

Een dynamische test voor cognitieve vaardigheden van anderstalige nieuwkomers in het secundair onderwijs in Vlaanderen

E. Vande Gaer, K. Verschueren, E. Buyse, V. Germeijs, en W. Magez¹

Samenvatting

In een steekproef van 462 leerlingen uit de Ont-
haalklas Anderstalige Nieuwkomers (OKAN)
werden de psychometrische eigenschappen
van de Covaar-II, een cognitieve-vaardigheden-
test, geëvalueerd. Deze evaluatie gebeurde
zowel voor de oorspronkelijke versie met uitge-
breide oefenfase als voor een versie met be-
perkte oefenfase. Het gebruik van een oefen-
fase is kenmerkend voor de Dynamische Test-
beweging, waar de Covaar-II bij aanleunt. De
uitgebreidheid van de oefenfase bleek geen
effect te hebben op de Covaar-II-scores. Ook
betrouwbaarheid en validiteit waren in beide
afnamecondities vergelijkbaar en goed. Beide
versies van de test kunnen bij deze doelgroep
dus worden afgenomen en de keuze zal afhan-
gen van andere dan psychometrische over-
wegingen.

1 Inleiding

Begeleiders in het onderwijs krijgen geregeld
vragen in verband met de cognitieve mogelijk-
heden van allochtone leerlingen. Een bijzonde-
re groep daarbinnen zijn de anderstalige
nieuwkomers. Met deze term verwijst men
naar leerlingen die heel recent uit het buiten-
land zijn gekomen en in het Nederlandstalige
onderwijs terecht komen. De reden waarom zij
hun moederland verlaten hebben, hun nationa-
liteit, sociaaleconomische achtergrond en
scholingsgraad zijn erg divers. In het secundair
of voortgezet onderwijs in Vlaanderen krijgen
anderstalige nieuwkomers gedurende een jaar
een Nederlands taalbad, om daarna naar een
gepaste vorm van regulier onderwijs over te
stappen. Op deze groep anderstalige minderja-
rige jongeren richt de huidige studie zich. Meer
bepaald gaat de studie de psychometrische
kwaliteiten na van een cognitieve vaardighe-
dentest voor anderstalige nieuwkomers in het
voortgezet of secundair onderwijs. Dergelijk

instrument maakt het mogelijk bepaalde aspec-
ten van het geboden onderwijs aan te passen
aan de mogelijkheden van de leerlingen.

In het Nederlandse taalonderwijs aan vol-
wassen anderstaligen werd enkele jaren gele-
den met dit doel de Covaar-II ontwikkeld
(Buyse, Verschueren, & Magez, 2005), een
cognitieve vaardighedentest voorafgegaan door
een uitgebreide oefenfase. Een dergelijke test
kon namelijk een bijdrage leveren aan de be-
slissing of een cursist beter zou starten in een
Centrum voor Basiseducatie (met als doelpu-
blik laaggeschoolde of traag lerende perso-
nen) of in een Centrum voor Volwassenenon-
derwijs (gericht op hooggeschoolde of vlot
lerende cursisten). Deze beslissing is in Vlaan-
deren wettelijk afhankelijk van het aantal jaren
scholing dat een persoon heeft genoten (minder
of meer dan 10 jaar). De benodigde informatie
hiervoor is echter niet in alle gevallen aanwe-
zig. Bovendien is het onderwijsniveau niet al-
tijd vergelijkbaar tussen de verschillende lan-
den. Een meer objectieve inschatting van de
mogelijkheden en de beginsituatie van de cur-
sist door middel van een cognitieve vaardighe-
dentest kon dus een nuttige aanvulling vormen
ter ondersteuning van deze doorverwijzing. De
Covaar-II is tot stand gekomen in verschil-
lende fasen en op basis van een reeks van uit-
gebreide kwalitatieve en kwantitatieve evalu-
aties binnen het volwassenenonderwijs (Buyse,
Vanhee, Magez, & Verschueren, 2003a, Buyse,
Vanhee, Magez, & Verschueren, 2003b; Verbo-
ven, Buyse, Magez, & Verschueren, 2004).
Hieruit bleek dat de Covaar-II een betrouwbaar
en valide instrument is voor de ondersteuning
van de doorverwijzing van kandidaat-cursisten
Nederlands als tweede taal naar een Centrum
voor Basiseducatie of een Centrum voor Vol-
wassenenonderwijs (Buyse et al., 2005).

Vraag is of deze test ook bij anderstalige
nieuwkomers in het secundair onderwijs de
cognitieve vaardigheden op betrouwbare en va-
lide wijze in kaart kan brengen. Een bijkomen-
de vraag betreft de invloed van de oefenfase op

de betrouwbaarheid en validiteit van de Covaar-II. Zoals verder aan bod zal komen, werd de oefenfase net ingevoerd om deze betrouwbaarheid en validiteit te verhogen. Maar oefent de oefenfase dit beoogde effect ook werkelijk uit? Tot dusver werd dit over de Covaar-II op basis van literatuur (zie verder) verondersteld, maar nog niet onderzocht.

1.1 Het testen van cognitieve vaardigheden bij anderstalige nieuwkomers

Uit verscheidene studies komt naar voren dat bepaalde niet-Westerse culturen over het algemeen, als groep, minder goed scoren op traditionele intelligentietesten (Gottfredson, 2000; Resing & Hessels, 2001; Reynolds & Kaiser, 2003; Rushton, 1998; Te Nijenhuis & Van der Flier, 1999). Dit verschil zou kunnen wijzen op bias of vertekening in cognitieve vaardigheidstesten bij gebruik in bepaalde culturele groepen. Vertekening verwijst naar iedere onbedoelde, systematische invloed op test scores (Van de Vijver, Fontaine, & Schittekatte, 2007) en vormt aldus een bedreiging voor de validiteit van test scores. Zeker in een groep recent geïmmigreerde allochtonen is de kans op vertekening reëel.

Er wordt in de literatuur een onderscheid gemaakt tussen interne en externe vertekening (Te Nijenhuis & Van der Flier, 1999; Van de Vijver et al., 2007). *Interne* vertekening heeft betrekking op de psychologische betekenis van een item of test. Drie vormen van interne bias worden doorgaans onderscheiden (Van de Vijver & Leung, 1997): constructbias, methodebias en itembias. Constructbias houdt in dat het door de test gemeten construct niet identiek is in de verschillende betrokken culturele groepen. Methodebias is een gevolg van methodische aspecten van de multiculturele diagnostiek, die een invloed hebben op de antwoorden op (bijna alle) items (Berry, Poortinga, Segall, & Dasen, 2002; Van de Vijver et al., 2007). Belangrijke bronnen van methodevertekening zijn een gebrek aan kennis van het Nederlands (taalbias) of een tekort aan testvaardigheid en testervaring (*test wiseness*) bij de geteste personen (Rogers & Yang, 1996). Men spreekt van itembias wanneer een vertekende factor slechts invloed heeft op de scores voor één of een beperkt aantal items (Berry et al., 2002). Een item wordt vertekend genoemd wanneer

groepen die gelijk zijn wat betreft het onderliggende construct toch niet dezelfde gemiddelde score behalen op het item (Van de Vijver & Phalet, 2004). Naast interne vertekening, kan er ook *externe* vertekening optreden. Hiervan spreekt men indien een test voor sommige groepen beter voorspelt dan voor andere (Resing, 2001; Van de Vijver et al., 2007). Externe vertekening heeft dus betrekking op de predictieve validiteit van een test.

1.2 De Covaar-II: Voorzorgen om bias te reduceren

Gezien het oorspronkelijke doelpubliek van de Covaar-II werd bij de opstelling van de test uitgebreide aandacht besteed aan de vermindering van eventuele bias die kan optreden bij afname van een cognitieve vaardigheidstest in cultureel verschillende groepen. Consistent met aanbevelingen in de literatuur werd, ten eerste, getracht de invloed van Nederlandse taalbeheersing (methodebias) te reduceren door invoering van een volledig non-verbale instructie, zoals eerder bijvoorbeeld ook bij de Leertest Etnische Minderheden gebeurde (LEM, Hessels & Hamers, 1993). Ook de antwoorden worden non-verbaal gegeven (namelijk via het omcirkelen van het juiste antwoordalternatief, zie Methode). Ten tweede, ten einde constructbias te minimaliseren, peilt de test naar het vermogen tot inductief redeneren, een deelaspect van 'fluïde intelligentie' (*Gf*; Carroll, 1993; Horn & Cattell, 1966; McGrew & Flanagan, 1998). Daar waar 'gekrystalliseerde intelligentie' (*Gc*) verwijst naar de breedte en diepte van iemands verworven kennis over een cultuur en de effectieve toepassing van deze kennis, behelst *Gf* de mentale operaties die een individu uitvoert wanneer hij of zij geconfronteerd wordt met een relatief nieuwe taak die hij of zij niet automatisch kan uitvoeren. De *Gf*-factor zou daarom minder onderhevig zijn aan voorafgaande (cultureel gemedieerde) leerervaringen dan de *Gc*-factor.

De meest opvallende maatregel om eventuele bias zoveel mogelijk te reduceren is echter de invoering van een uitgebreide oefenfase voorafgaand aan de eigenlijke testfase, zoals gesuggereerd door verschillende auteurs (Rogers & Yang, 1996; Te Nijenhuis & Van der Flier, 1999; Te Nijenhuis & Van de Vijver, 2001; Van de Vijver & Leung, 1997). Het ge-

bruik van een oefenfase om eventuele bias bij afname van testen bij cultureel verschillende groepen te reduceren is geen nieuw idee, maar wordt al jaren toegepast binnen de zogenaamde Dynamische Test (DT)-beweging (zie Grigorenko & Sternberg, 1998; Lidz, 1987; Resing, 2007; Sternberg & Grigorenko, 2002; Wiedl, Guthke, & Wingenfeld, 1995). Hamers en Sijtsma (1995) argumenteren dat bij klassieke testprocedures een verandering van testcores als gevolg van leren tijdens de test vaak gezien wordt als een bedreiging van de validiteit. Dergelijke verandering wordt bij dynamisch testen daarentegen gezien als een mogelijkheid tot verhoging van de validiteit. Volgens Sternberg en Grigorenko (1999) speelt de DT-beweging in op verscheidene maatschappelijke behoeften, waaronder de behoefte aan cultureel meer faire tests, die de integratie van immigranten in de samenleving kunnen ondersteunen.

Dynamische tests verschillen onderling sterk, onder andere in de testvorm en de wijze en het moment waarop hulp wordt geboden (voor een uitgebreide beschrijving zie onder andere Grigorenko & Sternberg, 1998; Lidz, 1987; Resing, 2007; Sternberg & Grigorenko, 2002; Wiedl, et al., 1995). De DT-benadering waarbij de Covaar-II het dichtst aanleunt, is de zogeheten Testing-the-Limits-benadering (Schmidt, 1977; ook vermeld in Wiedl et al., 1995), waarbij geen gebruik wordt gemaakt van een pretest, maar wel gezocht wordt naar de beste manier om een bepaalde doelgroep, in casu anderstalige nieuwkomers, te testen. In deze stroming wordt de invloed nagegaan van verschillende soorten aanpassingen in presentatie en context van de test op de prestaties van specifieke populaties. Dit is wat ook ons onderzoek beoogt te doen door het nagaan van de invloed van de oefenfase op de scores van anderstalige nieuwkomers (zie Vande Gaer, 2007 voor een uitgebreidere verantwoording).

Hoewel men aanneemt dat het gebruik van een oefen- of trainingsfase eventuele bias of vertekening kan reduceren bij afname in andere culturele groepen, is er geen eenduidige empirische evidentie voor die idee. Onderzoek binnen de DT-beweging leverde reeds enige evidentie dat de validiteit van bepaalde testen stijgt na een oefenfase (Day, Engelhardt, Maxwell, & Bolig, 1997; Elliot, 2003; Fabio, 2005;

Guthke & Beckmann, 2000; Guthke & Stein, 1996; Resing & Hessels, 2001; Sternberg et al., 2002), maar die evidentie wordt niet eenduidig vastgesteld (Grigorenko & Sternberg, 1998; Hamers, Hessels, & Pennings, 1996; Hamers & Sijtsma, 1995). Verder onderzoek is volgens verscheidene auteurs dan ook noodzakelijk vooraleer ondubbelzinnig kan worden besloten dat een oefenfase een verbeterde validiteit oplevert (Elliot, 2003; Fabio, 2005; Grigorenko & Sternberg, 1998; Guthke & Beckmann, 2000). In voorgaand onderzoek werd er meestal ook geen rechtstreekse vergelijking gemaakt tussen een test met beperkte of geen oefenfase en dezelfde test met uitgebreide oefenfase, maar werd het resultaat op een pretest (voor de oefenfase) vergeleken met het resultaat op een posttest (na de oefenfase) binnen een zelfde subject. Echter, op deze manier wordt niet louter het effect van de oefenfase gemeten, aangezien de pretest ook een invloed zal hebben op de posttest. Doel van het huidige onderzoek was dan ook om de vergelijking op een meer rechtstreekse manier maken, gebruik makend van een tussengroepdesign.

1.3 Doelstellingen

Een eerste doelstelling van het huidige onderzoek betrof de validering van de Covaar-II voor gebruik bij een nieuwe doelgroep, namelijk leerlingen in het secundair onderwijs die onthaalonderwijs voor anderstalige nieuwkomers volgen (OKAN). In dit onthaalonderwijs staat de verwerving van de Nederlandse taal centraal. Het huidige onderzoek had dus als doel het toetsen van de betrouwbaarheid en de validiteit (interne consistentie, begripsvaliditeit, convergente validiteit, criteriumvaliditeit) van de Covaar-II bij deze OKAN-leerlingen. Ter evaluatie van de criteriumvaliditeit werd de samenhang onderzocht tussen Covaar-II-scores in het begin van het schooljaar en scores voor een taaltest, leerkrachtinschattingen van leervaardigheid en vorderingen, oriënteringsadviezen en examenresultaten aan het einde van het schooljaar. Deze gegevens waren telkens beschikbaar voor een subgroep van de totale groep (grootte afhankelijk van het criterium, zie verder). In een kleine subgroep werd daarnaast ook het verband met scores voor een andere inductieve redeneertest onderzocht, ter evaluatie van de convergente validiteit.

Een tweede doelstelling was het nagaan van het effect van de oefenfase op de hoogte van de Covaar-II-scores, gebruik makend van een tussengroepdesign. De voorspelling is dat, wanneer er bij een bepaalde cognitieve vaardigheidentest (in casu de Covaar-II) inderdaad sprake is van bias ten nadele van de allochtone groep en deze door de oefenfase gereduceerd wordt, de score van een groep recent geïmmigreerde adolescenten (OKAN-leerlingen) na een uitgebreide oefenfase ('uitgebreide' afnameconditie) hoger zal liggen dan na slechts een minimale oefenfase ('minimale' afnameconditie). Deze minimale oefenfase komt overeen met de oefening zoals gegeven in de meeste traditionele niet-dynamische intelligentietests (zie Methode).

Ten derde kan men dan verwachten dat, aangezien de oefenfase verondersteld wordt testbias te reduceren, een verbetering van de validiteit van de Covaar-II zal optreden in de uitgebreide afnameconditie in vergelijking met de minimale afnameconditie. Een derde doelstelling van het onderzoek bestond er daarom in de impact na te gaan van de oefenfase op de betrouwbaarheid en validiteit van

de Covaar-II-testscores, eveneens gebruik maken van een tussengroepdesign.

2 Methode

2.1 Deelnemers

Er namen 462 OKAN-leerlingen deel aan het onderzoek. De participerende klassen waren afkomstig uit het hele land en het enige selectie criterium was de bereidheid van het begeleidende Centrum voor Leerlingenbegeleiding (CLB) om de afnames te coördineren. Hun verdeling over de minimale en uitgebreide afnameconditie staat in Tabel 1. De leerlingen werden willekeurig aan één van beide condities toegewezen (even versus oneven nummers in alfabetische klassenlijsten).

Tabel 1 biedt tevens een overzicht van de verdeling van de respondenten wat betreft hun geslacht, het opleidingsniveau van hun moeder en vader, hun thuistaal en hun land van herkomst, voor die leerlingen waarover de betreffende informatie voorhanden was. Dit was afhankelijk van het al dan niet beschikken over deze informatie door de bege-

Tabel 1

Frequentie van bepaalde achtergrondkenmerken van de deelnemers en de verdeling over de twee afnamecondities (minimaal en uitgebreid)

	Aantal leerlingen				Aantal leerlingen		
	M	U	Totaal		M	U	Totaal
CONDITIE	235	227	462	GESLACHT	213	207	420
				Mannelijk	118	112	230
				Vrouwelijk	95	95	190
THUISTAAL	198	198	396	OPLEIDING VADER	106	100	206
Nederlands	12	11	23	Hoogstens lager onderwijs	34	32	66
Een andere taal	126	131	257	Secundair onderwijs	44	42	86
Nederlands en andere	60	56	116	Hoger onderwijs	28	26	54
LAND V. HERKOMST	212	206	418	OPLEIDING MOEDER	113	92	205
België	1	0	1	Hoogstens lager onderwijs	54	33	87
Europa: Noord & West	1	7	8	Secundair onderwijs	39	35	74
Europa: Oost	34	35	69	Hoger onderwijs	20	24	44
Europa: Zuid	7	4	11				
voormalige USSR	33	28	61				
Afrika: Noord	21	19	40				
Afrika: overige	39	49	88				
Turkije	7	4	11				
Azië: Zuid en West	37	31	68				
Azië: Oost	19	16	35				
Amerika: Latijnsstalig	13	11	24				
Amerika: Engelstalig	0	1	1				
Oceanië	0	1	1				

Nota. M= minimale conditie; U= uitgebreide conditie

leidende CLB-medewerker of leerkracht.

De OKAN-leerlingen bleken gemiddeld 14,95 jaar oud ($n = 419$, $SD = 1,75$) en hadden er gemiddeld 7,55 jaar scholing opzitten ($n = 361$, $SD = 2,70$). Van 418 deelnemers verkregen we informatie over het land van herkomst. Deze groep bleek samengesteld uit 81 nationaliteiten, die voor deze studie werden onderverdeeld in 11 categorieën. De meest voorkomende categorieën waren: immigranten uit West-, Midden- en Zuid-Afrika (21%), Oost-Europa (17%) en Zuid- en West-Azië (16%). Gemiddeld verbleven de OKAN-leerlingen reeds gedurende 7,81 maanden ($n = 401$, $SD = 9,57$) in België.

2.2 Materiaal

Covaar-II

De Covaar-II is een cognitieve vaardigheden-test die bestaat uit twee delen: een oefenfase en een testfase. Het tweede deel, de eigenlijke test, omvat drie subtesten, elk bestaande uit 8 items: Verbale Analogieën, Figuratieve Reeksen en Analogieën met symbolen (zie Appendix 1). Deze drie subtesten bedoelen allen inductief redeneren te meten. De Covaar-II omvat zowel talig als figuratief materiaal. De talige items kunnen weliswaar ook opgelost worden zonder kennis van het Westerse alfabet, omdat de letters ook opgevat kunnen worden als symbolen. De test wordt over het algemeen afgenomen in groep met een tijdslimiet van 20 minuten. Bij eerder onderzoek bleek het merendeel van de volwassen participanten (86%) de test binnen deze tijdslimiet te kunnen oplossen (Verboven et al., 2004). De items zijn twee per twee per subtest afwisselend gerangschikt (omnibusvorm), waardoor vermeden werd dat sommige respondenten helemaal niet aan een bepaalde subtest zouden toe komen (als gevolg van de tijdslimiet). De Covaar-II is opgesteld volgens een meerkeuzeformaat waarbij er voor elk item vier antwoordalternatieven zijn. Het juiste antwoord dient omcirkeld te worden.

In de uitgebreide afnameconditie werd de Covaar-II gebruikt zoals gevalideerd in voorgaand onderzoek bij de doelgroep van volwassen anderstalige nieuwkomers. Dat wil zeggen dat aan de eigenlijke testfase een oefenfase voorafging. Tijdens deze oefenfase werd de leerling vertrouwd gemaakt met de verschillen

de subtesten en mogelijke oplossingsstrategieën. Voor elke subtest werden twee voorbeelditems geprojecteerd en klassikaal uitgelegd. Hierna losten de leerlingen zes voorbeelditems individueel op, twee voor elke subtest, waarna deze zes weer klassikaal besproken werden. Voor elk voorbeelditem werd er een volledige non-verbale vorm van uitleg uitgewerkt, maar extra verbale ondersteuning was toegelaten. Dergelijke extra ondersteuning mocht echter nooit ten koste gaan van de non-verbale uitleg, die altijd volledig diende gegeven te worden. In Appendix 2 is een voorbeeld van dergelijke non-verbale instructie opgenomen. Ook voor de andere voorbeelditems zijn in de handleiding gelijkaardige gedetailleerde non-verbale instructies opgenomen. Verdere individuele uitleg en hulp van medeleerlingen waren eveneens toegestaan en konden gebeuren in een andere dan de Nederlandse taal. De oefenfase nam minstens een half uur in beslag.

Voor de minimale afnameconditie werden de handleiding voor afname en het oefenboekje zodanig aangepast dat ze aansloten bij de wijze van uitleg die bij de meeste traditionele intelligentietesten gehanteerd wordt. Zo werd alleen de eerste trainingsoefening bij elke subtest klassikaal uitgelegd. De bespreking van de tweede oefening, die eerst individueel diende opgelost te worden, was beperkt tot het geven van de juiste oplossing. De zes extra oefeningen, wel aangeboden in de uitgebreide conditie, werden volledig weggelaten. Verder was het de proefleiders in de minimale conditie niet toegestaan de oefeningen op een andere dan de gestandaardiseerde manier uit te leggen, en mochten ze geen extra hulp verschaffen. Er werd benadrukt dat alle uitleg in het Nederlands diende te gebeuren en de leerlingen mochten elkaar ook niet helpen. Dit was wel toegestaan bij de uitgebreide afname. Voor de eigenlijke testafname werd in beide condities een tijdslimiet van 20 minuten opgelegd.

Achtergrondgegevens en opvolgingsfiches

Aan de hand van zelf opgestelde vragenlijsten werden achtergrondgegevens van de leerlingen verkregen. Er werd onder andere informatie opgevraagd aangaande leeftijd, geslacht, land van herkomst, verblijfsduur, jaren scholing en opleiding van de moeder en vader (zie Deelnemers). Daarnaast werd ook een leerkrachtin-

schatting gevraagd van vier deelvaardigheden van de kennis van het Nederlands op het moment van de testafname, in het begin van het schooljaar (spreken, begrijpen, lezen en schrijven; telkens een vijfpuntenschaal gaande van *zeer slecht* over *gemiddeld* tot *zeer goed*). Deze vier deelvaardigheden werden geaggregeerd tot een overkoepelende variabele *aanvankelijke kennis van het Nederlands*.

De opvolgingsfiche, door de leerkrachten ingevuld op het einde van het schooljaar, bevroeg informatie nodig voor het evalueren van de criteriumvaliditeit van de Covaar-II. Zo werden gegevens opgevraagd aangaande de leervaardigheid van de leerling (vijfpuntenschaal gaande van *heel zwak* over *gemiddeld* tot *heel goed*), de examenresultaten zowel procentueel als in vergelijking met klasgenoten (vijfpuntenschaal gaande van *heel zwak* over *gemiddeld* tot *heel goed*) en de vorderingen tijdens het jaar in vergelijking met medeleerlingen (vijfpuntenschaal gaande van *zeer klein* over *gemiddeld* tot *zeer groot*). Ook werd aan de hand van deze fiche informatie verkregen over het oriënteringsadvies op basis van de cognitieve mogelijkheden van de leerling (vijf opties binnen het secundair onderwijs genoemd: Algemeen Secundair Onderwijs (ASO), Technisch Secundair Onderwijs (TSO), Kunst Secundair Onderwijs (KSO), Beroeps Secundair Onderwijs (BSO), en *andere, namelijk...*). De onderwijsvorm ASO bereidt leerlingen voor op studies in het hoger onderwijs door het aanbieden van een brede, theoretische vorming. TSO biedt naast een algemene vorming ook technische en praktijkvakken aan; KSO is gericht op algemene vorming, theorie over kunsttechnieken en praktische kunstbeoefening. BSO bereidt leerlingen voor op het uitoefenen van een beroep door het bijbrengen van praktische beroepskennis. Een voorbeeld van een andere optie betreft bijvoorbeeld leerlingen die doorverwezen werden naar het buitengewoon onderwijs als gevolg van bijkomende problemen. Verder werd opnieuw informatie gevraagd over de kennis van het Nederlands maar nu aan het einde van het schooljaar (spreken, begrijpen, lezen en schrijven; telkens een vijfpuntenschaal gaande van *zeer slecht* over *gemiddeld* tot *zeer goed*). Ook nu werden de deelvaardigheden geaggregeerd tot een overkoepelende variabele (*kennis van het Neder-*

lands einde schooljaar). Verder werd gevraagd eventueel beschikbare andere objectieve testresultaten te noteren. We verkregen zo van een beperkt deel van de OKAN-leerlingen de scores op de Progressive Matrices (convergente validiteit) en de TASAN (criteriumvaliditeit; zie verder).

TASAN

De Taalvaardigheid Aanvang Secundair onderwijs Anderstalige Nieuwkomers (TASAN) Test van Gysen, Ramaut, en Sterckx (2002) is ontworpen door het Centrum voor Taal en Onderwijs, een universitair centrum verbonden aan de K.U.Leuven. De test is speciaal ontworpen voor afname bij anderstalige nieuwkomers tussen 12 en 16 jaar op het einde van het onthaaljaar, en bedoelt de luister-, lees- en schrijfvaardigheid te meten die men nodig heeft om op school te functioneren. De TASAN bestaat uit zeven typetaken (vergelijkbaar met taken die leerlingen moeten uitvoeren in de reguliere klas) die variëren qua moeilijkheid. Aan de hand van scoresleutels en beoordelingsmodellen kan de test verbeterd worden. Men verkrijgt op die wijze negen subscores, namelijk voor de vaardigheden luisteren, lezen en schrijven, voor receptieve en productieve taalvaardigheden, voor beschrijvend en structurerend verwerkingsniveau, en voor taken met veel of weinig extra contextuele ondersteuning. Elk van die subscores wordt op een vierpuntenschaal (*zwak, middelmatig zwak, middelmatig sterk en sterk*) gegeven. Betrouwbaarheid en validiteit van de test zijn uitvoerig onderzocht (Magez, Grysolle, Bos, & De Cleen, 2001). De TASAN-resultaten waren beschikbaar voor 138 tot 140 OKAN-leerlingen, afhankelijk van de subtest.

Progressive Matrices

De Progressive Matrices (PM, Raven, Court, & Raven, 1979) bestaat uit een reeks van 60 diagrammen waarbij steeds één deel ontbreekt. Onder elk diagram staat een aantal mogelijke aanvullingen waaruit die optie moet gekozen worden die het diagram op een logische manier vervolledigt. Door het optellen van het aantal juiste keuzes wordt een totale score op 60 bekomen. De test bedoelt g (algemene cognitieve vaardigheden) te meten (Raven, 2000) en werd uitgebreid gevalideerd (Evers, Van Vliet-Mul-

der, & Groot, 2000; Magez, 2003). Voor het huidige onderzoek werden, net als voor de Covaar-II, de ruwe scores gebruikt. Deze gegevens waren beschikbaar voor een subgroep van 33 OKAN-leerlingen, afkomstig uit twee scholen.

2.3 Procedure

De afname van de Covaar-II ging door in de periode september-november 2005. Deze gebeurde in 17 scholen en werd verzorgd door 10 CLB-medewerkers die de betreffende onthaalklassen ondersteunden. De afname werd door de onderzoekers in een trainingssessie gedemonstreerd aan de betrokken CLB-medewerkers. Er werd gevraagd dat de CLB-medewerkers zoveel als mogelijk zelf de afnames zouden begeleiden. Wanneer dit praktisch niet mogelijk bleek, kon een leerkracht de afname verzorgen. Voorwaarde hiervoor was echter dat deze leerkracht reeds een aantal afnames door een CLB-medewerker had bijgewoond. Bevraging nadien leerde dat in 81% van de gevallen de CLB-medewerker de test had afgenomen, in 12% van de gevallen was het de leerkracht geweest en voor 7% ontbrak deze informatie.

De achtergrondgegevens- en opvolgingsfiches werden ingevuld door de leerkracht of een CLB-medewerker die van deze gegevens op de hoogte was. Gevraagd werd deze bron ook te specificeren. De opvolgingsfiches werden steeds ingevuld aan het einde van het schooljaar. Op dat moment werden ook de resultaten voor de TASAN en de PM, indien afgenomen als onderdeel van de gebruikelijke schoolloopbaanbegeleiding in de betreffende OKAN-klassen, opgevraagd.

2.4 Analyseplan

De interne consistentie werd geëvalueerd aan de hand van Cronbachs α -coëfficiënten. De evaluatie van de begripsvaliditeit gebeurde door middel van de intersubtestcorrelaties en confirmatorische factoranalyses.

Gezien de data ordinaal zijn (de Covaar-II-itemscores hebben als mogelijke waarden 0, 1 en 2) en omwille van de beperkte steekproefgrootte werd voor de confirmatorische factoranalyses gebruik gemaakt van de maximum likelihood-methode, op basis van de polychorische correlatie- en asymptotische covariantiematrix (Jöreskog, 2005). Als fitmaten worden de Satorra-Bentler Scaled Chi-Square ($SBS\chi^2$),

de *RMSEA* (Root Mean Square Error of Approximation), de *CFI* (Comparative Fit Index) en de *SRMR* (Standardized Root Mean Square Residual) statistieken gerapporteerd. Voor de interpretatie van deze fitmaten werden de *cut-off*criteria van Hu en Bentler (1999) en Kline (2005) gehanteerd. Op basis van vorig onderzoek met de Covaar-II (Buyse et al., 2003b) werd verwacht dat een oplossing met één factor het beste bij de data zou passen, aangezien alle drie de subtesten de vaardigheid inductief redeneren bedoelen te meten. De fit van het a priori één-factormodel werd vergeleken met de fit van een drie-factorenmodel waarin de drie subtests als drie aparte factoren werden beschouwd. De Satorra-Bentler Scaled χ^2 difference test ($\Delta SBS\chi^2$; Satorra & Bentler, 1999) werd gebruikt voor de vergelijking van deze twee modellen.

Als indicatie voor de convergente validiteit werd de correlatie tussen de Covaar-II-scores en de scores op de Progressive Matrices (PM) berekend. Verder, omwille van het beperkte aantal OKAN-leerlingen ($n=33$) voor wie PM-scores beschikbaar waren, werden leerlingen in de minimale en de uitgebreide conditie samen genomen in de analyse.

De resultaten van de OKAN-leerlingen op de TASAN werden gecorreleerd met de Covaar-II-scores ter evaluatie van de criteriumvaliditeit. Daarnaast werd de criteriumvaliditeit onderzocht aan de hand van correlaties tussen de Covaar-II-scores en de leerkrachtinschattingen van leervaardigheid en vorderingen (steeds geëvalueerd aan de hand van een vijfpunten-schaal, waarbij een hoger cijfer een beter resultaat representeert) in de verschillende condities (minimale en uitgebreide afnameconditie) apart. Ook de examenresultaten werden gecorreleerd met de Covaar-II-scores. De samenhang van de Covaar-II-scores met het oriënteringsadvies aan het einde van het schooljaar (ASO, TSO en BSO) werd voor de beide condities apart nagegaan aan de hand van multinomiale logistische regressie-analyses. Omwille van het beperkte aantal leerlingen die naar KSO werden verwezen (resp. 3 en 0 in de minimale en uitgebreide conditie) werd deze categorie uit de analyses gehouden. Omwille van de heterogeniteit van de elementen in de categorie andere, werd ook deze uit de analyses geweerd. Verder werden zowel de correlatie tus-

sen de kennis van het Nederlands aan het einde van het schooljaar en de Covaar-II-scores berekend als de partiële correlatie uitgezuiverd voor de kennis van het Nederlands aan het begin van het schooljaar en dit voor de minimale en de uitgebreide conditie apart. Bovendien werd, waar mogelijk, de significantie van het verschil in correlaties tussen de minimale en de uitgebreide afnameconditie weergegeven om een zicht te krijgen op het verschil in predictieve waarde van de Covaar-II-scores als gevolg van de oefenfase. Hiertoe werden alle correlaties omgezet naar een Fischer Z-score. Vervolgens werd de significantie getoetst van de z-score van het verschil tussen beide z-scores van de correlaties.

Zoals eerder gezegd, waren deze criteriumgegevens slechts voor subgroepen (verschillend in grootte, afhankelijk van het criterium, zie verder) beschikbaar. Om de vergelijkbaarheid na te gaan van deze subgroepen met de overige respondenten aangaande bepaalde (achtergrond)variabelen, werden ze vergeleken aan de hand van t-toetsen. Dit leverde in totaal 77 vergelijkingen op (7 criteria en 11 (achtergrond)variabelen). Na toepassing van een Bonferroni-correctie voor meervoudige toetsing, bleek de groep waarover we geen gegevens hadden bij 5% van de 77 vergelijkingen significant te verschillen van de groep waarover wel gegevens beschikbaar waren. We kunnen dus besluiten dat, over het algemeen, de deelgroep waarover we geen gegevens hadden niet sterk afwijkt van de groep waarover we wel gegevens hadden.

Het effect van de oefenfase op de hoogte van de Covaar-II-scores werd getest aan de hand van een variantieanalyse (ANOVA) met afnameconditie (minimaal vs. uitgebreid) als onafhankelijke variabele.

3 Resultaten

3.1 Interne consistentie

De betrouwbaarheid (Cronbachs α) van de 24 items van de test is goed (Ter Laak & de Goede, 2005), zowel in de minimale ($\alpha = 0,94$, $n = 235$) als de uitgebreide afnameconditie ($\alpha = 0,93$, $n = 227$). Het geven van een uitgebreidere oefenfase (uitgebreide conditie) heeft dus geen effect op de betrouwbaarheid.

3.2 Begripsvaliditeit

Zowel de onderlinge correlaties tussen de subtestscores van de COVAAR-II in de minimale conditie ($r(225) = 0,67, 0,81, 0,68$, allen significant op 1%-niveau) als de intersubtestcorrelaties van de uitgebreide conditie ($r(227) = 0,72, 0,83, 0,70$, allen significant op 1%-niveau) waren hoog² (Cohen, 1988). Deze intersubtestcorrelaties verschilden bovendien niet significant tussen de minimale en uitgebreide conditie.

Aan de hand van confirmatorische factoranalyses werd de factorstructuur van de test onderzocht, en dit zowel voor de minimale als voor uitgebreide conditie. De resultaten bevestigen dat een éénfactoroplossing in beide condities goed bij de data past (minimale conditie: RMSEA= 0,06, CFI= 0,96 en SRMR= 0,04; Mediaan van itemladingen = 0,80; uitgebreide conditie: RMSEA= 0,05, CFI= 0,99 en SRMR= 0,07; Mediaan van itemladingen = 0,80). De passing van dit één-factormodel in de minimale conditie ($SBS\chi^2(252) = 527,76$, $p < 0,01$) en in de uitgebreide conditie ($SBS\chi^2(252) = 399,15$, $p < 0,01$) werd vervolgens vergeleken met deze van een model waarin de drie subtests als drie aparte factoren werden beschouwd. In de minimale conditie bleek het driefactormodel niet significant beter bij de data te passen dan een - spaarzamer - model met slechts één factor ($\Delta SBS\chi^2(3) = 4,41$, ns). Bovendien wezen de onderlinge correlaties tussen de drie factoren op een ongepaste oplossing (improper solution; Wothke, 1993; correlaties tussen factoren groter dan 1). In de uitgebreide conditie leverde het driefactormodel wel een betere passing ($\Delta SBS\chi^2(3) = 10,60$, $p < 0,05$), maar een eveneens ongepaste oplossing (nl. correlaties tussen factoren groter dan 1) op. Dit doet ons besluiten dat een éénfactormodel, zoals voorspeld, het beste bij de gegevens past en dit in beide condities.

3.3 Criteriumvaliditeit (predictieve validiteit) en convergente validiteit

Leervaardigheid

De correlaties tussen de scores voor leervaardigheid zoals beoordeeld door de leerkracht en de Covaar-II zijn in beide condities significant ($r(150) = 0,49$ ($p < 0,01$) en $r(139) = 0,47$ ($p < 0,01$)) en hoog. Er is bovendien geen signifi-

cant verschil tussen de correlatie voor de minimale versus de uitgebreide afnameconditie.

zijn de correlatieverschillen tussen condities niet significant.

Vorderingen

De correlaties tussen de leerkrachtinschattingen van de vorderingen van de leerlingen doorheen het schooljaar en de Covaar-II-scores zijn significant op 1%-niveau en matig tot hoog, zowel bij afname na minimale als na uitgebreide training ($r(153) = 0,39$ en $r(145) = 0,43$). Het verschil in correlaties tussen de minimale en de uitgebreide oefenfase is wederom niet significant.

Examenresultaten, procentueel

Het totale percentage dat de leerling behaalde op de eindexamens vormt een volgend criterium ter evaluatie van de predictieve validiteit. De volgende correlaties worden gevonden voor de minimale conditie: $r(82) = 0,47$ ($p < 0,01$) respectievelijk de uitgebreide conditie: $r(68) = 0,22$ (*ns*). Het verschil in correlatie tussen minimale en uitgebreide conditie is niet significant.

Examenresultaten, vergelijkend

Er werden matige tot hoge positieve correlaties vastgesteld tussen de leerkrachtinschatting van de examenresultaten in vergelijking met klasgenoten en de Covaar-II-scores: $r(82) = 0,44$ ($p < 0,01$) in de minimale conditie en $r(68) = 0,37$ ($p < 0,01$) in de uitgebreide conditie. Wederom

Oriëntering

Multinomiale logistische regressieanalyse wijst uit dat de resultaten van de OKAN-leerlingen op de Covaar-II aan het begin van het jaar de oriëntering van de leerlingen aan het einde van het jaar significant voorspellen (Tabel 2). De *odds ratios* in Tabel 2 geven aan dat met stijgende Covaar-II-scores, de waarschijnlijkheid vergroot dat een leerling naar het ASO of het TSO wordt georiënteerd in plaats van naar het BSO (de referentiecategorie), en dit zowel in de minimale als in de uitgebreide conditie. Wanneer we ASO en TSO vergelijken, blijkt uit de *odds ratios* dat met stijgende Covaar-II-scores de waarschijnlijkheid daalt dat de betreffende leerling naar het TSO wordt georiënteerd in plaats van naar het ASO (referentiecategorie), maar dit enkel wanneer de Covaar-II wordt afgenomen na een uitgebreide oefenfase.

Kennis van het Nederlands aan het einde van het schooljaar

De correlatie tussen de Covaar-II-scores en de kennis van het Nederlands aan het einde van het schooljaar is zowel in de minimale ($r(154) = 0,21$, $p < 0,01$) als in de uitgebreide afnameconditie significant ($r(144) = 0,34$, $p < 0,01$). Deze correlaties zijn matig en verschil-

Tabel 2

Odds Ratios van de Multinomiale Regressieanalyse voor de Voorspelling van het Oriënteringsadvies (ASO, TSO of BSO) aan het Einde van het Schooljaar met als Predictor de Covaar-II scores

Predictoren	χ^2 ^a	TSO vs. BSO ^c	ASO vs. BSO ^c	TSO vs. ASO ^d
		OR (95%BI) ^b	OR (95%BI)	OR (95%BI)
<i>Minimale conditie</i> Covaar-II scores	34,62***	1,08 (1,03-1,12)**	1,15 (1,07-1,23)***	0,94 (0,88-1,00)
<i>Uitgebreide conditie</i> Covaar-II scores	40,70***	1,06(1,03-1,10)**	1,18 (1,09-1,28)***	0,90 (0,83-0,98)*

Noof. De aantallen respondenten verwezen naar ASO, TSO en BSO waren voor de minimale en uitgebreide conditie resp. 31 en 27, 34 en 35, 52 en 54; OR= *odds ratio*, BI= betrouwbaarheidsintervallen.

^a χ^2 is het verschil in -2 log-likelihoods tussen het finale model en het model waarin het betreffende effect is weggelaten.

^bDe Wald-statistiek wordt gebruikt om te testen of de *odds ratios* significant verschillend zijn van 1.

^cBSO is de referentiecategorie.

^dASO is de referentiecategorie.

* $p < 0,05$ ** $p < 0,01$ *** $p < 0,001$

len niet significant van elkaar. Ook de partiële correlaties (pr) tussen de Covaar-II-scores van de OKAN-leerlingen en de kennis van het Nederlands aan het einde van het jaar, gecorrigeerd voor aanvankelijke kennis van het Nederlands, zijn zowel in de minimale afnameconditie ($pr(142) = 0,17, p < 0,05$) als in de uitgebreide afnameconditie significant ($pr(132) = 0,18, p < 0,05$). Dus hogere Covaar-II-scores gaan gepaard met een grotere kennis van de Nederlandse taal aan het einde van het schooljaar, ongeacht de uitgebreidheid van de oefenfase. Dit verband zwakt af, maar blijft significant wanneer gecorrigeerd wordt voor de aanvankelijke kennis van het Nederlands.

TASAN

De correlaties met de verschillende subscores van de TASAN zijn voor beide afnamecondities weergegeven in Tabel 3. Terwijl in de minimale afnameconditie alle correlaties significant zijn, ten minste op 5%-niveau, is dit niet voor alle subtestscores in de uitgebreide conditie het geval. Wanneer men de correlaties voor de minimale en de uitgebreide conditie vergelijkt, verschillen deze echter bij geen enkele subtest significant van elkaar.

Progressive Matrices

Ten slotte werd een correlatie van 0,80 ($p < 0,01, n = 33$) vastgesteld tussen de Covaar-II-scores en de scores op de Progressive Matrices, welke hoog kan genoemd worden.

3.4 Effect van de oefenfase op de Covaar-II-scores

Er wordt geen significant hoofdeffect gevonden van afnameconditie (minimaal versus uitgebreid) ($F(1, 461) = 0,12, ns$) op de Covaar-II-scores. De verwachting dat een uitgebreidere training zou leiden tot hogere gemiddelde scores, wordt dus niet bevestigd.

4 Discussie

De eerste doelstelling van dit onderzoek was het evalueren van de betrouwbaarheid en de validiteit van de Covaar-II bij een nieuwe doelgroep, namelijk OKAN-leerlingen. Dit onderzoek toont aan dat de betrouwbaarheid, in termen van interne consistentie, goed te noemen is.

Voor de evaluatie van de begripsvaliditeit werd gekeken naar de correlaties tussen de subtestscores en naar de factorstructuur. Uit deze intersubtestcorrelaties en de confirmatorische factoranalyse op de itemscores blijkt dat de test ook bij adolescenten één onderliggende vaardigheid meet, die gelabeld kan worden als inductief redeneren. De constructvaliditeit van de Covaar-II in de OKAN-groep toont zich ook in de hoge correlatie met de scores op een andere test voor inductief redeneren, namelijk de Progressive Matrices (PM). Opgemerkt dient wel dat slechts van een zeer beperkt aantal leerlingen ($n = 33$) de PM-scores voorhanden waren. Dat de samenhang tussen beide tests

Tabel 3

Correlaties tussen de scores op de subtests van de TASAN en de Covaar-II. Scores voor de minimale ($n = 74-75$) en de uitgebreide conditie ($n = 62-64$) en voor beide samen (totaal, $n = 138-140$)

TASAN-subtests	Conditie		
	Minimaal	Uitgebreid	Totaal
Luisteren	0,44**	0,57**	0,48**
Lezen	0,50**	0,47**	0,46**
Schrijven	0,47**	0,22	0,35**
Receptief	0,52**	0,55**	0,51**
Productief	0,47**	0,22	0,35**
Structurerend	0,56**	0,52**	0,53**
Beschrijvend	0,26*	0,24	0,23**
Ondersteuning, veel	0,40**	0,56**	0,45**
Ondersteuning, weinig	0,57*	0,49**	0,51**

** $p < 0,01$, * $p < 0,05$

hoog is maar niet perfect, is begrijpbaar. Beide tests kunnen niet als geheel equivalent worden beschouwd. Zo is de afnamewijze van de Covaar-II anders dan die van de PM en meer specifiek afgestemd op het gebruik bij groepen van anderstalige nieuwkomers (bv. non-verbale instructies, omvangrijke oefenfase in uitgebreide conditie).

De resultaten tonen dat de criteriumvaliditeit, zoals geëvalueerd aan de hand van examenresultaten en het oriënteringsadvies aan het einde van het jaar, goed te noemen is. Tevens worden de verwachte verbanden vastgesteld met de kennis van het Nederlands aan het einde van het jaar zoals beoordeeld door de leerkrachten en met scores op de TASAN, een taaltest specifiek afgestemd op de doelgroep adolescenten anderstalige nieuwkomers. In de onthaalklassen voor anderstalige nieuwkomers is het onderwijs vooral gericht op het aanleren van de Nederlandse taal. Bij het gebruik van de Covaar-II als hulpmiddel bij het geven van taalonderwijs op maat van deze leerlingen, is het natuurlijk aangewezen dat er een verband is met taalverwerving, wat in deze studie wordt teruggevonden. Wat betreft de kennis van het Nederlands aan het einde van het schooljaar zoals beoordeeld door de leerkracht zou men kunnen aanbrenge dat de partiële correlaties eerder beperkt zijn wat betreft effectgrootte. Dit suggereert dat ook andere factoren, naast cognitieve vaardigheden, voorspellend zijn voor vorderingen in het verwerven van de Nederlandse taal. Men dient hierbij echter ook in het achterhoofd te houden dat gecorrigeerd is voor aanvankelijke kennis van het Nederlands aan het begin van het schooljaar. Op dat moment verbleven de OKAN-leerlingen gemiddeld evenwel al bijna acht maanden in België. Mogelijk hadden leerlingen met sterkere cognitieve vaardigheden op dat moment reeds grotere kennis van het Nederlands verworven, waardoor er voor hen ook minder toename mogelijk was in de loop van het schooljaar. Dit fenomeen kan de partiële correlaties gedrukt hebben.

In dit onderzoek (doelstelling 2) gingen we eveneens na wat de invloed is van de oefenfase op de gemiddelde scores van OKAN-leerlingen. Bij een directe toetsing van het effect van deze oefenfase op de hoogte van de Covaar-II-scores, blijkt dit effect, tegen de oorspronkelijk-

ke hypothese in, niet significant te zijn. Een derde doelstelling handelde over de invloed van de oefenfase op de betrouwbaarheid en de validiteit (begripsvaliditeit, criteriumvaliditeit) van de Covaar-II-scores. Hier wijzen nagenoeg alle gegevens uniform in dezelfde richting: de oefenfase leidt niet tot een verbetering van de betrouwbaarheid of de validiteit.

Er zijn verschillende verklaringen mogelijk voor het niet vinden van de verwachte gunstige effecten van de (uitgebreide) oefenfase. Zo is het mogelijk dat de oefenfase zoals deze nu is geoperationaliseerd nog niet uitgebreid genoeg is om een belangrijk effect uit te oefenen op de validiteit van de test. Echter, hierbij dient opgemerkt dat de oefenfase nu al meer tijd inneemt (minimaal een half uur) dan de eigenlijke test. Als alternatieve verklaring zou men kunnen opperen dat de Covaar-II geen gebruik maakt van domein- en curriculumspecifiek materiaal. In andere onderzoeken (Guthke & Beckmann, 2000; Hamers & Sijtsma, 1995) werden er betere resultaten (o.a. verhoogde predictieve waarde) verkregen met dynamische tests die gebruik maken van dergelijk materiaal. Verder pleiten bepaalde onderzoekers voor 'dynamische' criteria in plaats van meer statische varianten (Grigorenko & Sternberg, 1998; Guthke & Beckmann, 2000; Hamers & Sijtsma, 1995). In deze studie werden naast de examenresultaten (statisch criterium) ook een leerkrachtinschatting van de vorderingen van de leerlingen tijdens het jaar en de leervaardigheden opgevraagd. Deze, toch meer dynamische, criteria zijn echter subjectief, en er werd geen meer 'objectief' dynamisch criterium opgenomen.

Al deze factoren bemoeilijken het formuleren van definitieve besluiten over de bijdrage van oefenfases (zoals gebruikt binnen de DT-beweging) tot een meer faire diagnostiek van cognitieve vaardigheden in het algemeen. De bevindingen nodigen daarentegen uit tot een verdere, kritische evaluatie van het effect van dergelijke oefenfases aan de hand van proefopzetten die gebruik maken van een vergelijking tussen respondenten, naast de traditionele pretest-oefenfase-posttestdesigns die het effect van de oefenfase nagaan binnen respondenten. Pas wanneer anderen op basis van onderzoek met andere cognitieve vaardigheidstests tot

dezelfde conclusies komen, zal met zekerheid besloten kunnen worden of de investering in het uitwerken van een uitgebreide oefenfase de moeite waard is en onder welke condities een oefenfase mogelijk wel te verkiezen is. Dit onderzoek plaatst voorlopig vraagtekens bij de meerwaarde van oefenfases, tenminste op psychometrisch vlak. De keuze voor een cognitieve test met een oefenfase kan uiteraard wel gemaakt worden op grond van andere overwegingen. Zo kan de observatie van de wijze waarop een leerling omgaat met de geboden hulp tijdens de oefenfase handelingsgerichte aanwijzingen bieden, met name voor toekomstige instructie op maat van de leerling. Ook bij afwezigheid van psychometrische voordelen kan dit voordeel de keuze voor een dynamische testvorm verantwoorden.

Noten

- 1 De auteurs wensen de verschillende deelnemende CLB's en de stuurgroep schoolloopbaanbegeleiding van de Vrije CLB Koepel te bedanken voor hun hulp bij de dataverzameling en voor de geleverde feedback in verband met het gebruik van de test.
- 2 Voor een beoordeling van de grootte van de correlaties werden de criteria van Cohen (1988) als vuistregels genomen: $r = 0,10$ is laag, $r = 0,30$ is matig en $r = 0,50$ is hoog.

Literatuur

Berry, J. W., Poortinga, Y. H., Segall, M. H., & Dasen, P. R. (2002). *Cross-cultural psychology: Research and applications (2nd Ed.)*. Cambridge, Verenigd Koninkrijk: Cambridge University Press.

Buyse, E., Vanhee, K., Magez, W., & Verschuere, K. (2003a). *Evaluatieonderzoek Covaar*. Rapport in opdracht van het Departement Onderwijs van het Ministerie van de Vlaamse Gemeenschap, afdeling Volwassenenonderwijs.

Buyse, E., Vanhee, K., Magez, W., & Verschuere, K. (2003b). *Ontwikkeling en evaluatie van de Covaar-II*. Eindrapport van een onderzoek in opdracht van het Departement Onderwijs van het Ministerie van de Vlaamse Gemeenschap, afdeling Volwassenenonderwijs. Brussel: De-

partement Onderwijs van het Ministerie van de Vlaamse Gemeenschap.

Buyse, E., Verschuere, K., & Magez, W. (2005). De ontwikkeling en validering van een cognitieve-vaardighedentest voor volwassen anders-taligen. *Nederlands Tijdschrift voor de Psychologie*, *60*, 15-25.

Carroll, J. B. (1993). *Human cognitive abilities: A survey of factor-analytic studies*. Cambridge, MS: Cambridge University Press.

Cohen, J. (1988). *Statistical power analysis for the behavioral sciences (revised edition)*. Hillsdale, NJ: Erlbaum.

Day, J. D., Engelhardt, S. E. M., Maxwell, J. L., & Bolig, E. E. (1997). Comparison of static and dynamic assessment procedures and their relation to independent performance. *Journal of Educational Psychology*, *89*, 358-368.

Elliott, J. (2003). Dynamic assessment in educational settings: Realising potential. *Educational Review*, *55*, 15-28.

Evers, A., Vliet-Mulder, J. C. van, & Groot, C. J. (2000). *Documentatie van tests en test-research in Nederland*. Amsterdam/Assen: NIP/van Gorcum.

Fabio, R. A. (2005). Dynamic assessment of intelligence is a better reply to adaptive behavior and cognitive plasticity. *The Journal of General Psychology*, *123*, 41-64.

Ganzeboom, H. B. G., Graaf, P. M. de, & Treiman, D. J. (1992). A standard international socioeconomic index of occupational status. *Social Science*, *21*, 1-56.

Gottfredson, L. S. (2000). Skills gaps, not tests, make racial proportionality impossible. *Psychology, Public Policy and Law*, *6*, 129-143.

Grigorenko, E. L., & Sternberg, R. J. (1998). *Dynamic testing*. *Psychological Bulletin*, *124*, 75-111

Guthke, J., & Beckmann, J. F. (2000). The learning test concept and its application in practice. In C.S. Lidz & J. Elliot (Eds.), *Advances in cognition and educational practice*. Vol. 6. *Dynamic assessment: Prevailing models and applications* (pp. 17-69). New York: Elsevier.

Guthke, J., & Stein, H. (1996). Are learning tests the better version of intelligence tests? *European Journal of Psychological Assessment*, *12*, 1-13.

Gysen, S., Ramaut, G., & Sterckx, M. (2002). *TASAN. Toets Aanvang Secundair Anderstalige Nieuwkomers*. Leuven, België: Centrum

- voor Taal en Migratie/Steunpunt Nederlands als Tweede Taal.
- Hamers, J. H. M., Hessels, M. G. P., & Pennings, A. H. (1996). Learning potential in ethnic minority children. *European Journal of Psychological Assessment, 12*, 183-192.
- Hamers, J. H. M., & Sijtsma, K. (1995). Trends in learning potential assessment. In J.S. Carlson (Ed.), *Advances in Cognition and Educational Practice, Volume 3. European Contributions to dynamic assessment* (pp. 83-115). Greenwich, CT: JAI Press.
- Hessels, M. G. P., & Hamers, J. H. M. (1993). A learning potential test for ethnic minorities. In J. Hamers, K. Sijtsma, & A. Ruijsenaars (Eds.), *Learning potential assessment: theoretical, methodological and practical issues* (pp. 285-311). Lisse, Nederland: Swets & Zeitlinger.
- Horn, J. L., & Cattell, R. B. (1966). Refinement and test of the theory of fluid and crystallized intelligence. *Journal of Educational Psychology, 57*, 253-270.
- Hu, L., & Bentler, P. M. (1999). Cutoff criteria for fit indexes in covariance structure analysis: Conventional criteria versus new alternatives. *Structural equation modeling, 6*, 1-55.
- Jöreskog, K. G. (2005). *Structural equation modeling with ordinal variables using LISREL*. Chicago: Scientific Software.
- Kline, R. B. (2005). *Principles and practice of structural equation modeling (2nd ed.)*. New York: The Guilford Press.
- Laak, J. J. F. ter, & Goede, P. M. de. (2005). *Psychologische diagnostiek. Inhoudelijke en methodologische grondslagen*. Amsterdam: Harcourt Assessment.
- Lidz, C. L. (1987). *Dynamic Assessment: An interactional approach to evaluating learning potential*. New York: Guildford Press.
- McGrew, K. S., & Flanagan, D. P. (1998). *The Intelligence Test Desk Reference (ITDR). Gf-Gc cross-battery assessment*. Boston: Allyn & Bacon.
- Magetz, W. (2003). *Diagnostiek bij allochtonen*. Schaarbeek, België: VCLB-Service.
- Magetz, W., Grysolle, R., Bos, A., & De Cleen, W. (2001). *Cap-vademecum van diagnostische instrumenten en methoden voor C.L.B., versie 2001*. Antwerpen, België: Coördinatieteam Antwerpen voor Psychodiagnostiek- CAP vzw.
- Nijenhuis, J. te, & Flier, H. van der. (1999). Bias research in The Netherlands: Review and implications. *European Journal of Psychological Assessment, 15*, 165-175.
- Nijenhuis, J. te, & Vijver, F. J. R. van der. (2001). Onderzoek naar partijdigheid. In N. Bleichrodt & F. van de Vijver (red.), *Diagnostiek bij allochtonen: mogelijkheden en beperkingen van psychologische tests* (pp. 61-88). Lisse, Nederland: Swets & Zeitlinger.
- Raven, J. (2000). The Raven's Progressive Matrices: Change and stability over culture and time. *Cognitive Psychology, 40*, 1-48.
- Raven, J. G., Court, J. H., & Raven, J. (1979). *Manual for Raven's progressive matrices and vocabulary scales*. London: HK Lewis.
- Resing, W. (2001). Het meten van de cognitieve mogelijkheden en het schoolgedrag van allochtone kinderen. In N. Bleichrodt & F. van de Vijver (red.), *Diagnostiek bij allochtonen: Mogelijkheden en beperkingen van psychologische tests* (pp. 89-118). Lisse, Nederland: Swets & Zeitlinger.
- Resing, W. (2007). Intelligentie en leervermogen: statisch en dynamisch testen. In K. Verschueren & H. Koomen (red.), *Handboek diagnostiek in de leerlingenbegeleiding* (pp. 105-123). Antwerpen, België: Garant.
- Resing, W. C. M., & Hessels, M. G. P. (2001). Het meten van de cognitieve mogelijkheden en het schoolgedrag van allochtone kinderen. In N. Bleichrodt & F. van de Vijver (red.), *Diagnostiek bij allochtonen: mogelijkheden en beperkingen van psychologische tests* (pp. 89-118). Lisse, Nederland Swets & Zeitlinger.
- Reynolds, C. R., & Kaiser, S. M. (2003). Bias in assessment of aptitude. In C. Reynolds & R. Kamphaus (Eds.), *Handbook of psychological & educational assessment of children. Intelligence, aptitude and achievement* (pp. 519-562). New York: Guildford New York.
- Rogers, T., & Yang, P. (1996). Test-wiseness: Its nature and application. *Journal of Psychological Assessment, 12*, 247-259.
- Rushton, J. P. (1998). The "Jensen effect" and the "Spearman-Jensen hypothesis" of black-white IQ differences. *Intelligence, 26*, 217-225.
- Satorra, A., & Bentler, P. M. (1999). *A scaled difference chi-square test statistic for moment structure analysis*. UCLA Statistics Series #260. Los Angeles: University of California.
- Schmidt, L. R. (1971). Testing the limits im leistungsverhalten: möglichen und grenzen [Testing the limits in performance behavior:

- possibilities and limitations]. In E. Duhn (Hrsg.), *Praxis der Klinischen Psychologie* (pp. 9-29). Göttingen, Duitsland: Hogrefe.
- Sternberg, R. J., & Grigorenko, E. L. (1999). Myths in psychology and education in the gene-environment debate. *Teachers College Record, 100*, 536-553.
- Sternberg, R. J., & Grigorenko, E. L. (2002). *Dynamic testing*. New York: Cambridge University Press.
- Sternberg, J. S., Grigorenko, E. L., Ngorosho, D., Tantufuye, E., Mbise, A., Nokes, C., Jukes, M., & Bundy, D. A. (2002). Assessing intellectual potential in rural Tanzanian school children. *Intelligence, 30*, 141-162.
- Vande Gaer, E. (2007). *De Covaar-II, een cognitieve vaardighedentest voor anderstalige nieuwkomers. Validatie van de test en differentieële effecten van de oefenfase bij beroepsleerlingen en anderstalige nieuwkomers in het secundair onderwijs*. Niet-gepubliceerde licentiaatsverhandeling. Katholieke Universiteit Leuven, Leuven, België.
- Verboven, K., Buyse, E., Magez, W., & Verschueren, K. (2004). *Normerings- en valideringsonderzoek Covaar-II*. Rapport van een onderzoek in opdracht van het Departement Onderwijs van het Ministerie van de Vlaamse Gemeenschap, Afdeling Volwassenenonderwijs. Brussel: Departement Onderwijs van het Ministerie van de Vlaamse Gemeenschap.
- Vijver, F. J. R. van de, Fontaine, J., & Schittekatte, M. (2007). Diagnostiek bij allochtone leerlingen. In K. Verschueren & H. Koomen, H. (red.). *Handboek diagnostiek in de leerlingenbegeleiding* (pp. 251-263). Antwerpen, België: Garant.
- Vijver, F. J. R. van de, & Leung, K. (1997). *Methods and data analysis for cross-cultural research*. Thousand Oaks, CA: Sage.
- Vijver, F. J. R. van de, & Phaet, K. (2004). Assessment in multicultural groups: The role of acculturation. *Applied Psychology: An international review, 53*, 215-236.
- Wiedl, K. H., Guthke, J., & Wingenfeld, S. (1995). Dynamic assessment in Europe: historical perspectives. In J. Carlson (Ed.), *Advances in cognition and educational practice, Volume 3. European contributions to dynamic assessment* (pp.33-82). Greenwich, CT: JAI Press.
- Wothke, W. (1993). Nonpositive definite matrices in structural modeling. In K.A. Bollen & J.S. Long (Eds.), *Testing structural equation models* (pp. 256-294). London: Sage.

Manuscript aanvaard: 26 mei 2009

Auteur

Eva Vande Gaer voerde dit onderzoek uit in het kader van haar masterproef aan het Centrum voor Schoolpsychologie o.l.v. prof. Karine Verschueren. **Karine Verschueren, Evelien Buyse, Veerle Germeijs** en **Walter Magez** zijn werkzaam aan het Centrum voor Schoolpsychologie en Ontwikkelingspsychologie van Kind en Adolescent, Katholieke Universiteit Leuven, België.

Correspondentieadres: Eva Vande Gaer, Katholieke Universiteit Leuven, Centrum voor Methodologie van het Pedagogisch Onderzoek, Andreas Vesaliusstraat 2 - bus 3762, 3000 Leuven, België. Email: eva.vandegaer@psy.kuleuven.be.

Abstract

Dynamic testing of cognitive abilities of non-Dutch speakers in secondary schools in Flanders

In a sample of 462 non-Dutch speaking adolescent newcomers in secondary schools in Flanders psychometric properties of the Covaar-II, a cognitive abilities test, have been evaluated. The original version with an extensive training phase has been evaluated as well as a version with a limited training. The use of a training phase characterizes the dynamic testing tradition, which the Covaar-II relates to. The extensiveness of the training phase did not impact upon the test scores. Reliability and validity of the test scores were comparable in both conditions and were satisfactory. Both versions of the test can, therefore, be applied to evaluate the cognitive abilities of this group of students. Hence, the choice between both versions of the test depends on other than psychometric considerations.

Appendix 1: Voorbeeld-items Covaar-II

Verbale analogieën

NADER : ? PADER : PAD	NADAR	NAD	ER	PAD
--------------------------	-------	-----	----	-----

Figuratieve reeksen

--	--

Analogieën met symbolen

+Δ O : +ж O L Δ O : ?	+Δ O	L Δ O	L ж O	+ж
--------------------------	------	-------	-------	----

Appendix 2: Voorbeeld van non-verbale instructie bij oefening ‘verbale analogieën’

LAN : LANOS

KAN : ?

KONOS KANOS KANAS LANAS

Instructies:

LAN : LANOS

KAN : ?

KONOS KANOS KANAS LANAS

Opgave oefening

1. Wijs de drie lettergroepen ('LAN' voor het dubbelpunt en 'LAN'en 'OS' na het dubbelpunt) aan in de eerste lijn van het voorbeeld.
2. Wijs vervolgens de lettergroep ('KAN') aan in de tweede lijn en schrijf het juiste antwoord ('KANOS') onder het vraagteken.
3. Geef op de eerste lijn met boogjes en een pijltje het behoud van de eerste lettergroep ('LAN') aan en geef vervolgens met een cirkeltje het toevoegen van de lettergroep 'OS' aan.
4. Doe hetzelfde in de tweede lijn: geef met boogjes en een pijltje het behoud van de eerste lettergroep ('KAN') aan en geef vervolgens met een cirkeltje het toevoegen van de lettergroep 'OS' aan.
5. Wijs de verschillende antwoordalternatieven één voor één aan en omcirkel vervolgens het correcte antwoordalternatief.
6. Geef aan dat de leerlingen dit antwoordalternatief ook mogen omcirkelen in hun oefenboekje.