105.11, SD = 25.92$) hoger scoren dan de jongens ($J = 93.26, SD = 24.83$).

Leeftijd.

Voor testleeftijd werden er significante positieve verbanden gevonden voor tegenstellingen ($r = .18, p = .01$), geheugen ($r = .21, p = .01$) en verwerkingssnelheid ($r = .30, p < .01$)

Effect van Acculturatiestrategie

Met behulp van lineaire regressie werd het verband tussen acculturatiestrategie en de somscores voor de 4 schalen nagegaan. Voor de subtesten van vloeiende intelligentie, gekristalliseerde intelligentie, kortetermijngeheugen en verwerkingssnelheid was het model niet significant (($F(2,39) = 1.719, p = .193$), ($F(2,39) = 1.495, p = .237$), ($F(2,39) = .551, p = .581$)) Voor tegenstellingen was het model wel significant; ($F(2,39), p = .011$). Onderstaande tabel geeft een overzicht van de regressieanalyse.

Tabel 12

Lineaire regressie tussen subtest “tegenstellingen” en acculturatie-oriëntatie

TEGTOT	b^*	R^2	$F(2,39)$	t	p
Behoud	-.09	.21	5.130	-.63	.53
Adaptatie	.45	.21	5.130	3.17	.01

Uit deze analyse blijkt dat $R^2 = .208$. Ongeveer 21 % van verschillen in de schaal-score van gekristalliseerde intelligentie voor een groep allochtonen wordt verklaard door de acculturatie-oriëntatie. De regressiecoëfficiënt b^* geeft de unieke bijdrage van elke voorspellende variabele weer. Enkel “culturele aanpassing” ($b = .45$, $p = .01$) is positief geassocieerd met de subtest van gekristalliseerde intelligentie; hoe meer men aangepast is aan de gastcultuur, hoe hoger de score.

Achtergrondvariabelen en acculturatie-strategie.

Er werd een meervoudige regressieanalyse uitgevoerd voor de scores op de 4 subtesten, met achtergrondvariabelen en acculturatie-strategie als onafhankelijke variabelen. Interactie tussen cultureel behoud en adaptatie was niet significant voor de vier scores ($p > .05$), bijgevolg werd er geen verdere onderverdeling gemaakt in de acculturatie-strategie. Achtergrondvariabelen en acculturatieoriëntatie werden in 2 aparte blokken geplaatst zodat de fit werd nagegaan voor 2 modellen waarbij in het 1^{ste} model de predictoren enkel bestonden uit achtergrondvariabelen en in het 2^{de} model de predictoren bestonden uit de achtergrondvariabelen en acculturatiestrategie.

Voor alle subtesten waren beide modellen niet significant. Voor geen enkele predictor was de regressiecoëfficiënt significant ($p > .05$).

Equivalentie

Indicatie van equivalentie werd nagegaan aan hand van principal axis factoring, eerst voor, daarna apart per groep om na te gaan of de interne structuur vergelijkbaar voor de allochtonen en de Vlamingen. Tabel 18 geeft een overzicht van de factormatrices voor de totaalgroep, de Vlaamse groep en de allochtone groep.

Tabel 18

Factormatrices voor de gehele groep, de Vlaamse groep en de allochtone groep.

Groep	Schaal	Factor	
		1	2
Gehele groep	Gf	.84	
	Gc	.39	
	Gsm	.20	
	Gs	.18	
Vlaamse groep	Gf	.87	
	Gc	.38	
	Gsm	.16	
	Gs	.32	
Allochtone groep	Gf	.22	.63
	Gc	-.10	.40
	Gsm	.83	-.09
	Gs	.27	-.10

Voor de gehele groep wordt er 1 factor gevonden waarbij vloeiende intelligentie (.84) en gekristalliseerde intelligentie (.39) de hoogste factorlading vertonen. Alle ladingen zijn positief; indien men goed scoort op 1 schaal, scoort men ook goed op de andere 3 schalen.

Net zoals de gehele groep wordt er voor de Vlaamse groep ook 1 factor gevonden, met de hoogste factorladingen voor vloeiende intelligentie en gekristalliseerde intelligentie, ook hier zijn alle factorladingen positief.

Voor de allochtone groep worden echter 2 factoren gevonden. Voor de eerste factor heeft kortetermijngeheugen de grootste lading (.83). Gekristalliseerde intelligentie heeft een negatieve lading van $-.10$. Dit zou betekenen dat indien men goed scoort op de subtests van vloeiende intelligentie, kortetermijngeheugen en verwerkingssnelheid, men laag scoort op de subtest van gekristalliseerde intelligentie. Op de tweede factor laden vloeiende (.63) en gekristalliseerde intelligentie (.40) het hoogst. Kortetermijngeheugen ($-.09$) en verwerkingssnelheid ($-.10$) hebben negatieve ladingen. Indien men hoge scores behaalt voor de subtests van vloeiende en gekristalliseerde intelligentie, behaalt men lagere scores voor de subtests van kortetermijngeheugen en verwerkingssnelheid. “Positive manifold” wordt dus niet teruggevonden voor de allochtone groep.

Discussie

Bespreking resultaten

Vooraleer de hypothesen nagegaan zijn, zijn eerst de interne consistenties van de schalen berekend aan de hand van Cronbach's alfa voor zowel de gehele groep als de allochtone groep en de Vlaamse groep apart. Uit de resultaten blijkt dat over het algemeen, de items van de schalen voor beide groepen redelijk tot vrij betrouwbaar zijn.

Culturele scoreverschillen.

Hypothese 1 stelt dat de niet-Europese groep significant lager scoort dan de Vlaamse groep op de subtests van vloeiende intelligentie, gekristalliseerde intelligentie en kortetermijngeheugen. De resultaten toonden significante verschillen aan voor gekristalliseerde intelligentie en kortetermijngeheugen, waarbij de niet-Europese groep gemiddeld lager scoort. Voor vloeiende intelligentie werd een marginaal significant verschil gevonden. Hypothese 1 wordt dus gedeeltelijk bevestigd. De significante verschillen voor gekristalliseerde intelligentie en kortetermijngeheugen zijn overeenkomstig met de resultaten uit eerder onderzoek van Bos et al (2014). Dat er geen significant verschil werd gevonden voor vloeiende intelligentie, is waarschijnlijk te wijten aan de kleine steekproef, de niet-Europese groep bestond uit slechts 22 kinderen.

Een kleine steekproef zorgt immers voor minder power waardoor significante effecten mogelijk niet gedetecteerd worden (Cohen, Cohen, West, & Aiken, 2013). Bij vergelijkingen tussen de Vlaamse groep en de allochtone groep (de niet-Euroese groep + Europese groep) werd er immers wel een significant verschil gevonden waarbij

Hypothese 2 die stelt dat Europese kinderen significant lager scoren dan de Vlaamse groep voor vloeiende intelligentie en gekristalliseerde intelligentie wordt bevestigd. Tussen de Europese en de niet-Europese kinderen werden geen significante verschillen gevonden, waardoor hypothese 3 ook bevestigd is. Deze resultaten liggen in lijn met de bevindingen uit onderzoek door Bos et al (2014)

Aangezien er geen significante verschillen werden gevonden tussen de Europese groep en de niet-Europese groep werd beslist om van de twee groepen, één groep te maken, die de allochtone groep genoemd werd. Vergelijkingen tussen de allochtone groep en Vlaamse groep toonden significant lagere scores aan voor vloeiende intelligentie, gekristalliseerde intelligentie en kortetermijngeheugen. Hypothese 4 die stelt dat de gemiddelde significante verschillen tussen de meerderheidsgroep en de minderheidsgroep het grootst zijn voor de subtests van vloeiende intelligentie en gekristalliseerde intelligentie werd niet bevestigd. Uit de resultaten bleek immers dat de gemiddelde scoreverschillen het grootst zijn voor subtests van gekristalliseerde intelligentie en kortetermijngeheugen. Een mogelijke verklaring hiervoor is dat de subtests van gekristalliseerde intelligentie kortetermijngeheugen vertekende items bevatten die ervoor zorgen dat de allochtone groep lager scoort. De subtests van gekristalliseerde intelligentie en kortetermijngeheugen zijn immers verbale subtests, die over het algemeen gevoeliger zijn aan vertekening dan performale subtests (Maes, Smeets, & Schittekatte, 2015; Sierens et al., 2006; te Nijenhuis & Van der Flier, 2003)

Itemvertekening.

Voor de subtest van vloeiende intelligentie werden geen vertekende items gevonden. Dit is niet verassend aangezien performale subtests minder gevoelig zijn aan vertekening (Maes, Smeets & Schittekatte, 2015).

Voor de subtest van gekristalliseerde intelligentie werden 5 vertekende items gedetecteerd, respectievelijk “kachel”, “wild”, “gift”, “aanwerven” en “schuw”. De uniforme vertekening voor “wild” en “gift” speelden in het nadeel van de allochtone groep. Dit is tegenstrijdig aan wat men zou verwachten. Beide woorden klinken namelijk in het Engels hetzelfde. De uniforme vertekening voor “aanwerven” en “schuw” speelden in het voordeel van de allochtonen. “Aanwerven” is een werk gerelateerd woord, waardoor de situatie van de ouders mogelijk anders is voor de allochtone kinderen en de Vlaamse kinderen. “Schuw” is een weinig gebruikt woord, waardoor het tegenstrijdig is dat allochtone kinderen hier hoger op scoren. Mogelijk zijn de tegenstrijdige bevindingen te wijten aan de kleine steekproef van allochtonen. Er moet ook in rekening gebracht worden dat bij de subtest van gekristalliseerde intelligentie, het juiste antwoord gekozen moest worden uit een reeks van gegeven woorden, waardoor het mogelijk is dat er gegokt werd.

Voor de subtest van kortetermijngeheugen werden 4 vertekende items gedetecteerd, respectievelijk “waterpas”, “roos”, “kano” en “zeppelin”. De uniforme vertekening voor “roos” speelden in het voordeel van de allochtonen. “Roos” klinkt zowel in het Engels als in het Frans hetzelfde, wat een mogelijke verklaring zou zijn. Voor “kano” en “zeppelin” speelden de uniforme vertekening in het nadeel van de allochtonen. Mogelijks woont het merendeel van allochtonen in een stad, wat ervoor zorgt dat zij niet of zelden in contact komen met een kano, aangezien er in een stad minder mogelijkheid is om met een kano te varen. Het woord “zeppelin” wordt vermeld in een Vlaams kinderliedje wat er mogelijks voor zorgt dat Vlaamse kinderen meer in contact zijn gekomen met het woord dan allochtone kinderen. Hypothese 5 die stelt dat de verbale subtests “Tegenstellingen” en “Geheugen A” vertekende items bevatten die in het nadeel spelen van de allochtone groep, wordt niet bevestigd aangezien er ook vertekende items werden gevonden die in het nadeel speelden voor de Vlaamse kinderen.

SES.

Hypothese 6 stelt dat SES significant positief gecorreleerd is met elk van de 4 schaalscores. Uit de resultaten blijkt dat SES enkel positief gecorreleerd is met de schaalscores van vloeiend intelligentie en gekristalliseerde intelligentie, waardoor hypothese 6 niet bevestigd kan worden. Dit resultaat gaat in tegen de bevinding dat kinderen met een lage SES lager scores behalen op algemene testen in het onderwijs (Frederickson & Petrides, 2008; Johnson, McGue, & Iacono, 2007). Een mogelijke verklaring waarom deze masterproef geen significante correlaties vindt tussen SES en de schaalscores van kortetermijngeheugen en verwerkingssnelheid, is dat SES geoperationaliseerd werd aan de hand van slechts 1 indicator, respectievelijk opleidingsniveau van de moeder. De beste indicatoren voor SES zijn opleidingsniveau, beroepsstatus en hoogte van het inkomen van de volwassen leden van een gezin (Winkleby et al., 1992; Van Berkel-van Schaik & Tax, 1990; Jeynes, 2002). Door enkel opleidingsniveau van de moeder te gebruiken, is er mogelijk geen goede indicatie is van SES.

Acculturatie-oriëntatie.

Deze masterproef is ook het verband tussen acculturatie-oriëntaties en de schaalscores nagegaan. Op basis van eerder onderzoek door Fontaine et al. (2007) werd hypothese 7 opgesteld, waar verwacht werd dat allochtone kinderen die meer aangepast zijn aan de gastcultuur, hogere scores zouden behalen dan allochtone kinderen die meer gericht zijn op de thuiscultuur. Dit effect zou het grootste moeten zijn voor de verbale subtests. Resultaten toonden enkel een significant verband aan voor de subtest van tegenstellingen waarbij culturele aanpassing een significant positief effect had op de schaalscore. Bij controle voor de achtergrondvariabelen viel het significant effect van acculturatie-oriëntatie op de schaalscore gekristalliseerde intelligentie weg. In dit model hadden de achtergrondvariabelen geslacht, leeftijd, schoolresultaten en SES ook geen significant effect meer. Mogelijks is ook deze tegenstrijdige bevinding te wijten aan de kleine steekproef van kinderen die een acculturatie-vragenlijst hebben ingevuld ($n = 42$).

Equivalentie.

Positive manifold, de robuuste bevinding dat subtesten van een IQ-batterij altijd positief gecorreleerd zijn, wordt in deze masterproef teruggevonden voor de gehele groep en de Vlaamse groep, maar niet voor de allochtone groep. Factoranalyse voor de allochtone groep resulteerde in 2 factoren waarbij de subtest van gekristalliseerde intelligentie een negatieve lading vertoont in de 1^{ste} factor en de subtests van kortetermijngeheugen en verwerkingssnelheid, negatieve lading vertoont in de 2^{de} factor. Indien de allochtone kinderen die van het verwachte lineaire verband afweken allemaal waren blijven zitten, was er een mogelijke verklaring voor de negatieve ladingen. Uit plots bleek dat dit echter niet het geval was. Aangezien de bevindingen voor de allochtone groep onverwachts en tegenstrijdig zijn met bevindingen uit de wetenschappelijke literatuur is het mogelijk dat het om toevals-fluctuaties gaat, te wijten aan een te kleine steekproef.

Beperkingen

De grootste beperking van dit onderzoek is de kleine steekproef voor de allochtone groep ($n = 77$). Een te kleine steekproef kan ervoor zorgen dat de power om effecten te detecteren te klein is de verschillen mogelijks te wijten zijn aan toevals-fluctuaties (Cohen et al., 2013). Bovendien lagen de deelnemende scholen allemaal in dezelfde regio, Vlaams-Brabant waardoor er geen geografische spreiding is.

Een andere beperking ligt in de methode van het onderzoek. De CoVaT-CHC onderzoekt 5 brede vaardigheden, waarvan 4 brede vaardigheden uit 2 subtests bestaan. Omwille van praktische redenen werden in deze masterproef 4 brede vaardigheden onderzocht aan de hand van 1 subtest, waardoor de schaalscores niet volledig zijn.

Binnen de steekproef was er een enorme spreiding was voor etnische afkomst en thuistaal, maar aangezien het om een vrij kleine steekproef gaat, was het niet mogelijk om veel onderverdelingen te maken voor de groepen. Dit heeft als gevolg dat Franstalige Belgische kinderen ook in de allochtone groep terecht kwamen, waardoor het moeilijk is om na te gaan of scoreverschillen te wijten zijn aan taalverschillen of culturele verschillen.

Om gebruik te kunnen maken van logistische regressieanalyse werden items binair gescoord. Hierdoor werd er afgeweken van de verbeter sleutels van de CoVaT-CHC waarbij items niet binair gescoord worden, wat kan gezien worden als een beperking.

Toekomstig onderzoek

Er wordt aanbevolen om culturele equivalentie en vertekening in de CoVaT-CHC verder te onderzoeken voor een grotere steekproef te onderzoeken. Indien het onderzoek gebeurt aan de hand van een masterproef wordt aangeraden om 3 of meer thesisstudenten in te schakelen, waardoor er meer participanten bereikt kunnen worden. De steekproef dient ook geografisch verspreid te liggen. Er wordt verder aanbevolen om scholen te zoeken die akkoord gaan met een volledige afname van de CoVaT-CHC die ongeveer 2 lesuren in beslag zou nemen.

Conclusie

Deze masterproef levert een bijdrage aan het onderzoek over culturele equivalentie en vertekening in de CoVaT-CHC. Er worden significante scoreverschillen gevonden voor vloeiende intelligentie, gekristalliseerde intelligentie en kortetermijngeheugen waarbij de allochtone groep significant lager scoort dan de Vlaamse groep. Er was sprake van culturele vertekening gevonden voor items van de verbale subtests, echter was de vertekening niet altijd in het nadeel van de allochtone groep. Verder werd binnen de allochtone groep, een positief significant verband gevonden tussen de subtest van gekristalliseerde intelligentie en culturele aanpassing; hoe hoger de score op culturele aanpassing, hoe hoger de score op de subtest. Voor de subtests van vloeiende intelligentie, kortetermijngeheugen en verwerkingsnelheid werden geen significante verbanden gevonden met acculturatie-oriëntatie. Er bestaat nog heel wat onduidelijkheid in verband met equivalentie aangezien voor de allochtone groep 2 factoren werden gevonden waarbij factorladingen negatief en positief zijn. Rekening houdend met de kleine steekproef als grootste beperking, is verder onderzoek naar culturele vertekening en equivalentie in de CoVaT-CHC aangewezen.

Referenties

- Abad, F. J., Colom, R., Juan-Espinosa, M., & Garcia, L. F. (2003). Intelligence differentiation in adult samples. *Intelligence*, 31(2), 157-166.
- Arends-Tóth, J. V. (2003). *Psychological acculturation of Turkish migrants in the Netherlands: Issues in theory and assesment*. Amsterdam: Dutch University Press.
- Berk, R. A. (1982). *Handbook of methods for detecting test bias*. Baltimore: Johns Hopkins University Press
- Berkel-van Schaik van, AB., & Tax, B. (1990). *Naar een standaard operationalisatie van sociaaleconomische status voor epidemiologisch en sociaal-medisch onderzoek*. Ministry of Welfare, National Health and Culture. Retrieved from <http://www.nationaalkompas.nl/bevolking/segv/wat-is-sociaaleconomische-status/>
- Berry, J. W. (1980). Acculturation as varieties of adaptation. In A. Padilla (Ed.), *Acculturation: Theory, models and findings* (pp. 9–25). Boulder: Westview.
- Berry, J. W. (1992). Acculturation and adaptation in a new society. *International Migration*, 30, 69–85.
- Berry, J. W., Kim, U., Power, S., Young, M., & Bujaki, M. (1989). Acculturation attitudes in plural societies. *Applied Psychology*, 38, 185–206
- Berry, J. W., & Sam, D. (1997). Acculturation and adaptation. In J. W. Berry, M. H. Segall, & C. Kagitcibasi (Eds.), *Handbook of cross-cultural psychology, Vol. 3, Social behavior and applications* (pp. 291–326). Boston: Allyn and Bacon
- Bos, A., Magez, W., Tierens, M., Van Parijs, K., & Decaluwé, V. (2014). Testing non-native Dutch speakers with the Cattell-Horn-Carroll-based Dutch Cognitive Ability Test. Conference of the international test commission. San Sebastian, Spain, 2-5 july, 2014
- Carroll, J. B. (1993). *Human cognitive abilities*. Cambridge: Cambridge University Press.
- Cattell, R. B. (1957). *Personality and motivation: Structure and measurement*. Yonkers-on-Hudson, NY: World Book Co.

- Cattell, R. B. (1971). *Abilities: Their structure, growth, and action*. Boston: Houghton Mifflin.
- Cattell, R. B. (1987). *Intelligence: Its structure, growth, and action*. New York: North-Holland.
- Choenni, C.E.S., & Van der Zwan, T.J. (1986). Allochtonen en psychologische test. *LBR-bulletin*, 2(5): 17-21.
- Cohen, J., Cohen, P., West, S. G., & Aiken, L. S. (2013). *Applied multiple regression/correlation analysis for the behavioral sciences*. Mahwah : Erlbaum.
- Colom, R., Abad, F. J., Garcia, L. F., & Juan-Espinosa, M. (2002). Education, Wechsler's Full Scale IQ, and g. *Intelligence*, 30(5), 449-462.
- Cronbach, L. J. (1951). Coefficient alpha and the internal structure of tests. *Psychometrika*, 16(3), 297-334.
- Deary, I. J., & Stough, C. (1996). Intelligence and inspection time: Achievements, prospects, and problems. *American Psychologist*, 51(6), 599-608.
- Deregowski, J., & Serpell, R. (1971). Performance on a sorting task: a cross-cultural experiment. *International Journal of Psychology*, 6, 273-281.
- Detterman, D. K. (1991) Reply to Deary and Pagliari: Is g intelligence or stupidity? *Intelligence*, 15(2), 251-255.
- Detterman, D. K., & Daniel, M. H. (1989). Correlations of mental tests with each other and with cognitive variables are highest in low IQ groups. *Intelligence*, 13, 349-360.
- Eckland, B. K. (1971). Social class structure and the genetic basis of intelligence. In: *"Intelligence: Genetic and Environmental Influences,"* (R.Cancro, ed.), pp. 65-76. Grune and Stratton, New York.
- Extra, G., & Verhoeven, L. (1985). Bilingualism and bilingual education at elementary school level. *Pedagogische Studiën*, 62, 3-24.

- Flanagan, D.P., & Harrison, P.L. (2012). *Contemporary intellectual assessment: Theories, tests, and issues*. New York: Guilford.
- Fontaine, J. R. J. (2005). Equivalence. *Encyclopedia of Social Measurement: Vol. I*. (pp. 803-813). San Diego, CA: Academic Press.
- Fontaine, J.R.J., Schittekatte, M., De Clercq, S., & Groenvynck, H. (2007). *The relationship between acculturation orientation and WISC-III performance among Flemish pupils from Moroccan and Turkish descent*. Manuscript in preparation.
- Frederickson, N. & Petrides, K. V. (2008). Ethnic, gender and socio-economic group differences in academic performance and secondary school selection: a longitudinal analysis. *Learning and Individual Differences, 18*, 144-151.
- Gass, S., & Varone, E. (1991). Miscommunication in normative speaker discourse. In: Coupland, N., Giles, H., Wiemann, J.M. (Eds.), *Miscommunication and Problematic Talk*. Sage, Newbury Park, CA, pp. 121– 145.
- Gottfredson, L. S. (1997). Why g matters; The complexity of everyday life. *Intelligence, 24*(1), 79-132.
- Gottfredson, L. S. (1998). Jensen, Jensenism, and the sociology of intelligence. *Intelligence, 26*(3), 291-299.
- Gottfredson, L. S. (2007) Innovation, fatal accidents, and the evolution of general intelligence. In M. J. Roberts (Ed.), *Integrating the mind; Domain general vs domain specific processes in higher cognition*. New York: Psychology Press, 387-425.
- Graves, T. (1967). Psychological acculturation in a tri-ethnic community. *South-Western Journal of Anthropology, 23*, 337-350.
- Gustafsson, J.-E., & Undheim, J. O. (1996). Individual differences in cognitive functions. In D. Berliner, & R. Calfee (Eds.), *Handbook of Educational Psychology*, New York: Macmillan, 1996, pp. 186–242.
- Helms-Lorenz, M., & Van de Vijver, F. J. R. (1995). Cognitive assessment in education in a multicultural society. *European Journal of Psychological Assessment, 11*, 158 – 169

- Helms-Lorenz, M. & Van de Vijver, F.J.R. (2007). *Can bias explain ethnic performance differences between immigrant and mainstream primary school pupils in the Netherlands* (manuscript ter publicatie aangeboden)
- Herrnstein, R. J., & Murray, C. (1994). *The Bell curve: Intelligence and class structure in American life*. New York: The Free Press.
- Ho, D. Y. F. (1996). Filial piety and its psychological consequences. In M. H. Bond (Ed.), *The handbook of Chinese psychology* (pp. 155-165). Hong Kong: Oxford University Press.
- Hofstee, W.K.B. (1990). Toepasbaarheid van psychologische tests bij allochtonen. *De Psycholoog*, 25, 291-294.
- Hofstee, W.K.B., Campbell, W.H., Eppink, A., Evers, A., Joe, R.C., Koppel, J.M.H., Van de Zweers, H., Choenni, C.E.S., & Van der Zwan, T.J. (1990). *Toepasbaarheid van psychologische tests bij allochtonen* (LBR-reeks, nr II.). Rotterdam: Landelijk Bureau Racismebestrijding.
- Holland, P. W., & Wainer, H. (1993). *Differential item functioning*. Hillsdale, NJ: Erlbaum
- Horn, J. L. (1985). Remodeling old models of intelligence? In B. B. Wolman (Ed.), *Handbook of intelligence: Theories, measurements, and applications?* New York: Wiley, pp. 267-300.
- Horn, J., & McArdle, J. (1992). A practical and theoretical guide to measurement invariance in aging research. *Experimental Aging Research*, 18(3), 117-144.
- Hough, L., Oswald, F., & Ployhart, R. (2001). Determinants, detection and amelioration of adverse impact in personnel selection procedures: Issues, evidence and lessons learned. *International Journal of Selection and Assessment*, 9(1), 152-194.
- Hui, C.H. and H.C. Triandis. 1989. Effects of cultural and response format on extreme response style. *Journal of Cross Cultural Psychology* 20(3): 296–309.

- Hunt, E. (1995) The role of intelligence in modern society. *American Scientist*, 83 (July-August), 356-368.
- Hunt, E. (2011). *Human intelligence*. Cambridge: Cambridge University Press.
- Hunter, J. E., & Hunter, R. E. (1984). Validity and utility of alternative predictors of job performance. *Psychological Bulletin*, 96, 72-98.
- Jensen, A. R. (1969). How much can we boost IQ and scholastic achievement? *Harvard Educational Review*, 39, 1–123.
- Jensen, A. R. (1998). *The g factor: The science of mental ability*. Westport, CT: Praeger Publishers/Greenwood Publishing Group.
- Jeynes, W. H. (2002). The challenge of controlling for SES in social science and education research. *Educational Psychology Review*, 14, 205-221
- Johnson, W., McGue, M., & Iacono, W. G. (2007). Socioeconomic status and school grades: placing their association in broader context in a sample of biological and adoptive families. *Intelligence*, 35, 526-541.
- Jodoin, M., & Gierl, M. (2001). Evaluating Type I error and power rates using an effect size measure with logistics regression procedure for DIF detection. *Applied Measurement in Education*, 14, 329-349.
- Keith, T. Z., & Reynolds, M. R. (2010). CHC theory and cognitive abilities: What we've learned from 20 years of research. *Psychology in the Schools*, 47, 635–650.
- Legree, P. J., Pfifer, M. E., & Grafton, F. C. (1996). Correlations among cognitive abilities are lower for higher ability groups. *Intelligence*, 23(1), 45-57.
- Lynn, R., Backhoff, E., & Contreras, L.A. (2005). Ethnic and racial differences on the standard progressive matrices in Mexico. *Journal of Biosocial Sciences*, 37, 107-113.
- Maes, B., Smeets, S., & Schittekatte, M. (2015). Diagnostisch onderzoek van intelligentie. In: Bosmans, G., Claes, L., Bijttebier, N., & Noens I. *Diagnostiek bij kinderen, jongeren en gezinnen* (pp; 13-36). Leuven, Acco.

- Magez, W., & De Cleen, W. (2007). *Intelligentiemeting in nieuwe banen: De integratie van het CHC-model in de psychodiagnostische praktijk*. Brasschaat: CAP vzw.
- Magez, W., & Stinissen, H. (2010). *Diagnostiek bij allochtonen. Schoolpsychologisch onderzoek met psychodiagnostische tests voor intelligentie en cognitieve vaardigheden*. Brussel: VCLB Service.
- Magez, W., & Tierens, M. (2015). *Introductie tot het project CoVaT-CHC Basisversie: de chronologie*. Antwerpen: Psychodiagnostisch Centrum, Thomas More.
- Magez, W., Tierens, M., Van Huynegem, J., Van Parijs, K., Decaluwé, V. & Bos, A. (2015). *CoVaT-CHC Basisversie*. Antwerpen: Psychodiagnostisch Centrum, Thomas More.
- Magez, W., Tierens, M., Van Huynegem, J., Van Parijs, K., Decaluwé, V. & Bos, A. (2015). *CoVaT-CHC Basisversie: Cognitieve vaardigheidstest volgens het CHC-model. Instructie Handleiding*. Antwerpen: Psychodiagnostisch Centrum en CAPvzw.
- Mascie-Taylor, C. G. N. (1984). Biosocial correlates of IQ. In C. J. Turner & H. J. Miles (Eds.), *The biology of human intelligence* (pp. 99-127). London: Proceedings of the Eugenics Society.
- McDermott, P. A., Fantuzzo, J. W., & Glutting, J. J. (1990). Just say no to subtest analysis: a critique on Wechsler theory and practice. *Journal of Psychoeducational Assessment*, 8, 290-302.
- McGrew, K. S., & Flanagan, D. P. (1998). *The intelligence test desk reference (ITDR): a Gf-Gc cross-battery approach to intelligence test interpretation*, Boston, MA: Allyn and Bacon.
- McGrew, K. S., Woodcock, R. W. (2001). *Technical manual: Woodcock Johnson III*. Itasca, IL: Riverside Press.
- Meiring, D., Van de Vijver, A.J.R., Rothmann, S., & Barrick, M.R. (2005). Construct, item, and method bias of cognitive and personality measures in South Africa. *South African Journal of Industrial Psychology*, 31(1), 1-8

- Mellenbergh, G.J. (1982). Contingency table models for assessing item bias. *Journal of Educational Statistics*, 7(2), 105-118.
- Neisser, U., Boodoo, G., Bouchard Jr., T. J., Boykin, A. W., Brody, N., Ceci, S. J., Halpern, D. F., Loehlin, J. C., Perloff, R., Sternberg, R.J., & Urbina, S. (1996) Intelligence: knowns and unknowns. *American Psychologist*, 51, 77-101.
- Redfield, R., Linton, R., & Herskovits, M. H. (1936). Memorandum on the study of acculturation. *American Anthropologist*, 38, 149-152.
- Ree, M. J., & Earles, J. A. (1992). Intelligence is the best predictor of job performance. *Current Directions in Psychological Science*, 1, 86-89.
- Rushton, J. P., Cvorovic, J., & Bons, T. A. (2007). General mental ability in South Asians: Data from three Roma (Gypsy) communities in Serbia. *Intelligence*, 35(1), 1-12.
- Rushton, J. P., & Skuy, M. (2000). Performance on Raven's Matrices by African and White university students in South Africa. *Intelligence*, 28 (4), 251-265.
- Rushton, J. P., Skuy, M. & Fridjhon, P. (2003). Performance on Raven's Advanced Progressive Matrices by African, East Indian, and White engineering students in South Africa. *Intelligence*, 31(2), 123-137.
- Sackett, P. R., & Wilk, S. L. (1994). Within-group norming and other forms of score adjustment in preemployment testing. *American Psychologist*, 49, 929-954.
- Sarnacki, R. E. (1979). An examination of test-wise in the cognitive test domain. *Review of Educational Research*, 49, 252-279.
- Schittekatte, M. & Hurks, P. (2015). De WISC-III-NL. In Resing, W.: *Handboek intelligentietheorie en Testgebruik*. Pearson: Amsterdam.
- Schmidt, F.L., & Hunter, J. E. (1998). The validity and utility of selection methods in personnel psychology: Practical and theoretical implications of 85 years of research findings. *Psychological Bulletin*, 124, 262-274.
- Serpell, R. (1979). How specific are perceptual skills? *British Journal of Psychology*, 70, 365-380.

- Shepard, L., Camilli, G., & Averill, M. (1981). Comparison of procedures for detecting test-item bias with both internal and external ability criteria. *Journal of Educational & Behavioral Statistics*, 6, 317-375.
- Sierens, S. (2006). Immigratiesamenleving, onderwijs en overheid in Vlaanderen: een gespannen driehoeksverhouding. In Sierens, S., Van Houtte, M., Loobuyck, P., Dellerue, K., Pelleriaux, K. (red), *Onderwijs in de immigratiesamenleving*, Academia Press, Gent, pp. 9-32.
- Spearman, C. (1904). General intelligence, objectively determined and measured. *American Journal of Psychology*, 15, 201-293.
- Super, C. (1983). Cultural variation in the meaning and uses of children's intelligence. In J. Deregowski, S. Dziurawiec, & R. Annis, *Explorations in crosscultural psychology* (pp.199-212). Lisse: Swets
- te Nijenhuis, J., & van der Flier, H. (2003). Immigrant-majority group differences in cognitive performance: Jensen effects, cultural effects, or both? *Intelligence*, 31, 443–459.
- Thurstone, L. L. (1938). *Primary mental abilities*. Chicago: University of Chicago Press.
- Van de Vijver, F. J. R. (2002). Cross-cultural assessment: Value for money? *Applied Psychology: An International Review*, 51, 545-566
- Van de Vijver, F. J. R., & Bleichrodt, N. (2001). Conclusies. In N. Bleichrodt & F. J. R. van de Vijver (Red.), *Het gebruik van psychologische tests bij allochtonen* (pp.237-243). Lisse, Nederland: Swets & Zeitlinger.
- Van de Vijver, F. J. R., Helms-Lorenz, M., & Feltzer, M. F. (1999). Acculturation and cognitive performance of migrant children in the Netherlands? *International Journal of Psychology*, 34, 149-162.
- Van de Vijver, F. J. R., & Leung, K. (1997a). Methods and data analysis of comparative research. second ed. In: Berry, J.W., Poortinga, Y.H., Pandey, J. (Eds.), *Handbook of Cross-cultural Psychology*, 1. Allyn & Bacon, Boston, pp. 257–300.

- Van de Vijver, F. J.R., & Leung, K. (1997b). *Methods and Data Analysis for Cross-cultural Research*. Sage, Newbury Park, CA.
- Van de Vijver, F. J. K., Poortinga, Y. H. (1997). Towards an integrated analysis of bias in cross-cultural assessment. *European Journal of Psychological Assessment* 13, 29–37.
- Van de Vijver, F. J. R., & Tanzer, N. K. (1997). Bias and equivalence in cross-cultural assessment: An overview. *European Review of Applied Psychology*, 47, 263-279
- Van de Vijver, F. J. R., & Tanzer, N. K. (2004). Bias and equivalence in cross-cultural assessment: An overview. *Revue Européenne de Psychologie Appliquée*, 54, 119-135.
- Van de Vijver, F. J. R., Willemse, G., & Van de Rijt, B. (1993). Het testen van cognitieve vaardigheden van kinderen. *De Psycholoog*, 152-159.
- Van Der Maas, H. L. K., Dolan, C. V.; Grasman, R. P., Wicherts, J. M., Huizenga, H. M., & Raijmakers, M. E. J. (2006). A dynamical model of general intelligence: The positive manifold of intelligence by mutualism. *Psychological Review*, 113 (4), 842-861.
- Verachtert, P., Van Damme, J., Onghena, P., & Ghesquière, P. (2009). A seasonal perspective on school effectiveness: evidence from a Flemish longitudinal study in kindergarten and first grade. *School Effectiveness and School Improvement*, 20, 215-233.
- Verschuere, K., & Koomen, H. (2007). *Handboek Diagnostiek in de leerlingenbegeleiding*. Antwerpen – Apeldoorn: Garant.
- Verweij, A. (2010). Wat is sociaaleconomische status? In: *Volksgezondheid Toekomst Verkenning, Nationaal Kompas Volksgezondheid*. Bilthoven: RIVM. Retrieved from <http://www.nationaalkompas.nl/bevolking/segv/wat-is-sociaaleconomische-status/>
- Winkleby, M. A., Jatulis, D. E., Frank, E., & Fortmann, S. P. (1992). Socioeconomic status and health: How education, income, and occupation contribute to risk factors for cardiovascular disease. *American Journal of Public Health*, 82, 816– 820.