KOHNSTAMM INSTITUUT

its

Doelgroepleerlingen in het basisonderwijs

JAAP ROELEVELD
GEERT DRIESSEN
GUUSKE LEDOUX
JORIS CUPPEN
JOOST MEIJER

Doelgroepleerlingen in het basisonderwijs

Historische ontwikkeling en actuele situatie

JAAP ROELEVELD
GEERT DRIESSEN
GUUSKE LEDOUX
JORIS CUPPEN
JOOST MEIJER

CIP-gegevens KONINKLIJKE BIBLIOTHEEK, DEN HAAG

Roeleveld, J., Driessen, G., Ledoux, G., Cuppen, J., Meijer, J.
Doelgroepleerlingen in het basisonderwijs. Historische ontwikkeling en actuele situatie.
Amsterdam: Kohnstamm Instituut.
(Rapport 857, projectnummer 20432)

ISBN 978-90-6813-920-4

Alle rechten voorbehouden. Niets uit deze uitgave mag worden vermenigvuldigd, opgeslagen in een geautomatiseerd gegevensbestand, of openbaar gemaakt, in enige vorm of op enige wijze, zonder voorafgaande schriftelijke toestemming van de uitgever.

All rights reserved. No part of this publication may be reproduced, stored in a retrieval system, or transmitted, in any form or by any means, electronic, mechanical, photocopying, or otherwise, without the prior written permission of the publisher.

Uitgave en verspreiding:
Kohnstamm Instituut
Plantage Muidergracht 24, Postbus 94208, 1090 GE Amsterdam
Tel.: 020-525 1226
www.kohnstamminstituut.uva nl
© Copyright Kohnstamm Instituut, 2011

Inhoudsopgave

Voorwoord 1

1 Inleiding 3
1.1 Achtergronden en probleemstelling 3
1.2 Veranderingen in de gewichtenregeling 9
1.3 Gegevens uit cohortonderzoeken 12
1.4 Indeling van leerlingen 16

2 De ontwikkeling van de onderwijspositie van OAB-doelgroepleerlingen tussen 1994/95 en 2007/08 23
2.1 Achtergronden 23
2.2 Probleemstelling 27
2.3 Methode 28
2.4 Resultaten 33
2.5 Samenvatting, conclusie en discussie 55

3 De cognitieve ontwikkeling van doelgroepleerlingen gedurende het basisonderwijs 61
3.1 Inleiding 61
3.2 Data 64
3.3 Ontwikkeling van taalprestaties 69
3.4 Ontwikkeling van rekenprestaties 76
3.5 Ontwikkeling van doelgroep- versus niet-doelgroepleerlingen 83
3.6 Conclusies en discussie 85

4 Over- of onderadvisering bij doelgroepleerlingen? 89
4.1 Het advies voortgezet onderwijs 89
4.2 Onderzoeksvragen 97
4.3 Methode 98
4.4 Resultaten 102
4.5 Conclusies 122

5 Cognitieve reserves bij doelgroepleerlingen 127
5.1 Inleiding 127
5.2 Methode en data 130
5.3 Resultaten 132
5.4 Alternatieve verklaringen? 141
5.5 Conclusies 143

6 Samenvatting 149
6.1 Achtergronden, vraagstellingen en opzet 149
6.2 Actuele taal- en rekenprestaties van de verschillende groepen in het basisonderwijs 152
6.3 Actuele positie van de verschillende groepen in het voortgezet onderwijs 154
6.4 Halen doelgroepleerlingen achterstanden in? 155
6.5 Krijgen doelgroepleerlingen en niet-doelgroepleerlingen verschillende adviezen voor het voortgezet onderwijs? 157
6.6 Cognitieve reserves bij doelgroepleerlingen? 159
6.7 Resumerend 160

Referenties 161

Recent uitgegeven Kohnstamm Instituut rapporten 171

Voorwoord

Sinds de jaren zeventig van de vorige eeuw wordt er door de Nederlandse overheid beleid gevoerd ter
bestrijding van onderwijsachterstanden die het gevolg zijn van sociale, economische en culturele factoren. Om vast te stellen of er verbetering optreedt in de onderwijskansen van leerlingen die behoren tot de doelgroepen van het achterstandenbeleid, wordt vanaf het schooljaar 1988/89 in het basisonderwijs en vanaf schooljaar 1989/90 in het voortgezet onderwijs cohortonderzoek uitgevoerd. Tot 2005 ging het om het LEO- het PRIMA-cohortonderzoek in het basisonderwijs en PRIMA-VO en VOCL in het voortgezet onderwijs. Daarna zijn deze cohorten geïntegreerd voortgezet onder de naam COOL$^{5\text{-}18}$. Dankzij deze onderzoeken zijn er uitgebreide databestanden opgebouwd, die informatie bevatten over meerdere leerlingcohorten in verschillende fasen van de schoolloopbaan in basis- en voortgezet onderwijs. Deze zijn al vele malen benut voor onderzoek naar de ontwikkeling van de prestaties en loopbanen van de doelgroepen van het onderwijsachterstandenbeleid.

Dit rapport bevat over dit onderwerp nieuwe gegevens, op basis van nieuwe analyses op de cohortdata, waaronder de meeste recente uit COOL$^{5\text{-}18}$. Twee hoofdonderwerpen komen aan de orde. In de eerste plaats wordt gerapporteerd over verschillen in leerprestaties in het basisonderwijs tussen de doelgroepen van het beleid en overige leerlingen, en over ontwikkelingen daarin in de tijd. In de tweede plaats wordt ingegaan op de vraag of er bij doelgroepleerlingen meer dan bij andere leerlingen sprake is van onbenut

talent en van discrepanties tussen leerprestaties en het advies voor het voortgezet onderwijs.

Het onderzoek is uitgevoerd door het Kohnstamm Instituut in Amsterdam en het ITS in Nijmegen, in opdracht van de Programmacommissie Beleidsgericht Onderzoek Primair Onderwijs (BOPO). Het maakt deel uit van de BOPO-programmalijn *Effecten van beleidsontwikkelingen in het onderwijs–achterstandenbeleid*.

Jaap Roeleveld, Guuske Ledoux, Joost Meijer (Kohnstamm Instituut)
Geert Driessen, Joris Cuppen (ITS)

N.B. Aan dit onderzoek is verder meegewerkt door H. Vierke en W. de Wit van het ITS.

1 Inleiding

Guuske Ledoux en Jaap Roeleveld

1.1 Achtergronden en probleemstelling

Bestrijding van onderwijsachterstanden is al sinds decennia een belangrijk thema in het onderwijsbeleid van de overheid. Het accent ligt op het basisonderwijs, maar de laatste jaren ook steeds meer op de voorschoolse periode. Er is al veel onderzoek op dit terrein gedaan, met name beleidsonderzoek. Een belangrijke opdrachtgever voor dit onderzoek was en is de Programmacommissie Beleidsgericht Onderzoek Primair Onderwijs (BOPO). Twee reviews die op verzoek van de BOPO zijn uitgevoerd, hebben de stand van zaken opgemaakt in de kennis verkregen uit onderzoek naar onderwijsachterstanden van de laatste twintig jaar (Meijnen, 2003; Doolaard & Leseman, 2008).
Een essentiële bron voor die kennis vormt het cohortonderzoek dat in de loop der jaren is uitgevoerd in het basis- en voortgezet onderwijs en dat ontwikkelingen zichtbaar maakt in de onderwijsloopbanen van kinderen en jongeren die tot de doelgroepen van het achterstandenbeleid behoren (bv. Tesser, 2003; Mulder e.a., 2005). Sinds 1994 (de start van het PRIMA-cohortonderzoek) is het mogelijk om de prestaties van leerlingen te volgen vanaf groep 2 tot en met groep 8 in het basisonderwijs, en voor een deel van deze leerlingen ook de loopbanen in het voortgezet onderwijs. In 2007 is PRIMA overgegaan in het nieuwe cohortonderzoek COOL$^{5\text{-}18}$, waardoor de mogelijkheden om leerlingen longitudinaal te volgen verder zijn uitgebreid.

Vanaf 2009 is daar nog een informatiebron bijgekomen, namelijk het cohortonderzoek voor- en vroegschoolse periode (Pre-COOL) voor kinderen van 2 tot 5 jaar.
Met de cohortdata kan periodiek een stand van zaken worden opgemaakt over de prestaties en de loopbanen van alle leerlingen in het Nederlands onderwijs, waaronder de doelgroepen van het onderwijsachterstandenbeleid. In drie eerdere BOPO-studies (Van Langen & Suhre, 2001; Mulder e.a., 2005; Guldemond & Bosker, 2006) is dat gedaan voor de perioden 1988-1996 en 1988-2002. Het accent in deze studies ligt op het volgen van ontwikkelingen in de tijd. De hoofdvraag daarbij was of en in hoeverre leerlingen uit de achterstandgroepen hun achterstand op de overige leerlingen inlopen. Het verminderen van die achterstand is de hoofddoelstelling van het onderwijsachterstandenbeleid.

De voorliggende studie vormt een vervolg op deze eerdere studies. Ditmaal is de periode 1994-2007 in beschouwing genomen. Opnieuw gaat het om de vraag hoe de prestaties en loopbanen van de doelgroepleerlingen zich verhouden tot die van de overige leerlingen. Vergeleken met de voorgaande studies zijn er echter wel enkele verschillen:

- De focus ligt nadrukkelijk op het basisonderwijs en minder op het voortgezet onderwijs. De belangrijkste reden daarvoor is dat de opzet van COOL$^{5\text{-}18}$ voor het voortgezet onderwijs afwijkt van de eerdere cohorten in deze sector; vergelijkingen in de tijd zijn daardoor voor de onderwijsposities en –resultaten in het voortgezet onderwijs op dit moment minder goed mogelijk.
- Er is vooral gebruik gemaakt van de data over leerprestaties van leerlingen en (in mindere mate) van leerkrachtoordelen over de leerlingen. Loopbaanvariabelen als zittenblijven, verwijzing naar het speciaal onderwijs zijn dit keer niet onderzocht.
- Nieuw daarentegen is dat er gebruik wordt gemaakt van een verder gedifferentieerde indeling in groepen leerlingen. Zowel binnen de groep doelgroepleerlingen als binnen de groep overige leerlingen zijn nieuwe subgroepen onderscheiden (deze worden later in dit hoofdstuk toegelicht). Dit maakt het mogelijk nauwkeuriger uitspraken te doen over de relatieve positie van de verschillende groepen achterstandsleerlingen ten opzichte van niet-achterstandsleerlingen en

ten opzichte van elkaar. Deze indeling in groepen is in elke analyse van dit onderzoek toegepast.

- Verder is voor een deel van de vraagstellingen gebruik gemaakt van nieuwe analysetechnieken. Hierdoor kunnen preciezere schattingen worden gegeven van bepaalde effecten en ontwikkelingen.

Een belangrijke overeenkomst is er echter ook: net als in de vorige studies zijn er zowel cross-sectionele analyses uitgevoerd als longitudinale analyses. Bij cross-sectionele analyses wordt de blik gericht op vergelijkingen tussen cohorten (bijvoorbeeld: presteren Turkse leerlingen in 2002 anders dan in 2007?) en bij longitudinale analyses op ontwikkeling van leerlingen gedurende hun schoolloopbaan (bijvoorbeeld: hoeveel leerwinst hebben leerlingen geboekt bij rekenen tussen groep 4 en groep 8 in het basisonderwijs?). Beide typen analyses komen in dit rapport aan bod.

Doel en kader van de studie

Het onderzoek is één van de deelprojecten in de BOPO-programmalijn *Effecten van beleidsontwikkelingen in het onderwijsachterstandenbeleid* (zie voor de lijnen in het BOPO-programma Doolaard & Mulder, 2009). De algemene vraagstelling van deze programmalijn luidt:

> *Welke gevolgen hebben actuele wijzigingen in het onderwijsachterstandenbeleid voor de organisatie, inrichting en effecten van achterstandenbestrijding in het basisonderwijs en de voorschoolse voorzieningen, en hoe verhoudt zich dit tot de ambities van de overheid?*

Het gehele programma omvat zes deelprojecten:

1. De rol van gemeenten en schoolbesturen bij voor- en vroegschoolse educatie
2. Voorschoolse opvang en educatie in peuterspeelzalen en kinderopvanginstellingen
3. Aansluiting tussen voor- en vroegschoolse educatie en tussen vroegschoolse educatie en verdere schoolloopbaan
4. Effecten van voor- en vroegschoolse educatie en opvang op latere ontwikkeling van kinderen
5. Prestaties en loopbanen van doelgroepleerlingen in het onderwijsachterstandenbeleid

6. Effecten van wijzigingen in de gewichtenregeling en andere OAB-faciliteiten

Twee van deze projecten hebben betrekking op het effect van het beleid op leerlingen, te weten project 4 en het onderhavige project, project 5. Project 4 gaat over de ontwikkeling van peuters en kleuters, project 5 over het basisonderwijs en (beperkt) het voortgezet onderwijs. In de titel van project 5 is het woord 'effect' niet gebruikt. De reden daarvoor is dat het directe effect van het onderwijsachterstandenbeleid als geheel niet gemeten kan worden. Dat heeft een aantal oorzaken:

- Het beleid bestaat uit een aantal maatregelen die *in onderlinge samenhang* moeten bijdragen aan het behalen van het beleidsdoel, het terugdringen van achterstanden. Een deel van die maatregelen is bovendien generiek, dat wil zeggen van toepassing op *alle* scholen in Nederland. Er kan dus niet experimenteel worden vastgesteld, met behulp van controlegroepen die niet aan het beleid worden blootgesteld, of het beleid effectief is in strikt methodologische zin (Tesser, 2003; Ledoux & Veen, 2009).
- Het beleid is bovendien steeds aan verandering onderhevig. Doelgroepen, speerpunten en beleidsinstrumenten zijn in de tijd diverse malen veranderd (Meijnen e.a., 2003; Mulder e.a., 2005; Ledoux & Veen, 2009). Bij onderzoek naar ontwikkeling in de tijd, zoals hier aan de orde is, speelt dus het probleem van een niet constant gehouden 'interventie'.

Een beleidseffectonderzoek in strikte zin is de onderhavige studie dus niet (vgl. Mulder, 1996). Maar wat deze studie wel kan, net als de eerdere onderzoeken naar ontwikkelingen in de tijd, is laten zien in hoeverre de leerlingen die tot de doelgroepen van het onderwijsachterstandenbeleid behoren hun onderwijsachterstand inhalen. Er kunnen dus wel uitspraken worden gedaan over het halen van de beleidsdoelen, ook al kunnen eventuele resultaten niet rechtstreeks aan het gevoerde beleid worden toegeschreven.

Beleidsdoelen en beleidsmaatregelen

Het hoofddoel van het onderwijsachterstandenbeleid is vermindering van de ongelijkheid in schoolprestaties die samenhangt met groepskenmerken, met name sociaal milieu en etnische herkomst. Lange tijd is dit doel niet nader geoperationaliseerd geweest. Bij de start van de decentralisatie van het

onderwijsachterstandenbeleid naar de gemeenten, in 1998, zijn voor het eerst meer precieze doelen geformuleerd in een landelijk beleidskader. Vier jaar later zijn deze doelstellingen verder aangescherpt in een poging ze beter meetbaar te maken. Er werden toen streefpercentages aan verbonden. Omdat deze percentages gebonden zijn aan een specifieke beleidsperiode (2002-2006), en we hier een langere periode onderzoeken, nemen we deze percentages hier niet als uitgangspunt. We vermelden alleen de ongespecificeerde doelen die sinds 1998 gesteld zijn voor het basisonderwijs (Ledoux & Veen, 2009):

- verbeteren van de startpositie van doelgroepleerlingen bij binnenkomst in de basisschool;
- terugdringen van de taalachterstand bij de doelgroepleerlingen;
- verhogen van de doorstroom van doelgroepleerlingen naar hogere vormen van voortgezet onderwijs.

Deze drie onderwerpen komen in deze studie aan de orde.

Om deze doelen te bereiken, zijn verschillende beleidsmaatregelen ingezet. Voor het eerste doel, verbeteren van de startpositie, is de Voor- en Vroegschoolse Educatie (VVE) het belangrijkste instrument. Deelname van doelgroepleerlingen aan voorschoolse voorzieningen (kinderopvang, peuterspeelzalen, voorscholen) die specifieke programma's aanbieden, doorloop van die programma's in de kleutergroepen en stimuleringsprogramma's gericht op ouders zijn hiervan de voornaamste componenten. Voor de beide andere doelen, verbetering van de prestaties van doelgroepleerlingen gedurende en aan het eind van het basisonderwijs, is in de loop van de tijd een hele reeks maatregelen ingezet. De oudste en naar financiële maatstaven de belangrijkste is de gewichtenregeling. Dit is een regeling die voorziet in extra formatie voor scholen met een substantieel aantal doelgroepleerlingen. Doelgroepleerlingen tellen in de formatieberekening extra zwaar mee en dit levert voor de school dus extra middelen op. De gewichtenregeling heeft in de loop der jaren enkele wijzigingen ondergaan, dit komt in de volgende paragraaf aan de orde. Naast de gewichtenregeling zijn er andere instrumenten ingezet: extra geld en scholing op het onderwerp Nederlands als tweede taal (NT2; voor de doelgroep allochtone leerlingen), voorzieningen voor het leren van de eigen taal (OETC, OET, OALT; idem), de Brede School, het Onderwijskansenbeleid (gericht op risicovolle scholen), schakelklassen, taal- en rekenpilots en (meest recent) projecten met verlengde leertijd. Op de inhoud van deze maatregelen gaan we hier niet in, die zijn en

worden op andere plaatsen uitvoerig beschreven (bv. Meijnen e.a., 2003; Ledoux e.a., 2005; Mulder e.a., 2008, 2009; Ledoux & Veen, 2009; www.bredeschool nl; www.schakel-klassen.nl; www.schoolaanzet nl). In de loop van de tijd zijn er ook qua organisatie en aansturing van het onderwijsachterstandenbeleid verschillende veranderingen geweest[1]. Deze zijn voor het onderwerp van deze studie niet direct relevant en worden daarom hier niet besproken.

De onderzoeksvragen

Zoals reeds aangegeven gaat het in deze studie in de eerste plaats om een actualisering van gegevens over de prestaties en loopbanen van leerlingen uit de doelgroepen van het onderwijsachterstandenbeleid. De centrale vraag hierbij is hoe de prestaties en loopbanen van de doelgroepleerlingen zich verhouden tot die van de overige leerlingen. Daarnaast is nog een andere vraag aan de orde, namelijk of en in hoeverre er bij leerlingen uit de doelgroepen sprake is van (verborgen) cognitief talent dat onvoldoende benut wordt. De eerste vraag komt aan bod in de hoofdstukken 2 en 3 van dit rapport, de tweede in de hoofdstukken 4 en 5. In deze vier hoofdstukken staan de volgende meer specifieke onderzoeksvragen centraal:

Hoofdstukken 2 en 3

1. Wat zijn de actuele taal- en rekenprestaties van de verschillende groepen doelgroepleerlingen en van de overige leerlingen in het basisonderwijs?
2. Welke positie nemen deze groepen leerlingen momenteel in in het voortgezet onderwijs?
3. In hoeverre is er bij de verschillende categorieën doelgroepleerlingen sprake van inhalen van onderwijsachterstanden in basis- en voortgezet onderwijs, wanneer cohorten worden vergeleken in de tijd?
4. Wat is de cognitieve ontwikkeling van de verschillende categorieën doelgroepleerlingen in het basisonderwijs? Wat valt vanuit dit (longitudinale) perspectief te zeggen over het inhalen van onderwijsachterstanden?

[1] Bijvoorbeeld de overgang van OVB naar GOA, het daarop volgende Onderwijskansenbeleid, en de latere verschuiving van de regierol van gemeenten naar schoolbesturen (zie Meijnen e.a., 2003; Mulder e.a., 2005; Ledoux & Veen, 2009).

Hoofdstukken 4 en 5

5. Krijgen doelgroepleerlingen en overige leerlingen verschillende adviezen voor het voortgezet onderwijs?
6. In hoeverre worden deze verklaard door cognitieve en niet-cognitieve competenties?
7. Krijgen sommige groepen leerlingen hogere of lagere adviezen dan verwacht?
8. Is er sprake van discrepantie tussen de cognitieve capaciteiten van leerlingen en hun schoolprestaties? Doet deze discrepantie zich meer/minder voor bij de verschillende groepen doelgroepleerlingen?

Om deze vragen te kunnen beantwoorden, is het nodig om eerst helder te krijgen welke leerlingen tot de doelgroepen van het beleid behoren. In principe is daarvoor een duidelijk criterium beschikbaar, namelijk het gewicht dat leerlingen krijgen volgens de zogenoemde gewichtenregeling. Deze gewichtenregeling heeft echter in de loop van de onderzochte periode een aantal veranderingen ondergaan. Dit heeft consequenties voor de wijze van indelen. In paragraaf 1.4 geven we aan voor welke indeling is gekozen. In de volgende paragraaf gaan we eerst nader in op de gewichtenregeling en de achtergronden van de veranderingen.

1.2 Veranderingen in de gewichtenregeling

Zoals gezegd is in het basisonderwijs de gewichtenregeling het meest centrale beleidsinstrument in de bestrijding van onderwijsachterstanden. De regeling zorgt ervoor dat scholen met kinderen uit de doelgroepen van het OAB over meer formatie (personeel) kunnen beschikken, omdat de doelgroepleerlingen zwaarder wegen in de berekening van de formatie. De beleidsverwachting daarbij is dat meer formatie het mogelijk maakt om meer gerichte hulp en aandacht te geven aan de doelgroepkinderen en de school dus meer mogelijkheden geeft om te compenseren voor ongunstige omgevingsfactoren[2]. De oorspronkelijke gewichtenregeling (ingevoerd in 1986) ging uit van vier categorieën leerlingen: leerlingen die niet tot een achterstandsgroep behoren

[2] Niet elke leerling met een leerlinggewicht telt mee voor extra formatie. De regeling kent een drempel; de school krijgt pas extra middelen als het aantal gewichtenleerlingen deze drempel overstijgt. De (ingewikkelde) rekenregels van de drempel laten we hier verder onbesproken.

(gewogen als 1.0-leerling, dus geen extra gewicht), autochtone achterstandsleerlingen (gewogen als 1.25, een kwart meer formatie per leerling), allochtone achterstandsleerlingen (gewogen als 1.9, bijna twee keer zo veel formatie per leerling), schipperskinderen (gewogen als 1.4) en woonwagen- en zigeunerkinderen (gewogen als 1.7).
Voor de 1.25- en de 1.9-leerlingen gold daarbij als criterium dat ten minste één van de ouders laag opgeleid moest zijn (niet hoger dan lager beroepsonderwijs), en/of dat de meeste verdienende ouder werkloos moest zijn dan wel een (ongeschoold) handarbeidersberoep moest uitoefenen. In 1993 is voor de 1.25-leerlingen het criterium aangescherpt. Voortaan moesten beide ouders laag opgeleid zijn en werd niet meer naar de beroepssituatie gekeken. Voor de 1.9-leerlingen bleef de oude definitie gelden. De reden voor de aanpassing voor de 1.25-leerlingen was dat na enige tijd bleek dat bijna de helft van alle leerlingen in het basisonderwijs onder de oorspronkelijke definitie viel, voor het overgrote deel autochtone leerlingen (Meijnen, 2003). Een vrij aanzienlijk deel van deze leerlingen presteerde op of boven het landelijk gemiddelde, zo werd duidelijk uit landelijke evaluaties. Op basis daarvan werd geconcludeerd dat de regeling voor deze groep te ruim was en is het criterium versimpeld (beroepssituatie geschrapt) en aangescherpt (beide ouders laag opgeleid). Het aandeel autochtone doelgroepleerlingen daalde daardoor tot circa 20% van alle leerlingen Ten tijde van de wijziging in 1993 is in discussie geweest of de aanscherping van het criterium ook voor de allochtone doelgroepleerlingen zou moeten gelden. Daar is toen echter niet voor gekozen.

Vanaf 2000 zijn door de Onderwijsraad (2001a,b) verkenningen verricht van nieuwe herzieningen van de gewichtenregeling, dit op verzoek van de toenmalige staatssecretaris van OCW. De Onderwijsraad werd gevraagd na te gaan of het onderscheid in criteria voor de 1.25- en 1.9-leerlingen nog gehandhaafd diende te blijven en of het mogelijk was om een nieuw systeem in te voeren dat gebaseerd zou zijn op de feitelijke (gemeten) onderwijsachterstand van een leerling, in plaats van op een groepskenmerk zoals opleiding ouders of etnische herkomst.
De Onderwijsraad adviseerde, mede op basis van studies waarin leerresultaten van verschillende groepen leerlingen werden geanalyseerd, om:

- niet over te gaan op een systeem gebaseerd op gemeten onderwijsachterstand, met als voornaamste argument dat dat zou

veronachtzamen dat er ook doelgroepleerlingen zijn die niet tot de categorie laagpresteerders behoren, maar bij wie toch sprake is van achterstand, omdat ook zij beïnvloed worden door ongunstige omgevingsfactoren (waardoor ze hun potentieel niet geheel realiseren);

- voortaan als basis voor de gewichtenregeling uit te gaan van het opleidingsniveau van de ouders, met een differentiatie naar twee niveaus: ouderlijk opleidingsniveau maximaal basisonderwijs en ouderlijk opleidingsniveau maximaal lbo/vbo;
- daarbovenop voor allochtone leerlingen niet meer te kijken naar het (groeps)criterium herkomstland (geboorteland ouders), maar naar (individuele) taalachterstand, te meten met taaltoetsen in de leerjaren 2, 4 en 8 van het basisonderwijs, en voor leerlingen met grote taalachterstanden een extra zwaar gewicht te gaan hanteren.

Het argument voor dit laatste was dat uit de verrichte studies bleek dat allochtone leerlingen nog steeds lagere prestaties behalen dan autochtone leerlingen, ook als rekening is gehouden met het opleidingsniveau van de ouders. Dit gold vooral voor leerlingen van Turkse en Marokkaanse afkomst. De verschillen waren substantieel. Allochtone leerlingen met ouders met een mbo- of hbo-opleiding bleken in het basisonderwijs lager te presteren op taal dan autochtone leerlingen met ouders met een lbo/vbo-opleiding. De Onderwijsraad concludeerde derhalve dat de factor etniciteit een aparte voorspellende waarde heeft voor later schoolsucces, bovenop de factor opleidingsniveau van de ouders. Maar omdat de Raad dit interpreteerde als een taalprobleem, kwam ze tot haar voorstel om een extra gewicht te gaan hanteren op gemeten taalachterstand en niet langer op herkomstland (Onderwijsraad, 2001a).

De Minister nam de voorstellen van de Onderwijsraad ten dele over. In 2006 werd een nieuwe regeling van kracht, gebaseerd de volgende besluiten:

- het opleidingsniveau van beide ouders is voortaan het hoofdcriterium en wordt verfijnd door twee niveaus te onderscheiden; hieraan worden twee gewichten gekoppeld, namelijk 0.3 voor kinderen van ouders met maximaal lbo/vbo-niveau en 1.2 voor kinderen van ouders met maximaal basisonderwijs[3];

[3] Met de wijziging werd ook de manier van aanduiden anders: 1.0-leerlingen waren voortaan 0.0-leerlingen, het nieuwe gewicht 1.3 werd 0.3, etc.

- leerlingen met extra grote taalachterstanden krijgen een extra zwaar gewicht van 2.4 (alleen toe te kennen aan kinderen in de groepen 1 tot en met 4);
- de aparte gewichten voor schipperskinderen en kinderen van woonwagenbewoners/zigeuners vervallen;
- etniciteit zal voortaan geen criterium meer zijn.

Het extra zware gewicht van 2.4 is er echter in de praktijk niet gekomen. Het baseren van gewichten op toetsresultaten bleek zowel methodologisch als praktisch niet goed uitvoerbaar. De wijziging van 2006 heeft dus per saldo geleid tot twee nieuwe doelgroepen: 0.3- en 1.2-leerlingen.

De consequentie van deze wijzigingen voor het onderhavige onderzoek is dat er verschillende categorieën doelgroepleerlingen moeten worden onderscheiden. In de eerste plaats de ex-doelgroepleerlingen (ex 1.25 en ex 1.9) en de huidige doelgroepleerlingen (0.3 en 1.2). Daarnaast vonden we het van belang om zowel binnen de categorie doelgroepleerlingen als binnen de categorie niet-doelgroepleerlingen onderscheid te blijven maken tussen allochtone en autochtone leerlingen. Niet alleen om nog vergelijkingen te kunnen maken met de groep ex 1.9-leerlingen, maar ook omdat de analyses van de Onderwijsraad lieten zien dat er rond 2000 nog steeds sprake was van ongelijkheid in prestaties naar etnische herkomst, ook na controle voor opleidingsniveau ouders. Voor het beleid is het relevant te weten of dit anno 2007 ook nog het geval is.
Zoals eerder vermeld presenteren we de gehanteerde indeling in paragraaf 1.4. Eerst gaan we echter in op de gebruikte data.

1.3 Gegevens uit cohortonderzoeken

Voor de beantwoording van de probleemstellingen voor dit onderzoek maken we gebruik van landelijke cohortgegevens, die gedurende een reeks van jaren verzameld zijn, mede met het oog op het monitoren van de positie van achterstandsleerlingen en het vaststellen van veranderingen daarin.
In de eerste plaats gaat het dan om gegevens uit het PRIMA-cohortonderzoek. Dit onderzoek was een vervolg op het eerdere onderzoek voor de Landelijke Evaluatie van het Onderwijsvoorrangsbeleid (LEO). Het was daarnaast ook een uitbreiding, omdat in PRIMA ook gegevens van jongere leerlingen (uit groep 2

van het basisonderwijs) werden verzameld[4]. De eerste meting van PRIMA vond plaats in schooljaar 1994/95. Daarna volgde om de twee jaar een nieuwe meting, tot aan een zesde meting in schooljaar 2004/05. Bij elke meting zijn gegevens verzameld van leerlingen uit de groepen 2, 4, 6 en 8 van het basisonderwijs. Met behulp van verschillende instrumenten is zowel informatie over een aantal cognitieve als over niet-cognitieve kenmerken van de leerlingen verkregen. De cognitieve kenmerken betreffen scores op verschillende landelijke genormeerde toetsen voor taal en rekenen. Bij de niet-cognitieve kenmerken gaat het om zelfoordelen van de oudere leerlingen (uit de groepen 6 en 8) met betrekking tot motivatie, schoolwelbevinden en dergelijke. Daarnaast zijn bij de leerkrachten van alle onderzochte groepen oordelen over hun leerlingen verzameld met betrekking tot gedrag, werkhouding en dergelijke.

Bij elke meting van PRIMA deden 50 tot 60 duizend leerlingen van ongeveer 600 basisscholen mee met het onderzoek. Een deel van die scholen vormt samen een representatieve steekproef uit alle Nederlandse basisscholen. De overige scholen vormden een aanvullende steekproef, waarin scholen met veel achterstandsleerlingen oververtegenwoordigd zijn.

Terwijl PRIMA om de twee jaar een meting kende, werd om de vier jaar een deel van de leerlingen uit groep 8 verder gevolgd in het voortgezet onderwijs. Dit aanvullende onderzoek staat bekend als PRIMA-VO en leverde informatie op over de VO-loopbaan van leerlingen tot en met het vierde jaar van het voortgezet onderwijs. Naast loopbaangegevens zijn ook rapportcijfers voor vakken als nederlands en wiskunde verzameld.

De laatste meting van PRIMA was in het schooljaar 2004/05. Vanaf schooljaar 2007/08 is het onderzoek, in gewijzigde vorm, voortgezet in het cohortonderzoek COOL$^{5\text{-}18}$. Een eerste belangrijkste wijzing is dat het COOL-onderzoek ook plaats vindt in het voortgezet onderwijs (en als zodanig ook opvolger is van de eerdere VOCL-cohorten in het voortgezet onderwijs). Een tweede wijziging in het basisonderwijsdeel betreft de frequentie van het onderzoek en daarmee ook de betrokken groepen: COOL$^{5\text{-}18}$ wordt niet om de twee maar om de drie jaar gehouden, en niet in de groepen 2, 4, 6 en 8, maar in de groepen 2, 5 en 8.

Aan het gedeelte van COOL in het basisonderwijs hebben ongeveer 38 duizend leerlingen van 550 basisscholen meegedaan. Ook hier is er weer sprake van een

[4] Bovendien werden in PRIMA ook gegevens verzameld bij scholen voor speciaal basisonderwijs. Deze gegevens worden in dit rapport verder niet gebruikt.

representatieve steekproef (van 400 scholen) en een aanvullende steekproef met oververtegenwoordiging van scholen waarop veel achterstandsleerlingen zitten. De gegevens die bij COOL in het basisonderwijs worden verzameld hebben, net als bij PRIMA, betrekking op cognitieve en niet-cognitieve kenmerken van de leerlingen. Er is wel sprake van een aantal wijzigingen in de gebruikte instrumenten. Zo wordt nog meer dan bij PRIMA gebruik gemaakt van toetsen uit het Cito-leerlingvolgsysteem, die veel van de scholen toch al afnemen. In groep 5 wordt een uitgebreidere Niet-Schoolse Cognitieve Capaciteiten Test (NSCCT) afgenomen, terwijl de leerlingen van groep 8 ook een lijst voor Burgerschapscompetenties invullen.
De verzamelde gegevens in de genoemde cohortonderzoeken maken het mogelijk om zowel cross-sectionele als longitudinale analyses uit te voeren. Beide mogelijkheden kunnen geïllustreerd worden aan de hand van het volgende schema, waarin de verschillende momenten staan afgebeeld, waarop informatie is verzameld bij een bepaalde groep leerlingen. De jaren, waarin gegevens zijn verzameld, staan in de rijen; de groepen waarbij die gegevens zijn verzameld staan in de kolommen.

Schema 1.1 Momenten van dataverzameling in PRIMA (P1–P6), PRIMA–VO en COOL (C–1)

	start schooljaar	groep 2	(3)	groep 4	groep 5	groep 6	(7)	groep 8	vo1	vo2	vo3
P1	1994	A		x		x		x			
	--		↘								
P2	1996	B		A		x		x			
	--		↘		↘				x		
P3	1998	C		B		A		x		x	
	--		↘		↘		↘				x
P4	2000	x		C		B		A	x		
	--				↘		↘			x	
P5	2002	x		x		C		B			x
	--						↘				
P6	2004	x		x		x		C			
	--										
	--										
C1	2007	x			x			x			x

In alle cellen, waarin een letter staat (A, B, C, x) zijn er, in het betreffende jaar en bij de betreffende groep, gegevens verzameld.
In een *cross–sectionele* analyse wordt voor een bepaalde kolom (dus een bepaalde groep) nagegaan of de situatie is veranderd over de loop der jaren. Bij groep 2 gaat het dan om veranderingen bij de instroom in het basisonderwijs; bij groep 8 om veranderingen aan het eind van het basisonderwijs. In Hoofdstuk 2 zullen dergelijke cross-sectionele analyses worden uitgevoerd voor groep 2, groep 8 en het derde jaar voortgezet onderwijs.
Bij een *longitudinale* analyse wordt niet de situatie van leerlingen in groep 2 van 1994 vergeleken met die van (andere) leerlingen in groep 2 van een ander jaar, maar worden *dezelfde* leerlingen door de tijd heen gevolgd van groep 2, via groep 4 en 6, tot in groep 8. Zo'n groep leerlingen wordt een *cohort* genoemd; in Schema 1.1 zijn de drie cohorten aangeduid met de letters A, B en

C, waarbij de pijltjes aangeven dat dezelfde groep gevolgd wordt. In Hoofdstuk 3 zal de ontwikkeling van schoolprestaties van leerlingen uit deze drie cohorten geanalyseerd worden.
In de verschillende onderzoeken zijn een groot aantal gegevens met betrekking tot de leerlingen verzameld. Bij de analyses in de volgende hoofdstukken zal steeds nader ingegaan worden op de variabelen die in die analyses gebruikt zijn. Bij alle analyses wordt wel steeds dezelfde indeling van leerlingen gebruikt. In de volgende paragraaf wordt deze indeling nader beschreven en beargumenteerd. We eindigen deze paragraaf met een laatste opmerking over het gebruik van de steekproeven.
Zowel bij PRIMA als bij COOL is er sprake van een steekproef die representatief is voor het gehele basisonderwijs in Nederland en van een aanvullende steekproef, waarin extra scholen zitten met veel achterstandsleerlingen. Wanneer er gemiddelde gegevens voor het hele basisonderwijs in Nederland nodig zijn, dient daarvoor de representatieve steekproef van de betreffende cohort gebruikt te worden. Wanneer het echter gaat om vergelijking van groepen (achterstands)leerlingen onderling, dan kan de grotere totale steekproef gebruikt worden, waarin zowel de representatieve als de aanvullende steekproef samen zijn genomen. Voor het vergelijken van gemiddelden is representativiteit in principe vereist, maar voor het bepalen van relaties tussen variabelen geldt dat niet (Hox, 1998). Belangrijk voordeel van het gebruik van de totale steekproef is natuurlijk dat daar grotere aantallen achterstandsleerlingen in zijn opgenomen, zodat er ook bij uitsplitsingen vrijwel steeds sprake is van een adequate celvulling. In de navolgende onderzoekingen worden daarom vrijwel steeds de totale steekproeven uit de verschillende cohorten gebruikt. Wanneer daarvan wordt afgeweken, wordt dit bij de betreffende analyse aangegeven.

1.4 Indeling van leerlingen

Zoals hiervoor al is opgemerkt, zijn er in de loop der tijd veranderingen geweest in de gewichtenregeling. In 1998 werd het criterium om in aanmerking te komen voor het 1.25-gewicht aangescherpt: niet ten minste één maar *beide* ouders mogen maximaal een vbo/lbo-niveau hebben. In eerder onderzoek naar onderwijsachterstanden (Mulder e.a., 2005) werd, vanwege zo'n verandering van die regeling in de negentiger jaren, een onderscheid gemaakt tussen 'echte' 1.25-leerlingen en ex 1.25-leerlingen.

Inmiddels is er in 2006 een nieuwe wijziging van de gewichtenregeling (fasegewijs) in gang gezet en moeten we de indeling van doelgroepleerlingen weer verder aanpassen. In de nieuwste gewichtenregeling wordt alleen nog naar het opleidingsniveau gekeken, en wel van beide ouders. Er wordt pas een gewicht toegekend als *beide* ouders ten hoogste een lbo-niveau hebben (zie kader).

Oude gewichten:
1.25 Leerling van wie beide ouders of verzorgers een schoolopleiding hebben genoten tot of tot en met niveau eindexamen voorbereidend beroepsonderwijs.
1.9 Leerling met een niet-Nederlandse culturele achtergrond die tevens voldoet aan één van de volgende voorwaarden:

1. de vader/verzorger heeft een schoolopleiding genoten tot het niveau eindexamen voorbereidend beroepsonderwijs, of
2. de moeder/verzorgster heeft een schoolopleiding genoten tot het niveau eindexamen voorbereidend beroepsonderwijs, of
3. de meest verdienende ouder/verzorger oefent een beroep uit in loondienst, waarin hij lichamelijke of handenarbeid verricht, of geniet geen inkomsten uit tegenwoordige arbeid.

(Incidenteel:
1.4 Leerling die verblijft in een internaat of pleeggezin en van wie de vader of moeder het schippersbedrijf uitoefent of heeft uitgeoefend.
1.7 Leerling van wie de ouders een trekkend bestaan leiden)

Nieuwe gewichten (vanaf augustus 2006 gefaseerd ingevoerd)
De gewichten zijn gekoppeld aan de opleidingscategorieën 1 en 2.

1. maximaal basisonderwijs of (v)so/zmlk
2. maximaal lbo/vbo, praktijkonderwijs of vmbo-basis of kaderberoepsgerichte leerweg
3. overig voortgezet onderwijs en hoger

0.3 Leerling van wie beide ouders of de ouder die belast is met de dagelijkse verzorging een opleiding uit categorie 2 heeft gehad.
1.2 Leerling van wie één van de ouders een opleiding heeft gehad uit categorie 1 en de ander uit categorie 1 óf 2.

Voor de autochtone doelgroepleerlingen (althans, de zogenaamde echte 1.25-leerlingen) is dat geen wijziging, voor de allochtone (1.9-) leerlingen wel. Voor het gewicht 1.9 was voldoende dat ten minste één van beide ouders een opleiding op hooguit het niveau van lager beroepsonderwijs (lbo) had. De andere ouder kon dus een hogere opleiding dan lbo hebben gevolgd. (En het gewicht 1.9 kon zelfs toegekend worden als beide ouders zo'n opleiding hoger dan lbo hadden: als de meest verdienende ouder een laag beroepsniveau (handenarbeid of geen inkomsten uit tegenwoordige arbeid) had.)
Een andere wijziging in de gewichtenregeling is dat wanneer beide ouders hooguit lbo-niveau hebben en ten minste één van beide alleen lager onderwijs (lo) heeft, er een extra zwaar gewicht (1.2) wordt toegekend. Binnen de groep van allochtone en autochtone doelgroepleerlingen wordt als het ware een groep 'zware gevallen' benoemd.
Naast deze veranderingen in de gewichtenregeling zullen we in dit onderzoek ook binnen de 'ongewogen' leerlingen nog nader onderscheid maken. Allereerst gaat het dan om het verschil allochtoon- autochtoon, maar ook om het onderscheid tussen ouders die middelbaar zijn opgeleid en ouders die hoger onderwijs hebben gevolgd.
In de verschillende cohorten zijn gegevens voorhanden over de opleiding en het geboorteland van beide ouders. Op basis daarvan kunnen de leerlingen nader ingedeeld worden in allochtoon versus autochtoon enerzijds en naar het opleidingsniveau van hun beide ouders anderzijds. Bij de indeling autochtoon-allochtoon hanteren we de CBS-definitie (zie kader); de westerse allochtonen worden bij de autochtone leerlingen gevoegd.
Het volgende schema geeft de uiteindelijke indeling van groepen leerlingen, die in de rest van dit onderzoek zal worden gehanteerd. Voor de korte aanduiding van de verschillende groepen houden we daarbij nog even vast aan de 'oude' gewichtenbenamingen 1.9, 1.25 en 1.0. Rechts in het schema staan de 'nieuwere' benamingen (vroeger 0.9 en 0.25, tegenwoordig 0.3 en 1.2).

Schema 1.2 Indeling van leerlingen naar opleiding ouders en etniciteit

		Leerlinggewicht		
Aanduiding	**Criteria**	**oud**	**na 1998**	**na 2006**
autochtoon				
1.25 zwaar	beide ouders max. lbo; ten minste 1 max. lo	0.25	0.25	1.2
1.25	beide ouders max. lbo	0.25	0.25	0.3
1.25 ex	één ouder max. lbo, de ander meer	0.25	0	0
1.0 mbo	beide ouders meer dan lbo; maar geen ho	-	-	-
1.0 ho	beide ouders meer dan lbo; minstens één ho	-	-	-
allochtoon				
1.9 zwaar	beide ouders max. lbo; ten minste 1 max. lo	0.9	0.9	1.2
1.9	beide ouders max. lbo	0.9	0.9	0.3
1.9 ex	één ouder max. lbo, de ander meer	0.9	0.9	0
1.0 mbo	beide ouders meer dan lbo; maar geen ho	-	-	-
1.0 ho	beide ouders meer dan lbo; minstens één ho	-	-	-

Samen vormt dit dus een uitgebreide tiendeling voor de in dit rapport te onderscheiden groepen leerlingen. In enkele analyses zal binnen de groep allochtone achterstandsleerlingen (de drie groepen met een 1.9-aanduiding) nog nader onderscheid gemaakt worden naar de volgende herkomstlanden:

- Turks;
- Marokkaans;
- Surinaams en Antilliaans;
- overig niet-westers.

CBS-definitie niet-westers allochtonen

Met ingang van augustus 1999 is CBS-breed de standaarddefinitie ingevoerd. Volgens deze nieuwe definitie wordt een persoon tot de allochtonen gerekend als ten minste één ouder in het buitenland is geboren. (…….)
Naast de nieuwe definitie is ook een standaardindeling van allochtonen geïntroduceerd waarbij onderscheid wordt gemaakt tussen westerse en niet-westerse allochtonen. Tot de westerse herkomstlanden worden gerekend alle landen in Europa (maar zonder Turkije), Noord-Amerika, Oceanië, Japan en Indonesië (met inbegrip van het voormalig Nederlands-Indië). De niet-westerse herkomstlanden zijn Turkije en alle landen in Afrika, Latijns-Amerika en Azië (met uitzondering van Japan en Indonesië).
De reden om dit onderscheid aan te brengen, is de verschillende sociaal-economische en culturele positie van westerse en niet-westerse allochtonen. Als een groep in sociaaleconomisch of cultureel opzicht sterk op de Nederlandse bevolking lijkt, wordt deze groep tot de westerse allochtonen gerekend.
De allochtonen met Indonesië als herkomstland vormen een bijzondere groep. Ze worden gerekend tot de westerse allochtonen. Veel van de personen uit Indonesië of voormalig Nederlands-Indië zijn van Nederlandse afkomst.
(Bron: CBS, Index No. 10 - november/december 2000).

Tot slot geven we nog een overzicht van hoe achterstandsleerlingen volgens de oorspronkelijke regeling uit 1986 verdeeld zijn over onze nieuwe indeling in ex, licht en zwaar gewicht. In de volgende tabel wordt dat weergegeven voor het eerste jaar van de periode die in dit onderzoek behandeld wordt, halverwege die periode en voor het laatste jaar.

Tabel 1.1 Verdeling van oorspronkelijke gewichtenleerlingen over de nieuwe indeling

	1994	2000	2007
'oude' 1.25 leerlingen			
1.25 – ex	36.2%	47.4%	57.2%
1.25 licht (0.3)	52.7%	45.2%	36.7%
1.25 zwaar (1.2)	11.1%	7.5%	6.1%
'oude' 1.9 leerlingen			
1.9 – ex	17.5%	23.2%	27.8%
1.9 licht (0.3)	27.3%	21.7%	17.9%
1.9 zwaar (1.2)	55.2%	55.1%	54.4%

Uit deze tabel blijkt dat de oorspronkelijke autochtone achterstandsleerlingen in toenemende mate helemaal geen gewicht meer krijgen. En in 2007 krijgt nog slechts 6.1% het gewicht 1.2. Bij de allochtone leerlingen is het aandeel dat helemaal geen gewicht krijgt beduidend lager. In alle jaren krijgt rond de 55% het gewicht 1.2.

Literatuur

Van de verschillende metingen in de cohortonderzoeken zijn steeds technische rapporten verschenen. We noemen hier de technische rapporten van:

PRIMA–1

Langen, A. van, Vierke, H. & Robijns, M. (1996) *Veldwerkverslag basisonderwijs en speciaal onderwijs*. Nijmegen/Amsterdam: ITS/SCO-Kohnstamm Instituut

Jungbluth, P., Langen, A. van, Peetsma, Th. & Vierke, H. (1996) *Leerlinggegevens basisonderwijs en speciaal basisonderwijs*. Nijmegen/Amsterdam: ITS/SCO-Kohnstamm Instituut

PRIMA–6

Driessen, G., Langen, A. van & Vierke, H. (2006) *Basisonderwijs: veldwerkverslag, leerlinggegevens en oudervragenlijsten*. Nijmegen: ITS

COOL$^{5-18}$ basisonderwijs

Driessen, G., Mulder, L., Ledoux, G., Roeleveld, J., & Veen, I. van der (2009). *Cohortonderzoek COOL$^{5-18}$. Technisch rapport basisonderwijs, eerste meting 2007/08*. Nijmegen: ITS/ Amsterdam: SCO-Kohnstamm Instituut.

COOL$^{5-18}$ voortgezet onderwijs

Zijsling, D., Keuning, J., Kuyper, H., Batenburg, T. van, & Hemker, B. (i.v.). *Technisch rapport. Eerste meting van COOL$^{5-18}$ in het derde leerjaar van het voortgezet onderwijs.* Groningen/Arnhem: RUG/Cito.

2 De ontwikkeling van de onderwijspositie van OAB-doelgroepleerlingen tussen 1994/95 en 2007/08

Geert Driessen, m m.v. Hermann Vierke en Wouter de Wit

2.1 Achtergronden

2.1.1 Het achterstandenbeleid

De overheid voert sinds de zeventiger jaren van de vorige eeuw gericht beleid om de onderwijskansen van autochtone en allochtone kinderen uit achterstandssituaties te verbeteren (Driessen & Dekkers, 2008). Gaandeweg heeft het beleid verschillende wijzigingen ondergaan, met name qua organisatie en financiering; ook zijn er nieuwe beleidsterreinen, thema's, doelen en doelgroepen gedefinieerd (Ledoux & Veen, 2009). De kern van het beleid – tenminste in financiële zin – vormt de toekenning van extra faciliteiten aan scholen met een bepaald aandeel aan doelgroepleerlingen. Dit beleidsinstrument staat te boek als 'de gewichtenregeling'. Hoe ongunstiger de thuissituatie, hoe hoger het gewicht en hoe meer geld de school krijgt. De indicatoren die worden gehanteerd voor het identificeren van doelgroepleerlingen, en daarmee voor het leerlinggewicht, zijn enkele malen gewijzigd (Ledoux & Veen, 2009).

2.1.2 Monitoring van achterstanden

Uniek in de wereld is dat het Nederlandse achterstandenbeleid van meet af aan vergezeld is gegaan van grootschalige evaluaties. Bij de start van het Onderwijsvoorrangsbeleid in 1985 is meteen een grootschalige cohortstudie opgezet, de Landelijke Evaluatie van het Onderwijsvoorrangsbeleid (LEO). In

1994 is deze na drie tweejaarlijkse metingen voortgezet via de cohortstudie Primair Onderwijs (PRIMA), die zes eveneens tweejaarlijkse metingen heeft gekend. In 2007 is de cohortstudie COOL[5-18] van start gegaan. Hadden de eerder genoemde cohortstudies alleen betrekking op het primair onderwijs, voor COOL is er feitelijk sprake van een integratie van de PRIMA-cohortstudie en de VOCL-cohortstudie, die tot dan toe in het voortgezet onderwijs werd uitgevoerd.
De vele gegevens die zijn verzameld via de verschillende cohortstudies leverden niet alleen in belangrijke mate de input voor de evaluatie van het gevoerde beleid, maar zijn tevens zeer frequent gebruikt voor de monitoring van de onderwijspositie van de doelgroepen. In het eerste geval gaat het om het nagaan of het beleid effectief is geweest, respectievelijk in welke mate de gestelde doelen ook daadwerkelijk zijn bereikt. In het tweede geval is er geen rechtstreekse koppeling naar het beleid, maar ligt het accent op het in kaart brengen van de ontwikkelingen van bepaalde groepen van leerlingen met betrekking tot hun onderwijssucces. In het onderhavige onderzoek staat deze laatste functie centraal.[5]
In een uitgebreide serie van studies is in voorgaande jaren aandacht besteed aan de ontwikkelingen in schoolloopbanen van achterstandsleerlingen. Dat is op de eerste plaats gedaan door de onderzoeksinstituten die ook verantwoordelijk waren, respectievelijk zijn voor de uitvoering van de verschillende cohortstudies, i.c. het ITS, het (SCO-)Kohnstamm Instituut en het GION. Daarnaast hebben ook diverse andere instituten en onderzoekers deze data benut voor het beschrijven van de ontwikkelingen. In dit verband kan met name het SCP genoemd worden, dat voor de Rapportage minderheden en het Jaarrapport integratie steeds cohortgegevens heeft geanalyseerd.

2.1.3 Eerder onderzoek

Om een indruk te geven van wat deze studies tot nu toe aan inzichten hebben opgeleverd, zetten we hierna de belangrijkste bevindingen op een rijtje. We maken daarvoor vooral gebruik van de samenvatting van Ledoux en Veen (2009), die zich baseren op Mulder, Roeleveld, Van der Veen en Vierke (2005), Tesser en Iedema (2001), Dagevos, Gijsberts en Van Praag (2003), Vogels en Bronnemans-Helmers (2003), Dagevos en Gijsberts (2007), Roeleveld, Van der Veen en Ledoux (2009) en Driessen (2009). Daarnaast kunnen nog de wat

[5] Dit sluit overigens niet uit dat er van daaruit ook evaluatieve opmerkingen richting beleid gemaakt kunnen worden. Beide functies vallen deels samen.

oudere studies van Mulder (1996) en Van Langen en Suhre (2001) en de meer recente studie van Gijsberts en Dagevos (2009) worden genoemd. Al deze studies zijn gebaseerd op analyses van de databestanden van de LEO-, PRIMA- en COOL-cohortonderzoeken, allemaal grootschalige, landelijke onderzoeken naar loopbanen en prestaties van leerlingen in het basisonderwijs die bij elkaar de periode van 1988 tot en met 2008 bestrijken. Bij dit overzicht dient overigens te worden aangetekend dat, zoals we eerder al aangaven, de definiëring van de doelgroepen in de loop van de jaren enkele malen gewijzigd is, zodat bij een vergelijking over de jaren heen soms enige terughoudendheid geboden is.

2.1.3.1 Aanvangsverschillen

Bij aanvang van het basisonderwijs zijn de prestatieverschillen tussen de doelgroepleerlingen en de niet-doelgroepleerlingen groot. Vooral Turkse en Marokkaanse leerlingen starten hun schoolloopbaan met een grote achterstand, zowel op het gebied van taal als - zij het in minder mate - rekenen. In de meest recente cohorten is er sprake van een afname van de aanvangsverschillen. Turkse en Marokkaanse kinderen en autochtone kinderen van laagopgeleide ouders groeien iets toe naar het aanvangsniveau van de niet-doelgroepleerlingen, en dat geldt voor zowel taal als voor rekenen. De kloof ten opzichte van de niet-doelgroepleerlingen is echter nog lang niet overbrugd.

2.1.3.2 Ontwikkeling allochtone doelgroepleerlingen

Was er bij de eerste cohortonderzoeken sprake van geen of hooguit geringe prestatie-ontwikkelingen, de laatste jaren blijkt dat de allochtone doelgroepleerlingen in de loop van het basisonderwijs de aanvangsachterstanden op de niet-doelgroepleerlingen enigszins inlopen. In groep 8 zijn de niveauverschillen voor taal kleiner geworden (alhoewel nog altijd duidelijk aanwezig) en voor rekenen grotendeels verdwenen (vooral Marokkaanse leerlingen hebben volgens de laatste PRIMA- en COOL-gegevens hun achterstand op niet-doelgroepleerlingen bij rekenen vrijwel ingehaald). Deze winst wordt met name behaald in de hogere groepen van het basisonderwijs.

2.1.3.3 Ontwikkeling autochtone doelgroepleerlingen
Bij de autochtone doelgroepleerlingen doet zich een dergelijke positieve ontwikkeling niet voor. Sinds de striktere definiëring van deze doelgroep is de prestatie-achterstand van de betreffende leerlingen ten opzichte van de niet-doelgroepleerlingen vergroot in plaats van verkleind; de achterstand is bij rekenen groter dan bij taal. Vergeleken met de allochtone doelgroepleerlingen presteren de autochtone doelgroepleerlingen momenteel nog wel beter bij taal, maar niet meer bij rekenen.

2.1.3.4 Doorstroom naar hogere vormen van voortgezet onderwijs
In 2007 zat circa 20 procent van de autochtone en allochtone doelgroepleerlingen in een havo/vwo- of vwo-klas; in 2002 was dit nog 15 procent. Er is dus sprake van een stijging. Bij de allochtone doelgroepleerlingen komt die stijging vooral voor rekening van de Surinamers en de Marokkanen. De niet-doelgroepleerlingen gaan echter nog ruim twee keer zo vaak naar deze hogere schooltypen dan de doelgroepleerlingen. Van de niet-doelgroepleerlingen ging in 2002 al bijna de helft naar het havo/vwo of vwo. Dit percentage is in de loop der jaren verder gestegen en de stijging in deze groep ging sneller dan die bij de doelgroepleerlingen. Op het doorstroomcriterium is dus voor de doelgroepen wel vooruitgang waarneembaar, maar is er nog weinig sprake van een inhalen van de achterstand.
Het bovenstaande samenvattend, is er over de periode 1988/89-2007/08 een positieve ontwikkeling te zien wat betreft zowel de taal- als rekenprestaties van allochtone doelgroepleerlingen, hoewel die in groep 8 voor taal nog altijd de laagste scores behalen (ook lager dan de autochtone achterstandsleerlingen). Voor autochtone doelgroepleerlingen zijn er in het basisonderwijs geen positieve ontwikkelingen in prestaties. Zij halen hun achterstand niet in, althans niet de groep die na de striktere definiëring nog tot deze doelgroep behoort. De doorstroom naar hogere vormen van voortgezet onderwijs verbetert wel voor alle doelgroepleerlingen, zowel de autochtone als allochtone, maar hun achterstand op dit gebied blijft bestaan, omdat ook niet-doelgroepleerlingen vaker naar hogere vormen van voortgezet onderwijs gaan.

2.2 Probleemstelling

Bij het volgen van de ontwikkeling van de onderwijspositie van leerlingen kunnen twee benaderingen worden onderscheiden, een cross-sectionele en een longitudinale. Bij een cross-sectionele benadering wordt de positie van per meetmoment *verschillende* steekproeven van leerlingen over de tijd gevolgd. Zo worden bijvoorbeeld de leerlingen die in 1994 in groep 4 zaten vergeleken met de leerlingen die in 1998 in groep 4 zaten en de leerlingen die in 2002 in groep 4 zaten. Bij een longitudinale benadering wordt de positie van een steekproef van *dezelfde* leerlingen gevolgd; dit is de cohort-benadering. Er wordt bijvoorbeeld nagegaan hoeveel leerwinst de leerlingen die in 1994 in groep 4 zaten twee jaar later in groep 6 en weer twee jaar later in groep 8 relatief hebben behaald. In het onderhavige hoofdstuk wordt uitgegaan van de *cross-sectionele* benadering; in een volgend hoofdstuk wordt de longitudinale benadering uitgewerkt.

Als vervolg op twee eerdere NWO/BOPO-overzichtsstudies (Van Langen & Suhre, 2001; Mulder e.a., 2005), die gingen over de periode 1988/89 – 1997/98, respectievelijk 1988/89 – 2002/03, wordt in het onderhavige onderzoek nagegaan hoe zich de onderwijspositie van OAB-doelgroepleerlingen in de periode 1994/95 – 2007/08 heeft ontwikkeld. De centrale vraag luidt of de achterstand van de verschillende doelgroepen ten opzichte van een referentiecategorie in de loop van de jaren verkleind is. Om deze vraag te beantwoorden schetsen we allereerst op basis van de meest recente grootschalige gegevens de stand van zaken met betrekking tot de taal- en rekenvaardigheid van leerlingen in groep 2 en 8 basisonderwijs, en de onderwijspositie en de taal- en wiskundevaardigheid van leerlingen in leerjaar 3 voortgezet onderwijs. Vervolgens gaan we na of er wat deze kenmerken betreft verschillen in ontwikkelingen zijn tussen de onderscheiden categorieën van leerlingen.

Het onderhavige onderzoek betreft dus een belangrijke update van de eerdere studies. Hiermee wordt namelijk inzicht verkregen in hoe het de verschillende categorieën van leerlingen in de periode sinds de laatste studie (die van Mulder c.s.) is vergaan. Zoals hierboven al opmerkt, is de zogenoemde gewichtenregeling, en daarmee de definiëring van doelgroepleerlingen, de afgelopen jaren wezenlijk veranderd. Door gebruik te maken van de speciaal voor het onderhavige onderzoek ontwikkelde doelgroepenindeling is het toch

mogelijk op een zeer gedetailleerde wijze de ontwikkeling van de verschillende categorieën te blijven volgen.

2.3 Methode

2.3.1 Steekproeven

Voor het in kaart brengen van de ontwikkelingen in prestaties en onderwijspositie maken we gebruik van gegevens die zijn verzameld bij de verschillende metingen van de PRIMA-cohortstudies en de eerste meting van de COOL-cohortstudie. PRIMA is in het schooljaar 1994/95 gestart met een eerste meting in de groepen 2, 4, 6 en 8 van het basisonderwijs, en is vervolgens tweejaarlijks herhaald tot en met 2004/05. Aan PRIMA hebben per meting in totaal steeds circa 60.000 leerlingen van 600 scholen meegewerkt. Van de leerlingen die in schooljaar 1996/97 en 2000/01 in groep 8 zaten, is ongeveer de helft verder gevolgd in het voortgezet onderwijs. In het schooljaar 2007/08 is PRIMA opgevolgd door COOL, waarin zowel het primair als voortgezet onderwijs zijn vertegenwoordigd. In het basisonderwijs is informatie verzameld in de groepen 2, 5 en 8, en in het voortgezet onderwijs in leerjaar 3. In het basisonderwijs gaat het om ruim 38.000 leerlingen van 550 scholen, en in het voortgezet onderwijs om 9.000 leerlingen van 81 scholen.
Binnen de totale PRIMA- en COOL-steekproef in het basisonderwijs wordt een steekproef van scholen onderscheiden die representatief is voor alle basisscholen in Nederland en een aanvullende steekproef van scholen met een oververtegenwoordiging van achterstandsleerlingen. Er is om twee redenen voor die opzet gekozen. Op de eerste plaats is het van belang om voldoende achterstandsleerlingen in het onderzoek te krijgen en dat daarbij ook kleine categorieën van leerlingen met voldoende aantallen vertegenwoordigd zijn. Op de tweede plaats zitten achterstandsleerlingen meestal geconcentreerd op bepaalde scholen, vooral name allochtone achterstandsleerlingen. Door te zorgen voor een sterke vertegenwoordiging van deze scholen in de steekproef kan een goed beeld worden verkregen van de ‘typische’ achterstandsleerling. Omdat het bij de onderhavige vraagstelling niet gaat om het geven van een representatief beeld, maar om het bepalen van ontwikkelingen *binnen* OAB-categorieën, hebben we de analyses steeds uitgevoerd op de totale steekproeven van leerlingen. Omdat we bij onze uitsplitsingen steeds een betrekkelijk groot aantal leerlingcategorieën onderscheiden, draagt deze keuze

er toe bij dat er ook in de kleine categorieën voldoende leerlingen zitten (vgl. Luyten, Cremers-Van Wees & Bosker, 2001).
Binnen de PRIMA-cohorten in het voortgezet onderwijs is het – vanwege het uitwaaieren van de basisschoolleerlingen over zeer veel scholen voor voortgezet onderwijs - niet mogelijk een representatieve steekproef te onderscheiden. Het COOL-cohort in het voortgezet onderwijs heeft een andere opzet dan die in het basisonderwijs. Op leerlingniveau is er wat de verdeling over de verschillende onderwijstypes sprake van een ondervertegenwoordiging van de lagere onderwijsniveaus en een oververtegenwoordiging van de hogere niveaus (Zijsling e.a., i.v.). Voor het onderhavige onderzoek is bovendien van belang dat van circa 40% van de leerlingen geen informatie beschikbaar is over het opleidingsniveau en geboorteland van hun ouders en dat het bij de leerlingen waarvan die informatie wel beschikbaar is om een selectieve steekproef gaat (waarbij kinderen van allochtone en laagopgeleide ouders ondervertegenwoordigd zijn).6 Voor de vergelijking van de ontwikkelingen binnen OAB-categorieën heeft dat wellicht minder ernstige consequenties.
Voor de analyses die betrekking hebben op het basisonderwijs maken we gebruik van de gegevens die verzameld zijn in 1994/95, 1998/99, 2002/03 (PRIMA) en 2007/08 (COOL). Daarbinnen selecteren we de leerlingen in groep 2 en groep 8, ofwel bij de start en op het eind van het basisonderwijs. De analyse van de gegevens uit het voortgezet onderwijs heeft betrekking op leerjaar 3. Bij de beide PRIMA-vo cohorten gaat het om leerlingen die vanuit groep 8 doorgestroomd zijn naar een zeer groot aantal vo-scholen. We beschikken alleen over de gegevens van *deze* leerlingen op het moment dat ze in leerjaar 3 zitten en dus niet van *alle* leerlingen in de klassen waarin ze zitten. In het COOL-vo cohort ligt de situatie anders; dat cohort is namelijk gestart in leerjaar 3 en zijn er gegevens verzameld bij hele klassen. Qua samenstelling van de leerlingensteekproef is er ook een verschil tussen PRIMA-vo en COOL-vo. De hier gebruikte PRIMA-vo steekproef omvat alleen leerlingen die gedurende de eerste drie jaren in het vo niet zijn blijven zitten; COOL-vo omvat ook leerlingen die in die fase wel zijn blijven zitten. Voor uitvoerige informatie over de beide cohortonderzoeken verwijzen we naar Claassen en Mulder (2003, 2006), Claassen, Mulder en Hulsen (2005), Driessen, Van Langen en Vierke (2002, 2004), Driessen e.a. (2009), Jungbluth e.a. (1996) en Zijsling e.a. (i.v.).

---

[6] Bij PRIMA en COOL-po is de informatie over deze kenmerken zowel verkregen van de scholen als van de ouders zelf en stemmen de gegevens van beide groepen van informanten vrij goed overeen. Bij COOL-vo is er over deze kenmerken alleen informatie beschikbaar die door de ouders is verstrekt.

2.3.2 Instrumenten en kenmerken

In het onderzoek staan drie soorten gegevens centraal, namelijk de sociaal-etnische achtergronden van de leerlingen, hun taal- en rekenvaardigheid in het basisonderwijs, en hun onderwijspositie en taal- en wiskundevaardigheid in het voortgezet onderwijs.

Wat betreft de achtergronden onderscheiden we de OAB-categorie (zie hoofdstuk 1 van deze rapportage) waartoe de leerling behoort en de etnische herkomst.[7] Het eerste kenmerk is gebaseerd op de combinatie van etnische herkomst (i.c. het geboorteland) en opleidingsniveau van de ouders, het tweede op de etnische herkomst van de ouders. De betreffende gegevens zijn voor PRIMA en COOL-po afkomstig uit de schooladministraties, voor COOL-vo uit de vragenlijsten die de ouders hebben ingevuld.

Voor de beschrijving van de ontwikkelingen van de taal- en rekenvaardigheid in het *basisonderwijs* is gebruik gemaakt van gegevens uit het basisonderwijsdeel van de cohortstudies PRIMA en COOL. De taal- en rekenvaardigheid zijn in beide cohortstudies gemeten met toetsen die alle door het Cito zijn ontwikkeld. De resultaten op deze toetsen zijn uitgedrukt in zogenoemde vaardigheidsscores (zie Vierke, 1995; de Cito toetshandleidingen).

In Tabel 2.1 geven we een overzicht van de afgenomen toetsen. Een deel daarvan is speciaal door het Cito voor PRIMA ontwikkeld ('PRIMA'); het betreft de PRIMA taal- en rekentoetsen die in groep 8 zijn gebruikt. Deze toetsen meten algemene taal- en rekenvaardigheden. De overige toetsen zijn afkomstig uit het Cito Leerlingvolgsysteem. De toets Rekenen/Wiskunde voor groep 8 bestrijkt een aantal verschillende deelvaardigheden. In COOL zijn in groep 8

---

[7] In eerder, soortgelijk onderzoek zijn we ook nagegaan of er verschillen bestonden die te maken hadden met de sekse van de leerlingen en hun verblijfsduur in Nederland (Mulder e.a., 2005). Daar hebben we nu vanaf gezien. In de COOL-po steekproef is nog slechts 3.7% van de leerlingen niet in Nederland geboren en bovendien verblijft tweederde van die leerlingen ook nog eens langer dan 5 jaar in Nederland. Uit de eerdere analyses is gebleken dat vooral kinderen die korter dan 4 jaar in Nederland verbleven wat lager scoorden op taal en rekenen (vooral in groep 8). Deze categorie is bij de laatste metingen qua aantallen echter nog slechts minimaal in de steekproef vertegenwoordigd. Met betrekking tot mogelijke sekseverschillen in prestaties hebben we onlangs twee uitvoerige trendstudies verricht (Van Langen & Driessen, 2006; Driessen & Van Langen, 2010). In beide studies zijn we nagegaan of er in het basis- en voortgezet onderwijs verschillen bestaan tussen jongens en meisjes ten aanzien van onder meer hun taal- en rekenvaardigheid en of er zich daarbij sinds 1994 (tot 2004, resp. 2007) ontwikkelingen hebben voorgedaan. Voor deze analyses hebben we dezelfde data gebruikt als in het onderhavige onderzoek (en daarnaast nog verschillende andere bronnen). De conclusie luidde dat meisjes wat beter zijn qua taal en jongens wat beter qua rekenen; de verschillen zijn echter zeer gering en er is ook geen sprake van ontwikkelingen over de tijd. In de meeste recente studie (uit 2010) is bovendien specifiek onderzocht of er *binnen* sociaal-etnische groepen (de combinatie van etnische herkomst en opleiding ouders; vgl. de hier geanalyseerde OAB-categorieën) verschillen zijn tussen jongens en meisjes (m.a.w. of er sprake is van een interactie-effect). Dat bleek niet het geval te zijn: er zijn geen sociaal-etnische categorieën waar het (sowieso al geringe) sekseverschil in prestaties systematisch afwijkt van dat in andere categorieën.

voor taal en lezen drie verschillende toetsen afgenomen, te weten Leeswoordenschat, Drie Minuten Toets, en Begrijpend lezen. Omdat Begrijpend lezen doorgaans het sterkste samenhangt met een extern criterium, zoals het advies voortgezet onderwijs (vgl. Driessen & Cuppen, zie hoofdstuk 4 van dit rapport), hebben we hier deze toets als indicator voor taalvaardigheid genomen.

Tabel 2.1 Overzicht afgenomen toetsen taal- en rekenvaardigheid in PRIMA en COOL, naar meting en groep

	1994/95	1998/99	2002/04	2007/08
Groep 2				
taal	Begrippen	Begrippen	Taal voor kleuters	Taal voor kleuters
rekenen	Ordenen	Ordenen	Ordenen	Ordenen
Groep 8				
taal	PRIMA	PRIMA	PRIMA	Begrijpend lezen
rekenen	PRIMA	Rekenen/Wiskunde	Rekenen/Wiskunde	Rekenen/Wiskunde

Met uitzondering van de bepaling van de rekenvaardigheid in groep 2 (m.b.v. de toets Ordenen) zijn er gedurende de bestudeerde periode verschillende toetsen gebruikt. Voor de analyses die we hier presenteren is dat geen bezwaar. We kijken hier immers niet naar verschillen in toetsscore tussen cohorten, maar naar de relatieve positie (verschillende categorieën van leerlingen ten opzichte van elkaar) en eventuele veranderingen daarin in de tijd. Daarvoor maakt het niet zoveel uit of de toetsen precies hetzelfde meten, als maar alle leerlingen steeds dezelfde toetsen hebben gemaakt.
Voor de beschrijving van de ontwikkelingen in het *voortgezet onderwijs* hebben we gegevens gebruikt die door de scholen zijn aangeleverd. We concentreren ons op de onderwijspositie en de taal- en wiskundevaardigheid. De onderwijspositie bepalen we in termen van hun positie op de zogenoemde leerjarenladder (Bosker & Van der Velden, 1989; Claassen, Mulder & Hulsen, 2005). Deze ladder loopt in principe (afhankelijk van het leerjaar) van 1 tot en met 12. Score 1 – de laagste trede op de ladder – staat voor leerlingen in leerjaar 1 van het praktijkonderwijs, score 12 voor leerlingen in leerjaar 6 van het vwo. In Tabel 2.2 geven we een overzicht van de onderwijstypes in leerjaar 3 en de bijbehorende leerjarenladderscores.

Tabel 2.2 Overzicht onderwijstypes en bijbehorende leerjarenladderscores voor leerjaar 3 voortgezet onderwijs

Type*	Score	Type	Score
pro	3	gl/tl en tl	6
pro/lwoo	3.5	tl/havo	6.5
lwoo	4	havo	7
bbl	4.5	havo/vwo	7.5
bbl/kbl en kbl	5	vwo	8
kbl/gl en gl	5.5		

* pro=praktijkonderwijs; lwoo=leerwegondersteunend onderwijs; bbl=basisberoepsgerichte leerweg; kbl=kaderberoepsgerichte leerweg; gl=gemengde leerweg; tl=theoretische leerweg.

Om de taal- en wiskundevaardigheid in kaart te brengen maken we gebruik van twee soorten gegevens, rapportcijfers en toetsscores. In PRIMA-vo zijn geen toetsen afgenomen, maar zijn voor enkele vakken de gemiddelde rapportcijfers bij de scholen opgevraagd (die in principe variëren van 0 tot 10). Het gaat hier om de brede domeinen Nederlands en wiskunde. In COOL-vo zijn geen rapportcijfers verzameld, maar zijn toetsen voor de kernvakken afgenomen. Om de lijn vanuit het basisonderwijs te kunnen doortrekken hebben we hieruit de toetsen Begrijpend lezen en Wiskunde geselecteerd. De resultaten hierop zijn door de onderzoekers omgezet in zogenoemde bankscores. Bankscores liggen tussen de 0 en 100%, en kunnen worden geïnterpreteerd in termen van beheersingsniveau. Een score van 70 wil zeggen dat een leerling naar verwachting 70% van de alle items van de toets Begrijpend lezen correct kan maken (Zijsling e.a., i.v.).
Omdat er in beide cohorten voor de taal- en wiskundevaardigheid twee verschillende instrumenten zijn gebruikt, kan wel de ontwikkeling tussen de beide PRIMA-vo metingen worden beschreven, maar kan de taal- en wiskundevaardigheid zoals gemeten bij PRIMA-vo niet vergeleken worden met die zoals gemeten bij COOL-vo.

2.3.3 Analyse–opzet

Om de ontwikkelingen van de leerlingen in kaart te brengen, zetten we de gemiddelde scores van elk van de onderscheiden categorieën van leerlingen steeds af tegen die van een referentiegroep. Voor deze rapportage is gekozen voor de kinderen van autochtone ouders met een mbo-opleiding (hierna kortheidshalve aangeduid als ‘aut. 1.0 mbo’). Het betreft feitelijk de ouders met

mavo, havo, vwo of mbo. Om een indicatie te geven van de relevantie van het verschil maken we, waar mogelijk, gebruik van zogenoemde effect sizes (*ES* of effectgroottes; vgl. Thompson, 1998). Effect sizes hebben als voordeel dat ze niet afhankelijk zijn van de steekproefomvang en bovendien dat, doordat het een gestandaardiseerde coëfficiënt betreft, ze voor verschillende indicatoren rechtstreeks met elkaar kunnen worden vergeleken (Coe, 2002). Een ander groot voordeel is dat onderzoeksresultaten op een zeer compacte wijze kunnen worden samengevat. In zijn meest simpele vorm is een *ES* het verschil tussen de gemiddelden (bv. toetsscores) van twee groepen gedeeld door de standaarddeviatie (deze *ES* wordt ook wel aangeduid als Cohen's *d*). Wat de interpretatie van een *ES* betreft, wordt doorgaans de suggestie van Cohen (1988) overgenomen, die een coëfficiënt van 0.20 als 'klein' bestempelt, die van 0.50 als 'middelmatig' en die van 0.80 als 'groot'. Deze classificatie dient echter niet al te rigide te worden opgevat; Cohen zelf noemt het een vuistregel.

2.4 Resultaten

Hierna presenteren we in paragraaf 2.4.1 eerst op basis van de meeste recente gegevens de stand van zaken in het schooljaar 2007/08. Het gaat daarbij om de taal- en rekenvaardigheid in het basisonderwijs en de onderwijspositie en taal- en wiskundevaardigheid in het voortgezet onderwijs. Vervolgens beschrijven we in paragraaf 2.4.2 de ontwikkelingen in deze gegevens in de periode 1994/05 tot en met 2007/08.

2.4.1 Stand van zaken 2007/08

2.4.1.1 Basisonderwijs

Taalvaardigheid

In Tabel 2.3 geven we allereerst een overzicht van de taalvaardigheid in groep 2 en groep 8 uitgesplitst naar OAB-categorie. De gemiddelden voor groep 2 en groep 8 in deze tabel kunnen overigens niet met elkaar worden vergeleken (dat is wel het geval bij de beschrijving van de ontwikkelingen in de volgende paragraaf). Ze geven per groep wel een goede indruk van de onderlinge verschillen tussen de categorieën. Onder in de tabel presenteren we de overall effect size (*ES*) en de mate waarin de verschillen in vaardigheidsscores kunnen worden toegeschreven aan de OAB-indeling (% verklaarde variantie).

Tabel 2.3 De taalvaardigheid in groep 2 en 8 in 2007/08, naar OAB–categorie (gemiddelden, standaarddeviaties en aantallen)

	Groep 2			Groep 8		
	gem.	sd	n	gem.	sd	n
aut. 1.0 ho	76.3	10.0	2744	64.6	20.0	2103
aut. 1.0 mbo	74.4	9.3	2616	57.7	15.1	1946
aut. 1.25 ex	73.1	9.2	1844	53.9	16.3	1801
aut. 1.25 0.3	71.4	9.4	1270	49.2	14.2	1435
aut. 1.25 1.2	68.5	9.2	317	46.9	14.5	265
all. 1.0 ho	69.8	10.0	413	57.8	15.8	274
all. 1.0 mbo	66.6	9.5	568	51.6	15.5	376
all. 1.9 ex	66.8	9.2	749	50.9	14.6	527
all. 1.9 0.3	65.9	8.6	389	46.4	13.9	380
all. 1.9 1.2	64.4	8.7	1370	46.4	14.3	1157
Totaal	71.8	10.3	12280	54.7	17.4	10264
ES (% verkl. var.)	0.87 (15.8%)			0.77 (13.0%)		

De gegevens voor groep 2 laten zien dat er zowel binnen de totale groep van autochtonen als binnen de totale groep van allochtonen een nagenoeg monotoon verloop is van de taalscores over de OAB-categorieën Ook wordt duidelijk dat er grote verschillen zijn tussen autochtone en allochtone leerlingen met dezelfde opleidingsachtergrond. Om enkele opvallende voorbeelden te geven. Tussen autochtone en allochtone kinderen van hoogopgeleide ouders bestaat een verschil van 6.5 punt, ofwel tweederde standaarddeviatie, wat tussen een 'middelmatig' en 'groot' verschil in ligt. Tussen autochtone en allochtone 'lichte' (0.3') achterstandskinderen bestaat een verschil van 5.5 punt, ofwel ruim een halve standaarddeviatie of middelmatig verschil. De overall effect size bedraagt 0.87 en bijna 16% van de verschillen in taalvaardigheid kan worden toegeschreven aan de OAB-categorie waartoe de leerlingen behoren.
In groep 8 is de overall effect size minder groot (0.77) en kan er ook minder van de taalverschillen door de OAB-categorie worden verklaard (13.0%).[8] De

[8] We willen er hier nog eens aan herinneren, dat het hier om cross-sectionele analyses gaat en de twee steekproeven (groep 2 en groep 8) onafhankelijk van elkaar zijn. Dat betekent dus dat - anders dan bij echte longitudinale analyses, waar alleen dezelfde leerlingen worden gevolgd – groep 8 een dwarsdoorsnee betreft (met bijvoorbeeld ook zittenblijvers). Dit houdt in dat de gemiddelden niet worden vertekend door (selectieve) tussentijdse uitval van laagpresterende zittenblijvers.

verschillen tussen de autochtone en vergelijkbare allochtone OAB-categorieën zijn (in termen van standaarddeviaties) ook veel minder groot. Tussen de autochtone en allochtone kinderen van hoogopgeleide ouders bestaat nog een verschil van 6.8 punt of vier tiende standaarddeviatie. Tussen autochtone en allochtone 'lichte' achterstandskinderen bedraagt het verschil 2.8 punt of een zevende standaarddeviatie. Het lijkt er al met al op dat de sociaal-etnische achtergrond (i.c. OAB-categorie) er voor de taalvaardigheid in groep 8 wat minder toe doet dan in groep 2.
In Tabel 2.4 splitsen we de taalvaardigheid van de drie 1.9-categorieën verder uit naar herkomstland. Hierbij onderscheiden we Turkije, Marokko, Suriname en de Antillen, en de overige niet-westerse landen. Onder in de tabel vermelden we ter vergelijking ook nog een keer het gemiddelde van de referentiecategorie, de kinderen van autochtone ouders met een mbo-niveau.

Tabel 2.4 De taalvaardigheid van de 1.9–leerlingen in groep 2 en 8 in 2007/08, naar herkomstland (gemiddelden, standaarddeviaties en aantallen)

	Groep 2			Groep 8		
	gem.	sd	n	gem.	sd	n
ex Turkije	65.3	8.3	173	48.8	13.0	154
ex Marokko	67.0	8.9	170	51.5	15.6	117
ex Suriname/Antillen	66.8	9.0	54	51.6	12.2	47
ex Overig niet-westers	65.2	8.1	112	54.2	14.0	80
0.3 Turkije	65.1	8.6	71	43.7	13.1	89
0.3 Marokko	65.6	8.0	73	44.9	10.5	48
0.3 Suriname/Antillen	66.3	8.6	86	45.8	12.7	120
0.3 Overig niet-westers	65.7	9.3	83	49.7	16.1	67
1.2 Turkije	63.8	8.6	394	44.4	13.2	425
1.2 Marokko	65.0	8.3	429	46.8	14.6	366
1.2 Suriname/Antillen	63.6	9.8	50	43.3	14.2	52
1.2 Overig niet-westers	64.7	9.3	369	49.2	14.8	237
aut. 1.0 mbo	74.4	9.3	2616	57.7	15.1	1946

Tabel 2.4 laat zien dat er in groep 2 qua taalvaardigheid tussen herkomstlanden binnen de drie 1.9-categorieën weliswaar verschillen zijn, maar dat deze verschillen doorgaans vrij gering zijn en bovendien dat er geen duidelijke lijn in zit. Opvallend is dat er tussen de ex 1.9-leerlingen en de 0.3-

leerlingen nauwelijks verschillen zijn. Beide categorieën scoren wel wat hoger dan de 1.2-leerlingen. In groep 8 zijn de verschillen tussen de ex 1.9-leerlingen en de 0.3-leerlingen geprononceerder, terwijl de 0.3- en de 1.2-leerlingen juist meer op elkaar lijken.

Rekenvaardigheid

In Tabel 2.5 staan de cijfers voor de rekenvaardigheid in groep 2 en 8.

Tabel 2.5 De rekenvaardigheid in groep 2 en 8 in 2007/08, naar OAB–categorie (gemiddelden, standaarddeviaties en aantallen)

	Groep 2			Groep 8		
	gem.	sd	n	gem.	sd	n
aut. 1.0 ho	62.1	13.8	2770	120.7	8.8	2011
aut. 1.0 mbo	58.9	13.1	2599	117.5	9.2	1904
aut. 1.25 ex	57.4	12.6	1835	116.1	8.8	1767
aut. 1.25 0.3	54.9	12.8	1254	113.6	9.1	1390
aut. 1.25 1.2	52.1	12.1	313	113.0	11.8	246
all. 1.0 ho	55.8	12.9	417	118.7	8.6	264
all. 1.0 mbo	51.7	12.2	566	114.6	9.4	363
all. 1.9 ex	52.6	11.7	745	115.3	8.9	514
all. 1.9 0.3	50.6	11.7	388	112.3	9.5	373
all. 1.9 1.2	50.5	11.7	1355	113.5	9.5	1109
Totaal	56.8	13.4	12242	116.4	9.5	9941
ES (% verkl. var.)	0.63 (9.0 %)			0.59 (8.0%)		

Wat ten aanzien van rekenen op de eerste plaats opvalt, is dat het effect van OAB-categorie (en het percentage verklaarde variantie) beduidend geringer is dan voor taal. Dit is een fenomeen dat al eerder is geconstateerd. Het heeft waarschijnlijk te maken met het gegeven dat rekenen iets is wat vooral op school wordt geleerd, terwijl taal veel meer - en zeker op jongere leeftijd - onder invloed staat van de sociaal-etnische omstandigheden van het gezin waarin het kind opgroeit. Het feit dat door de invoering van het realistisch rekenen ook rekenen veel taliger is geworden doet daar dan weinig aan af.
In groep 2 is er binnen de totale groep autochtonen een duidelijk scoreverloop over de OAB-categorieën: naarmate kinderen zich in een zwaardere achterstandspositie bevinden halen ze ook lagere scores. Binnen de totale

groep allochtonen ontbreekt een dergelijke trend. In groep 8 zien we een vergelijkbaar patroon

In Tabel 2.6 volgt de nadere differentiatie in rekenscores naar geboorteland.

Tabel 2.6 De rekenvaardigheid van de 1.9–leerlingen in groep 2 en 8 in 2007/08, naar herkomstland (gemiddelden, standaarddeviaties en aantallen)

	Groep 2			Groep 8		
	gem.	sd	n	gem.	sd	n
ex Turkije	51.3	10.9	170	116.1	8.0	152
ex Marokko	53.1	11.4	169	114.0	8.5	114
ex Suriname/Antillen	53.6	12.8	54	113.3	7.6	46
ex Overig niet-westers	52.6	11.5	113	119.3	9.6	78
0.3 Turkije	49.6	11.8	68	113.1	9.7	90
0.3 Marokko	49.9	15.3	74	113.1	8.1	46
0.3 Suriname/Antillen	50.5	9.7	85	109.5	9.9	114
0.3 Overig niet-westers	51.0	12.4	85	115.2	8.0	69
1.2 Turkije	50.2	11.3	389	113.9	9.0	412
1.2 Marokko	50.6	11.7	425	112.8	9.0	360
1.2 Suriname/Antillen	50.3	18.0	50	108.6	12.8	46
1.2 Overig niet-westers	50.7	11.4	366	114.6	9.5	219
aut. 1.0 mbo	58.9	13.1	2599	117.5	9.2	1904

In groep 2 ligt de rekenvaardigheid van de ex 1.9-leerlingen wat hoger dan die van de 0.3-leerlingen. Tussen de 1.9- en 0.3-leerlingen zijn er maar zeer geringe verschillen; de Turkse en Marokkaanse 1.2-leerlingen scoren zelfs iets hoger dan de Turkse en Marokkaanse 0.3-leerlingen. In groep 8 zien we dat er slechts minimale verschillen zijn tussen 0.3- en 1.2-leerlingen, maar dat beide categorieën wel redelijk wat lager scoren dan de ex 1.9-leerlingen.

2.4.1.2 Voortgezet onderwijs

Onderwijspositie leerjaar 3

In Tabel 2.7 staan de gemiddelde leerjarenladderscores in leerjaar 3 van het voortgezet onderwijs. Zoals eerder toegelicht geven deze scores een indicatie van het niveau van het onderwijstype waarin de kinderen zitten. We willen er

trouwens op attenderen dat de aantallen leerlingen in enkele categorieën vrij klein zijn.

Tabel 2.7 Positie op de leerjarenladder in leerjaar 3 in 2007/08, naar OAB–categorie (gemiddelden, standaarddeviaties en aantallen)

	gem.	sd	n
aut. 1.0 ho	7.2	1.0	1991
aut. 1.0 mbo	6.4	1.1	1437
aut. 1.25 ex	6.2	1.2	809
aut. 1.25 0.3	5.7	1.1	269
aut. 1.25 1.2	5.1	1.2	30
all. 1.0 ho	7.0	1.0	162
all. 1.0 mbo	6.4	1.3	136
all. 1.9 ex	6.2	1.3	87
all. 1.9 0.3	6.0	1.3	33
all. 1.9 1.2	5.6	1.1	106
Totaal	6.6	1.2	5060
ES (% verkl. var.)	0.94 (18.1%)		

Tabel 2.7 laat zien dat er tussen de leerlingen met de hoogste en die met de laagste score op leerjarenladder een verschil bestaat van twee volledige treden. Meer concreet: de autochtone leerlingen van ouders met een hbo- of wo-opleiding (aut. 1.0 ho) zitten gemiddeld in het havo, terwijl de eveneens autochtone leerlingen van laagopgeleide ouders (aut. 1.25 1.2) gemiddeld in de kbl zitten. Opmerkelijk is dat de huidige autochtone achterstandsleerlingen lager scoren dan hun allochtone jaargenoten, met name de kinderen van zeer laagopgeleide ouders. Bij de overige categorieën ontbreken dergelijke verschillen.
Bij de presentatie van de gegevens over het basisonderwijs hebben we steeds ook de scores van elk van de 1.9-categorieën uitgesplitst naar herkomstland. De aantallen leerlingen in de vo-bestanden zijn echter beduidend kleiner dan die in de basisonderwijsbestanden. Om die reden presenteren we in Tabel 2.8 de gegevens voor de totale categorie van 1.9-leerlingen. Hieruit volgt dat de Marokkaanse achterstandsleerlingen de laagste onderwijspositie innemen en de overige niet-westerse achterstandsleerlingen de hoogste. Aangezien ook hier de aantallen klein zijn, blijft het nodige voorbehoud op z'n plaats.

Tabel 2.8 Positie op de leerjarenladder van de 1.9–leerlingen in leerjaar 3 in 2007/08, naar herkomstland (gemiddelden, standaarddeviaties en aantallen)

	gem.	sd	n
Turkije	5.7	1.2	51
Marokko	5.6	1.2	50
Suriname/Antillen	6.2	1.4	15
Overig niet-westers	6.7	1.2	41
aut. 1.0 mbo	6.4	1.1	1437

Taal– en wiskundevaardigheid

In Tabel 2.9 volgen de scores op de taal- en wiskundetoets, uitgesplitst naar OAB-categorie.

Tabel 2.9 De taal– en wiskundevaardigheid in leerjaar 3 in 2007/08, naar OAB–categorie (gemiddelden, standaarddeviaties en aantallen)

	Taal			Wiskunde		
	gem.	sd	n	gem.	sd	n
aut. 1.0 ho	71.5	16.8	1930	74.1	15.6	1915
aut. 1.0 mbo	65.5	18.1	1388	66.1	16.7	1391
aut. 1.25 ex	62.9	18.7	794	64.3	17.4	796
aut. 1.25 0.3	58.4	19.3	265	59.2	16.9	264
aut. 1.25 1.2	52.3	19.2	28	52.2	17.5	27
all. 1.0 ho	68.4	16.4	158	67.0	18.4	156
all. 1.0 mbo	61.0	18.7	131	63.2	19.2	134
all. 1.9 ex	59.0	17.7	85	59.6	20.9	87
all. 1.9 0.3	57.6	22.0	31	55.9	21.5	31
all. 1.9 1.2	55.9	17.6	106	53.8	17.5	110
Totaal	66.6	18.3	4916	68.0	17.5	4911
ES (% verkl. var.)	0.53 (6.5%)			0.66 (9.9%)		

Wat zowel de taal- als wiskundevaardigheid betreft, blijkt dat naarmate de achterstandspositie van de leerlingen ongunstiger wordt, zij ook lager scoren op de toetsen. Er is in dat opzicht een duidelijke parallel tussen autochtone en allochtone leerlingen, zij het dat de allochtone leerlingen doorgaans lager presteren dan de autochtone leerlingen. Er is echter één uitzondering: de

‘zwaarste’ autochtone achterstandscategorie scoort lager dan de ‘zwaarste’ allochtone achterstandscategorie. (Maar ook hier geldt weer het voorbehoud van het geringe aantal leerlingen in die categorie.)
Vergelijken we deze bevindingen met die in het basisonderwijs, dan vallen enkele dingen op. In het basisonderwijs was het effect van de OAB-categorie op de taalvaardigheid groter dan op de rekenvaardigheid. Als verklaring daarvoor wezen we op het grotere belang van de thuissituatie (ofwel sociaal-etnische achtergrond) op de taalontwikkeling. We zagen ook dat de invloed van die thuissituatie op de taal- en rekenvaardigheid in groep 8 wat geringer is dan die in groep 2. In leerjaar 3 voortgezet onderwijs is de invloed van de thuissituatie op de taalvaardigheid nog weer veel geringer dan in groep 8 en bovendien is de invloed van de thuissituatie op de wiskundevaardigheid nu groter dan die op de taalvaardigheid. Een mogelijke verklaring voor dit laatste zou kunnen zijn dat de kinderen van hoogopgeleide ouders thuis meer ondersteuning krijgen dan kinderen van laagopgeleide (allochtone) ouders, bijvoorbeeld bij het maken van het huiswerk voor wiskunde. Een rol daarbij kan spelen dat wiskunde in het voortgezet onderwijs vaak zeer abstract en talig is.
Ter informatie presenteren we in Tabel 2.10 ook nog de taal- en wiskundevaardigheid uitgesplitst naar herkomstland. Die tabel laat zien dat de Surinaamse en Antilliaanse leerlingen het laagst scoren qua taal. Dat is tegen de verwachting in, omdat beide groepen in het herkomstland vaak al Nederlands spreken. Mogelijk is het geringe aantal leerlingen in deze categorie debet aan dit laag gemiddelde. De overige niet-westerse leerlingen scoren het hoogst op taal, en ook op wiskunde. De Marokkaanse leerlingen hebben de geringste wiskundevaardigheid.

Tabel 2.10 De taal- en wiskundevaardigheid van de 1.9-leerlingen in leerjaar 3 in 2007/08, naar herkomstland (gemiddelden, standaarddeviaties en aantallen)

	Taal			Wiskunde		
	gem.	sd	n	gem.	sd	n
Turkije	56.6	17.7	50	56.1	18.5	55
Marokko	54.2	19.4	52	51.6	18.6	52
Suriname/Antillen	52.8	15.5	13	60.0	20.0	13
Overig niet-westers	62.9	16.7	38	66.3	16.9	40
aut. 1.0 mbo	65.5	18.1	1388	66.1	16.7	1391

2.4.2 Ontwikkelingen 1994/95 – 2007/08

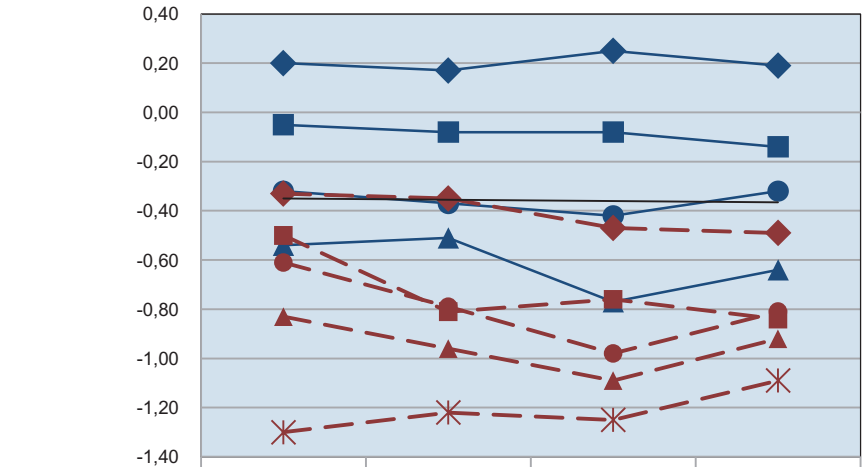
In de vorige paragraaf lag het accent op de verschillen tussen de onderscheiden OAB-categorieën bij de meest recente meting in 2007/08. In deze paragraaf staan de ontwikkelingen van elk van de OAB-categorieën ten opzichte van de referentiecategorie tussen 1994/95 en 2007/08 centraal. We presenteren daarom niet de gemiddelden qua taal- en rekenvaardigheid en onderwijspositie, maar de bijbehorende effect sizes.

2.4.2.1 Basisonderwijs

Taalvaardigheid

In Figuur 2.1 en 2.2 volgen de effect sizes met betrekking tot de ontwikkeling in taalvaardigheid, eerst voor groep 2 en dan voor groep 8.

Figuur 2.1 De ontwikkeling van de taalvaardigheid in groep 2 tussen 1994/95 en 2007/08, naar OAB–categorie (effect sizes; referentiecategorie: autochtoon, mbo)



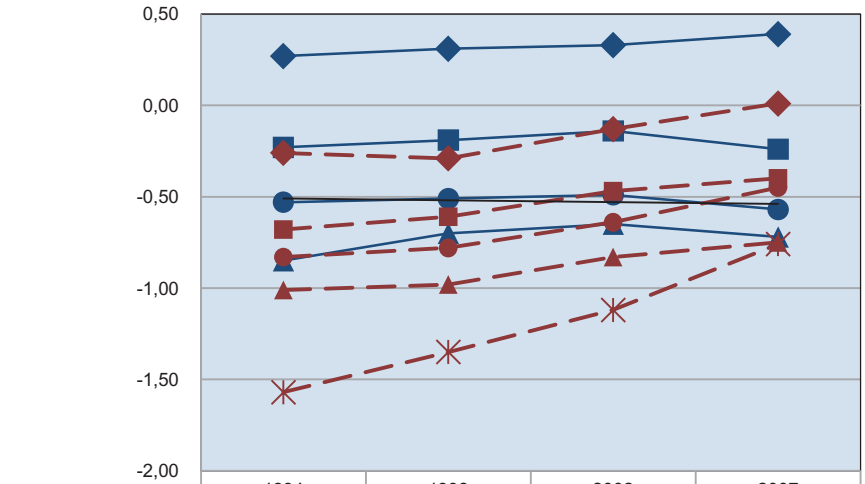
	1994	1998	2002	2007
aut. 1.0 ho	0,20	0,17	0,25	0,19
aut. 1.25 ex	-0,05	-0,08	-0,08	-0,14
aut. 1.25 0.3	-0,32	-0,37	-0,42	-0,32
aut. 1.25 1.2	-0,54	-0,51	-0,77	-0,64
all. 1.0 ho	-0,33	-0,35	-0,47	-0,49
all 1.0 mbo	-0,50	-0,81	-0,76	-0,84
all 1.9 ex	-0,61	-0,79	-0,98	-0,81
all 1.9 0.3	-0,83	-0,96	-1,09	-0,92
all 1.9 1.2	-1,30	-1,22	-1,25	-1,09

Bij wijze van toelichting bij deze en volgende grafieken. In de eerste rij van de tabel zien we dat in 1994 het gestandaardiseerde verschil (de *ES*) tussen de autochtone leerlingen met hoogopgeleide ouders (aut. 1.0 ho) en de referentiecategorie (de autochtone leerlingen met middelbaar opgeleide ouders: aut. 1.0 mbo) 0.20 bedraagt. Dit is een klein *positief* effect: de kinderen van autochtone hoogopgeleide ouders doen het wat beter dan de kinderen uit de referentiecategorie. De gegevens over 1998, 2002 en 2007 maken duidelijk dat de verschillen tussen beide categorieën ongeveer dezelfde omvang houden. In de tweede rij van de tabel worden de leerlingen die vroeger tot de autochtone achterstandscategorie werden gerekend (aut. 1.25 ex) vergeleken met de referentiecategorie. Voor 1994 blijkt dat de *ES* -0.05 is, ofwel een minimaal

negatief effect: de voormalige achterstandsleerlingen doen het iets (maar minimaal) slechter dan de referentiecategorie. In de drie navolgende jaren blijft het effect vrij stabiel. Alleen in 2007 is het negatieve effect toegenomen tot -0.14. Los van het feit dat het nog steeds om een zeer klein effect handelt, betekent dit wel dat de afstand tot de referentiecategorie groter is geworden en dat de betreffende leerlingen dus verder achterop raken.

Wat de taalvaardigheid in groep 2 betreft is er over de onderzochte periode eigenlijk nauwelijks sprake van (min-of-meer rechtlijnige) ontwikkelingen. Soms zien we tussen jaren enig verschil, maar dat zet vervolgens niet door. Alleen voor de huidige 'zware' allochtone achterstandscategorie (all. 1.9 1.2) zien we bij de laatste meting een redelijke vooruitgang ten opzichte van eerdere metingen; de vraag is echter of dat bij een volgende meting doorzet.

Figuur 2.2 De ontwikkeling van de taalvaardigheid in groep 8 tussen 1994/95 en 2007/08, naar OAB-categorie (effect sizes; referentiecategorie: autochtoon, mbo)



	1994	1998	2002	2007
aut. 1.0 ho	0,27	0,31	0,33	0,39
aut. 1.25 ex	-0,23	-0,19	-0,14	-0,24
aut. 1.25 0.3	-0,53	-0,51	-0,49	-0,57
aut. 1.25 1.2	-0,85	-0,70	-0,65	-0,72
all. 1.0 ho	-0,26	-0,29	-0,13	0,01
all 1.0 mbo	-0,68	-0,61	-0,47	-0,40
all 1.9 ex	-0,83	-0,78	-0,64	-0,45
all 1.9 0.3	-1,01	-0,98	-0,83	-0,75
all 1.9 1.2	-1,57	-1,35	-1,12	-0,76

In groep 8 zien we voor alle allochtone categorieën vooruitgang, met name de sterk positieve ontwikkeling van de 'zware' allochtone achterstandscategorie valt op. In de onderzochte periode is hun achterstand gehalveerd.
In Tabel 2.11 splitsen we de gegevens van de drie categorieën van allochtone achterstandsleerlingen verder uit naar de vier onderscheiden herkomstlanden.

Tabel 2.11 De ontwikkeling in de taalvaardigheid van de 1.9–leerlingen in groep 2 en 8 tussen 1994/95 en 2007/08, naar herkomstland (effect sizes; referentiecategorie: autochtoon, mbo)

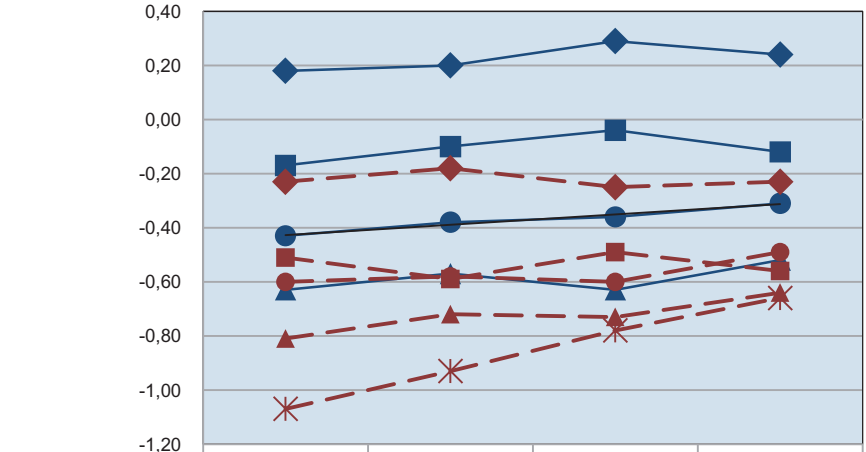
	Groep 2				Groep 8			
	1994	1998	2002	2007	1994	1998	2002	2007
ex Turkije	-1.01	-1.16	-1.21	-0.98	-1.30	-1.31	-1.00	-0.59
ex Marokko	-0.97	-0.83	-0.88	-0.79	-1.47	-0.88	-0.82	-0.41
ex Suriname/Antillen	-0.47	-0.72	-0.94	-0.82	-0.99	-0.82	-0.47	-0.40
ex Overig niet-westers	-0.53	-0.77	-1.21	-0.99	-0.90	-0.80	-0.63	-0.23
0.3 Turkije	-1.27	-1.27	-1.43	-1.00	-1.35	-1.41	-1.24	-0.93
0.3 Marokko	-0.92	-0.99	-1.06	-0.95	-1.11	-0.84	-0.99	-0.85
0.3 Suriname/Antillen	-0.76	-0.90	-1.02	-0.87	-1.10	-0.93	-0.89	-0.79
0.3 Overig niet-westers	-0.58	-0.81	-1.11	-0.93	-1.03	-0.91	-0.73	-0.52
1.2 Turkije	-1.41	-1.28	-1.39	-1.15	-1.79	-1.53	-1.34	-0.90
1.2 Marokko	-1.15	-1.20	-1.10	-1.02	-1.46	-1.25	-1.00	-0.73
1.2 Suriname/Antillen	-0.87	-1.02	-1.06	-1.15	-1.49	-1.09	-0.90	-0.95
1.2 Overig niet-westers	-1.14	-1.09	-1.19	-1.04	-1.26	-1.08	-0.96	-0.57

In groep 2 zien we dat voor met name de drie onderscheiden categorieën Turkse leerlingen het verschil met de referentiecategorie kleiner wordt; hun achterstand wordt dus kleiner. In groep 8 ligt dat wat anders. Daar is voor verschillende categorieën de achterstand afgenomen, vooral tussen de voorlaatste en laatste meting.

Rekenvaardigheid

In Figuur 2.3 en 2,4 presenteren we de ontwikkelingen qua rekenvaardigheid in achtereenvolgens groep 2 en 8. Daarna volgt in Tabel 2.12 de nadere differentiatie naar herkomstland.

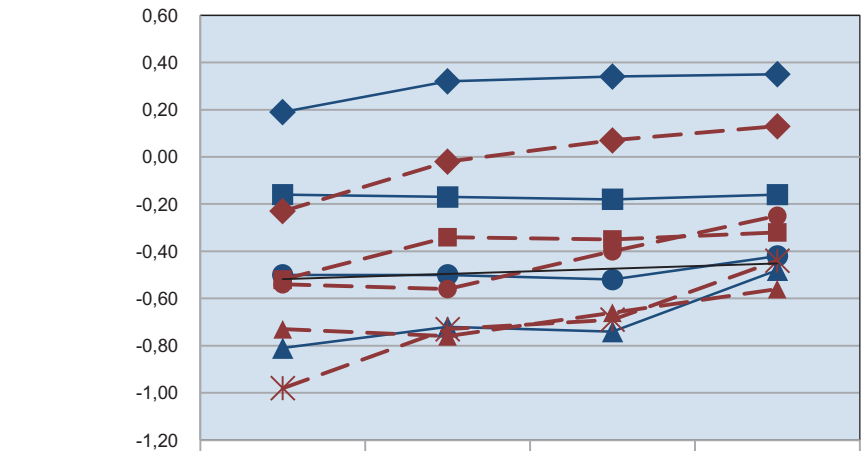
Figuur 2.3 De ontwikkeling van de rekenvaardigheid in groep 2 tussen 1994/95 en 2007/08, naar OAB–categorie (effect sizes; referentiecategorie: autochtoon, mbo)



	1994	1998	2002	2007
aut. 1.0 ho	0,18	0,20	0,29	0,24
aut. 1.25 ex	-0,17	-0,10	-0,04	-0,12
aut. 1.25 0.3	-0,43	-0,38	-0,36	-0,31
aut. 1.25 1.2	-0,63	-0,57	-0,63	-0,52
all. 1.0 ho	-0,23	-0,18	-0,25	-0,23
all 1.0 mbo	-0,51	-0,59	-0,49	-0,56
all 1.9 ex	-0,60	-0,58	-0,60	-0,49
all 1.9 0.3	-0,81	-0,72	-0,73	-0,64
all 1.9 1.2	-1,07	-0,93	-0,78	-0,66

Ook ten aanzien van de rekenvaardigheid in groep 2 blijkt dat met name de zware allochtone achterstandscategorie hun positie flink hebben weten te verbeteren. Mogelijk dat de vele VVE-programma's die de laatste jaren beschikbaar zijn hier een positief effect bij hebben gehad.

Figuur 2.4 De ontwikkeling van de rekenvaardigheid in groep 8 tussen 1994/95 en 2007/08, naar OAB-categorie (effect sizes; referentiecategorie: autochtoon, mbo)



	1994	1998	2002	2007
aut. 1.0 ho	0,19	0,32	0,34	0,35
aut. 1.25 ex	-0,16	-0,17	-0,18	-0,16
aut. 1.25 0.3	-0,50	-0,50	-0,52	-0,42
aut. 1.25 1.2	-0,81	-0,72	-0,74	-0,48
all. 1.0 ho	-0,23	-0,02	0,07	0,13
all 1.0 mbo	-0,52	-0,34	-0,35	-0,32
all 1.9 ex	-0,54	-0,56	-0,40	-0,25
all 1.9 0.3	-0,73	-0,76	-0,66	-0,56
all 1.9 1.2	-0,98	-0,73	-0,69	-0,44

Wat de rekenvaardigheid in groep 8 aangaat, valt vooral de vooruitgang van de lichte en zware allochtone en autochtone achterstandscategorieën op.

Tabel 2.12 De ontwikkeling in de rekenvaardigheid van de 1.9–leerlingen in groep 2 en 8 tussen 1994/95 en 2007/08, naar herkomstland (effect sizes; referentiecategorie: autochtoon, mbo)

	Groep 2				Groep 8			
	1994	1998	2002	2007	1994	1998	2002	2007
ex Turkije	-0.96	-0.78	-0.66	-0.58	-0.63	-0.64	-0.36	-0.15
ex Marokko	-0.83	-0.71	-0.59	-0.45	-0.84	-0.32	-0.54	-0.38
ex Suriname/Antillen	-0.50	-0.51	-0.73	-0.41	-0.92	-0.77	-0.55	-0.46
ex Overig niet-westers	-0.62	-0.57	-0.66	-0.48	-0.41	-0.55	-0.29	0.19
0.3 Turkije	-1.08	-0.85	-0.82	-0.71	-0.75	-0.74	-0.64	-0.48
0.3 Marokko	-0.85	-0.70	-0.75	-0.68	-0.76	-0.65	-0.57	-0.48
0.3 Suriname/Antillen	-0.83	-0.67	-0.76	-0.64	-1.05	-0.84	-0.90	-0.87
0.3 Overig niet-westers	-0.51	-0.70	-0.67	-0.61	-0.60	-0.76	-0.46	-0.25
1.2 Turkije	-1.14	-0.87	-0.81	-0.67	-0.92	-0.70	-0.65	-0.40
1.2 Marokko	-1.00	-0.96	-0.72	-0.64	-1.03	-0.83	-0.75	-0.51
1.2 Suriname/Antillen	-0.75	-0.78	-0.87	-0.65	-1.22	-0.96	-0.96	-0.96
1.2 Overig niet-westers	-1.02	-0.83	-0.70	-0.63	-0.83	-0.44	-0.62	-0.31

In groep 2 is er wat de rekenvaardigheid betreft voor alle categorieën een lichte trend waarneembaar dat met het verstrijken van de jaren de achterstand afneemt, in het bijzonder voor de ex 1.9-leerlingen en voor alle groepen Turkse leerlingen in groep 2. Voor groep 8 laat de tabel wat meer diversiteit zien; vooral de gunstige ontwikkeling van de overige niet-westerse achterstandsleerlingen valt in positieve zin op.

Reductie taal– en rekenachterstand

In dit rapport is als insteek steeds gekozen voor een vergelijking van groepen van leerlingen die zijn gevormd op basis van binnen de gewichtenregeling gehanteerde criteria. Omdat deze criteria in de loop der jaren enkele malen zijn gewijzigd, hebben we voor een sterk gedifferentieerde indeling van tien categorieën gekozen. In het Onderwijsachterstandenbeleid zijn bepaalde doelstellingen geformuleerd in termen van een reductie van de achterstanden van doelgroepleerlingen ten opzichte van de niet-doelgroepleerlingen (zie MinOCW, 2010). Zo zou de taalachterstand op het eind van de basisschool in de periode 2002-2014 met 40% moeten zijn gereduceerd (waarbij wordt uitgegaan van de oude gewichtencriteria). En in de periode 2008-2011 zou de

taalachterstand met 20% moeten zijn gereduceerd (waarbij wordt uitgegaan van de nieuwe gewichtencriteria).
Ter afsluiting van deze paragraaf over het basisonderwijs presenteren we in Tabel 2.13 de resultaten van analyses die een indicatie geven van de reductie in achterstanden in groep 8. We volgen daarbij de strategie die eerder door Driessen (2009) werd toegepast, namelijk door het relatieve verschil tussen de effect sizes van de eerste en laatste meting te berekenen. We doen dat voor de volledigheid niet alleen voor de taalvaardigheid, maar ook voor de rekenvaardigheid. In de tabel maken we verschillende vergelijkingen: we zetten de ontwikkelingen van de totale doelgroep (dus autochtoon plus allochtoon) af tegen die van de niet-doelgroep, en vervolgens splitsen we de totale doelgroep nog uit naar de autochtone en de allochtone doelgroep. We gaan na hoe groot de reductie was tussen 1994 en 2007 en tussen 2002 en 2007. Dat doen we op basis van de oude gewichtencriteria (alle 1.25- en/of 1.9-leerlingen versus alle 1.0-leerlingen) en de nieuwe criteria (de 1.25 en/of 1.9 0.3- en 1.2-leerlingen versus de overige leerlingen).

Tabel 2.13 Verschillen tussen doelgroep- en niet-doelgroepleerlingen in taal- en rekenvaardigheid in groep 8 in 1994, 2002 en 2007 (effect sizes), en reductie achterstand tussen 1994 en 2007, respectievelijk 2002 en 2007 (in %)

Totale doelgroep vs. niet-doelgroep	1994	2002	2007	1994-2007	2002-2007
Taalvaardigheid					
- oude criteria	-0.77	-0.66	-0.62	19%	6%
- nieuwe criteria	-0.77	-0.71	-0.61	21%	14%
Rekenvaardigheid					
- oude criteria	-0.60	-0.59	-0.47	22%	20%
- nieuwe criteria	-0.62	-0.61	-0.47	24%	23%
Autochtone doelgroep vs. niet-doelgroep					
Taalvaardigheid					
- oude criteria	-0.53	-0.45	-0.52	2%	-16%
- nieuwe criteria	-0.50	-0.47	-0.53	-6%	-13%
Rekenvaardigheid					
- oude criteria	-0.47	-0.51	-0.42	11%	19%
- nieuwe criteria	-0.49	-0.55	-0.45	8%	18%
Allochtone doelgroep vs. niet-doelgroep					
Taalvaardigheid					
- oude criteria	-1.34	-1.03	-0.75	44%	27%
- nieuwe criteria	-1.24	-0.96	-0.67	46%	30%
Rekenvaardigheid					
- oude criteria	-0.88	-0.74	-0.55	38%	26%
- nieuwe criteria	-0.82	-0.67	-0.49	40%	27%

Als we eerst naar het bovenste paneel van de tabel kijken, waarin de totale doelgroep wordt vergeleken met de niet-doelgroep, dan blijkt dat er in de periode 1994-2007 een flinke reductie heeft plaatsgevonden van de achterstanden. Volgens zowel de oude als de nieuwe gewichtencriteria zijn de doelgroepen er ongeveer 20% op vooruitgegaan ten opzichte van de niet-doelgroep. Voor de rekenvaardigheid is de reductie nog wat groter dan voor de taalvaardigheid. In de periode 2002-2007 is de reductie voor de taalvaardigheid geringer, met name als daarbij wordt uitgegaan van de oude criteria. Met betrekking tot de rekenvaardigheid bedraagt in deze periode de reductie ruim 20%.

Wanneer we de totale doelgroep opsplitsen in de autochtone en de allochtone doelgroep, dan ontstaat echter een veel genuanceerder beeld. De allochtone doelgroep heeft een forse inhaalslag gemaakt: in de periode 1994-2007 van rond de 40%, en in de periode 2002-2007 van ruim 25%. Voor de autochtone doelgroep is de ontwikkeling echter lang niet zo rooskleurig. In de periode 1994-2007 is er qua taalvaardigheid sprake van stilstand of zelfs een lichte

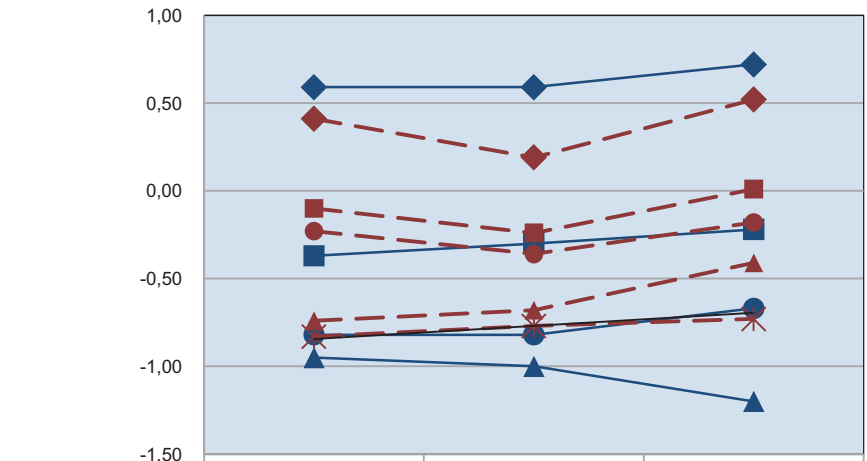
verslechtering; in de periode 2002-2007 gaat het om een verslechtering van rond de 15%. Met betrekking tot de rekenvaardigheid doet zich wel een positieve ontwikkeling voor: in de periode 1994-2007 zijn de achterstanden met rond de 10% gereduceerd, en in de periode 2002-2007 met bijna 20%. Terzijde zij opgemerkt dat het bij deze analyses niet zo bijster veel uitmaakt welke gewichtencriteria worden gehanteerd, de oude of de nieuwe. Met name bij de allochtone doelgroep is het verschil bijzonder gering.

2.4.2.2 Voortgezet onderwijs

Onderwijspositie leerjaar 3

In Figuur 2.5 presenteren we allereerst voor de OAB-categorieën de ontwikkeling qua onderwijstype, dit in termen van de positie op de leerjarenladder. In Tabel 2.14 volgt een uitsplitsing van de gegevens van de achterstandsleerlingen naar herkomstland.

Figuur 2.5 De ontwikkeling van de onderwijspositie in leerjaar 3 tussen 1999/2000 en 2007/08, naar OAB-categorie (effect sizes; referentiecategorie: autochtoon, mbo)



	1999	2003	2007
aut. 1.0 ho	0,59	0,59	0,72
aut. 1.25 ex	-0,37	-0,30	-0,22
aut. 1.25 0.3	-0,82	-0,82	-0,67
aut. 1.25 1.2	-0,95	-1,00	-1,20
all. 1.0 ho	0,41	0,19	0,52
all 1.0 mbo	-0,10	-0,24	0,01
all 1.9 ex	-0,23	-0,36	-0,18
all 1.9 0.3	-0,74	-0,68	-0,41
all 1.9 1.2	-0,83	-0,77	-0,73

Tabel 2.14 De ontwikkeling van de onderwijspositie van de 1.9-leerlingen in leerjaar 3 tussen 1999/2000 en 2007/08, naar herkomstland (effect sizes; referentiecategorie: autochtoon, mbo)

	1999	2003	2007
Turkije	-0.87	-0.77	-0.63
Marokko	-0.82	-0.80	-0.78
Suriname/Antillen	-0.71	-0.70	-0.21
Overig niet-westers	-0.52	-0.45	0.20

Wat de onderwijspositie betreft lijkt er vooral bij de allochtone categorieën sprake van vooruitgang. En meer in het bijzonder zijn vooral de Surinaamse/Antilliaanse en overige niet-westerse leerlingen er op vooruitgegaan.

Taalvaardigheid

In Tabel 2.15 volgen de gegevens over de taalvaardigheid per OAB-categorie en in Tabel 2.16 die per herkomstland. Zoals eerder al aangegeven, kunnen de taal- en wiskundegegevens uit 1999 en 2003 niet worden vergeleken met die uit 2007. Om die reden beperken we ons tot de gegevens uit 1999 en 2003.

Tabel 2.15 De taalvaardigheid in leerjaar 3 in 1999/2000 en 2003/04, naar OAB-categorie (effect sizes; referentiecategorie: autochtoon, mbo)

	1999	2003
aut. 1.0 ho	0.20	0.12
aut. 1.25 ex	0.00	0.02
aut. 1.25 0.3	-0.09	-0.05
aut. 1.25 1.2	-0.18	-0.04
all. 1.0 ho	-0.06	-0.12
all. 1.0 mbo	0.09	0.02
all. 1.9 ex	-0.07	-0.13
all. 1.9 0.3	0.09	-0.22
all. 1.9 1.2	-0.17	-0.22

Tabel 2.16 De taalvaardigheid van de 1.9-leerlingen in leerjaar 3 in 1999/2000 en 2003/04, naar herkomstland (effect sizes; referentiecategorie: autochtoon, mbo)

	1999	2003
Turkije	-0.08	-0.40
Marokko	-0.15	-0.05
Suriname/Antillen	-0.09	-0.17
Overig niet-westers	-0.06	-0.14

Uit Tabel 2.15 volgt allereerst dat er qua taalvaardigheid in 1999 en 2003 doorgaans maar kleine verschillen zijn tussen de onderscheiden categorieën. Daarnaast blijkt dat tussen 1999 en 2003 met name de positie van de zware

autochtone achterstandsleerlingen verbeterd is, terwijl die van de lichte allochtone achterstandsleerlingen juist verslechterd is.
Tabel 2.16 maakt duidelijk dat het vooral de Turkse achterstandsleerlingen zijn van wie de positie is verslechterd.

Wiskundevaardigheid
In Tabel 2.17 volgt de ontwikkeling van de wiskundevaardigheid naar OAB-categorie en in Tabel 2.18 die naar herkomstland.

Tabel 2.17 De wiskundevaardigheid in leerjaar 3 in 1999/2000 en 2003/04, naar OAB–categorie (effect sizes; referentiecategorie: autochtoon, mbo)

	1999	2003
aut. 1.0 ho	0.08	0.11
aut. 1.25 ex	-0.04	0.03
aut. 1.25 0.3	-0.08	0.01
aut. 1.25 1.2	-0.24	-0.18
all. 1.0 ho	-0.02	-0.13
all. 1.0 mbo	-0.13	-0.22
all. 1.9 ex	-0.26	-0.17
all. 1.9 0.3	-0.19	-0.15
all. 1.9 1.2	-0.20	-0.14

Tabel 2.18 De wiskundevaardigheid in leerjaar 3 in 1999/2000 en 2003/04, naar herkomstland (effect sizes; referentiecategorie: autochtoon, mbo)

	1999	2003
Turkije	-0.18	-0.01
Marokko	-0.19	-0.29
Suriname/Antillen	-0.24	-0.12
Overig niet-westers	-0.31	-0.14

Ook wat betreft de wiskundevaardigheid zijn er blijkens Tabel 2.17 in 1999 en 2003 in het algemeen maar kleine verschillen tussen de onderscheiden categorieën. En bovendien is er tussen beide jaren slechts sprake van geringe ontwikkelingen.

En uit Tabel 2.18 volgt dat met name de Marokkaanse en Turkse leerlingen laag scoren.

2.5 Samenvatting, conclusie en discussie

2.5.1 Samenvatting

2.5.1.1 Onderzoeksopzet

In deze deelstudie is nagegaan hoe de onderwijsachterstand van OAB-doelgroepen in de periode 1994/95 – 2007/08 zich heeft ontwikkeld. Daartoe zijn de taal- en reken/wiskundevaardigheid van de doelgroepen in de groepen 2 en 8 basisonderwijs en in leerjaar 3 voortgezet onderwijs afgezet tegen die van een referentiegroep, te weten autochtone kinderen van middelbaar opgeleide ouders. Eenzelfde vergelijking is gemaakt wat betreft het onderwijsniveau in leerjaar 3 voortgezet onderwijs.
Voor deze analyses is gebruik gemaakt van gegevens uit de cohortstudies PRIMA en COOL. De data die zijn verzameld bij de metingen van PRIMA-PO en (het vervolg daarop) de eerste meting van COOL-PO laten zich goed vergelijken. Ten gevolge van verschillen in opzet is dat voor de twee metingen van PRIMA-VO en de eerste meting van COOL-VO veel minder het geval. Wat betreft de vergelijkbaarheid van de taal- en wiskundevaardigheid is er tussen PRIMA-VO en COOL-VO sprake van een breuk. Hierdoor is het niet mogelijk deze twee vaardigheden over de drie metingen heen te vergelijken.

2.5.1.2 Stand van zaken in 2007/08

In groep 2 van het basisonderwijs zijn er tussen de onderscheiden leerlingcategorieën grote verschillen in *taalvaardigheid*. Opvallend is dat allochtone leerlingen steeds veel lager presteren dan autochtone leerlingen met 'formeel' dezelfde achtergronden. In groep 8 zijn deze verschillen minder groot, wat niet wegneemt dat er tussen OAB-doelgroepleerlingen en niet-doelgroepleerlingen nog een flinke afstand bestaat. De verschillen tussen leerlingen uit verschillende herkomstlanden zijn vrij gering.
De verschillen in *rekenvaardigheid* zijn veel minder groot dan die in taalvaardigheid, en ook geldt dat de verschillen in groep 8 weer wat minder groot zijn dan die in groep 2. De invloed van de thuissituatie is voor rekenen van minder belang dan die voor taal en neemt gedurende de

onderwijsloopbaan ook nog eens flink af. Maar ook hier geldt dat er in groep 8 nog steeds flinke verschillen bestaan tussen OAB-doelgroepleerlingen en niet-doelgroepleerlingen.
In leerjaar 3 van het voortgezet zijn er zeer grote verschillen in gerealiseerd *onderwijsniveau.* De doelgroepleerlingen nemen een flink lagere positie in dan de niet-doelgroepleerlingen. Verschillen tussen autochtone en allochtone leerlingen met dezelfde milieu-achtergronden zijn echter betrekkelijk gering.
In leerjaar 3 zijn er ook grote verschillen qua *taal– en wiskundevaardigheid*. In het algemeen presteren de allochtone doelgroepleerlingen wat lager dan de autochtone doelgroepleerlingen. De verschillen voor wiskunde zijn overigens in deze fase groter dan voor taal.

2.5.1.3 Ontwikkelingen 1994/95 – 2007/08
In groep 2 van het basisonderwijs is er in de onderzochte periode met betrekking tot de taalvaardigheid nauwelijks sprake van ontwikkelingen. In groep 8 zijn echter met name de allochtone doelgroepleerlingen flink vooruit gegaan.
Wat betreft de rekenvaardigheid hebben de allochtone doelgroepleerlingen hun positie behoorlijk weten te verbeteren. Dit geldt zowel voor groep 2 als voor groep 8.
Met betrekking tot de situatie in groep 8 van het basisonderwijs is nagegaan in hoeverre de doelgroepen hun achterstanden hebben weten te reduceren. Dat is gebeurd voor de periode 1994-2007 en de periode 2002-2007. De resultaten van de analyses wijzen op grote verschillen tussen de autochtone en de allochtone doelgroep. De allochtone doelgroep heeft een forse inhaalslag gemaakt: in de periode 1994-2007 van rond de 40%, en in de periode 2002-2007 van ruim 25%. Voor de autochtone doelgroep is de ontwikkeling echter lang niet zo rooskleurig. In de periode 1994-2007 is er qua taalvaardigheid sprake van stilstand of zelfs een lichte achteruitgang; in de periode 2002-2007 gaat het om een verslechtering van rond de 15%. Met betrekking tot de rekenvaardigheid doet zich wel een positieve ontwikkeling voor: in de periode 1994-2007 zijn de achterstanden met rond de 10% gereduceerd, en in de periode 2002-2007 met bijna 20%.
Als het gaat om het *onderwijsniveau* in leerjaar 3 van het voortgezet onderwijs is er vooral onder de allochtone leerlingen sprake van vooruitgang.

Met betrekking tot de *taal– en wiskundevaardigheid* in leerjaar 3 van het voortgezet onderwijs doen er zich tussen 1999/2000 en 2003/04 nagenoeg geen ontwikkelingen voor.

2.5.2 Conclusie en discussie

De centrale vraag van deze studie luidde of de achterstand van de verschillende doelgroepen van het onderwijsachterstandenbeleid de afgelopen jaren is afgenomen. De hierboven samengevatte bevindingen liggen in lijn met die uit eerdere overzichtsstudies (zie bv. Driessen, 2009; Mulder e.a., 2005; Van Langen & Suhre, 2001; Vogels & Bronnemans-Helmers, 2003). Grofweg samengevat luidt de conclusie dat er bij de start in het basisonderwijs grote verschillen zijn tussen achterstandsleerlingen en niet-achterstandsleerlingen. Met name allochtone doelgroepleerlingen hebben een flinke taalachterstand. In het laatste jaar van het basisonderwijs is die achterstand weliswaar verminderd, maar nog steeds fors. Vooral allochtone doelgroepleerlingen hebben de afgelopen jaren hun achterstand in het basisonderwijs voor een belangrijk deel weten te verkleinen, in de periode 1994-2007 zelfs met rond de 40%. Een deel van de autochtone doelgroepleerlingen heeft eveneens een positieve ontwikkeling doorgemaakt, maar deze is toch veel minder groot dan die van de allochtone doelgroepleerlingen. Bovendien is er voor de periode 2002-2007 met betrekking tot de taalvaardigheid sprake van een relatieve verslechtering van de positie van de autochtone doelgroepleerlingen. Ook in het voortgezet onderwijs is de onderwijspositie van de autochtone doelgroepleerlingen verslechterd.

Voor de positieve ontwikkeling van de allochtone doelgroepleerlingen kunnen uiteenlopende verklaringen worden gegeven. Op de eerste plaats is er louter het gegeven dat de verblijfduur van allochtonen is toegenomen en er ook veel van hen al tot de tweede of zelfs derde generatie behoren (vgl. Mulder, 1996). Dergelijke demografische ontwikkelingen hebben in het algemeen een positief effect op de Nederlandse taalvaardigheid van niet alleen de ouders maar ook van hun kinderen (Driessen, 2010, 2011). Allochtone ouders van latere generaties zijn ook hoger opgeleid dan die van de eerste generatie. Verwacht mag worden dat zij over meer ‘cultureel kapitaal’ beschikken, beter de weg weten te vinden in de Nederlandse samenleving en hun kinderen ook beter kunnen ondersteunen met betrekking tot schoolse zaken. Daarnaast is het ongetwijfeld ook van belang dat leerkrachten in de loop van de jaren meer

ervaring hebben gekregen in hun omgang met allochtone kinderen en het onderwijs steeds beter op hun specifieke situatie hebben leren afstemmen. Mogelijk hebben zij ook hun verwachtingen en ambities in de loop van de tijd naar boven bijgesteld (Van den Bergh e.a., 2010). Alhoewel overtuigend bewijs hiervoor vooralsnog ontbreekt, zou een deel van de verklaring ook gelegen kunnen zijn in het effect van het onderwijsachterstandbeleid en dan in het bijzonder van de vele programma's die worden uitgevoerd in het kader van de Voor- en Vroegschoolse Educatie (VVE) en die vooral gericht zijn op de verbetering van de taalvaardigheid van allochtone doelgroepleerlingen.
Het is overigens de vraag of de positieve ontwikkeling van allochtone doelgroepleerlingen zich de komende tijd door zal zetten. De wijziging in de gewichtenregeling die enige jaren geleden heeft plaatsgevonden hield in dat de etnische component is komen te vervallen. Voor de toekenning van extra faciliteiten aan scholen betekent dit dat het niet meer uitmaakt of de ouders van het kind in het buitenland zijn geboren; het enige criterium is nu de hoogte van het ouderlijk opleidingsniveau. Onduidelijk is vooralsnog welke consequenties dit heeft voor de extra aandacht voor de betreffende kinderen in de klas. In een andere BOPO-studie wordt nagegaan welke herverdeeleffecten er zijn opgetreden ten gevolge van de wijzigingen in de gewichtenregeling (Claassen, i.v.).

Vergeleken met die van de allochtone doelgroepleerlingen verloopt de ontwikkeling van de autochtone doelgroepleerlingen minder voorspoedig of is er zelfs sprake van een relatieve achteruitgang. Op deze negatieve ontwikkeling is in het verleden al vaker gewezen. Ook is regelmatig de vraag gesteld of al die beleidsaandacht voor de allochtone doelgroepleerlingen niet ten koste is gegaan van de aandacht voor de autochtone doelgroepleerlingen. Voor Vogels en Bronneman-Helmers (2003) reden om deze groep als 'een vergeten groep' te labelen. Zij voeren enkele mogelijke verklaringen op voor de relatieve achteruitgang van de schoolprestaties van autochtone achterstandsleerlingen, waarbij overigens moet worden aangetekend dat deze zich empirisch nauwelijks laten toetsen. Een verklaring zou kunnen zijn dat er onder allochtonen vanwege hun migratiegeschiedenis nog sprake was, respectievelijk is van een potentieel en onderbenutting, terwijl het reservoir aan onbenut talent onder autochtonen waarschijnlijk veel geringer is. Het vroegere onbenutte talent onder de autochtone arbeiderskinderen heeft volgens deze verklaring geleidelijk aan plaatsgemaakt voor uitgeput talent: qua erfelijke

aanleg is het plafond bereikt. Een andere verklaring wijst op het geringe ambitieniveau van ouders in plattelandsgebieden (vgl. Van Ruijven, 2003), dat in schril contrast zou staan met het hoge (maar vaak onrealistische) ambitieniveau van allochtone ouders (Smit, Driessen & Doesborgh, 2005). Als dit inderdaad het geval zou zijn, dan kan het overigens hooguit een deel van de verklaring vormen. Uit onderzoek blijkt namelijk dat er tussen plattelandsgebieden grote verschillen bestaan in het prestatieniveau van de leerlingen: de noordelijke gebieden (Friesland, Groningen, Drenthe) doen het steevast slecht, terwijl Limburg goed scoort (Driessen & Withagen, 1999). Een verklaring die in dit verband regelmatig wordt opgevoerd, wijst op een mogelijk negatief effect van het spreken van een dialect of streektaal. Hiervoor bestaat echter slechts zwakke empirische evidentie, en ook hierbij geldt dat er zich grote regionale verschillen voordoen (voor Limburg is er zelfs sprake van een positief effect; vgl. Driessen, 2006). Volgens de Inspectie van het Onderwijs (2001) hebben verschillen tussen noordelijke en zuidelijke (plattelands)gebieden vooral te maken met verschillen in de kwaliteit van het leerstofaanbod op de - veelal kleine - noordelijke plattelandsscholen (zie ook Van Ruijven, 2003). Daaraan gekoppeld is het op dergelijke kleine scholen ontbreken van mogelijkheden om voldoende extra zorg te bieden aan leerlingen die dat nodig hebben, zoals remedial teaching en interne begeleiding). En juist op die scholen zou meer zorg nodig zijn, omdat er minder verwezen wordt naar het speciaal onderwijs. Een laatste verklaring die door Vogels en Bronneman-Helmers (2003) wordt geopperd betreft het geringe extra budget dat de plattelandsscholen ontvangen vanuit de gewichtenregeling. Het argument is dat niet alleen het extra bedrag per autochtone achterstandsleerling veel geringer is dan dat per allochtone achterstandsleerling, maar dat door de ingebouwde drempel van 9% veel scholen niets of nauwelijks iets ontvangen. Dergelijke kritiek heeft er mede toe geleid dat er in 2006 een aanscherping van de gewichtenregeling heeft plaatsgevonden, waarbij de etnische component is geschrapt en de drempel is verlaagd van 9 naar 6%. Daarnaast is met ingang van schooljaar 2009/10 de Regeling impulsgebieden in werking getreden op grond waarvan scholen in achterstandsgebieden extra middelen krijgen toegekend voor alle gewichtsleerlingen. De verwachting is dat er een verschuiving van financiële middelen op zal treden van de grote steden (met vooral allochtone achterstandsleerlingen) naar het platteland (met vooral autochtone achterstandsleerlingen). Het is echter nog afwachten welke gevolgen deze

beleidswijzingen hebben voor de onderwijspraktijk en of het mogelijk is het tij te keren voor autochtone achterstandsleerlingen zonder dat dit ten koste gaat van de positie van de allochtone achterstandsleerlingen. Bij dit alles is het overigens maar de vraag in hoeverre extra financiële middelen zich ook rechtsreeks vertalen in betere prestaties (vgl. Mulder, 1996).

3 De cognitieve ontwikkeling van doelgroepleerlingen gedurende het basisonderwijs

Jaap Roeleveld

3.1 Inleiding

In het vorige hoofdstuk is een beschrijving gegeven van de positie van de diverse groepen (achterstands)leerlingen bij het begin (groep 2) en aan het eind van het basisonderwijs (groep 8) en in leerjaar 3 van het voortgezet onderwijs. Door gegevens van verschillende jaren te gebruiken konden ook de ontwikkelingen door de tijd aan de orde komen.
Bij deze *cross-sectionele* benadering, waarbij verschillende groepen leerlingen in de tijd met elkaar vergeleken worden, is het mogelijk dat de opeenvolgende groepen leerlingen (de groepen 2 respectievelijk 8 uit 1994, 1998, 2002 en 2007) onderling verschillen in samenstelling, bijvoorbeeld wat betreft het aandeel allochtonen, maar ook wat betreft mogelijk andere relevante, maar niet in het onderzoek gemeten, kenmerken. Ook kunnen er verschillen zijn opgetreden in de mate waarin er in de eerdere loopbaan sprake is geweest van uitval, bijvoorbeeld door verwijzing naar het speciaal (basis)onderwijs. Daarom zullen we in deze bijdrage gebruik maken van een *longitudinale* benadering, waarbij in principe dezelfde leerlingen door te tijd heen gevolgd wordt. Hierdoor wordt het mogelijk om bij deze leerlingen de leerwinst gedurende het basisonderwijs te bepalen. In de literatuur wordt zo'n leerwinstbenadering van het onderwijs geprefereerd boven benaderingen, waarbij alleen gecontroleerd

wordt voor leerlingkenmerken (zie OECD, 2008; Peschar, 2007; Roeleveld, Van der Veen & Ledoux, 2009).
De cross-sectionele analyses uit het vorige hoofdstuk hebben laten zien hoe de positie van de verschillende groepen (achterstands)leerlingen is tussen verschillende generaties. De longitudinale analyses uit dit hoofdstuk zijn gericht op de ontwikkeling binnen één bepaalde generatie leerlingen en laten zien in hoeverre aanvangsachterstanden gedurende de loopbaan door het basisonderwijs af- of toenemen.
Bij eerdere longitudinale analyses van gegevens uit het PRIMA-cohort (Mulder e.a., 2005) werd opgemerkt dat deze analyses op een niet-representatieve groep leerlingen betrekking hadden, namelijk alleen op leerlingen die onvertraagd het basisonderwijs doorlopen. Leerlingen die zijn blijven zitten of zijn verwezen naar het speciaal (basis)onderwijs konden niet in de analyses worden opgenomen, doordat er van hen geen gegevens over de gehele loopbaan beschikbaar zijn. Daarmee bestaat het gevaar dat er een te positief beeld van de ontwikkeling van leerlingen gedurende het basisonderwijs wordt verkregen, omdat het gebaseerd is op een selectieve groep: alleen de niet-vertraagde leerlingen. In de navolgende analyses zullen we technieken gebruiken die het mogelijk maken de groei (leerwinst) van leerlingen te schatten, waarbij wél zoveel mogelijk gebruik kan worden gemaakt van gegevens van alle leerlingen, ook van de uitvallers.
In dit hoofdstuk willen we nagaan hoe de longitudinale ontwikkeling van de leerlingen gedurende het basisonderwijs verloopt. Daarmee kan duidelijk worden welke verschillen in groei van leerprestaties er bestaan tussen de verschillende groepen (achterstands)leerlingen en of er gedurende de loopbaan sprake is van het inlopen van achterstanden. De analyses zullen worden uitgevoerd met gegevens van drie opeenvolgende cohorten leerlingen, zodat ook kan worden nagegaan of de groei van leerprestaties verandert voor deze opeenvolgende ‘generaties’ van basisschoolleerlingen.

Groeimodellen
De taal- en rekenprestaties worden geanalyseerd door middel van latente groeicurve analyse. In zo’n analyse wordt een, liefst uitgebreide, reeks metingen door de tijd samengevat in een veel kleiner aantal parameters, die samen de ontwikkeling door de tijd beschrijven (zie voor voorbeelden Stoel, 2003; Guldemond & Bosker, 2006; Roeleveld, Van der Veen & Ledoux, 2009). Het voordeel van deze analysetechniek is dat het ontwikkelingspatroon van

elke leerling kan worden weergegeven in de vorm van een unieke groeicurve. Er wordt rekening mee gehouden dat leerlingen een verschillend prestatieniveau hebben aan de start en ook dat individuele leerlingen verschillen in de mate waarin hun prestaties toenemen. In de analyse worden het niveau van een variabele tijdens de eerste meting (het intercept van de lijn) en de mate van groei in een variabele (de helling/richtingscoëfficiënt van de lijn) als latente variabelen beschouwd. De ontwikkeling van de prestaties van de leerlingen, zoals gemeten in de opeenvolgende leerjaren, zijn de direct geobserveerde variabelen, waarmee schattingen van deze latente variabelen kunnen worden verkregen. Bij drie meetmomenten kan alleen lineaire groei, dus gelijkmatig over de hele onderzochte periode, worden gemodelleerd. Deze groei wordt volledig beschreven door een startniveau (intercept) en de daarna volgende groei over de jaren (veelal aangeduid met de Engelse term 'slope'). Als er nog meer meetmomenten zijn kunnen ook complexere modellen, waarbij de groei varieert over de jaren, worden geschat (zie bijvoorbeeld Guldemond & Bosker, 2006). In zo'n geval kan dus de groeicurve in het begin sterk stijgen, terwijl de groei in latere jaren afvlakt. Of juist andersom: eerst zwakke groei, gevolgd door een versnelling in latere jaren.

Dit type analyse kan worden uitgevoerd met behulp van programma's voor het analyseren van structurele modellen. Voor de navolgende analyses is het programma Mplus (Muthén & Muthén, 2004) gebruikt.

Het voordeel van het modelleren van het intercept en de helling als latente variabelen is dat deze variabelen in relatie gebracht kunnen worden met elkaar en met andere latente of controlevariabelen. Een negatieve samenhang tussen het startniveau en de mate van groei betekent bijvoorbeeld dat een lager startniveau overeenkomt met een hoger groeiniveau en andersom. Een ander voorbeeld is een herkomst-effect op de latente variabelen. We kunnen dan bijvoorbeeld vaststellen of en hoeveel Turkse leerlingen meer of minder vooruit zijn gegaan in taalprestaties van groep 4 naar 8 dan Marokkaanse leerlingen.

In Nederland zijn in het onderzoek naar de positie van achterstandsleerlingen groeimodellen eerder gebruikt door Guldemond en Bosker (2006). Ook bij hen was echter sprake van de analyse van alleen leerlingen die onvertraagd in groep 8 terecht zijn gekomen. Zoals reeds gezegd zullen wij hier de modellering van de groei van leerlingen baseren op gegevens van *alle* leerlingen.

Opzet van dit hoofdstuk
In de volgende paragraaf gaan we nader in op de data, waarmee de analyses zullen worden uitgevoerd, en op het probleem van uitval (mortaliteit) en ontbrekende gegevens. Ook presenteren we een aantal beschrijvende gegevens. Daarna worden achtereenvolgens de uitkomsten voor de ontwikkeling van taalprestaties (paragraaf 3.3) en van rekenprestaties (paragraaf 3.4) besproken. We eindigen met een aantal conclusies en een discussie van de uitkomsten.

3.2 Data

Bij de analyses in dit hoofdstuk maken we steeds gebruik van de representatieve steekproeven uit de verschillende relevante PRIMA-metingen. We beperken ons hier tot de representatieve steekproef, omdat er in eerste instantie een groeimodel geschat gaat worden, dat geldig is voor de 'gemiddelde' leerling in het Nederlandse basisonderwijs. Voor dat doel is de representatieve steekproef het meest geëigend. Pas na de schatting van zo'n groeimodel komen in de analyses de verschillen tussen de diverse onderscheiden groepen (achterstands)leerlingen aan de orde.
De eerste PRIMA-meting vond plaats in schooljaar 1994/1995 en de laatste in schooljaar 2004/2005, steeds met dataverzameling in de groepen 2, 4, 6 en 8. Daardoor zijn complete loopbanen (van groep 2 tot en met groep 8) beschikbaar voor een drietal opeenvolgende cohorten (zie ook Schema 1.1 in hoofdstuk 1):

- cohort A: groep 2 in 1994/1995 → groep 8 in 2000/2001
- cohort B: groep 2 in 1996/1997 → groep 8 in 2002/2003
- cohort C: groep 2 in 1998/1999 → groep 8 in 2004/2005

Door het verschil in frequentie (PRIMA om de 2 jaar en COOL om de 3 jaar) zijn er geen gegevens van een cohort van recentere datum beschikbaar. Pas na de volgende meting van COOL (schooljaar 2010/2011) kan weer de ontwikkeling van een heel cohort van groep 2, via groep 5 naar groep 8 worden geanalyseerd.

Vaardigheidsschalen
Om de cognitieve ontwikkeling (groei) van de leerlingen weer te kunnen geven, dienen er in elk van de groepen toetsen te zijn afgenomen die op één en dezelfde (vaardigheids)schaal geplaatst kunnen worden. Pas dan kan goed

duidelijk worden hoeveel een leerling tussen twee meetmomenten vooruit is gegaan in een bepaalde vaardigheid.
Aan deze eis voldoet de PRIMA-taaltoets voor de groepen 4, 6 en 8 in elk van de drie cohorten. Bij rekenen is vanaf PRIMA-3 (schooljaar 1998/1999) de PRIMA-rekentoets vervangen door rekentoetsen uit het Cito-leerlingvolgsysteem. Via een 'omrekenformule' zijn scores op beide type rekentoetsen wel op een gelijke schaal uit te drukken (Kamphuis e.a., 1998)[9].
De kleutertoetsen (voor rekenen: Ordenen; voor taal: eerst de Begrippentoets, later vervangen door Taal voor Kleuters) zijn echter niet goed op dezelfde schaal te plaatsen. En daarmee kan dan ook niet goed de ontwikkeling (groei) bepaald worden tussen de groepen 2 en 4. Daarom zal in de analyses de groei van de leerlingen bepaald worden aan de hand van scores op toetsen in de groepen 4, 6 en 8 en worden de scores op de kleutertoetsen als covariaat in de analyse opgenomen[10]. We beschrijven dus de groei tussen groep 4 en groep 8, waarbij er rekening wordt gehouden met verschillen in aanvangsscores op de kleutertoetsen.
Doordat we gegevens van drie meetmomenten hebben kan in de groeimodellen alleen lineaire groei worden onderzocht; om groeiversnellingen te kunnen onderzoeken zijn ten minste gegevens van vier meetmomenten nodig. Dat hoeft hier geen probleem te zijn, omdat het er ons vooral om gaat of achterstandgroepen er in slagen hun aanvankelijke achterstand te verminderen door snellere groei. Een lineair groeimodel is voldoende om die vraag te kunnen beantwoorden.
Uitgangspunt van de analyses worden daarmee de leerlingen in groep 4 van de genoemde drie cohorten. Van deze leerlingen worden (voor zover bekend) hun scores op de kleutertoetsen als covariaat meegenomen en worden de scores op toetsen in de groepen 4, 6 en 8 gebruikt om de groeimodellen op te ontwerpen.

---

[9] Maar zie ook verderop in dit hoofdstuk de analyses van de rekenontwikkeling.
[10] Vanwege de veranderingen bij de kleutertoetsen doen we dit in de vorm van (in de referentiesteekproeven gestandaardiseerde) Z-scores.

Uitval

Zoals alle cohortonderzoek kent ook PRIMA een (aanzienlijke) mortaliteit. In eerder onderzoek naar de uitval van leerlingen uit PRIMA (Roeleveld & Portengen, 1998; Roeleveld & Vierke, 2003) wordt onderscheid gemaakt tussen twee vormen van uitval:

- de uitval van hele groepen leerlingen, doordat hun school niet langer wil of kan deelnemen aan het onderzoek (*schooluitval*); en
- uitval van leerlingen op scholen die wel blijven deelnemen aan PRIMA (*individuele uitval*); daarbij kan het gaan om leerlingen die verhuizen, maar ook om zaken als zittenblijven op dezelfde school en om verwijzing naar het speciaal (basis)onderwijs.

Uit de genoemde onderzoeken komt naar voren dat de schooluitval niet of nauwelijks selectief is, maar dat het bij individuele uitval gaat om een groep leerlingen met gemiddeld lagere prestaties. Deze individuele uitval is dus wel selectief.

Voor dit onderzoek hebben we een nieuwe uitvalsanalyse uitgevoerd. Daarvoor zijn een aantal scores in groep 4 van beide typen van uitgevallen leerlingen vergeleken met die van niet-uitgevallen leerlingen. Dat zijn in de eerste plaats de cognitieve prestaties: de vaardigheidsscores voor taal en rekenen. Daarnaast kijken we naar een drietal schalen, die gebaseerd zijn op items uit het zogenaamde Leerlingprofiel. Hiermee geven leerkrachten een oordeel over hun leerlingen wat betreft hun (cognitief) zelfvertrouwen, werkhouding en thuisklimaat.

In Tabel 3.1 geven we in de eerste kolom, als voorbeeld, de gemiddelden in cohort A. Vervolgens worden in de volgende kolommen voor alle drie de cohorten de verschillen tussen uitgevallen leerlingen en ‘blijvers’ uitgedrukt in effectgroottes (verschil in gemiddelde, gedeeld door de overall standaarddeviatie).

Tabel 3.1 Verschillen in toets- en schaalscores naar soort uitval; in drie cohorten

		gem.	effectgrootte			gemiddelde
	uitval	coh A	coh A	coh B	coh C	effectgrootte
vaardigheidsscore taal	niet	1044. 2				
	individueel	1030.0	-0.39	-0.36	-0.48	-0.41
	school	1042.3	-0.05	-0.02	-0.03	-0.03
vaardigheidsscore rekenen	niet	69.3				
	individueel	58.3	-0.49	-0.56	-0.62	-0.56
	school	67.9	-0.06	-0.15	-0.09	-0.10
schaal ll-profiel: zelfvertrouwen	niet	3.54				
	individueel	3.34	-0.28	-0.33	-0.18	-0.26
	school	3.51	-0.05	-0.08	0.06	-0.03
schaal ll-profiel: werkhouding	niet	3.46				
	individueel	3.12	-0.41	-0.46	-0.44	-0.44
	school	3.36	-0.12	-0.15	-0.04	-0.10
schaal ll-profiel: thuisklimaat	niet	3.77				
	individueel	3.52	-0.37	-0.46	-0.31	-0.38
	school	3.71	-0.09	-0.11	0.01	-0.06

Uit de tabel blijkt overduidelijk dat individueel uitgevallen leerlingen een zwakkere groep vormen, die vaak een halve standaarddeviatie lager scoort dan de niet-uitgevallen leerlingen. De schooluitval wijkt veel minder af: doorgaans 0.05 tot 0.10 sd lager dan de niet uitgevallen leerlingen. In geen enkel geval bereikt het verschil de 0.20 sd ('small effect' bij Cohen, 1988) en incidenteel is het verschil zelfs positief.

Wanneer we de groei van leerlingen door het basisonderwijs alleen baseren op gegevens van de onvertraagde, niet uitgevallen leerlingen, zullen we daarom een vertekend beeld krijgen: de individueel uitgevallen, minder presterende leerlingen ontbreken daarbij immers. Voor een volledig beeld zouden we het liefst een uitspraak doen over hoe de gemiddelde groep 4-leerling zich tot en met groep 8 ontwikkeld heeft, dus *inclusief* de leerlingen die uitvallen. Om dit toch zo goed mogelijk te kunnen benaderen, wordt uitgegaan van de representatieve groep 4 leerlingen aan de start en worden vervolgens de scores van de kinderen waarvan we in latere metingen geen gegevens hebben, geschat. Daarbij wordt gebruik gemaakt van de (achtergrond-)kenmerken en scores van

de latere uitvallers. De missende gegevens zijn geschat met behulp van 'Full-Information Maximum Likelihood estimation' (FIML)[11].
In Tabel 3.2 geven we een overzicht van de aantallen leerlingen uit groep 4 van de representatieve steekproef, die wel of niet zijn uitgevallen in de loopbaan naar groep 8.

Tabel 3.2 Aantallen leerlingen naar soort uitval; in drie cohorten

	Cohort A		Cohort B		Cohort C	
	N	%	N	%	N	%
geen uitval	3586	35.4	3766	33.6	4126	40.1
individuele uitval	1595	15.7	2248	20.1	1484	14.4
schooluitval	4953	48.9	5184	46.3	4667	45.4
totaal	10134	100.0	11198	100.0	10277	100.0

Ruim een derde van de leerlingen uit groep 4 kan helemaal gevolgd worden tot aan groep 8. Bijna de helft van de leerlingen valt uit doordat de school niet langer deelneemt aan het onderzoek. De individuele uitval betreft 15 tot 20 procent van de leerlingen.
Om de aantallen ontbrekende waarden in de analyses niet al te groot te laten worden is besloten om de schooluitval buiten beschouwing te laten. Leerlingen die individueel zijn uitgevallen blijven wel in de analyse. In elk van de drie cohorten betekent dit dus dat *de school* waarop de leerlingen in groep 4 zitten, mee blijft doen met PRIMA tot de onvertraagde leerlingen in groep 8 zitten. Op die scholen worden vervolgens alle leerlingen bij de analyses betrokken, ook als ze later uitvallen of als er geen scores op de kleutertoetsen van groep 2 bekend zijn.
De PRIMA-cohortbestanden omvatten naast een representatieve steekproef ook een aanvullende steekproef met extra achterstandsleerlingen. Zoals gezegd geven we er de voorkeur aan om alleen gebruik te maken van de representatieve steekproeven, om een algemeen groeimodel voor de 'gemiddelde' Nederlandse leerling te schatten. We moeten dan nog wel nagaan of de aantallen voldoende zijn, wanneer we de leerlingen nader onderverdelen naar

[11] FIML is gebaseerd op de aanname dat missende waarden 'missing at random' zijn, wat wil zeggen dat missende waarden voorspeld kunnen worden uit de beschikbare gegevens. Het helemaal verwijderen van leerlingen met missende waarden (listwise deletion') is gebaseerd op de striktere aanname dat missende gegevens 'completely at random', dus volledig willekeurig zijn. Aan deze aanname is hier zeker niet voldaan.

het type (achterstands)leerling. In de volgende tabel geven we de aantallen leerlingen per cohort, waarover de analyses kunnen worden uitgevoerd.

Tabel 3.3 Aantallen leerlingen naar type leerling en soort uitval; in drie cohorten

	Uitval cohort A			Uitval cohort B			Uitval cohort C		
	niet	indivi -dueel	totaal	niet	indivi -dueel	totaal	niet	indivi -dueel	totaal
onbekend	102	130	232	113	224	337	42	66	108
auto 1.0, ho	631	211	842	797	302	1099	996	275	1271
auto 1.0, mbo	580	217	797	659	300	959	757	194	951
auto 1.25-ex	722	237	959	764	311	1075	910	221	1131
auto 1.25 (nu 0.3)	862	360	1222	733	442	1175	683	288	971
auto 1.25-zwaar (nu 1.2)	158	109	267	114	122	236	95	70	165
allo 1.0, ho	61	30	91	71	49	120	70	49	119
allo 1.0, mbo	58	36	94	62	63	125	74	36	110
allo 1.9-ex	80	41	121	95	70	165	106	70	176
allo 1.9 (nu 0.3)	83	62	145	97	119	216	107	55	162
allo 1.9-zwaar (nu 1.2)	249	162	411	261	246	507	286	160	446
totaal	3586	1595	5181	3766	2248	6014	4126	1484	5610
	69.2%	30.8%	100%	62.6%	37.4%	100%	73.5%	26.5%	100%

De totale aantallen per type leerling komen bijna overal (ruim) boven de 100. Alleen in het eerste cohort A (start groep 4 in 1996) zijn de beide groepen allochtone leerlingen zonder OAB-gewicht iets minder dan 100 in aantal (dat is inclusief de individueel uitgevallen leerlingen, die als gezegd wel in de analyses meedoen). Mede omdat we drie cohorten tegelijk analyseren zijn de aantallen groot genoeg om alleen de referentiesteekproeven te gebruiken.

3.3 Ontwikkeling van taalprestaties

We beginnen de analyses met de ontwikkeling van een groeimodel. Hierbij starten we met een model waarin veel variatie kan zitten tussen de opeenvolgende cohorten. In een aantal volgende stappen wordt steeds nagegaan of een spaarzamer model, waarin er steeds meer gelijkheid tussen de

cohorten wordt verondersteld, nog steeds een goede *fit* bij de data laat zien. Op basis daarvan wordt dan een basis-groeimodel vastgesteld.
Dit model vormt dan het uitgangspunt voor de volgende analyses, waarin nagegaan wordt welke verschillen er tussen de onderscheiden categorieën (achterstands)leerlingen bestaan in de groei van de taalprestaties tijdens het basisonderwijs.

3.3.1 Modelontwikkeling

We beschikken over gegevens van drie opeenvolgende cohorten, met in elk cohort metingen van de taalvaardigheid in de groepen 4, 6 en 8 en met een meting van aanvangsprestaties voor taal in groep 2. De ontwikkeling van de taalvaardigheid van groep 4 naar groep 8 modelleren we met een (lineair) groeimodel met een beginniveau in groep 4 (*Intercept*) en een groei over de jaren (*Slope*).
In een 0-model mogen deze Intercepten en Slopes, én de correlatie tussen die beide, variëren tussen de drie cohorten. Vervolgens is nagegaan welke gelijkheidsrestricties opgelegd konden worden:

- in alle drie cohorten kan de correlatie tussen Intercept en Slope op 0 gesteld worden: een hoger startniveau (Intercept) hangt dus niet samen met meer of minder groei (Slope);
- het startniveau van cohort B en C verschilt niet van elkaar; dat van cohort A is significant lager;
- de groei in cohort B en C verschilt ook niet van elkaar; de groei in cohort A is significant hoger.

Aan het model met de bovenstaande restricties (model 1) zijn vervolgens de gestandaardiseerde scores van de taaltoets uit groep 2 toegevoegd. Deze Z-scores kunnen (per cohort verschillende) effecten hebben op het beginniveau in groep 4 (Intercept) en op de verdere groei (Slope). Vervolgens zijn weer gelijkheidsrestricties opgelegd. In alle drie de cohorten kan het effect op de Slope op 0 gesteld worden. En het effect op het Intercept kan in alle drie cohorten gelijkgesteld worden (maar is wel ongelijk aan 0). Dit betekent dus dat de prestaties in groep 2 alleen een relatie hebben met het niveau in groep 4, maar niet met de groei in scores over de jaren.
Daarmee is het basis-groeimodel (model 2) voor taal vastgesteld en kan vervolgens nagegaan worden op welke punten er verschillen optreden tussen de tien onderscheiden (achterstand)groepen. We doen dat door het toevoegen van negen dummy-variabelen, waarbij de categorie ‘autochtoon, ouders met

middelbare opleiding' als referentie wordt gebruikt. We onderzoeken dus hoe de groei van de andere groepen leerlingen is ten opzichte van deze referentiegroep.
In een verdere toetsing van de modellen aan de data blijkt dat de effecten van deze negen dummy-variabelen op het beginniveau en op de groei in de drie cohorten aan elkaar gelijkgesteld kunnen worden. De onderscheiden groepen leerlingen ontwikkelen zich in de drie cohorten dus op vergelijkbare wijze.

3.3.2 Uitkomsten

De volgende tabel geeft de parameterschattingen van het uiteindelijke model voor de ontwikkeling van de taalprestaties. Daarin verschilt cohort A wat betreft beginniveau en groei van de cohorten B en C. Maar de effecten van de verschillende variabelen op beginniveau en groei zijn gelijk in de drie cohorten.

Tabel 3.4 Parameterschattingen van het groeimodel voor taal

	coh A	coh B	coh C	
beginniveau	1048.0	1051.5	1051.5	
groei	19.1	17.7	17.7	
	effecten op beginniveau			**effecten op groei**
taalscore in groep 2		10.67		0.00
auto 1.0, ho		5.33		1.07
auto 1.0, mbo *(referentie)*		0		0
auto 1.25-ex		-3.84		-0.14
auto 1.25 (nu 0.3)		-10.06		-1.04
auto 1.25-zwaar (nu 1.2)		-15.66		-0.62
allo 1.0, ho		-7.30		3.65
allo 1.0, mbo		-14.92		1.34
allo 1.9-ex		-17.68		1.33
allo 1.9 (nu 0.3)		-22.20		0.85
allo 1.9-zwaar (nu 1.2)		-30.15		1.70

Vergeleken met de cohorten B en C zien we in cohort A een lager beginniveau maar wel een sterkere groei over de jaren. Het effect van de eerder in groep 2 behaalde toetsscore voor taal op het beginniveau is 10.67: wie één standaarddeviatie hoger scoorde op die toets begint in groep 4 met ruim 10 punten meer op de taalvaardigheid. Het effect op de latere groei is in alle cohorten 0.

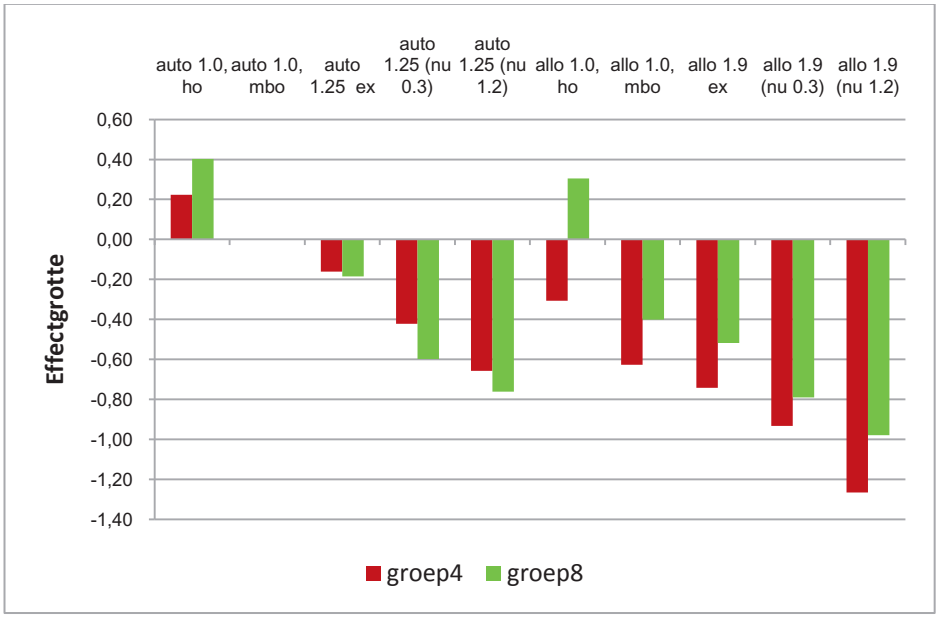
De groep autochtone leerlingen met hoger opgeleide ouders heeft een hoger startniveau dan de referentiegroep; alle andere groepen hebben een lager startniveau. Bij de effecten op de jaarlijkse groei zien we dat de drie groepen 1.25-leerlingen minder groeien dan de referentiegroep; de overige groepen groeien meer.
Op basis van deze parameterschattingen kunnen we startscores in groep 4 (intercept) en de (lineaire) groei naar Eindscores in groep 8 berekenen; voor de hele referentiesteekproef en voor de afzonderlijke groepen leerlingen.

Tabel 3.5 Schattingen van taalvaardigheid in groep 4, groep 8 en de groei in taalvaardigheid;naar type leerling

	Cohort A			Cohort B&C		
type leerling	**groep4**	**groep8**	**groei**	**groep4**	**groep8**	**groei**
auto 1.0, ho	1053.3	1134.0	80.6	1056.8	1131.9	75.0
auto 1.0, mbo *(referentie)*	1048.0	1124.4	76.4	1051.5	1122.3	70.8
auto 1.25-ex	1044.1	1120.0	75.8	1047.7	1117.9	70.2
auto 1.25 (nu 0.3)	1037.9	1110.1	72.2	1041.5	1108.0	66.6
auto 1.25-zwaar (nu 1.2)	1032.3	1106.2	73.9	1035.9	1104.1	68.3
allo 1.0, ho	1040.7	1131.6	90.9	1044.2	1129.6	85.3
allo 1.0, mbo	1033.1	1114.8	81.7	1036.6	1112.7	76.1
allo 1.9-ex	1030.3	1112.0	81.7	1033.8	1109.9	76.1
allo 1.9 (nu 0.3)	1025.8	1105.5	79.8	1029.3	1103.5	74.2
allo 1.9-zwaar (nu 1.2)	1017.8	1101.0	83.2	1021.4	1098.9	77.6
totaal	1040.2	1117.2	77.0	1045.3	1117.2	71.9

We zien dat de groei bij allochtone leerlingen doorgaans groter is dan bij de autochtone leerlingen. Daarmee maken zij een deel van hun achterstand in beginniveau ongedaan. Dit wordt nog duidelijker door de uitkomsten van het groeimodel uit te drukken in termen van (verandering in) achterstand ten opzichte van onze referentiegroep, de middelbaar opgeleide autochtonen. Deze verandering in achterstand is in alle drie cohorten hetzelfde. In Figuur 3.1 staan de resultaten, uitgedrukt in effectgroottes.

Figuur 3.1 Achterstanden in taalvaardigheid in groep 4 en groep 8; naar type leerling; referentie: autochtoon met mbo



Te zien valt dat alle drie groepen 1.25-leerlingen in groep 8 verder achterliggen dan in groep 4: hun achterstand wordt tijdens hun gang door het basisonderwijs dus groter. De allochtone groepen zien hun achterstand juist kleiner worden en die van de kleine groep allochtone leerlingen waarvan de ouders hoger onderwijs hebben gehad komen zelfs op een voorsprong uit. Bij de andere groepen allochtone leerlingen worden de achterstanden verkleind, maar ze zijn in groep 8 nog steeds aanzienlijk.
Overigens merken we wel op dat, ondanks het feit dat de allochtone 1.9-leerlingen een deel van hun achterstand inlopen en de autochtone 1.25-leerlingen verder op achterstand raken ten opzichte van de referentiegroep, de gemiddelden voor de 1.9-leerlingen in groep 8 nog altijd lager zijn dan die van de 1.25-leerlingen.

3.3.3 Herkomst allochtone leerlingen

De allochtone (achterstands)leerlingen vormen natuurlijk een heterogene groep, waarvan de ouders afkomstig zijn uit verschillende landen. We maken daarom bij de drie groepen allochtone achterstandsleerlingen (de 1.9 groepen) nog een nader onderscheid naar hun herkomst:

- Turkije;
- Marokko;
- Suriname of de Antillen;
- overige niet-westerse landen.

Hiermee kunnen dus 12 groepen allochtone achterstandsleerlingen worden onderscheiden. (De aantallen allochtone leerlingen met middelbaar of hoger opgeleide ouders zijn te klein om nog zinvol nader naar herkomst ingedeeld te worden.)

In een volgende analyse zijn de effecten van dummy-variabelen voor deze 12 groepen op Intercept en Slope van het groeimodel geschat. Opnieuw is daarbij de groep Nederlandse leerlingen met middelbaar opgeleide ouders de referentiegroep. (De andere groepen zijn hier buiten de analyse gehouden.)

Tabel 3.6 Parameterschattingen van het groeimodel voor taal; drie groepen 1.9–leerlingen naar etnische herkomst

	coh A	coh B	coh C
beginniveau	1048.3	1051.1	1051.1
groei	19.1	17.6	17.6

	effecten op beginniveau	effecten op groei
taalscore in groep 2	11.75	0
ex Turkije	-30.61	3.01
ex Marokko	-16.43	0.53
ex Sur/Ant	-9.70	-0.85
ex Overig nw	-11.36	1.42
0.3 Turkije	-34.65	2.51
0.3 Marokko	-18.34	0.79
0.3 Sur/Ant	-15.56	-0.17
0.3 Overig nw	-17.82	0.61
1.2 Turkije	-37.71	2.61
1.2 Marokko	-25.87	1.10
1.2 Sur/Ant	-19.56	0.69
1.2 Overig nw	-22.87	1.98
auto 1.0, mbo *(referentie)*	0	0

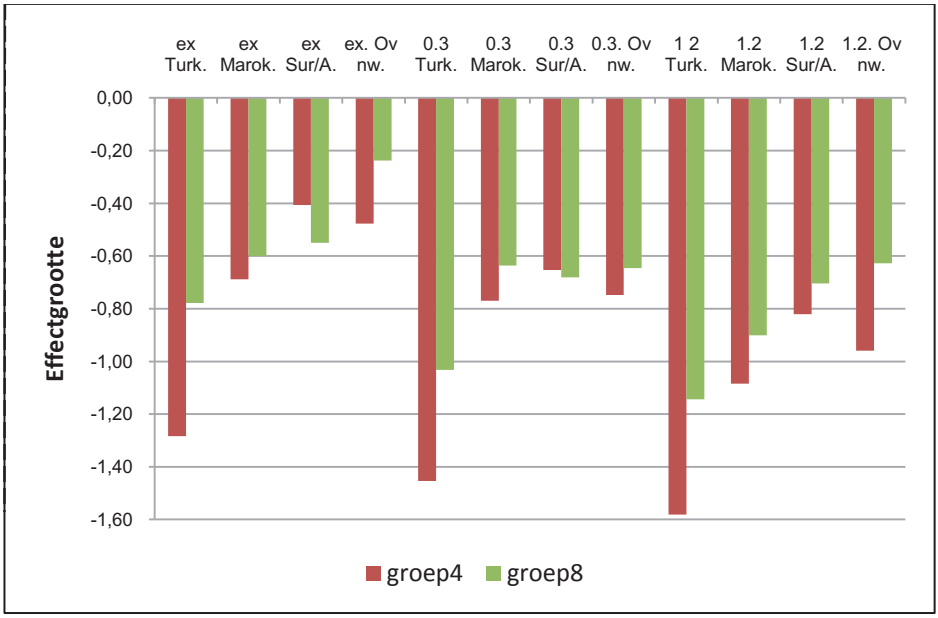
Alle allochtone groepen hebben een lager startniveau in groep 4 dan de referentiegroep. Maar over het algemeen groeien zij wat meer gedurende de volgende schooljaren. De combinatie hiervan wordt duidelijker in de volgende tabel, waarin we, op basis van bovenstaand model, de startscores in groep 4 en groei naar Eindscores in groep 8 hebben weergegeven.

Tabel 3.7 Schattingen van de taalvaardigheid in groep 4, groep 8 en de groei in taalvaardigheid; drie groepen 1.9–leerlingen naar etnische herkomst

	cohort A			Cohort B&C		
	groep 4	groep 8	groei	groep 4	groep 8	groei
auto 1.0, mbo *(referentie)*	1048.2	1124.7	76.5	1051.1	1121.5	70.5
ex Turkije	1017.6	1106.2	88.5	1020.5	1103.0	82.5
ex Marokko	1031.8	1110.4	78.6	1034.7	1107.2	72.6
ex Sur/Ant	1038.5	1111.6	73.1	1041.4	1108.4	67.1
ex Overig nw	1036.9	1119.0	82.2	1039.7	1115.9	76.1
0.3 Turkije	1013.6	1100.1	86.5	1016.4	1096.9	80.5
0.3 Marokko	1029.9	1109.5	79.6	1032.7	1106.4	73.6
0.3 Sur/Ant	1032.7	1108.5	75.8	1035.5	1105.3	69.8
0.3 Overig nw	1030.4	1109.3	78.9	1033.3	1106.1	72.9
1.2 Turkije	1010.5	1097.4	86.9	1013.4	1094.3	80.9
1.2 Marokko	1022.4	1103.2	80.9	1025.2	1100.1	74.9
1.2 Sur/Ant	1028.7	1107.9	79.2	1031.5	1104.8	73.2
1.2 Overig nw	1025.4	1109.7	84.4	1028.2	1106.6	78.4

Wat een en ander betekent voor de achterstand van deze 12 allochtone groepen ten opzichte van de referentiegroep wordt in de Figuur 3.2 weergegeven. Daarin worden de achterstanden weer uitgedrukt in effectgroottes.

Figuur 3.2 Achterstanden in taalvaardigheid in groep 4 en groep 8; 1.9–leerlingen naar etnische herkomst; referentie: autochtoon met mbo



Uit deze figuur wordt duidelijk dat de meeste groepen achterstanden inlopen. Alle drie de Turkse groepen beginnen met de grootste achterstand, maar lopen die ook het meeste in. Toch hebben zij in groep 8 nog steeds de grootste achterstand van alle allochtone groepen. De Surinaamse (en Antilliaanse) leerlingen beginnen met de kleinste achterstand, maar alleen de 'zwaarste' groep (met huidig gewicht 1.2) loopt iets in; beide andere Surinaamse groepen groeien minder dan de referentiegroep en komen in groep 8 dus iets verder op achterstand.

3.4 Ontwikkeling van rekenprestaties

Ook bij rekenen proberen we allereerst weer een 'basis-groeimodel' te ontwikkelen, dat zo spaarzaam mogelijk is (zo veel mogelijk identiek over de cohorten), maar wel een goede *fit* bij de data vertoont.

3.4.1 Modelontwikkeling

Net als bij taal beschikken we ook hier over gegevens van drie opeenvolgende cohorten, met in elk cohort metingen van de rekenvaardigheid in de groepen 4, 6 en 8 en met een meting van aanvangsprestaties voor rekenen in groep 2. En we proberen de ontwikkeling van de rekenvaardigheid van groep 4 naar groep 8 weer te modelleren met een (lineair) groeimodel met een beginniveau in groep 4 (*Intercept*) en een helling (*Slope*).

Het blijkt echter niet mogelijk om een groeimodel op te stellen dat goed bij de data van alle drie de cohorten past (er zijn convergentieproblemen en de software waarschuwt voor onmogelijke waarden bij sommige parameterschattingen). Nadere inspectie laat zien dat het probleem zit in de data van Cohort A; over de gegevens van de cohorten B en C is wel een plausibel groeimodel te fitten. We veronderstellen dat de problemen bij cohort A komen door de omrekening van de scores op de, in dit cohort in groep 4 nog gebruikte, PRIMA-rekentoets naar een score op de, na die tijd in groep 6 en 8 en de beide andere cohorten gebruikte, Cito-rekentoetsen. Deze omrekening levert in elk geval veel grotere spreiding in scores bij rekenen dan we vinden bij de (rechtstreekse) scores op Cito-rekentoetsen. Ook eerder zijn er reeds vraagtekens gezet bij de hele omreken-procedure (Koopman, 2002). Voor rekenen zullen we daarom alleen de ontwikkeling van leerlingen analyseren in het tweede en derde cohort.

We starten weer met een 0-model waarin Intercept, Slope en de samenhang tussen beide mogen variëren tussen de cohorten. Om een goed fittend model te krijgen blijkt een aanpassing van de schattingen van de tijdpunten nodig. Standaard wordt de groei over de groepen 4-6-8 geschat met parameters gefixeerd op 0-2-4 en het probleem blijkt in het vaste tijdpunt voor groep 6 te zitten. Een model met vrije schatting van dit punt fit aanzienlijk beter. Het tijdpunt voor groep 6 wordt dan geschat op 2.15 in plaats van op 2.0. Dit betekent dus dat de helft van de groei tussen groep 4 en groep 8 niet na 2 jaar wordt bereikt, maar pas na 2.15 jaar. Omdat het ons hier gaat om de ontwikkeling van onderwijsachterstanden over het hele traject van de basisschool is deze aanpassing van het groeimodel voor rekenen verder geen probleem.

In het groeimodel van rekenen is er, in tegenstelling tot wat we bij taal vonden, een significante negatieve samenhang van het beginniveau (Intercept) en de groei (Slope): naarmate leerlingen hoger beginnen is er sprake van minder

groei. De schattingen voor beginniveau, groei én de genoemde negatieve samenhang tussen die twee verschillen niet tussen de twee cohorten.
Aan dit model wordt vervolgens het effect van de (gestandaardiseerde) score op de rekentoets uit groep 2 (toets Ordenen) toegevoegd. Deze variabele heeft een positief effect op het beginniveau in groep 4 en een negatief effect op de verdere groei: hoe beter een leerling in groep 2 was bij de toets Ordenen, des te hoger zijn of haar startniveau voor rekenen in groep 4, maar des te lager ook de verdere groei. Beide effecten verschillen weer niet tussen de twee cohorten.
Tot slot worden de dummy-variabelen voor de verschillende groepen (achterstands-)leerlingen in het groeimodel opgenomen. Ook de effecten van deze variabelen blijken gelijkgesteld te kunnen worden tussen beide cohorten.

3.4.2 Uitkomsten

De volgende tabel geeft de parameterschattingen van het uiteindelijke model voor de ontwikkeling van de rekenprestaties.

Tabel 3.8 Parameterschattingen van het groeimodel voor rekenen

Cohort B en C		
beginniveau	69.21	
groei	17.26	
beginniveau X groei	-0.08	
	effecten op beginniveau	**effecten op groei**
rekenscore in groep 2	5.19	-0.05
auto 1.0, ho	1.98	0.20
auto 1.0, mbo *(referentie)*	0	0
auto 1.25-ex	-1.05	-0.11
auto 1.25 (nu 0.3)	-2.84	-0.40
auto 1.25-zwaar (nu 1.2)	-4.77	-0.56
allo 1.0, ho	-0.48	0.43
allo 1.0, mbo	-3.15	0.28
allo 1.9-ex	-2.92	-0.01
allo 1.9 (nu 0.3)	-4.65	-0.02
allo 1.9-zwaar (nu 1.2)	-4.91	0.12

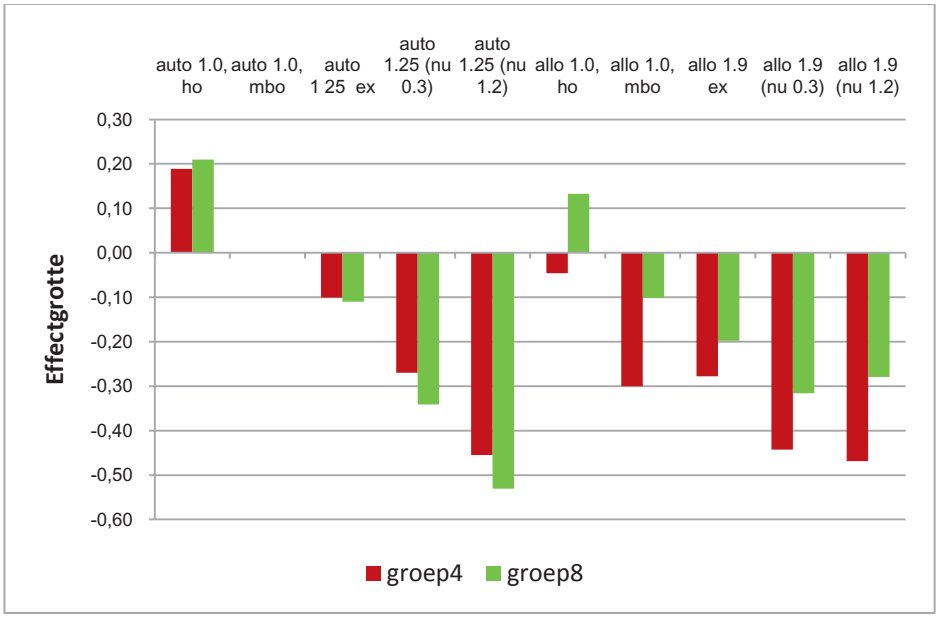
De uitkomsten uit deze tabel zijn lastig rechtstreeks te interpreteren, ook al vanwege het negatieve effect van een hogere startpositie (Intercept) op de groei (Slope). In Tabel 3.9 geven we daarom de op basis van dit model geschatte startscores in groep 4 en de groei naar Eindscores in groep 8.

Tabel 3.9 Schattingen van rekenvaardigheid in groep 4, groep 8 en de groei in rekenvaardigheid; naar type leerling

	Cohort B en C		
	groep 4	groep 8	groei
auto 1.0, ho	71.2	119.4	48.2
auto 1.0, mbo *(referentie)*	69.2	117.2	48.0
auto 1.25-ex	68.2	116.1	47.9
auto 1.25 (nu 0.3)	66.4	113.6	47.3
auto 1.25-zwaar (nu 1.2)	64.4	111.6	47.2
allo 1.0, ho	68.7	118.6	49.9
allo 1.0, mbo	66.1	116.2	50.1
allo 1.9-ex	66.3	115.1	48.8
allo 1.9 (nu 0.3)	64.6	113.9	49.3
allo 1.9-zwaar (nu 1.2)	64.3	114.3	50.0
totaal	68.1	116.2	48.1

De verschillende groepen allochtone leerlingen groeien meer dan de autochtone leerlingen. De drie onderscheiden groepen 1.25-leerlingen groeien het minst.
De uitkomsten worden nog duidelijker door ze weer uit te drukken in termen van (verandering in) achterstand ten opzichte van onze referentiegroep, de middelbaar opgeleide autochtonen. Figuur 3.3 geeft de uitkomsten daarvan.

Figuur 3.3 Achterstanden in rekenvaardigheid in groep 4 en groep 8; naar type leerling; referentie: autochtoon met mbo



De drie groepen autochtone (ex-)achterstandsleerlingen (1.25) zien hun achterstand ten opzichte van de referentiegroep van groep 4 naar groep 8 oplopen; de groepen met een licht (0.3) of zwaar (1.2) gewicht hebben in groep 8 de laagste scores voor rekenen. Bij de allochtone leerlingen zien we juist een inlopen van achterstand. Bij de groep allochtone leerlingen met hoger opgeleide ouders verandert de kleine achterstand in groep 4 zelfs in een voorsprong in groep 8.

3.4.3 Herkomst allochtone leerlingen

Ook bij de rekenontwikkeling hebben we weer nader onderscheid gemaakt binnen de allochtone achterstandsleerlingen (de 1.9-groepen) naar hun herkomst: Turkije, Marokko, Suriname/Antillen en overig niet-westers. (De aantallen allochtone leerlingen met middelbaar of hoger opgeleide ouders zijn te klein om nog nader ingedeeld te worden.)
In een volgende analyse zijn de effecten van dummy-variabelen voor de resulterende 12 groepen op Intercept en Slope van het groeimodel geschat. Opnieuw is daarbij de groep Nederlandse leerlingen met middelbaar opgeleide

ouders de referentiegroep. (De andere groepen zijn hier buiten de analyse gehouden.)

Tabel 3.10 Parameterschattingen van het groeimodel voor rekenen; drie groepen 1.9–leerlingen naar herkomst

Cohort B & C		
beginniveau	69.15	
groei	17.92	
beginniveau X groei	-0.09	
	effecten op beginniveau	**effecten op groei**
rekenscore in groep 2	5.31	-0.15
ex Turkije	1.24	0.07
ex Marokko	-3.71	-0.32
ex Sur/Ant	-4.87	-0.09
ex Overig nw	-3.24	0.06
0.3 Turkije	-2.68	0.11
0.3 Marokko	-7.74	0.11
0.3 Sur/Ant	-5.14	-0.42
0.3 Overig nw	-2.19	-0.31
1.2 Turkije	-3.39	-0.03
1.2 Marokko	-5.76	-0.20
1.2 Sur/Ant	-8.25	0.18
1.2 Overig nw	-3.50	0.06
auto 1.0, mbo*(referentie)*	0	0

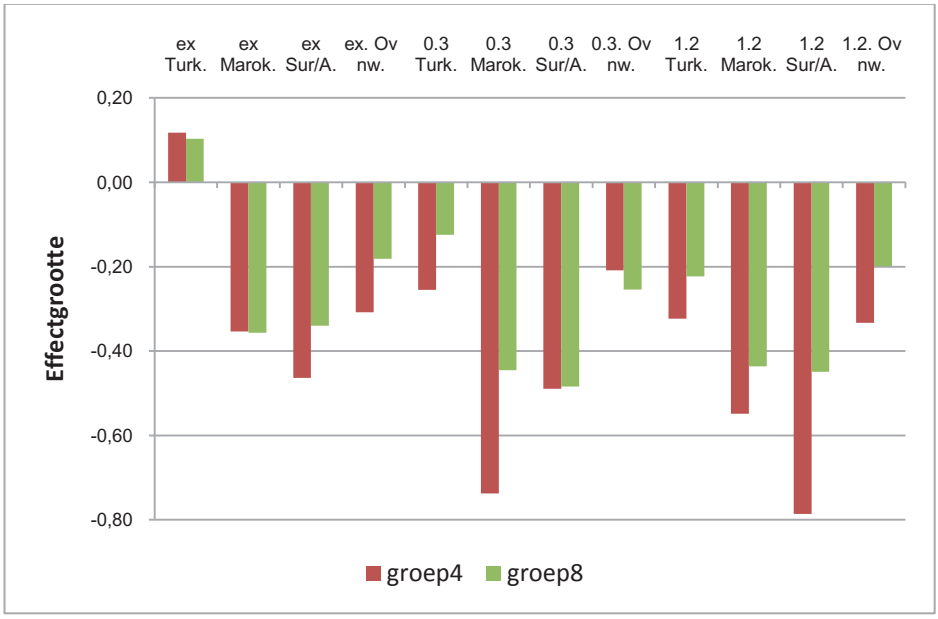
Ook hier kan de interpretatie van het model, met een negatieve samenhang tussen startniveau (Intercept) en groei (Slope), vergemakkelijkt worden door te kijken naar de op basis van het model berekende scores in groep 4 en groep 8 (en daarmee de groei tussen die tijdstippen).

Tabel 3.11 Schattingen van rekenvaardigheid in groep 4, groep 8 en de groei in rekenvaardigheid; drie groepen 1.9–leerlingen naar herkomst

	Cohort B&C		
	groep 4	groep 8	groei
auto 1.0, mbo *(referentie)*	69.2	117.3	48.2
ex Turkije	70.4	118.4	48.0
ex Marokko	65.4	113.6	48.1
ex Sur/Ant	64.3	113.8	49.5
ex Overig nw	65.9	115.4	49.5
0.3 Turkije	66.5	116.0	49.5
0.3 Marokko	61.4	112.7	51.2
0.3 Sur/Ant	64.0	112.3	48.2
0.3 Overig nw	67.0	114.7	47.7
1.2 Turkije	65.8	115.0	49.2
1.2 Marokko	63.4	112.8	49.4
1.2 Sur/Ant	60.9	112.6	51.7
1.2 Overig nw	65.7	115.3	49.6

De meeste van de verschillende herkomstgroepen groeien meer in rekenen dan de referentiegroep. Maar ze beginnen in groep 4 dan ook vrijwel allemaal op een lager niveau. Wat een en ander betekent voor de achterstanden van deze groepen ten opzichte van onze referentiegroep kan uit de volgende figuur blijken.

Figuur 3.4 Achterstanden in rekenvaardigheid in groep 4 en groep 8;1.9–leerlingen naar herkomst; referentie: autochtoon met mbo



Ook hieruit blijkt dat voor de meeste herkomstgroepen de achterstand in rekenvaardigheid afneemt. Dat geldt nog het sterkst voor de groepen die aanvankelijk de grootste achterstand hadden (de Marokkaanse 0.3 en 1.2 leerlingen en de Surinaams/Antilliaanse 1.2 leerlingen). Maar er zijn ook groepen die geen achterstand inlopen: de Marokkaanse ex 1.9-leerlingen en de Surinaamse en overig niet-westerse leerlingen met gewicht 0.3.

3.5 Ontwikkeling van doelgroep– versus niet–doelgroepleerlingen

In de voorgaande paragrafen zijn de ontwikkeling van de taal- en rekenprestaties van verschillende groepen leerlingen geanalyseerd volgens de uitgebreide indeling naar (achterstands)leerling. Tot slot van dit hoofdstuk gaan we nog na hoe deze ontwikkeling eruit ziet wanneer we de hele groep allochtone achterstandsleerlingen en de hele groep autochtone achterstandsleerlingen vergelijken met alle leerlingen zonder gewicht. Dat kan meer zicht geven op het al dan niet halen van beleidsdoelen met betrekking tot het inhalen van onderwijsachterstanden.

Omdat de hier onderzochte cohorten leerlingen al in de jaren negentig van de vorige eeuw zijn begonnen aan hun loopbaan in het basisonderwijs gebruiken we daarvoor de oude gewichtenindeling. We vergelijken de hele groep autochtone achterstandsleerlingen (ex-1.25 en de huidige lichte en zware categorie) en de hele groep allochtone achterstandsleerlingen (ex-1.9 en de huidige lichte en zware categorie) met alle leerlingen zonder gewicht (allochtoon en autochtoon, met mbo en met ho).
In de volgende tabel geven we de schattingen van de taal- en rekenvaardigheid van de drie onderscheiden groepen in de groepen 4 en 8, gebaseerd op de eerder gebruikte groeimodellen voor de taal en rekenvaardigheid.

Tabel 3.12 Schattingen van taal– en rekenvaardigheid in groep 4 en groep 8, en de groei van vaardigheden ; voor leerlingen zonder gewicht en allochtone en autochtone doelgroepleerlingen

	taalvaardigheid (cohorten B en C)			rekenvaardigheid (cohorten B en C)		
	groep 4	groep 8	groei	groep 4	groep 8	groei
gemiddelden						
zonder gewicht	1053.1	1127.0	73.9	70.0	118.2	48.2
autochtone doelgroep	1043.9	1112.5	68.6	67.1	114.5	47.5
allochtone doelgroep	1026.1	1102.6	76.6	64.9	114.3	49.5
achterstand (effect size)						
autochtone doelgroep	-0.38	-0.61		-0.28	-0.35	
allochtone doelgroep	-1.13	-1.02		-0.49	-0.37	

De cijfers in deze tabel bevestigen het beeld dat reeds eerder in dit hoofdstuk werd verkregen. Binnen deze generatie(s) leerlingen neemt de achterstand van de (gehele) autochtone doelgroep ten opzichte van niet-doelgroepleerlingen toe gedurende het basisonderwijs. Dat geldt in het bijzonder voor taal en in mindere mate voor rekenen. De achterstand van de (gehele) allochtone doelgroep wordt juist wat kleiner. Deze neemt bij taal af van 1.13 naar 1.02 standaarddeviatie en bij rekenen van 0.49 naar 0.37 standaarddeviatie.

3.6 Conclusies en discussie

3.6.1 Conclusies

In dit hoofdstuk zijn we nagegaan hoe de schoolprestaties van een instroomcohort leerlingen zich ontwikkelen gedurende hun loopbaan door het basisonderwijs en of die ontwikkeling verschillend is voor de hier onderscheiden groepen leerlingen. Bovendien zijn we nagegaan of er op deze punten sprake is van veranderingen voor opeenvolgende instroomcohorten van leerlingen.
Vergeleken met het vorige hoofdstuk, waarin per jaargroep werd nagegaan welke verschillen in prestaties er door de tijd heen zijn tussen de onderscheiden groepen leerlingen, is hier gekozen voor een aanpak waarbij eenzelfde (instroom)cohort leerlingen longitudinaal gevolgd wordt gedurende hun loopbaan door het basisonderwijs. Daarbij zijn ook de gegevens van uitvallende leerlingen benut.

Ontwikkeling taalvaardigheid

De groei van de taalvaardigheid van de leerlingen blijkt niet gelijk te zijn over alle drie de onderzochte cohorten. In het oudste cohort is er sprake van een wat lager startniveau en een hoger groeitempo dan in de beide jongere cohorten. Het eindniveau in groep 8 is in alle drie de cohorten wel gelijk. Ook in eerder onderzoek werd gevonden dat in de lagere groepen het prestatieniveau in de jaren negentig stijgt, terwijl het niveau in groep 8 min of meer gelijk blijft (Mulder e.a., 2005). Mogelijk hangt het wat lagere beginniveau bij het oudste cohort samen met het in die tijd nog minder ‘gewend’ zijn aan het afnemen van toetsen in de lagere groepen van het basisonderwijs.
Wat betreft het beginniveau kunnen we concluderen dat alle doelgroepleerlingen (zowel volgens de oude als volgens de nieuwe regelingen) lager starten dan de overige leerlingen. Binnen elke categorie van de doelgroepleerlingen (de ex-groep, de huidige lichte en zware gewichtengroepen) starten de allochtone leerlingen lager dan de autochtonen. Bij de groei van de taalvaardigheid gedurende het basisonderwijs is er bij alle allochtone groepen sprake van een sterkere groei dan bij de referentiegroep (autochtone leerlingen met middelbaar opgeleide ouders) en doorgaans ook sterker dan de groei bij de autochtone leerlingen met hoog opgeleide ouders. De verschillende onderscheiden groepen autochtone achterstandsleerlingen groeien minder dan de referentiegroep en komen daarmee verder op

achterstand. Dat geldt in het bijzonder voor de groepen met huidig gewicht 0.3 en 1.2.
Als we binnen de allochtone doelgroepleerlingen nader onderscheid maken naar hun herkomst, dan zien we dat de leerlingen van Turkse komaf het laagste beginniveau hebben. Deze leerlingen lopen weliswaar meer achterstand in dan de allochtone leerlingen met andere herkomst, maar ook in groep 8 hebben ze nog steeds het laagste niveau. De leerlingen van Surinaamse of Antilliaanse herkomst beginnen in groep 4 met de minste achterstand op de referentiegroep, maar lopen deze achterstand doorgaans niet in.

Ontwikkeling rekenvaardigheid
Bij de groei van de rekenvaardigheid van de leerlingen gedurende het basisonderwijs is er een negatief verband tussen hun startniveau en de daarna volgende groei: leerlingen die aan het begin al relatief vaardig zijn in rekenen gaan minder vooruit dan leerlingen die minder vaardig zijn begonnen. Dit patroon wordt in beide hier onderzochte cohorten teruggevonden.
Ook bij de rekenvaardigheid beginnen alle doelgroepleerlingen met een lager niveau dan de overige leerlingen. En binnen elk van de onderscheiden groepen starten de allochtone leerlingen lager dan de autochtone leerlingen met vergelijkbaar opgeleide ouders, hoewel het verschil bij de leerlingen met een 'zwaar' gewicht (1.2) verwaarloosbaar is.
Alle allochtone leerlingen, ook die uit de 1.0-groepen, gaan meer vooruit in rekenvaardigheid dan alle onderscheiden groepen autochtone leerlingen. De groei van de 1.25-groepen, vooral die met de huidige gewichten 0.3 en 1.2, is minder dan die van de andere groepen leerlingen. Het gevolg is dat hun aanvangsachterstand op de referentiegroep gedurende het basisonderwijs verder toeneemt. Bij alle groepen allochtone leerlingen neemt die achterstand juist af (en wordt bij de groep met de hoogst opgeleide ouders zelfs een voorsprong). Binnen de allochtone doelgroepleerlingen is sprake van een wisselend beeld. De groepen die starten met de meeste achterstand op de referentiegroep lopen daar een flink deel van in (de Marokkaanse 0.3- en 1.2-leerlingen en de Surinaams/Antilliaanse 1.2-leerlingen). Maar er zijn ook groepen die hun aanvankelijke achterstand blijven behouden.

3.6.2 Discussie

In het vorige hoofdstuk is de ontwikkeling van achterstanden van de verschillende groepen achterstands(leerlingen) aan de orde gekomen door vergelijkingen *tussen* verschillende generaties leerlingen. De prestaties in groep 2 van schooljaar 1994/95 werden vergeleken met die van de groepen 2 in een reeks van volgende jaren. En ook voor groep 8 werden de scores in verschillende jaren met elkaar vergeleken. De globale conclusies daaruit waren dat de allochtone leerlingen een flinke inhaalslag hebben gemaakt over deze generaties heen. Bij de autochtone leerlingen is daar veel minder sprake van en bij taalvaardigheid bleek zelf sprake van een relatieve verslechtering.
In dit hoofdstuk is de ontwikkeling van achterstandsgroepen *binnen* een generatie leerlingen aan de orde geweest. Hoe ontwikkelen de leerlingen zich, die met aanvankelijke, soms grote, achterstanden in het onderwijs beginnen?
In het algemeen kunnen we concluderen dat de ontwikkeling van alle allochtone leerlingen relatief gunstig is: ze halen gedurende hun loopbaan door het basisonderwijs een deel van hun aanvangsachterstand in. Wel blijft hun achterstand in vaardigheden (vooral bij taal) aan het eind van het basisonderwijs nog aanzienlijk.
De ontwikkeling van de autochtone achterstandsleerlingen (met huidig gewicht 0.3 en 1.2) is zorgelijk. Zij beginnen ook met achterstand aan het basisonderwijs, maar komen gedurende hun loopbaan nog verder op achterstand.
De relatief gunstige ontwikkeling van de allochtone leerlingen kwam ook in eerder onderzoek (bv. Mulder e.a., 2005) naar voren en bleek ook uit de analyses in het vorige hoofdstuk van dit rapport. De minder gunstige ontwikkeling van de autochtone achterstandsleerlingen is ook eerder geconstateerd (Vogels & Bronneman-Helmers, 2003; Mulder e.a., 2005; Mulder e.a., 2007) , maar komt in de cross-sectionele analyses in het vorige hoofdstuk minder duidelijk naar voren. Wat taal betreft is de relatieve positie van de autochtone achterstandsleerlingen in groep 8 min of meer onveranderd (zie Figuur 2.2 in het vorige hoofdstuk); wat rekenen betreft geldt ongeveer hetzelfde (zie Figuur 2.4), zij het dat er bij de recentste meting van 2007 sprake is van vermindering van achterstand.
Bij verschillende analyses in dit hoofdstuk zagen we dat de groepen leerlingen, die met de grootste achterstanden aan het basisonderwijs beginnen, daar ook relatief het meeste van weten in te lopen. Dat betekent dat het basisonderwijs er, in elk geval voor een deel, in slaagt om zulke leerlingen extra vooruit te

helpen, vergeleken met leerlingen met middelbaar opgeleide ouders. Natuurlijk valt er bij de groepen met de meeste achterstand ook de meeste winst te behalen. Maar dat dat niet vanzelf gaat blijkt uit de minder gunstige ontwikkeling van de autochtone achterstandsleerlingen. Het, in elk geval deels, inlopen van achterstanden bij groepen met veel achterstand lijkt ons dan ook een compliment voor de basisscholen in Nederland.

4 Over- of onderadvisering bij doelgroepleerlingen?

Geert Driessen en Joris Cuppen

4.1 Het advies voortgezet onderwijs

4.1.1 De overgang van basis- naar voortgezet onderwijs

Wanneer kinderen na het basisonderwijs overstappen naar het voortgezet onderwijs, is dat waarschijnlijk het meest bepalende transitiemoment in hun schoolloopbaan. Zeker wanneer de mogelijkheden tot het zogenoemde stapelen beperkt zijn,[12] is het van belang meteen een optimale keuze te maken, dat wil zeggen de keuze voor een type voortgezet onderwijs dat het beste aansluit bij de capaciteiten van de leerling. Deze keuze is slechts tot op zekere hoogte een vrije keuze; ze wordt in sterke mate vastgelegd met het advies dat de leerkracht van groep 8 (of directeur van de basisschool) geeft. Dat advies is weer gebaseerd op een inschatting van vooral de cognitieve capaciteiten, maar daarnaast ook op zaken als motivatie en interesses en mogelijke ondersteuning in de thuissituatie. Dit laatste is weer sterk verweven met het sociale en etnische herkomstmilieu van de leerling: in hogere en autochtone milieus zijn er doorgaans meer mogelijkheden tot ondersteuning.

Bij het kiezen van een school voor voortgezet onderwijs wordt, naast het advies, ook gebruik gemaakt van een 'tweede gegeven'. Dat kan de uitkomst van een schoolvorderingenonderzoek zijn, of van een intelligentie-onderzoek,

[12] In de jaren '90 zijn er vanuit bezuinigingsoverwegingen maatregelen genomen om het stapelen (bv. na mavo verder gaan met havo) in te dammen. Recentelijk is de beleidskoers echter gewijzigd en gaat de overheid er weer vanuit dat leerlingen een eventueel verkeerde keuze moeten kunnen bijstellen (Tweede Kamer, 2010).

of van een toelatingsexamen. Op de meeste basisscholen wordt gebruik gemaakt van een eindtoets, doorgaans die van het Cito.
Bij de totstandkoming van het advies en de uiteindelijke keuze zijn behalve de leerling zelf en de groepsleerkracht nog andere partijen betrokken. De ontvangende scholen voor voortgezet onderwijs stellen bijvoorbeeld soms eisen wat betreft de minimumscore op een eindtoets; ouders hebben specifieke wensen met betrekking tot wat voor hen nog een acceptabel minimumniveau is; en vriendjes en vriendinnetjes oefenen groepsdruk uit om samen naar een school te gaan die 'cool' is (Driessen e.a., 2005).

4.1.2 Discrepanties tussen prestaties en adviezen

Het advies van de basisschool wordt dus door meerdere factoren bepaald. Een gevolg daarvan is dat de (statistische) samenhang tussen de cognitieve capaciteiten, zoals bijvoorbeeld gemeten met een eindtoets, en het advies niet perfect is.[13] Driessen en Doesborgh (2005) en Driessen en Smeets (2007, 2011) hebben op basis van grootschalige, landelijke gegevens berekend dat ruim tweederde van de verschillen in advieshoogte kan worden toegeschreven aan de resultaten op taal-, reken- en leestoetsen of de eindtoets. Eenderde komt dus op het conto van andere invloeden. Het Cito verrichtte onderzoek onder scholen voor voortgezet onderwijs naar de relatie tussen het advies van de basisschool en het advies op basis van de Cito-score (Van der Lubbe, Verhelst, Heuvelmans & Staphorsius, 2005). Het bleek dat beide voor 86% van de leerlingen overeenstemden. Wanneer er sprake was van een discrepantie, ging 8% van de scholen voor voortgezet onderwijs louter af op het advies van de basisschool en minder dan 1% op het Cito-advies; de helft van de scholen voor voortgezet onderwijs overlegde met de basisschool en als die een aannemelijk verklaring had, werd het advies van de basisschool gevolgd. Luyten en Bosker (2004) concludeerden aan de hand van een reeks studies naar de relatie tussen de score op de Cito Eindtoets en het advies voortgezet onderwijs dat er een trend waarneembaar is waarbij de gemeten capaciteiten een steeds prominentere rol zijn gaan spelen. Eerder constateerden Tesser en Iedema (2001) ook al dat de adviezen in steeds sterkere mate op de gemeten prestaties werden gebaseerd. Niettemin zijn er bij de advisering soms forse discrepanties

---

13 Overigens is de samenhang tussen advies en uiteindelijk gerealiseerd onderwijsniveau evenmin perfect. Ook bij een perfecte samenhang tussen capaciteiten en advies staat de loopbaan in het voortgezet onderwijs onder invloed van allerlei factoren die er voor kunnen zorgen dat het beoogde niveau niet wordt gerealiseerd (Inspectie van het Onderwijs, 2007).

tussen prestaties en advies, zowel bij individuele leerlingen als bij bepaalde categorieën van leerlingen. Twee soorten afwijkingen doen zich voor, namelijk 'overadvisering' en 'onderadvisering'. In het eerste geval ontvangt een leerling een advies dat hoger ligt dan dat van leerlingen met vergelijkbare prestaties; in het tweede geval krijgt een leerling een lager advies dan leerlingen met vergelijkbare prestaties.

4.1.3 Van over– naar onderadvisering?

Eind jaren '80 wees De Jong (1987) op basis van Rotterdamse gegevens als eerste op het bestaan van overadvisering. Hij liet zien dat allochtone leerlingen bij gelijke prestaties een hoger advies kregen dan hun autochtone jaargenoten. Analyses van Driessen (1991), die gebruik maakte van landelijke gegevens, bevestigden dit beeld. In deze en de daarna rond dit thema uitgevoerde analyses ging het meestal niet om de vraag of er überhaupt sprake was van overadvisering, maar of er een systematische relatie was met de etnische herkomst van de leerlingen (i.c. hogere adviezen bij allochtone leerlingen). Voor het bestaan van een dergelijke etnische overadvisering werden verschillende verklaringen gegeven, waarbij drie niveaus kunnen worden onderscheiden, namelijk het niveau van de leerling, van de klas en school, en van de bredere context.

Een verklaring op *het niveau van de leerling* die aanvankelijk vaak werd gegeven is dat leerkrachten bij hun advisering nadrukkelijk negatieve effecten van de migratiegeschiedenis van de kinderen verdisconteren. Zij schatten dan in dat die leerlingen ondanks hun nog ontoereikende Nederlandse taalvaardigheid toch over voldoende intellectuele capaciteiten[14] beschikken om met succes een hoger niveau van voortgezet onderwijs aan te kunnen dan hun toetsprestaties suggereren. De leerkrachten laten dus het door hen bij de leerlingen veronderstelde potentieel zwaarder meewegen dan hun feitelijke resultaten op een taaltest (De Jong, 1987). Leerkrachten kunnen ook niet-cognitieve aspecten als motivatie, inzet en houding zwaarder laten wegen dan de toetsuitslag. Zij geven de betreffende leerlingen het voordeel van de twijfel, omdat ze er vanuit gaan dat die het met hard werken en steun uit de omgeving uiteindelijk waarschijnlijk wel zullen redden (Kerkhoff, 1988). Een andere verklaring is dat leerkrachten bang zijn van discriminatie te worden beticht

[14] En wellicht een goede motivatie en extra ondersteuning vanuit de eigen omgeving en school voor voortgezet onderwijs.

wanneer ze veel allochtone leerlingen lage adviezen zouden geven (Driessen & Doesborgh, 2005).
Verklaringen op *het niveau van de klas of school* hebben meestal te maken met de samenstelling van de leerlingenpopulatie. Achterliggend idee is dat het oordeel van de leerkracht over een individuele leerling wordt beïnvloed door de relatieve positie die die leerling inneemt in de klas, het zogenoemde 'frog-pond'-effect (Davis, 1966). Volgens deze verklaring houden leerkrachten bij de beoordeling van hun leerlingen een bepaalde rangorde aan. Daardoor krijgen iets betere leerlingen in een klas met een voornamelijk laag prestatieniveau gemakkelijker een relatief hoog advies dan leerlingen met vergelijkbare prestaties in een klas met een hoog niveau (Driessen, 2007). Het prestatieniveau van de klas wordt op zijn beurt sterk bepaald door de sociale en etnische samenstelling van de leerlingenpopulatie. Allochtone en lager milieu kinderen presteren gemiddeld lager dan autochtone kinderen uit hogere sociaal-economische milieus. Naarmate er sprake is van een concentratie van de eerstgenoemde categorie van kinderen, ligt ook het prestatieniveau van die klas lager en zal er ten gevolge van het 'frog-pond'-effect eerder overadvisering plaatsvinden.
Op *het niveau van de bredere context* wordt er op gewezen dat overadvisering ook een typisch grootstedelijk fenomeen zou kunnen zijn. Dronkers, Van Erp, Robijns en Roeleveld (1998) toonden aan dat, ook nadat rekening is gehouden met relevante leerling- en schoolkenmerken, vooral in de grote steden sprake is van overadvisering. Volgens hen kan dit mede worden toegeschreven aan de specifieke leefstijl en het assertievere klimaat in die steden en de daaruit voortvloeiende pressie van ouders. Leerkrachten hebben er minder gezag en ouders komen eerder op voor hun rechten. Ouders nemen er tegenwoordig niet zonder meer genoegen mee wanneer hun kind een vmbo-advies krijgt en eisen dat de leerkracht een hoger advies geeft (Mulder & Suhre, 1995). Verschillen in advisering zouden daarnaast ook kunnen worden teruggevoerd op het kleinere aandeel laaggeschoolde autochtone ouders en het grotere aandeel allochtone ouders in de grote steden.

Uit enkele vanaf de tweede helft van de jaren '90 uitgevoerde studies bleek dat de overadvisering van allochtone leerlingen substantieel was verminderd, dan wel geheel was verdwenen (Dagevos, Gijsberts & Van Praag, 2003). Claassen en Mulder (2003) concludeerden dat de in 1988 en 1992 nog geconstateerde overadvisering in 2000 was omgeslagen in onderadvisering. Tesser en Iedema

(2001) vermoedden dat dit het gevolg was van procedure-afspraken die ondertussen in de grote steden waren gemaakt om het advies beter af te stemmen op de prestaties. Maar er zijn natuurlijk nog andere verklaringen mogelijk. Een afname valt bijvoorbeeld ook te verwachten omdat er steeds minder allochtone kinderen (en ouders) zijn die niet in Nederland zijn geboren. Leerkrachten hebben in de loop der jaren ook steeds meer ervaring met allochtone leerlingen opgedaan en daarom een betere kijk op hun reële capaciteiten en mogelijkheden gekregen. Verder is er het afgelopen decennium het een en ander veranderd in het politieke klimaat en in de publieke discussie, met name over onderwerpen op het gebied van immigratie en integratie die eerder nog in de taboesfeer lagen. Wellicht zijn leerkrachten daardoor minder geneigd om bij allochtone leerlingen anders te adviseren dan bij autochtone leerlingen. Een laatste punt is dat, omdat scholen voor voortgezet onderwijs steeds meer in het openbaar worden afgerekend op hun rendement, zij ook wat kritischer zijn geworden in hun toelatingsbeleid (Driessen, 2006; Kuyper, Suhre, Jansen & Pijl, 2000). Dat kan consquenties hebben voor de wijze waarop basisscholen adviseren. Wellicht zijn zij zelf door het aangescherpte toelatingsbeleid ook wat kritischer geworden wat betreft de hoogte van de adviezen die zij geven en varen ze meer op de feitelijke prestaties dan op het potentieel van leerlingen.

Zoals al eerder opgemerkt, richt zich de discussie rond de discrepantie tussen cognitieve capaciteiten en advies voornamelijk op de categorie allochtone leerlingen (vgl. Koeslag & Dronkers, 1994). Mulder, Roeleveld, Van der Veen en Vierke (2005) wezen er echter op dat, terwijl de overadvisering van allochtone leerlingen in 2002 verdwenen was, er op dat moment wel sprake was van onderadvisering en dat die zich bij autochtone achterstandsleerlingen sterker voordeed dan bij allochtone achterstandsleerlingen (zie ook Claassen & Mulder, 2002). In een latere studie van Claassen en Mulder (2006) werd dezelfde conclusie getrokken.
In 2007 werd er in de media uitvoerig aandacht besteed aan een onderzoek waaruit zou blijken dat er in de gemeente Amsterdam sprake was van onderadvisering van bepaalde groepen allochtone leerlingen (Gemeente Amsterdam, 2007a, 2007b; Babeliowsky & Den Boer, 2007). Hoewel al snel bleek dat het om een wat uit z'n verband gerukte conclusie ging en het hooguit een beperkte en selecte groep van allochtonen betrof, vormde deze overtrokken aandacht toch aanleiding tot het instellen van een nader

onderzoek. De hoofdvraag van dat onderzoek (Driessen & Smeets, 2007) luidde of er bij de overgang van het basis- naar het voortgezet onderwijs sprake is van een correcte advisering, ofwel of leerlingen een advies krijgen dat aansluit bij hun capaciteiten zoals gemeten via toetsen. Een belangrijke aanvullende vraag luidde of er daarbij eventueel verschillen waren tussen Amsterdam, de overige grote gemeenten en Nederland als geheel. Voor de analyses werd gebruik gemaakt van grootschalige, landelijke gegevensbestanden. De resultaten van die studie lieten zien dat er verschillen zijn in advieshoogte tussen gemeentetypen: in de grotere steden worden gemiddeld genomen lagere adviezen gegeven. Wanneer echter rekening wordt gehouden met verschillen in toetsprestaties verdwijnen deze verschillen. Met uitzondering van de zeer heterogene categorie 'overig allochtone leerlingen' werden, gegeven de toetsprestaties, allochtone leerlingen ook niet anders geadviseerd dan autochtone leerlingen. Het antwoord op de hoofdvraag luidde dan ook dat er geen sprake was van systematische onderadvisering van onderscheiden etnische groepen in verschillende gemeentetypen.
In het betreffende onderzoek werd tevens ingegaan op niet-deelname van scholen en leerlingen aan de Cito Eindtoets. Er waren namelijk aanwijzingen dat sommige scholen potentieel laagpresterende leerlingen ten onrechte niet lieten deelnemen, om zo niet het schoolgemiddelde omlaag te halen (vgl. Luyten & Bosker, 2004). Het bleek dat het voor de hoogte van het advies niet uitmaakte of *hele* scholen niet of wel deelnamen aan de Cito-toets. Wanneer echter de school wel deelnam, maar leerlingen van die school niet, trad er wel een negatief effect op voor die leerlingen. Het bleek namelijk dat dergelijke niet-deelnemende leerlingen een lager advies kregen dan andere wel-deelnemende leerlingen met een vergelijkbare taal-, reken- en leesvaardigheid. Niet-deelnemende leerlingen leken dus te worden ondergeadviseerd.

4.1.4 Gevolgen van en bezwaren tegen over– en onderadvisering

Zowel over- als onderadvisering worden in het algemeen als een ongewenst iets beschouwd, omdat ze tot negatieve gevolgen zouden kunnen leiden. Vergeleken met leerlingen die wel in overeenstemming met hun prestaties zijn geadviseerd, starten bijvoorbeeld overgeadviseerde leerlingen in principe meteen al met een achterstand in het voortgezet onderwijs. Tesser en Iedema (2001) zagen overadvisering daarom als een belangrijke oorzaak voor hun lagere rapportcijfers en hogere drop-out percentage in het voortgezet onderwijs. Andere onderzoekers zijn echter minder pessimistisch; volgens hen

zijn er ook aanwijzingen dat overadvisering op de wat langere termijn niet altijd negatief hoeft uit te pakken. Het kan voor leerlingen ook een extra uitdaging en stimulans zijn om hun ambities waar te maken en boven hun niveau uit te stijgen (Hustinx, 2002; Koeslag & Dronkers, 1994). Ook nuanceerden Mulder (1993) en Mulder en Suhre (1995) de conclusie van Tesser en Iedema (2001). Uit hun analyse bleek namelijk dat de ongunstiger loopbaan van allochtone leerlingen niet zozeer het gevolg was van een te hoog advies, maar van een te hoge schoolkeuze. Die keuze wordt gemaakt door ouders die vaak onvoldoende zicht hebben op wat er bij de kinderen zelf en in hun thuissituatie voor nodig is om een dergelijke keuze ook te kunnen waarmaken (Smit, Driessen & Doesborgh, 2005; Van der Veen, 2001). Veel leerlingen die te hoog hebben gekozen, blijven zitten. Een te hoge schoolkeuze kan ook aanleiding geven tot motivatieverlies en ondermaats presteren. Van der Werf en Kuyper (2004) concludeerden op basis van gegevens uit de VOCL-cohortstudies uit 1989, 1993 en 1999 dat na de invoering van het vmbo basisscholen hogere adviezen zijn gaan geven. Zij veronderstellen dat dit zou kunnen komen door druk van de ouders die per se niet willen hebben dat hun kind naar het vmbo gaat. Van der Werf en Kuyper signaleren dat de betreffende leerlingen als brugklasser lager scoren op toetsen dan hun voorgangers. De scholen passen hun lesniveau echter niet aan: in het derde leerjaar zijn de prestaties namelijk even hoog als vóór de invoering van het vmbo. Dat dit niveau even hoog blijft wordt veroorzaakt doordat de leerlingen tussentijds vaker dan voorheen zijn uitgevallen en afgestroomd naar een lager schooltype (Kuyper & van der Werf, 2005). Dit duidt dus op negatieve gevolgen van overadvisering.

Er zijn echter ook aanwijzingen voor negatieve gevolgen van onderadvisering. De Boer, Bosker en Van der Werf (2007) lieten recentelijk zien dat deze gevolgen voor het succes in het voortgezet onderwijs fors kunnen zijn. Volgens deze onderzoekers bedraagt het verschil tussen de categorie sterkst ondergeadviseerde en de categorie sterkst overgeadviseerde leerlingen met betrekking tot het bereikte niveau voortgezet onderwijs 1.35 onderwijsjaar, waarvan -0.73 onderwijsjaar voor rekening komt van de ondergeadviseerde leerlingen en +0.63 voor rekening van de overgeadviseerde leerlingen. Hierbij doen zich overigens geen verschillen voor tussen allochtone en autochtone leerlingen. Leerlingen die worden ondergeadviseerd worden daarmee blijvend op achterstand gezet.

Onderadvisering kan ook gezien worden als een vorm van onderbenutting of verspilling van talent (Mulder, Roeleveld & Vierke, 2007; Driessen, Mooij & Doesborgh, 2007). Zeker in een land als Nederland dat sterk afhankelijk is van een bloeiende kenniseconomie is het belangrijk zo efficiënt en effectief mogelijk een uitgebreid kenniskader op te bouwen en in stand te houden. Uit het onderzoek van Mulder e.a. (2007) bleek dat er bij rond de 5% van de leerlingen sprake is van onderbenutting in termen van onderadvisering (dat wil zeggen een lager advies dan wat past bij de toetsscores). Dat betekent dan niet alle potentieel wordt benut.

Men kan ook vanuit een politiek-filosofisch perspectief bezwaar hebben tegen over- en onderadvisering. Dit kan namelijk worden opgevat als een afwijking van het meritocratische principe. Volgens dat principe zou het innemen van een bepaalde (maatschappelijke) positie louter bepaald mogen worden door iemands capaciteiten, talenten en inzet, kortom iemands competenties; factoren als sekse, sociaal milieu en etnische herkomst zouden daarbij dan geen rol mogen spelen (Luyten & Bosker, 2004; Meijnen, 2004). Er kunnen hier twee interpretaties van dit principe worden onderscheiden: een brede en een beperkte. Bij de brede interpretatie worden behalve cognitieve competenties, ook niet-cognitieve competenties meegenomen als indicatoren voor het bepalen van het advies. De omschrijving van 'niet-cognitieve competenties' is ruim, maar vaak worden er aspecten als inzet en motivatie mee bedoeld. Bij de beperkte interpretatie wordt louter gekeken naar het effect van de cognitieve competenties, in casu toetsprestaties, op het advies (Tesser & Iedema, 2001). Volgens die laatste benadering zou de samenhang tussen toetsprestaties en advies dus perfect moeten zijn; volgens de eerste benadering hoeft dat niet het geval te zijn en is er ook een samenhang met bijvoorbeeld inzet en motivatie. Wanneer de samenhangen tussen de cognitieve (en eventueel niet-cognitieve) competenties enerzijds en het advies anderzijds niet perfect zijn, wordt gesproken van een afwijking van het meritocratische principe.

4.2 Onderzoeksvragen

In het bovenstaande hebben we ons gericht op ontwikkelingen op het gebied van over- en onderadvisering, mogelijke oorzaken daarvoor, en gevolgen daarvan. In dit hoofdstuk willen we op basis van de meest recente landelijke gegevens opnieuw de stand van zaken opmaken. We kijken daarbij naar zowel individuele kenmerken van leerlingen als naar kenmerken van de klas en de context. Van belang is ook dat we nadrukkelijk een onderscheid aanbrengen tussen cognitieve en niet-cognitieve leerlingkenmerken. Daarbij maken we gebruik van enkele nieuwe niet-cognitieve leerlingkenmerken, die in eerder onderzoek nog niet beschikbaar waren. Centraal bij dit alles staan uiteraard de doelgroepen van het onderwijsachterstandenbeleid (OAB), over hen gaat immers dit rapport. Nieuw in dit kader is dat we meer dan voorheen allerlei verschillende doelgroepen kunnen onderscheiden (zie het inleidende hoofdstuk).

We willen de volgende onderzoeksvragen beantwoorden:

- In hoeverre verschillen de adviezen voortgezet onderwijs voor de categorieën van leerlingen die in het kader van het onderwijsachterstandenbeleid worden onderscheiden?
- Welke verschillen zijn er tussen deze leerlingcategorieën ten aanzien van een aantal cognitieve competenties (toetsprestaties) en niet-cognitieve competenties (attitudes, gedrag, karaktertrekken)?
- Hoe verhouden zich deze cognitieve en niet-cognitieve competenties tot elkaar bij de verklaring van verschillen in adviezen?
- Welke relatie bestaat er met klas- en contextuele kenmerken (klassesamenstelling en -niveau, gemeentegrootte)?
- Is er ten aanzien van sommige categorieën van leerlingen sprake van onder- of overadvisering, ofwel krijgen sommige categorieën lagere dan wel hogere adviezen dan op basis van de prestaties verwacht zou worden?
- Wijken de adviezen van leerlingen en scholen die niet deelnemen aan de Cito Eindtoets basisonderwijs af van leerlingen en scholen die wel deelnemen?

4.3 Methode

4.3.1 Steekproef

Voor de beantwoording van de onderzoeksvragen hebben we gebruik gemaakt van gegevens die in het schooljaar 2007/08 zijn verzameld bij de eerste meting van het COOL[5-18] cohortonderzoek (Driessen e.a., 2009). Aan deze meting hebben 550 basisscholen deelgenomen met in totaal ruim 38000 leerlingen in de groepen 2, 5 en 8. Binnen de totale COOL-steekproef kan een onderscheid worden gemaakt naar een landelijk representatieve steekproef van 400 scholen en een aanvullende steekproef van 150 scholen met een oververtegenwoordiging van leerlingen uit achterstandssituaties. Omdat het in het onderhavige onderzoek niet gaat om het geven van een representatief beeld, maar om het bepalen van samenhangen tussen prestaties en advies en het verklaren van verschillen in adviezen, zijn de analyses uitgevoerd op de totale steekproef van leerlingen uit groep 8. Deze keuze garandeert ook voor relatief kleine categorieën van leerlingen een adequate celvulling (vgl. Luyten, Cremers-Van Wees & Bosker, 2001).

De totale COOL-steekproef telt 11609 leerlingen uit groep 8. Voor 89% van deze leerlingen hebben de scholen gegevens doorgegeven over het advies voortgezet onderwijs en – voor zover afgenomen – de Cito-scores. Deze respons bleek niet selectief te zijn met betrekking tot de sociaal-etnische achtergrond van de leerlingen. Met andere woorden: er zijn geen verschillen in respons tussen categorieën van leerlingen met verschillende sociaal-etnische achtergronden (Driessen e.a., 2009). Het oorspronkelijke bestand telde een – per kenmerk variërend – aantal ontbrekende waarden. We hebben deze zoveel mogelijk per samenhangend blok van kenmerken aangevuld door middel van imputatie van schattingen die zijn verkregen via Missing Values Analysis (MVA). In het geval een leerling niet heeft deelgenomen aan de Cito-toets zijn diens Cito-scores echter niet aangevuld. Ook hebben we de leerlingen waarvan de OAB-categorie niet kon worden bepaald niet meegenomen. In de analyses onderscheiden we twee niveaus, namelijk leerlingen binnen klassen. We hebben klassen onderscheiden en niet scholen (i.c. de complete groep 8 binnen een school), omdat we ervan uitgaan dat de adviesprocedure die een leerkracht volgt kan variëren naar gelang de specifieke samenstelling van zijn of haar klas (en dat er zich dus verschillen tussen klassen binnen een school kunnen voordoen; vgl. Hanushek e.a., 2003). In verband met de betrouwbaarheid van de resultaten hebben we ten slotte klassen met minder dan vijf leerlingen uit het bestand

verwijderd (vgl. Harker & Tymms, 2004). Deze bewerkingen leidden tot twee databestanden, een bestand dat wordt gebruikt om eventuele onder- of advisering op te sporen, en een bestand waarmee daarnaast nog wordt nagegaan wat de invloed is van al-dan-niet deelname aan de Cito Eindtoets. Dit laatste bestand bevat gegevens van 9099 leerlingen in 544 klassen (van 440 scholen, dus gemiddeld 1.2 klassen per school). We zijn nog nagegaan of er verschillen zijn tussen de scores op alle kenmerken in het oorspronkelijke bestand en die in het via imputatie van ontbrekende waarden aangevulde bestand. Daaruit bleek dat er zich op nagenoeg geen enkel kenmerk significante verschillen ($p<0.01$) voordeden en voor de paar kenmerken waarbij dat wel het geval was, ging het steeds om slechts rond de 1% verschil in aantallen. We kunnen dus concluderen dat het uiteindelijke voor de analyses gebruikte bestand een goede afspiegeling vormt van het oorspronkelijke bestand. Het eerste (beperktere) bestand dat we gebruiken, en dat de basis vormt voor de analyse van onder- en overadvisering, bevat alleen leerlingen die aan de Cito-toets hebben deelgenomen. Dat bestand telt 7269 leerlingen in 443 klassen van 366 scholen.

4.3.2 Instrumenten en kenmerken

De gegevens zijn verzameld met behulp van uiteenlopende instrumenten; het betreft de volgende soorten informatie:

- het advies voortgezet onderwijs en de scores op de Cito Eindtoets basisonderwijs, verkregen via de scholen;
- de taal-, lees- en rekenvaardigheid, die met behulp van toetsen bij de leerlingen zijn gemeten;
- niet-cognitieve leerlingkenmerken (zoals attitude en gedrag), verkregen via leerkrachtbeoordelingen;
- achtergrondgegevens van de leerlingen, afkomstig uit de schooladministratie;
- klas-, school- en contextuele kenmerken, verkregen via aggregatie van leerlingkenmerken, respectievelijk via de scholen en het ministerie van Onderwijs.

In Tabel 4.1 volgt een korte aanduiding van elk van de kenmerken die worden geanalyseerd.

Tabel 4.1 Overzicht instrumenten en variabelen advies, cognitieve en niet-cognitieve competenties, en achtergrondkenmerken van leerlingen, klassen en scholen

Advies	vmbo pro-bbl, (2) vmbo bbl-kbl, (3) vmbo kbl-tl, (4) vmbo tl-havo, (5) havo/vwo-vwo [1]
Cognitieve competenties	
Cito Eindtoets	200 items; gestandaardiseerde totaalscore
Cito Eindtoets Taal	100 items; somscore
Cito Eindtoets Rekenen	60 items; somscore
Cito Eindtoets Studievaardigheden	40 items; somscore
COOL Taal LWS	Cito Leeswoordenschat M8; 32 items, vaardigheidsscore
COOL Taal DMT	Cito Drie Minuten Toets, kaart 3; 150 items, aantal goed
COOL Lezen BL	Cito Begrijpend lezen 4; 50 items, vaardigheidsscore
COOL Rekenen RW	Cito Rekenen/Wiskunde M8; 50 items, vaardigheidsscore
Niet-cognitieve competenties [2]	
Onderpresteren	Schaal; 3 items: (1) beslist onwaar, (2) onwaar, (3) niet onwaar/niet waar, (4) waar, (5) beslist waar [3]
Gedrag	Schaal; 4 items: (1) - (5)
Werkhouding	Schaal; 3 items: (1) - (5)
Afhankelijkheid	Schaal Leerling-Leerkracht Vragenlijst; 5 items: (1) - (5)
Conflict	Schaal Leerling-Leerkracht Vragenlijst; 5 items: (1) - (5)
Nabijheid	Schaal Leerling-Leerkracht Vragenlijst; 5 items: (1) - (5)
Ouderbetrokkenheid [4]	Schaal; 3 items: (1) - (5)
Extravert vs. introvert	Five Factor Personality Inventory; 5-puntsschaal [5]
Onverdraagzaam vs. verdraagzaam	
Ordelijk vs. niet-ordelijk	
Emotioneel vs. niet-emotioneel	
Autonoom vs. niet-autonoom	
Achtergronden leerling	
OAB-categorie [6]	(1) autochtoon 1.0, met ho [hoger onderwijs], (2) autochtoon 1.0, mbo, (3) autochtoon 1.25-ex, (4) autochtoon 1.25 (nu 0.3), (5) autochtoon 1.25-zwaar (nu 1.2), (6) allochtoon 1.0, met ho, (7) allochtoon 1.0, met mbo, (8) allochtoon 1.9-ex, (9) allochtoon 1.9 (nu 0.3), (10) allochtoon 1.9-zwaar (nu 1.2)
Sekse	(0) jongen, (1) meisje
Loopbaanvertraging	Indicatie zittenblijven [7]: (0) nee, (1) ja
Deelname Cito Eindtoets	(1) school wel, leerling ook, (2) school wel, leerling niet, (3) school niet, leerling niet
Achtergronden klas en school	
Aandeel autochtone achterstandsleerlingen	Het % autochtone achterstandsleerlingen in de klas [8]
Aandeel allochtone achterstandsleerlingen	Het % allochtone achterstandsleerlingen in de klas [9]
Cognitief niveau klas	Gemiddelde prestaties COOL-toetsen [10]; gemiddelde z-score
Gemeentegrootte	Van plaats van vestiging school: (1) G3 [11], (2) G32, (3) overig

Noten bij Tabel 4.1

[1] Pro=praktijkonderwijs, lwoo=leerwegondersteunend onderwijs, bbl=basisberoepsgerichte leerweg, kbl=kaderberoepsgerichte leerweg, gl=gemengde leerweg, tl=theoretische leerweg. Deze vijfdeling is een indikking van oorspronkelijk 15 categorieën: vmbo-pro + vmbo-pro/lwoo + vmbo-lwoo + vmbo-lwoo/bbl = 1; vmbo-bbl + vmbo-bbl/kbl + vmbo-kbl = 2; vmbo-kbl/gl + vmbo-gl + vmbo-gl/tl+ vmbo-tl = 3; vmbo-tl/havo + havo = 4; havo/vwo + vwo = 5. Ze vertoont een perfect-lineaire relatie met de taal-, lees- en rekenprestaties; het verschil tussen de nominaal-metrische en metrisch-metrische samenhangen (*eta*, resp. *r*) bedraagt maximaal 0.02.

[2] Van de niet-cognitieve competenties zijn alleen de eerste drie in eerder onderzoek gebruikt; alle andere (Afhankelijkheid t/m Autonoom) zijn nieuw.

[3] De zeven schalen binnen dit blok zijn zodanig geconstrueerd dat een hoge score duidt op het in sterke mate van toepassing zijn van het betreffende kenmerk. Een hoge score op 'onderpresteren' wil bijvoorbeeld zeggen dat er in sterke mate sprake is van onderpresteren; een hoge score op 'gedrag' en 'werkhouding' duidt op positief gedrag, respectievelijk een goede werkhouding.

[4] Eigenlijk is 'ouderbetrokkenheid' van een andere orde dan de voorafgaande zes kenmerken; het is namelijk een gezinskenmerk, terwijl de andere zes leerlingkenmerken zijn. Bij de bepaling van de hoogte van het advies kan de leerkracht rekening houden met zijn of haar inschatting van de mate van ondersteuning die het kind thuis waarschijnlijk zal ontvangen.

[5] De leerkracht is gevraagd het kind te typeren tussen de twee voorgegeven polen. Het middelste punt is daarbij aangeduid als 'tussenin'. Een score onder de 3 betekent voor bijvoorbeeld de eerste karaktertrek dat een kind meer extravert dan introvert is, terwijl een score ver boven de 3 betekent dat een kind duidelijk meer introvert dan extravert is.

[6] Voor een verantwoording van dit kenmerk, zie hoofdstuk 1 van deze rapportage.

[7] Op basis van leeftijd per 1 oktober van het schooljaar 2007/08.

[8] Geaggregeerd; gebaseerd op de OAB-categorieën 3, 4 en 5.

[9] Geaggregeerd; gebaseerd op de OAB-categorieën 8, 9 en 10.

[10] Geaggregeerde scores van de vier COOL-toetsen.

[11] De gemeenten Amsterdam, Rotterdam en Den Haag; er namen geen scholen uit de gemeente Utrecht deel aan de eerste COOL-meting.

4.3.3 Analyse–opzet

Bij de analyse en rapportage onderscheiden we een beschrijvend en een verklarend deel. Eerst geven we een beschrijving van de bivariate samenhangen tussen de centrale kenmerken. Hiervoor gebruiken we variantie-analyse. Daarbij worden allereerst de samenhangen berekend tussen het advies en de cognitieve en niet-cognitieve competenties, en vervolgens worden advies en competenties gerelateerd aan de achtergrondkenmerken van leerlingen, klassen en scholen. Deze beschrijvende analyses zijn bedoeld om inzicht te verkrijgen in de verschillende samenhangen en vormen tevens de opmaat voor het opstellen van modellen voor de verklarende analyses. Voor deze verklarende (multivariate) analyses maken we gebruik van multilevel-analyse. Daarbij proberen we verschillen in de hoogte van het advies te verklaren uit de uiteenlopende competenties en achtergrondkenmerken, om zo meer zicht te krijgen op een eventuele over- en onderadvisering bij de onderscheiden categorieën van leerlingen.

4.4 Resultaten

4.4.1 Beschrijvende analyses

4.4.1.1 Adviezen en cognitieve en niet–cognitieve competenties

In Tabel 4.2 relateren we de verschillende competenties aan het advies voortgezet onderwijs. Uit deze gegevens kan bijvoorbeeld worden afgeleid of leerlingen met een laag advies anders scoren qua toetsprestaties en houding en gedrag dan leerlingen met een hoog advies.[15] De tabel geeft ook inzicht in de deelname aan de Cito Eindtoets. We vermelden het percentage leerlingen in de totale steekpoef dat de toets heeft gemaakt, en het percentage leerlingen dat de toets niet heeft gemaakt terwijl op de school de toets wel wordt afgenomen. In de tabel staan ook het aantal leerlingen per adviescategorie en de gemiddelden en standaarddeviaties van elk van de kenmerken van de totale steekproef vermeld. Om een indicatie te geven van de sterkte van de samenhangen tussen het advies en de overige variabelen, vermelden we de *eta*, waarbij we als ondergrens voor relevantie uitgaan van een coëfficiënt van minimaal 0.15.

[15] We gaan hierbij niet uit van een strikt causale relatie (vgl. Pijl, 1994). Advies en competenties zijn ongeveer tegelijkertijd gemeten, waardoor het meer verweven componenten zijn dan dat de een een verandering in de ander *veroorzaakt*. Dat leerlingen met hoge prestaties in het algemeen ook een hoog advies krijgen, komt volgens die zienswijze dus gewoon omdat het allebei aspecten zijn van een goed-functionerende leerling. Het gaat er dan ook meer om te bepalen of een bepaald advies *consistent* is met een bepaald prestatieniveau, en welke factoren een eventuele inconsistentie kunnen verklaren (vgl. Dronkers e.a., 1998).

Tabel 4.2 Cognitieve en niet-cognitieve competenties en Cito-deelname naar advies (gemiddelden)

	Advies							
	vmbo pro-bbl	vmbo bbl-kbl	vmbo kbl-tl	vmbo tl-havo	havo-vwo	totaal	*sd*	*eta*
Cito Eindtoets	515.9	523.3	530.6	537.6	544.3	532.8	10.4	0.86
Cito Taal	50.6	58.8	67.1	75.8	84.5	70.1	14.2	0.75
Cito Rekenen	26.1	33.2	39.9	46.5	52.5	41.9	11.2	0.73
Cito Studievaardigheden	19.0	22.9	27.2	31.1	34.9	28.3	6.7	0.75
COOL Taal LWS	99.7	104.5	109.3	114.1	122.3	111.6	14.0	0.51
COOL Taal DMT	76.0	82.1	84.1	89.4	94.5	86.6	15.5	0.37
COOL Lezen BL	36.3	43.1	50.4	58.6	71.1	54.4	16.0	0.70
COOL Rekenen RW	105.1	110.3	114.5	119.3	124.9	116.3	9.4	0.66
Onderpresteren	2.5	2.7	2.7	2.4	2.1	2.5	0.9	0.28
Gedrag	3.4	3.5	3.6	3.7	3.9	3.6	0.9	0.19
Werkhouding	2.8	3.0	3.3	3.6	4.0	3.4	0.9	0.41
Afhankelijkheid	2.5	2.3	2.2	2.0	1.7	2.1	0.8	0.29
Conflict	2.0	1.9	1.9	1.7	1.6	1.8	0.8	0.19
Nabijheid	3.4	3.4	3.5	3.5	3.5	3.5	0.7	0.07
Ouderbetrokkenheid	2.9	3.1	3.4	3.7	4.0	3.5	0.9	0.39
Extravert - introvert	2.7	2.6	2.6	2.6	2.7	2.6	1.2	0.04
Onverdraagzaam - verdraagzaam	3.3	3.3	3.4	3.6	3.8	3.5	1.1	0.15
Ordelijk - niet-ordelijk	3.3	3.2	3.0	2.7	2.3	2.8	1.1	0.31
Emotioneel - niet-emotioneel	2.6	2.8	2.8	3.0	3.1	2.9	1.0	0.15
Autonoom - niet-autonoom	3.4	3.1	2.8	2.5	2.1	2.7	1.0	0.40
Loopbaanvertraging (%)	58.7	37.7	30.3	15.6	5.9	25.4	43.5	0.36
Deelname Cito schl.+ ll.+ (%)	86.1	84.2	82.5	79.9	81.3	82.3	38.2	0.05
Deelname Cito schl.+ ll.- (%)	14.0	2.9	2.8	0.7	0.2	2.9	16.8	0.23
n totaal	890	1580	2397	2202	2030	9099		
%	9.8	17.4	26.3	24.2	22.3	100.0		
n Cito-toets	659	1293	1922	1748	1647	7269		
%	9.1	17.8	26.4	24.0	22.7	100.0		

De tabel maakt allereerst duidelijk dat er een sterk monotone samenhang bestaat tussen de vijf adviescategorieën en de meeste van de competenties. We zijn nagegaan of deze samenhang lineair is (door de *eta* met de *r* te vergelijken) en dat blijkt voor alle kenmerken het geval te zijn (één uitzondering vormt onderpresteren). Dit betekent dat naarmate leerlingen een hogere score op de competenties hebben (zowel de cognitieve als niet-cognitieve), ze ook een hoger advies hebben ontvangen. Hierop zijn er twee uitzonderingen: omdat de kenmerken ordelijk en autonoom 'negatief' zijn geformuleerd, gaat hier een hogere score samen met een lager advies. De samenhang met de Cito-toets[16] is met een correlatie van 0.86 zeer sterk; de correlaties met de deeltoetsen zijn wat zwakker. Bij de COOL-toetsen is de samenhang het sterkst met Begrijpend lezen en het zwakst met de Drie Minuten Toets. Dit laatste ligt voor de hand, aangezien het bij die toets om technisch lezen gaat en daarvan mag worden uitgegaan dat het merendeel van de leerlingen in groep 8 die vaardigheid onder de knie hebben. Binnen het eerste blok van de niet-cognitieve competenties (het leerlingenprofiel) zijn vooral de inschatting van de leerkracht van de werkhouding van de leerling en betrokkenheid van diens ouders bij het onderwijs van belang. Binnen het tweede blok (de vijf persoonlijkheidstrekken) is er vooral een samenhang met het oordeel van de leerkracht over de mate van autonomie en ordelijkheid. In totaal heeft ongeveer een kwart van de leerlingen vertraging opgelopen (door m.n. doubleren). Voor leerlingen met een pro-bbl advies ligt het percentage doubleerders echter ongeveer tien keer zo hoog als voor leerlingen met een havo-vwo advies.

Ruim 82% van de leerlingen in de steekproef heeft de Cito-toets gemaakt. Van de leerlingen met een vmbo pro-bbl advies heeft een relatief groot deel de toets niet gemaakt, terwijl de betreffende school die toets wel afneemt (14.0% tegenover 0.2% van de leerlingen met een havo-vwo advies). Daar kunnen legitieme redenen voor zijn. De categorie pro-bbl advies bestaat uit leerlingen die een pro- of lwoo-advies hebben. Voor hen geldt dat ze moeten voldoen aan een criterium voor leerachterstand en daarom vaak al getoetst zijn voordat een eindtoets wordt afgenomen. Scholen laten deze leerlingen dan dikwijls niet meer meedoen met een eindtoets omdat ze dat overbodig vinden (waarom een kind lastig vallen met (nog) een toets als het advies al vaststaat?). Een andere

[16] De score op de Cito-toets van alle leerlingen van alle deelnemende scholen bedroeg 535.4 (Cito, 2008). Dat de score in onze steekproef wat lager ligt spreekt voor zich, omdat in COOL achterstandsscholen met opzet zijn oververtegenwoordigd.

reden kan zijn dat scholen vinden dat de eindtoets te moeilijk is voor de betreffende leerlingen en hen de frustratie willen besparen.

Voor enkele kenmerken zijn we nagegaan of er binnen de groep van scholen die de Cito-toets af heeft genomen verschillen zijn tussen leerlingen waarvan geen Cito-scores beschikbaar en leerlingen waarvan wel scores beschikbaar zijn. De resultaten van deze analyses staan in Tabel 4.3.

Tabel 4.3 Advies, COOL–toetsprestaties en loopbaanvertraging naar deelname aan Cito–toets (gemiddelden) (steekproef van scholen die hebben deelgenomen aan de Cito–toets; n=7486)

	Cito–deelname			
	ja	nee	totaal	*eta*
Advies:				
· vmbo pro-bbl (%)	9.1	49.3	10.2	0.22
· vmbo bbl-kbl (%)	17.8	17.5	17.8	0.00
· vmbo kbl-tl (%)	26.4	25.8	26.4	0.00
· vmbo tl-havo (%)	24.0	5.5	23.5	0.07
· havo-vwo (%)	22.7	1.8	22.1	0.08
gemiddeld advies	3.3	1.9	3.3	0.19
COOL Taal LWS	111.6	102.5	111.3	0.11
COOL Taal DMT	86.9	74.4	86.6	0.14
COOL Lezen BL	54.5	39.1	54.1	0.16
COOL Rekenen	116.7	106.3	116.4	0.19
Loopbaanvertraging (%)	25.2	51.6	26.0	0.10

Uit Tabel 4.3 kan worden afgelezen dat er tussen leerlingen met en zonder Cito-score van scholen die de Cito-toets hebben afgenomen enkele verschillen bestaan. Het belangrijkste is dat van de leerlingen zonder Cito-score bijna de helft een vmbo pro-bbl advies heeft gekregen en slechts een fractie van hen een havo- of vwo-advies. Ook zijn de COOL-toetsscores van leerlingen zonder Cito-score redelijk wat lager, vooral qua rekenvaardigheid. Ten slotte is er een fors verschil met betrekking tot loopbaanvertraging: van degenen zonder Cito-score heeft ruim de helft vertraging tegen een kwart van degenen met een Cito-score.

Met name de verschillen qua advieshoogte en loopbaanvertraging vormen een ondersteuning van de stelling dat niet-deelname goed te verklaren is.[17]

4.4.1.2 Adviezen, competenties en achtergrondkenmerken

Hierna presenteren we de resultaten van een serie analyses waarbij we de samenhangen tussen enerzijds het advies en de cognitieve en niet-cognitieve competenties en anderzijds enkele *leerlingkenmerken* hebben onderzocht. Bij deze vergelijkingen gaan we na of er relevante verschillen zijn tussen de onderscheiden categorieën. We noemen een verschil relevant bij een *eta* van minimaal 0.15.

Centraal in deze rapportage staan de gegevens van de hier onderscheiden OAB-categorieën. In Tabel 4.4 geven we per categorie een overzicht van de scores op het advies en de competenties.

[17] We konden geen betrouwbare toetsing uitvoeren op verblijfsduur. Het feit dat er vooral verschillen zijn qua taalvaardigheid biedt echter wel ondersteuning voor de aanname dat niet-deelname terecht was.

Tabel 4.4 Advies, cognitieve en niet-cognitieve competenties en Cito-deelname naar OAB-categorie (gemiddelden)

	OAB autochtoon					OAB allochtoon					tot.	*eta*
	1.0 ho	1.0 mbo	1.25 ex	1.25 0.3	1.25 1.2	1.0 ho	1.0 mbo	1.9 ex	1.9 0.3	1.9 1.2		
Advies (%)												
vmbo pro-bbl	1.9	4.5	8.9	16.1	19.6	5.8	9.8	13.3	18.0	19.9	9.8	0.22
vmbo bbl-kbl	5.2	12.7	19.7	27.9	28.6	9.1	20.0	20.0	26.3	24.2	17.4	0.21
vmbo kbl-tl	17.5	26.9	29.8	29.7	28.6	17.8	26.2	28.3	30.7	30.9	26.3	0.11
vmbo tl-havo	29.7	30.1	23.6	17.8	12.9	34.0	28.6	23.2	15.2	16.3	24.2	0.14
havo-vwo	45.6	25.8	18.0	8.5	10.2	33.2	15.4	15.2	9.7	8.7	22.3	0.32
gemiddeld	4.1	3.6	3.2	2.7	2.7	3.8	3.2	3.1	2.7	2.7	3.3	0.41
Cito Eindtoets	539.4	535.1	532.3	529.0	526.5	535.3	531.3	530.5	527.9	527.8	532.8	0.40
Cito Taal	78.9	73.7	69.5	65.9	62.8	73.2	68.0	66.5	64.1	62.5	70.1	0.39
Cito Rekenen	47.5	43.3	41.5	38.3	36.0	44.5	40.5	40.6	37.9	38.9	41.9	0.30
Cito Studievaard.	32.1	29.7	28.1	26.3	25.0	29.9	27.5	27.2	25.5	25.3	28.3	0.36
COOL Taal LWS	118.3	114.9	111.2	109.5	107.8	110.7	109.1	106.8	104.6	104.0	111.6	0.34
COOL Taal DMT	89.0	87.1	85.3	83.7	83.8	90.9	87.9	88.0	86.1	85.8	86.6	0.12
COOL Lezen BL	63.8	57.4	53.4	48.8	47.6	57.3	51.9	50.3	46.8	47.1	54.4	0.37
COOL Reken RW	120.5	117.6	115.9	113.5	112.7	118.4	114.7	114.9	112.4	113.9	116.3	0.28
Onderpresteren	2.4	2.4	2.4	2.5	2.6	2.5	2.6	2.5	2.7	2.6	2.5	0.11
Gedrag	3.8	3.7	3.7	3.5	3.5	3.6	3.6	3.6	3.5	3.5	3.6	0.12
Werkhouding	3.6	3.5	3.4	3.2	3.2	3.6	3.4	3.4	3.3	3.3	3.4	0.15
Afhankelijkheid	1.9	2.0	2.1	2.2	2.2	2.0	2.1	2.2	2.3	2.2	2.1	0.13
Conflict	1.7	1.7	1.8	1.9	1.9	1.8	1.9	1.8	2.0	1.9	1.8	0.12
Nabijheid	3.5	3.6	3.5	3.5	3.5	3.4	3.5	3.4	3.4	3.3	3.5	0.10
Ouderbetrokenh	4.1	3.8	3.6	3.3	2.9	3.8	3.4	3.3	3.0	2.8	3.5	0.44
Extrovert - introvert	2.7	2.6	2.6	2.6	2.5	2.7	2.6	2.7	2.6	2.6	2.6	0.04
Onverdraagzaam - verdraagzaam	3.7	3.6	3.5	3.4	3.2	3.5	3.5	3.5	3.4	3.4	3.5	0.12
Ordelijk - niet-ordelijk	2.7	2.8	2.9	3.1	3.1	2.6	2.8	2.7	2.8	2.9	2.8	0.11
Emotioneel - niet-emotioneel	3.0	2.9	2.8	2.7	2.8	2.9	3.0	3.0	2.9	3.0	2.9	0.10
Autonoom - niet-autonoom	2.4	2.6	2.8	2.9	3.0	2.4	2.6	2.8	2.8	2.8	2.7	0.18
Loopbaanvertr. %	13.7	16.5	21.0	31.5	36.9	20.3	29.5	36.9	40.2	45.8	25.4	0.25
Deeln. Cito schl.+ ll+ (%)	76.1	78.3	81.4	82.1	88.2	84.2	90.5	90.7	90.0	90.0	82.3	0.14
Deeln. Cito schl.+ ll.- (%)	1.1	1.9	2.0	4.2	7.1	1.5	4.8	2.6	3.1	5.6	2.9	0.10

Als het gaat om de adviezen zien we dat er redelijk grote verschillen bestaan tussen de onderscheiden categorieën. Afgaande op de *eta* doen zich die verschillen vooral voor bij het hoogste en de twee laagste adviezen. De *eta* voor het verschil tussen de gemiddelde adviezen bedraagt 0.41. Deze gemiddelde scores laten enkele interessante (en verrassende) conclusies toe. Op de eerste plaats zijn er – gemiddeld genomen – geen verschillen tussen de nieuwe autochtone en allochtone gewichtscategorieën; de lichte en zware achterstandscategorieën (0.3 en 1.2) hebben allemaal een gemiddelde van 2.7. Er doen zich daarentegen wel verschillen voor tussen de autochtonen en allochtonen met het gewicht 1.0 (voor ho 4.1 vs. 3.8, resp. voor mbo 3.6 vs. 3.2). Het verschil tussen de 1.25 ex en 1.9 ex categorie is met 3.2 vs. 3.1 marginaal. De grootste verschillen worden veroorzaakt doordat er onder de beide 1.0-categorieën zeer grote verschillen zijn tussen autochtonen en allochtonen. Krijgt bijvoorbeeld binnen de autochtone 1.0 ho-categorie 1.9% een pro-advies, binnen de allochtone tegenhanger gaat het om 5.8%. En krijgt binnen de autochtone 1.0 ho-categorie 5.2% een bbl-advies, binnen de allochtone tegenhanger is dat 9.1%. Aan de andere kant krijgt 45.6% van de autochtone 1.0 ho-categorie een havo-vwo advies, tegen slechts 33.2% van allochtone 1.0 ho-categorie. Deze bevindingen maken duidelijk dat er enerzijds binnen sterk verschillende achterstandscategorieën (0.3 en 1.2, zowel autochtoon als allochtoon) slechts relatief geringe verschillen bestaan in adviezen, terwijl er anderzijds tussen gelijke niet-achterstandscategorieën (1.0) sterke (etnische) verschillen bestaan in adviezen.
Verder zien we hier nogmaals (vergelijk hoofdstuk 2 over de ontwikkeling van de onderwijspositie elders in dit rapport) dat de scores op de Cito Eindtoets binnen zowel de autochtone als de allochtone categorieën monotoon teruglopen, en tevens dat daarbij de allochtonen (m.u.v. de categorie 1.9 1.2) steeds lager scoren dan de autochtonen. Een soortgelijk beeld ontstaat ten aanzien van de COOL-toetsscores. De kleinste verschillen doen zich overigens voor op het gebied van het technische lezen, zoals gemeten met de Drie Minuten Toets.
Binnen de twee blokken met niet-cognitieve competenties zijn de verschillen tussen de OAB-categorieën in het algemeen betrekkelijk klein. Bij het eerste blok niet-cognitieve competenties treden opvallend grote verschillen op met betrekking tot de inschatting van de leerkracht over de ouderbetrokkenheid (*eta* = 0.44). Het verschil tussen de hoogste en laagste autochtone OAB-categorie bedraagt liefst 1.3 standaarddeviatie; voor de allochtonen gaat het

om 1.1 standaarddeviatie. Bij het tweede blok zijn er verschillen wat betreft het kenmerk autonomie (*eta* = 0.18). Volgens hun leerkrachten zijn autochtone en allochtone niet-achterstandskinderen wat autonomer dan achterstandskinderen.
Er zijn wel grote verschillen in loopbaanvertraging: naarmate de achterstand groter is, is de vertraging ook groter, en bovendien is de vertraging onder allochtonen veel groter dan onder autochtonen. Bij de autochtonen heeft de ongunstigste categorie bijna drie keer zo vaak vertraging opgelopen als de gunstigste categorie; bij de allochtonen is het ruim twee keer zo vaak.
Wat betreft Cito-deelname zien we dat relatief veel 0.3- en 1.2-leerlingen de toets niet hebben gemaakt terwijl die toets op hun school wel wordt afgenomen; dit speelt meer nog voor de autochtone dan de allochtone leerlingen. Verrassend is het relatief hoog aandeel allochtone 1.0 mbho-leerlingen dat de toets niet hebben gemaakt; het betreffende percentage ligt veel hoger dan dat van hun autochtone categorie-genoten.
Een van de verklaringen die in het verleden werden aangedragen voor de overadvisering van allochtone leerlingen was dat de groepsleerkrachten in groep 8 toen nog te weinig ervaring hadden met allochtone leerlingen en nog niet precies duidelijk hadden wat hun capaciteiten waren. Als gevolg van deze onzekerheid zouden de leerkrachten allochtone leerlingen vaker een dubbeladvies geven dan hun autochtone jaargenoten. Om deze verklaring te toetsen zijn we nagegaan of er zich ten aanzien van het aantal dubbeladviezen verschillen voordoen tussen de OAB-categorieën. In Tabel 4.5 staan de verdelingen van de oorspronkelijk door de scholen aangeleverde 15 adviescategorieën. In de onderste rij hebben we de percentages dubbeladviezen opgeteld voor de range vmbo-bbl/kbl – havo/vwo.

Tabel 4.5 Enkel- en dubbeladviezen naar OAB-categorie (in %; verticaal gepercenteerd)

	OAB autochtoon					OAB allochtoon					tot.
	1.0 ho	1.0 mbo	1.25 ex	1.25 0.3	1.25 1.2	1.0 ho	1.0 mbo	1.9 ex	1.9 0.3	1.9 1.2	
vmbo-pro	0.1	0.1	0.4	1.3	2.4	0.4	0.9	0.6	0.8	1.2	0.6
vmbo-pro/lwoo	0.0	0.1	0.0	0.2	0.4	0.0	0.0	0.2	0.0	0.1	0.1
vmbo-lwoo	1.8	4.2	7.9	13.9	16.5	5.4	8.9	12.2	16.6	18.0	8.8
vmbo-lwoo/bbl	0.0	0.2	0.6	0.6	0.4	0.0	0.0	0.2	0.6	0.6	0.3
vmbo-bbl	0.9	2.6	3.7	6.4	9.8	2.5	4.0	6.5	7.8	7.4	4.2
vmbo-bbl/kbl	0.4	1.4	2.4	2.6	0.8	0.0	2.2	1.1	3.3	1.7	1.6
vmbo-kbl	3.9	8.7	13.6	18.9	18.0	6.6	13.8	12.4	15.2	15.1	11.6
vmbo-kbl/gl	0.5	0.6	0.8	1.7	7.5	0.4	0.3	1.1	0.3	1.5	1.1
vmbo-gl	2.2	4.9	5.5	9.7	5.1	3.3	3.1	3.6	12.2	10.8	5.9
vmbo-gl/tl	1.5	2.1	2.5	2.0	2.7	2.1	7.1	6.3	1.7	1.2	2.3
vmbo-tl	13.3	19.3	21.0	16.4	13.3	12.0	15.7	17.3	16.6	17.3	17.0
vmbo-tl/havo	7.9	10.6	9.6	7.7	3.5	17.4	10.5	10.8	6.9	8.5	9.1
havo	21.8	19.4	13.9	10.1	9.4	16.6	18.2	12.4	8.3	7.8	15.1
havo/vwo	17.7	13.0	10.9	5.1	6.7	17.8	8.0	7.6	6.1	5.4	10.8
vwo	27.9	12.8	7.1	3.4	3.5	15.4	7.4	7.6	3.6	3.3	11.5
totaal dubbeladviezen	28.1	27.9	26.8	19.9	22.0	37.8	28.0	27.2	18.8	19.1	25.3

Uit de tabel kan worden afgelezen dat ruim een kwart van de leerlingen een dubbeladvies krijgt. Als we naar de aantallen per OAB-categorie kijken zien we dat zich tussen de autochtone en allochtone categorieën eigenlijk twee verschillen voordoen. De autochtone 1.2-categorie krijgt wat vaker een dubbeladvies dan de allochtone 1.2-categorie. Het grootste verschil zit echter in de 1.0 ho-categorie, met 28.1% dubbeladviezen voor de autochtonen en 37.8% voor de allochtonen. Dit verschil wordt geheel veroorzaakt door het hogere aandeel vmbo-tl/havo-adviezen voor allochtone 1.0 ho-leerlingen.
Op het niveau van de leerlingen hebben we niet alleen gekeken naar verschillen wat betreft de OAB-categorieën, maar ook naar verschillen met betrekking tot de *sekse* van de leerling. We besteden alleen aandacht aan kenmerken waarvoor *eta* voldoet aan het relevantiecriterium van ≥ 0.15. (We zullen hierna kenmerken die niet voldoen aan dit criterium niet in de tabel opnemen en ook niet bespreken in de tekst.) Jongens scoren iets hoger op de Cito rekentoets

(*eta* = 0.17). Volgens hun groepsleerkrachten is er bij hen ook wat vaker sprake van onderpresteren (*eta* = 0.17), zijn ze wat minder sociaal in hun gedrag (*eta* = 0.20), is hun werkhouding duidelijk minder goed (*eta* = 0.29), en is er bij hen vaker sprake van een conflictueuze en ook minder positief-emotionele relatie met de leerkracht (*eta* = 0.18, resp. 0.19). Het grootste verschil is dat leerkrachten jongens minder ordelijk vinden dan meisjes (*eta* = 0.31).
Behalve leerlingkenmerken, zijn ook *klaskenmerken* geanalyseerd. Voor de beschrijvende analyses hebben we de betreffende drie kenmerken elk in drie categorieën verdeeld. Wat betreft de aandelen autochtone en allochtone achterstandleerlingen zijn dat de categorieën: 0%, 1-24% en ≥25%. Ten aanzien van het cognitief niveau zijn drie categorieën gevormd van ongeveer gelijke omvang: laag (33.3%), midden (33.5%) en hoog (33.2%). In de tabellen hierna presenteren we alleen díe kenmerken, waarop er relevante verschillen (eta ≥ 0.15) zijn.
Wat het aandeel autochtone achterstandsleerlingen in de klas betreft kunnen we kort zijn: op geen enkel kenmerk zijn er relevante verschillen (de maximale eta bedraagt 0.09). Met betrekking tot het aandeel allochtone achterstandsleerlingen in de klas zijn er wel enkele relevante verschillen; deze staan in Tabel 4.6.

Tabel 4.6 Advies, cognitieve en niet-cognitieve competenties naar aandeel allochtone achterstandsleerlingen (gemiddelden; alleen verschillen met eta ≥ 0.15)

	Aandeel allochtone achterstandsleerlingen				
	0%	1–24%	≥25%	totaal	*eta*
Advies:					
· havo-vwo (%)	28.6	22.5	14.2	22.3	0.15
gemiddeld advies	3.6	3.3	3.0	3.3	0.21
Cito Eindtoets	535.4	533.1	529.7	532.8	0.23
Cito Taal	74.0	70.6	65.7	70.1	0.25
Cito Rekenen	43.9	42.0	39.7	41.9	0.16
Cito Studievaardigheden	29.9	28.6	26.5	28.3	0.21
COOL Taal LWS	114.8	111.5	107.7	111.6	0.21
COOL Lezen BL	58.2	54.2	49.7	54.4	0.22
COOL Rekenen RW	117.8	116.4	114.3	116.3	0.16
Ouderbetrokkenheid	3.8	3.6	3.2	3.5	0.27
Loopbaanvertraging (%)	17.7	23.4	36.8	25.4	0.19
Deelname Cito school (%)	72.7	87.2	89.9	82.3	0.21

De gegevens in Tabel 4.6 laten zich eenvoudig samenvatten: naarmate het aandeel allochtone achterstandsleerlingen in een klas groter is, is hun positie qua advies, Cito- en COOL-toetsprestaties, ouderbetrokkenheid en loopbaanvertraging ongunstiger.[18] Voor al deze samenhangen geldt overigens dat de relatie lineair is. Verder blijkt dat op scholen met geen allochtone achterstandsleerlingen relatief vaak de Cito-toets niet wordt afgenomen.
In Tabel 4.7 staan de relevante verschillen naar cognitief niveau van de klas. Hiervoor gelden dezelfde conclusies als voor het aandeel allochtone achterstandsleerlingen.

[18] We willen hier benadrukken dat het bij deze analyses om *ongecorrigeerde* uitkomsten gaat. Hierdoor kan het lijken dat er sterke effecten van klaskenmerken optreden. Voor een belangrijk deel zijn die effecten natuurlijk voorspelbaar gezien de samenhang met individuele achtergrondkenmerken (als er bijvoorbeeld veel laagpresterende kinderen in een klas zitten, zal er ook een negatief effect van het prestatieniveau van de klas optreden). In een volgende paragraaf, waarin de resultaten van de verklarende analyses worden gepresenteerd, zullen we corrigeren voor individuele achtergrondkenmerken en de zuivere effecten van klaskenmerken presenteren.

Tabel 4.7 Advies, cognitieve en niet-cognitieve competenties naar cognitief niveau van de klas (gemiddelden; alleen verschillen met eta ≥ 0.15)

	Cognitief niveau klas				
	laag	midden	hoog	totaal	*eta*
Advies:					
· vmbo pro-bbl (%)	17.0	8.6	3.7	9.8	0.19
· havo-vwo (%)	11.3	21.9	33.7	22.3	0.22
gemiddeld advies	2.9	3.3	3.8	3.3	0.29
Cito Eindtoets	528.5	532.9	536.9	532.8	0.33
Cito Taal	64.3	70.1	76.1	70.1	0.33
Cito Rekenen	38.1	42.2	45.4	41.9	0.26
Cito Studievaardigheden	26.0	28.4	30.7	28.3	0.28
COOL Taal LWS	105.8	111.8	117.2	111.6	0.33
COOL Taal DMT	83.2	86.1	90.5	86.6	0.19
COOL Lezen BL	47.7	54.5	60.9	54.4	0.34
COOL Rekenen RW	112.8	116.5	119.8	116.3	0.30
Ouderbetrokkenheid	3.3	3.5	3.8	3.5	0.22
Loopbaanvertraging (%)	33.3	25.0	17.7	25.4	0.15

Behalve leerling- en klaskenmerken analyseren we ook een *contextkenmerk*, namelijk de gemeentegrootte. De betreffende gegevens staan in Tabel 4.8.

Tabel 4.8 Advies, cognitieve en niet-cognitieve competenties naar gemeentegrootte (gemiddelden; alleen verschillen met eta ≥ 0.15)

	Gemeentegrootte				
	G3	G32	overig	totaal	*eta*
Gemiddeld advies	3.0	3.1	3.5	3.3	0.15
Cito Eindtoets	530.8	530.9	534.1	532.8	0.15
Cito Taal	68.0	67.4	72.0	70.1	0.15
Ouderbetrokkenheid	3.3	3.4	3.7	3.5	0.18
Loopbaanvertraging (%)	33.7	32.2	20.4	25.4	0.15

Het aantal relevante verschillen naar gemeentegrootte is zeer beperkt; bovendien voldoen deze ook nog maar net aan het criterium voor relevantie. Buiten de grote steden worden wat hogere adviezen gegeven, ligt de score op de Cito Eindtoets en taaltoets wat hoger, zijn ouders wat meer betrokken, en is er minder sprake van loopbaanvertraging. De verschillen tussen de G3 en G32 zijn daarbij overigens maar minimaal.

4.4.2 Verklarende analyses

In de vorige paragraaf hebben we een beschrijving gegeven van de bivariate samenhangen tussen de verschillende kenmerken uit dit onderzoek. In deze paragraaf analyseren we deze relaties multivariaat. We maken daarvoor gebruik van het multilevel-programma MLwiN 2.20 (Rasbash, Steele, Browne & Goldstein, 2009).

De analyses worden uitgevoerd op twee bestanden, namelijk een bestand met alleen gegevens van leerlingen waarvan de Cito-scores beschikbaar zijn, en een bestand met gegevens van alle leerlingen. In de eerste serie analyses staan de totaalscores op de Cito Eindtoets centraal, daarna die op de drie afzonderlijke Eindtoetsonderdelen. De reden hiervoor is dat effecten mogelijk verschillen per toetsonderdeel. In de tweede serie analyses richten we ons op de scores op de vier COOL-toetsen. We voeren afzonderlijke analyses uit met de COOL-toetsen, omdat die door *alle* leerlingen zijn gemaakt, terwijl de Cito Eindtoets door een deel van leerlingen niet is gemaakt.

In onze analyses gaan we ook na of er een effect is van het niet-deelnemen van individuele leerlingen en het niet-deelnemen van hele scholen aan de Cito-toets. In de vorige paragraaf hebben we al enkele bivariate samenhangen met deze twee kenmerken beschreven. Voor de multivariate analyses doet zich echter het probleem voor dat we deze effecten niet zonder meer kunnen toetsen, omdat de Cito-scores juist ontbreken. Luyten en Bosker (2004) hebben dit probleem opgelost door deze ontbrekende scores te imputeren via schattingen op basis van de scores op andere toetsen die waren afgenomen. Omdat wij de Cito-scores zo zuiver mogelijk willen houden, hebben wij de effecten van niet-deelname getoetst in analyse van de modellen met de COOL-toetsen.

Het bovenstaande leidt tot toetsing van drie series van modellen, namelijk:

1a alleen leerlingen die aan de Cito-toets hebben deelgenomen, met de totaalscores;

1b idem, maar dan met de drie Cito-deeltoetsscores;

2 alle leerlingen, met de vier COOL-toetsscores.

Binnen de eerste twee series met de Cito-toetsen worden telkens 8 modellen getoetst die als volgt worden opgebouwd:

0 Welk deel van de verschillen heeft te maken met verschillen tussen leerlingen, en welk met verschillen tussen klassen?
1 Wat zijn de verschillen in advies naar OAB-categorie?
2 Resteren er nog verschillen in advies tussen de OAB-categorieën nadat rekening wordt gehouden met cognitieve competenties, i.c. toetsprestaties (1a: Cito totaal; 1b: Cito-deeltoetsen)?
3 Als 2, maar nu nadat rekening wordt gehouden met niet-cognitieve competenties.
4 Als 2, maar nu met zowel cognitieve als niet-cognitieve competenties samen.
5 Als 4, maar nu met toevoeging van enkele leerlingkenmerken.
6 Als 5, maar nu met toevoeging van enkele klaskenmerken.
7 Als 6, maar nu met toevoeging van het contextkenmerk gemeentegrootte.

In de derde serie analyses met de COOL-toetsen wordt tot en met model 5 dezelfde opbouw aangehouden. Daarna wordt in model 6 deelname aan de Cito Eindtoets toegevoegd, en in model 7 zowel de drie klaskenmerken als gemeentegrootte.

De resultaten van de eerste serie analyses met de Cito totaalscores presenteren we in Tabel 4.9. De tabel vermeldt de ongestandaardiseerde regressiecoëfficiënten. Wanneer een coëfficiënt statistisch significant is op 0.001-niveau is die vet gezet, en wanneer die significant is op 0.01-niveau is die cursief gezet. In verband met het grote aantal leerlingen dient hier eigenlijk $p<0.001$ als criterium te worden aangehouden (vgl. Luyten & Bosker, 2004; Luyten, 2004). Bij de dichotome en gedummificeerde kenmerken hebben we steeds aangegeven wat de referentiecategorie ('ref.') is. Onder in de tabel is steeds aangegeven hoeveel procent van de variantie in het advies wordt verklaard, dan wel additioneel wordt verklaard door toevoeging van extra variabelen in het model. Door de specifieke opbouw (in model 2 alleen cognitieve competenties, in model 3 alleen niet-cognitieve competenties) gaat het om de volgende modelvergelijkingen: model 1 ten opzichte van 0; 2 ten opzichte van 1; 3 ten opzichte van 1; 4 ten opzichte van 2; 5 ten opzichte van 4; 6 ten opzichte van 5; 7 ten opzichte van 6.

Om qua coëfficiënten tot een eenvormige presentatie te komen, hebben we de scores voor de percentages autochtone en allochtone achterstandsleerlingen door 10 gedeeld, zodat deze coëfficiënten de verandering in advies per 10 procent verandering in het aandeel autochtone, respectievelijk allochtone achterstandsleerlingen uitdrukt.

Tabel 4.9 Resultaten multilevel-analyse advies voortgezet onderwijs, met de Cito Eindtoets (ongestandaardiseerde regressiecoëfficiënten en verklaarde varianties; 7269 leerlingen in 443 klassen)

	Model							
	0	1	2	3	4	5	6	7
Regressiecoëfficiënten								
Intercept	**3.33**	**3.62**	**–51.95**	**3.35**	**–47.77**	**–46.64**	**–46.60**	**–46.62**
OAB-categorie								
· aut. 1.0 ho		**0.51**	*0.07*	**0.31**	*0.07*	**0.07**	**0.07**	**0.07**
· aut. 1.0 mbo (= ref.)		ref.	ref.	ref.	ref.	ref.	ref.	ref.
· aut. 1.25 ex		**–0.37**	**–0.08**	**–0.19**	*–0.07*	*–0.07*	*–0.06*	*–0.06*
· aut. 1.25 0.3		**–0.82**	**–0.19**	**–0.42**	**–0.16**	**–0.15**	**–0.14**	**–0.14**
· aut. 1.25 1.2		**–0.92**	-0.08	**–0.45**	-0.05	-0.04	-0.04	-0.04
· all. 1.0 ho		0.17	*0.13*	0.08	*0.11*	*0.11*	0.10	0.11
· all. 1.0 mbo		**–0.34**	-0.02	**–0.20**	-0.02	-0.01	-0.03	-0.02
· all. 1.9 ex		**–0.50**	-0.07	**–0.29**	-0.06	-0.04	-0.06	-0.05
· all. 1.9 0.3		**–0.83**	**–0.13**	**–0.49**	*–0.12*	*–0.11*	*–0.12*	*–0.12*
· all. 1.9 1.2		**–0.83**	**–0.11**	**–0.45**	*–0.09*	*–0.08*	*–0.09*	*–0.09*
Cito Eindtoets (totaal)			**0.10**		**0.10**	**0.09**	**0.09**	**0.09**
Onderpresteren				-0.02	0.02	0.02	0.02	0.02
Gedrag				-0.05	0.01	0.01	0.01	0.01
Werkhouding				**0.34**	**0.07**	**0.07**	**0.08**	**0.08**
Afhankelijkheid				**–0.20**	*–0.04*	*–0.04*	*–0.04*	*–0.04*
Conflict				*0.06*	0.02	0.02	0.02	0.02
Nabijheid				**–0.13**	-0.02	-0.01	-0.01	-0.01
Ouderbetrokkenheid				**0.25**	**0.05**	**0.04**	**0.04**	**0.04**
Extrovert – introvert				0.03	0.00	0.00	0.00	0.00
Onverdraagzaam – verdr.				-0.00	-0.02	-0.02	-0.02	-0.02
Ordelijk – niet-ordelijk				**–0.06**	**–0.03**	**–0.03**	**–0.03**	**–0.03**
Emotioneel – niet-emot.				-0.00	0.00	-0.00	-0.00	-0.00
Autonoom – niet-auton.				**–0.30**	**–0.07**	**–0.07**	**–0.07**	**–0.07**
Loopbaan (ref.=onvertraagd)						**–0.14**	**–0.14**	**–0.14**
Sekse (ref.=jongen)						**–0.05**	**–0.05**	**–0.05**

Vervolg Tabel 4.9

	Model							
	0	1	2	3	4	5	6	7
Cognitief niveau klas							-0.06	-0.06
% Autochtone achterst.lln. /10							*–0.02*	*–0.02*
% Allochtone achterst.lln. /10							-0.00	0.00
Gemeentegrootte								
· G3								0.00
· G32 (=ref.)								ref.
· overig								0.05
Verklaarde varianties (%)								
Klasniveau	11.9	40.1	67.7	35.1	68.6	69.1	69.7	69.7
Leerlingniveau	88.1	16.6	74.2	39.1	75.0	75.3	75.4	75.4
Totaal		19.4	73.4	38.6	74.3	74.5	74.7	74.7
Toename verklaarde variantie (%)								
+ Klasniveau			27.6	-5.0	1.0	0.4	0.7	0.0
+ Leerlingniveau			57.6	22.5	0.8	0.3	0.1	0.0
+ Totaal			54.0	19.2	0.8	0.3	0.2	0.0

Cursief: $p<0.01$, vet: $p<0.001$

Onder het 0-model zien we dat ongeveer 12% van de verschillen in advieshoogte gebonden is aan verschillen tussen klassen en 88% aan verschillen tussen leerlingen.
Onder model 1 staan de adviezen per OAB-categorie weergegeven in termen van afwijkingsscores ten opzichte van de referentiecategorie (i.c. autochtoon 1.0 mbo).[19] Op de categorie allochtoon 1.0 hoger onderwijs na zijn alle effecten sterk significant. Opnieuw wordt duidelijk dat er tussen de huidige lichte en zware achterstandscategorieën (0.3 en 1.2, ongeacht de etnische herkomst) geen of hooguit beperkte verschillen zijn in advies, maar dat er juist wel verschillen zijn tussen de autochtone en allochtone niet-achterstandscategorieën (1.0). De sociaal-etnische achtergrond van de leerlingen (i.c. OAB-categorie) verklaart 40% van de variantie in adviezen op klasniveau en 17% van de variantie op leerlingniveau, in totaal 19%.

[19] Deze gemiddelden kunnen iets afwijken van de gemiddelden zoals die in Tabel 4.4 zijn gepresenteerd doordat in de multilevel-analyses expliciet rekening is gehouden met de geneste structuur (leerlingen binnen klassen) van de data.

Bij model 2 worden de totaalscores op de Cito Eindtoets toegevoegd. Rekening houdend met verschillen in OAB-categorie betekent bijvoorbeeld 10 punten verschil op de Eindtoets een advies dat één niveau hoger ligt. We zien dat doordat nu de cognitieve competenties worden verdisconteerd de effecten van de sociaal-etnische achtergrond van sterk significant naar licht significant of niet-significant gaan en de verschillen in advies tussen de categorieën grotendeels verdwijnen. Dit neemt niet weg dat het advies van de lichte en zware achterstandscategorieën toch nog significant afwijkt van dat van de referentiecategorie. Dit houdt in dat wanneer rekening wordt gehouden met de cognitieve competenties zoals gemeten met de Eindtoets, met name de huidige 0.3-categorie, en daarbinnen de autochtone leerlingen meer nog dan de allochtone leerlingen, in lichte mate wordt ondergeadviseerd. Toevoeging van de Eindtoetsscores leidt ertoe dat 54% extra variantie in het advies wordt verklaard; samen met de OAB-categorie wordt nu bijna driekwart van de variantie verklaard.

Bij model 3 worden in plaats van de cognitieve competenties (de Cito Eindtoets) de twee blokken met niet-cognitieve competenties ingevoerd. Werd met de cognitieve competenties nog 54% extra variantie verklaard, met de 12 niet-cognitieve competenties samen gaat het om niet meer dan 19% extra verklaarde variantie. Binnen het eerste blok blijken vooral de werkhouding en ouderbetrokkenheid positief en afhankelijkheid en nabijheid negatief bij te dragen aan het door de leerkracht gegeven advies. Binnen het tweede blok is het vooral de mate van autonomie. Als we de modellen 3 en 4 vergelijken, dan blijkt dat de cognitieve competenties bijna drie maal zoveel extra variantie verklaren als de niet-cognitieve competenties.

In model 4 worden de cognitieve en niet-cognitieve competenties tegelijkertijd ingevoerd. Daardoor worden de effecten van de niet-cognitieve competenties beduidend zwakker. We zien ook dat de niet-cognitieve competenties minder dan 1% extra variantie verklaren bovenop de cognitieve competenties. In dit en de volgende modellen blijkt dat er, ook nadat rekening is gehouden met de cognitieve en niet-cognitieve competenties en met de leerling-, klas- en contextkenmerken, nog één OAB-categorie resteert waarbij sprake is van onderadvisering, namelijk die van de autochtone lichte achterstandsleerlingen. Daarnaast wordt nog een categorie licht overgeadviseerd, namelijk de kinderen van autochtone hoogopgeleide ouders.

Bij model 5 worden nog twee leerlingkenmerken toegevoegd, loopbaanvertraging en sekse. Beide zijn sterk significant. Vertraagde leerlingen

en meisjes krijgen een iets lager advies. Hoewel deze twee effecten significant zijn, voegen ze praktisch niets aan de verklaringskracht van het vorige model toe.
Uit model 6 volgt dat ook de drie klaskenmerken, de aandelen autochtone en allochtone achterstandsleerlingen en het prestatieniveau van de klas, nagenoeg niets aan verklaringskracht toevoegen.
Model 7 laat zien dat er bij dit alles geen verschillen zijn tussen scholen in onderscheiden gemeentetypen. In de drie grote steden wordt, gegeven de verschillen in competenties en achtergronden, dus niet anders geadviseerd dan in de kleinere steden of in de rest van Nederland.
Alle modellen overziend blijkt dat met de eerste twee kenmerken, de sociaal-etnische achtergrond van de leerlingen en hun Cito Eindtoetsprestaties driekwart van de verschillen in adviezen wordt verklaard. Toevoeging van 18 andere kenmerken levert niet meer dan 1% extra verklaringskracht op. Als het gaat om mogelijke over- of onderadvisering, dan blijkt dat de huidige autochtone lichte achterstandscategorie iets wordt ondergeadviseerd, terwijl de autochtone categorie van hoogopgeleiden iets wordt overgeadviseerd. Daarbij dient overigens te worden aangetekend dat het in statistische zin weliswaar om significante effecten gaat, maar dat deze empirisch gezien maar zeer beperkt zijn.

In Tabel 4.9 keken we naar het effect van de Cito totaalscore. We hebben vergelijkbare analyses verricht met de drie deeltoetsscores in plaats van de totaalscore. Veel verschillen zijn er echter niet. Om die reden laten we weergave in een tabel achterwege. Twee opmerkingen volstaan. Met de deeltoetsen wordt wat minder variantie verklaard dan met de totaaltoets. In het slotmodel worden nu de autochtone lichte en de allochtone lichte en zware achterstandscategorie in beperkte mate ondergeadviseerd en de autochtone categorie van hoger opgeleiden enigszins overgeadviseerd.
In Tabel 4.10 staan de effecten van het al-dan-niet deelnemen van scholen en leerlingen aan de Citotoets centraal. In plaats van de scores op de Citotoets analyseren we nu binnen de totale steekproef de scores op de COOL-toetsen. Bovendien introduceren we het kenmerk Deelname Cito, waarbij we drie categorieën onderscheiden: (1) school wel, leerling ook; (2) school wel, leerling niet; en (3) school niet en leerling niet. De eerste categorie fungeert als referentiecategorie.

Tabel 4.10 Resultaten multilevel-analyse advies voortgezet onderwijs, met de vier COOL-deeltoetsen (ongestandaardiseerde regressiecoëfficiënten en verklaarde varianties; 9099 leerlingen in 544 klassen)

	Model							
	0	1	2	3	4	5	6	7
Regressiecoëfficiënten								
Intercept	**3.31**	**3.59**	**-6.19**	**3.35**	**-5.16**	**-4.59**	**-4.55**	**-4.52**
OAB-categorie								
· aut. 1.0 ho		**0.50**	**0.14**	**0.31**	**0.12**	**0.13**	**0.13**	**0.12**
· aut. 1.0 mbo (= ref.)		ref.	ref.	ref.	ref.	ref.	ref.	ref.
· aut. 1.25 ex		**-0.37**	**-0.11**	**-0.19**	***-0.08***	*-0.08*	*-0.08*	-0.06
· aut. 1.25 0.3		**-0.83**	**-0.30**	**-0.44**	**-0.23**	**-0.22**	**-0.22**	**-0.20**
· aut. 1.25 1.2		**-0.89**	**-0.28**	**-0.41**	**-0.18**	**-0.16**	**-0.16**	*-0.14*
· all. 1.0 ho		*0.20*	*0.13*	0.10	0.10	0.11	0.11	0.11
· all. 1.0 mbo		**-0.36**	-0.08	**-0.21**	-0.07	-0.06	-0.05	-0.06
· all. 1.9 ex		**-0.51**	**-0.15**	**-0.29**	*-0.12*	-0.09	-0.09	-0.09
· all. 1.9 0.3		**-0.82**	**-0.21**	**-0.47**	**-0.17**	*-0.14*	**-0.14**	*-0.14*
· all. 1.9 1.2		**-0.87**	**-0.30**	**-0.48**	**-0.24**	**-0.20**	**-0.19**	**-0.19**
COOL Taal LWS			**0.01**		**0.01**	**0.01**	**0.01**	**0.01**
COOL Taal DMT			**0.01**		**0.01**	**0.01**	**0.01**	**0.01**
COOL Lezen BL			**0.03**		**0.02**	**0.02**	**0.02**	**0.02**
COOL Rekenen RW			**0.05**		**0.05**	**0.04**	**0.04**	**0.04**
Onderpresteren				-0.02	0.02	0.02	0.02	0.02
Gedrag				*-0.05*	-0.00	-0.00	-0.00	0.00
Werkhouding				**0.34**	**0.13**	**0.13**	**0.13**	**0.13**
Afhankelijkheid				**-0.20**	**-0.07**	**-0.06**	**-0.06**	**-0.06**
Conflict				0.06	0.03	0.03	0.03	0.03
Nabijheid				**-0.12**	*-0.04*	-0.03	-0.03	*-0.04*
Ouderbetrokkenheid				**0.24**	**0.10**	**0.08**	**0.08**	**0.08**
Extrovert - introvert				*0.03*	0.01	0.01	0.01	0.01
Onverdraagzaam - verdr.				-0.01	-0.01	-0.02	-0.02	-0.02
Ordelijk - niet-ordelijk				**-0.04**	**-0.03**	**-0.04**	**-0.04**	**-0.04**
Emotioneel - niet-				-0.00	0.00	0.00	0.00	-0.00
Autonoom - niet-				**-0.31**	**-0.12**	**-0.12**	**-0.12**	**-0.12**
Loopbaan (ref.=onvertraagd)						**-0.25**	**-0.25**	**-0.25**
Sekse (ref.=jongen)						**-0.08**	**-0.08**	**-0.08**
Deelname Cito								
. school + ll + (= ref.)							ref.	ref.
. school + ll -							*-0.16*	**-0.18**
. school - ll -							0.04	0.02

Vervolg Tabel 4.10

	Model							
	0	1	2	3	4	5	6	7
Cognitief niveau klas								**–0.21**
% Autochtone achterst.lln. /10								**–0.04**
% Allochtone achterst.lln. /10								*–0.02*
Gemeentegrootte								
· G3								0.03
· G32 (=ref.)								ref.
· overig								0.06
Verklaarde varianties (%)								
Klasniveau	12.6	39.9	60.0	32.0	60.7	61.6	62.0	64.7
Leerlingniveau	87.4	16.1	62.3	37.8	64.8	65.5	65.6	66.1
Totaal	100.0	19.1	62.0	37.0	64.3	65.0	65.1	65.9
Toename verklaarde variantie (%)								
+ Klasniveau			20.1	-7.9	0.7	0.9	0.3	2.8
+ Leerlingniveau			46.2	21.6	2.5	0.8	0.1	0.5
+ Totaal			42.9	17.9	2.2	0.8	0.1	0.8

Cursief: $p<0.01$, vet: $p<0.001$

De analyseresultaten met de COOL-toetsen tot en met model 5 zijn in grote lijnen vergelijkbaar met die met de Cito-deeltoetsen (niet in tabelvorm gepresenteerd). In model 6 wordt het extra effect van deelname aan de Cito Eindtoets bepaald. Vergeleken met leerlingen die die toets hebben gemaakt, krijgen leerlingen op dezelfde scholen die die toets niet hebben gemaakt een iets lager advies. Deze leerlingen[20] worden dus bij vergelijkbare competenties en achtergronden licht ondergeadviseerd. Dit effect blijft bestaan, en wordt zelfs iets sterker, als in model 7 nog de klas- en contextkenmerken worden toegevoegd. Wat deze kenmerken betreft valt het negatieve effect van het cognitieve niveau van de klas op. Rekening houdend met de andere kenmerken krijgen leerlingen in klassen die gemiddeld hoog presteren een significant lager advies dan leerlingen in klassen die gemiddeld lager presteren.

[20] We zijn in deze modellen niet nagegaan of er bij deze onderadvisering een samenhang is met de OAB-categorie, bijvoorbeeld via het opnemen van interactietermen. De reden hiervoor is dat het aantal leerlingen dat de toets niet heeft gemaakt terwijl die op hun school wel wordt afgenomen relatief beperkt is (n=217).

4.5 Conclusies

4.5.1 Samenvatting

In de inleiding hebben we een aantal vragen gesteld. Deze zullen we hierna puntsgewijs beantwoorden.

Verschillen in adviezen tussen OAB–categorieën

Er zijn redelijk grote verschillen tussen de hier onderscheiden tien OAB-categorieën. In de beschrijvende analyses, waarin nog geen rekening is gehouden met de invloed van uiteenlopende leerling-, klas- en contextfactoren, valt op dat er zich geen verschillen voordoen tussen de nieuwe autochtone en allochtone gewichtscategorieën: de lichte en zware achterstandscategorieën (0.3 en 1.2) hebben gemiddeld genomen allemaal eenzelfde advies. Er doen zich daarentegen wel verschillen voor tussen de autochtonen en allochtonen met het gewicht 1.0, c.q. de kinderen van middelbaar en hoger opgeleide ouders. De autochtonen krijgen daarbij een hoger advies dan de allochtonen. Deze bevindingen maken duidelijk dat er enerzijds binnen sterk verschillende achterstandscategorieën (0.3 en 1.2, zowel autochtoon als allochtoon) slechts relatief geringe verschillen bestaan in adviezen, terwijl er anderzijds tussen gelijke niet-achterstandscategorieën (1.0) sterke (etnische) verschillen bestaan in adviezen.

Verschillen in adviezen tussen OAB–categorieën

Er zijn redelijk grote verschillen tussen de hier onderscheiden tien OAB-categorieën. In de beschrijvende analyses, waarin nog geen rekening is gehouden met de invloed van uiteenlopende leerling-, klas- en contextfactoren, valt op dat er zich geen verschillen voordoen tussen de nieuwe autochtone en allochtone gewichtscategorieën: de lichte en zware achterstandscategorieën (0.3 en 1.2) hebben gemiddeld genomen allemaal eenzelfde advies. Er doen zich daarentegen wel verschillen voor tussen de autochtonen en allochtonen met het gewicht 1.0, c.q. de kinderen van middelbaar en hoger opgeleide ouders. De autochtonen krijgen daarbij een hoger advies dan de allochtonen. Deze bevindingen maken duidelijk dat er enerzijds binnen sterk verschillende achterstandscategorieën (0.3 en 1.2, zowel autochtoon als allochtoon) slechts relatief geringe verschillen bestaan in adviezen, terwijl er anderzijds tussen gelijke niet-achterstandscategorieën (1.0) sterke (etnische) verschillen bestaan in adviezen.

Samenhangen OAB–categorie met cognitieve en niet–cognitieve competenties
Wat betreft de cognitieve competenties, i.c. de resultaten op de Cito Eindtoets en de COOL-toetsen, blijkt dat binnen zowel de autochtone als de allochtone categorieën de scores monotoon teruglopen naargelang de zwaarte van de achterstandscategorie, maar dat daarbij de allochtonen vrijwel steeds lager scoren dan de autochtonen. Ook hier doen zich dus etnische verschillen voor. Met betrekking tot een eerste serie niet-cognitieve competenties, met name gedrag en attitudes, is er een zeer lichte trend dat naarmate de achterstand groter is, de score op de competentie iets ongunstiger is, waarbij de positie van de allochtonen in het algemeen iets ongunstiger is dan die van de autochtonen. De verschillen binnen een tweede serie niet-cognitieve competenties, i.c. karaktertrekken, zijn betrekkelijk klein.

De verhouding cognitieve versus niet–cognitieve competenties
Bij de verklaring van verschillen in adviezen wegen de cognitieve competenties ongeveer drie maal zo zwaar mee als de niet-cognitieve competenties. Met de OAB-categorie en de cognitieve competenties wordt bijna 75% van de verschillen in advies verklaard, met de OAB-categorie en niet-cognitieve competenties bijna 40%.

Invloed klas– en contextuele kenmerken
Wanneer er rekening is gehouden met verschillen die te maken hebben met kenmerken van leerlingen, is de invloed van de klas- en contextkenmerken praktisch nihil: in geen van de modellen wordt meer dan 1% extra variantie verklaard.

Onder– of overadvisering
Er kunnen twee interpretaties van onder- en overadvisering worden onderscheiden, een beperkte en een brede. Bij de beperkte interpretatie wordt alleen rekening gehouden met de cognitieve competenties, hier de scores op de Cito Eindtoets. In dat geval blijkt uit de analyses dat met name de huidige lichte achterstandscategorie (0.3) iets worden ondergeadviseerd, de autochtone leerlingen daarbij meer dan de allochtone leerlingen. Bij de brede interpretatie van onder- en overadvisering wordt behalve met de cognitieve competenties ook nog rekening gehouden met niet-cognitieve competenties. Deze kunnen worden opgevat als verklaringen voor de verschillen die nog resteren na verdiscontering van de cognitieve competenties. De analyses maken duidelijk

dat deze nauwelijks enige aanvullende verklaringskracht hebben. Wanneer vervolgens ook nog andere leerling-, klas- en contextkenmerken worden verdisconteerd, dan blijkt uiteindelijk dat alleen nog de autochtone lichte achterstandscategorie (0.3) iets wordt ondergeadviseerd, en de autochtone hoogopgeleide niet-achterstandscategorie (1.0) iets wordt overgeadviseerd. De verschillen zijn bij dit alles weliswaar in statistische zin significant, maar empirisch gezien zijn ze zeer klein. Ze verschillen overigens ook niet naar gemeentegrootte.

Verschillen naar deelname Cito Eindtoets
Vergeleken met leerlingen die de Cito Eindtoets hebben gemaakt, krijgen leerlingen op dezelfde scholen die die toets niet hebben gemaakt een iets lager advies. Deze leerlingen worden dus bij vergelijkbare competenties en achtergronden licht ondergeadviseerd.

4.5.2 Discussie: terugblik
Eind jaren '80 werd er in onderzoek voor het eerst op geattendeerd dat verschillende categorieën van leerlingen bij een gelijk prestatieniveau verschillende adviezen voor voortgezet onderwijs kregen. Uit die en enkele latere studies kwam naar voren dat er sprake was van overadvisering van allochtone leerlingen: bij gelijke prestaties op de Cito Eindtoets ontvingen zij een hoger advies dan autochtone leerlingen. Begin 21e eeuw bleek dat die overadvisering van allochtone leerlingen verdwenen was en dat die plaats had gemaakt voor onderadvisering van autochtone achterstandsleerlingen. Enkele jaren later lieten analyses zien dat bij gelijke eindtoetsprestaties allochtone leerlingen niet anders werden geadviseerd dan autochtone leerlingen.
In al deze studies stond het onderscheid allochtoon – autochtoon centraal. In de onderhavige studie is er meer differentiatie aangebracht en is er binnen de allochtone en autochtone leerlingencategorieën nog een nader onderscheid gemaakt naar mate van achterstand op basis van het ouderlijke opleidingsniveau. Het bleek dat bijna driekwart van de verschillen in advies herleid kan worden tot verschillen in eindtoetsprestaties. Voor zover er zich toch nog discrepanties tussen prestaties en advies voordeden, ging het om een geringe onderadvisering van de lichte autochtone achterstandscategorie en om een nog geringere overadvisering van de autochtone niet-achterstands-categorie.

In de loop der jaren zijn verschillende verklaringen geopperd voor over- dan wel onderadvisering. Gewezen kan worden op procedure-afspraken in de grote gemeenten met betrekking tot de overgang van basis- naar voortgezet onderwijs, op de grotere ervaring die leerkrachten met het onderwijs aan allochtone leerlingen hebben gekregen, op het veranderde, minder-tolerante politieke klimaat en op de gevolgen van de marktwerking binnen het onderwijs. Dat momenteel de autochtone achterstandsleerlingen (licht) worden ondergeadviseerd heeft mogelijk te maken met het feit dat deze categorie de laatste decennia toch relatief weinig aandacht heeft gekregen, met name ten gevolge van de sterke aandacht voor de allochtone achterstandsleerlingen (Vogels & Bronneman-Helmers, 2003). Of de grotere aandacht voor deze laatste groep ook feitelijk ten koste is gegaan van de aandacht van de autochtone achterstandsleerlingen is overigens de vraag.
De discussie rond over- en onderadvisering kan ook in een ruimer perspectief worden geplaatst. De vraag waar het dan uiteindelijk om draait is wat de grootste voorspellende waarde heeft voor het niveau dat de leerlingen in het voortgezet onderwijs (en wellicht daarna) weten te realiseren. Is dat het advies, dat deels is gebaseerd op subjectieve inschattingen en waarbij ook uiteenlopende individuele leerlingkenmerken en contextuele factoren een rol spelen, of is dat een objectieve meting van de capaciteiten of het prestatieniveau via een (eind)toets? Deze vraag is momenteel erg actueel, gelet op het feit dat in de politiek het plan is gelanceerd om voor *alle* leerlingen een eindtoets verplicht te stellen, ook dus voor de zwakkere leerlingen die nu nog verstek kunnen laten gaan omdat ze een pro- of lwoo-advies hebben. Recent longitudinaal onderzoek waarin de relatieve invloed van advies en toetsscores op de langere termijn wordt bepaald is echter niet beschikbaar. Relevant in dit verband zijn wel de bevindingen uit het onderzoek van De Boer e.a. (2007) waaruit blijkt dat leerlingen die worden ondergeadviseerd, zowel de autochtone als allochtone, in hun verdere schoolloopbaan blijvend op achterstand worden gezet.

5 Cognitieve reserves bij doelgroepleerlingen

Joost Meijer en Jaap Roeleveld

5.1 Inleiding

Een achterliggende veronderstelling bij het onderwijsachterstandenbeleid is dat de doelgroepen van dat beleid, door allerlei oorzaken, minder goed presteren in het onderwijs dan eigenlijk zou kunnen. Door de inzet van extra middelen zouden de scholen meer mogelijkheden moeten krijgen om de achterstanden weg te werken; oftewel: om meer uit deze leerlingen te halen dan er tot dan toe uit is gekomen. Een halve eeuw geleden werd dit wel aangeduid met “verborgen talent” (Van Heek, 1968). En meer recent met “onderbenutting van capaciteiten” (Mulder e.a., 2007) en “onderpresteren” (Onderwijsraad, 2007).
In het onderzoek van Mulder e.a. (2007), verricht in opdracht van de Onderwijsraad, werd nagegaan of er aanwijzingen zijn voor de onderbenutting van capaciteiten in het basis- en het voortgezet onderwijs door te zoeken naar discrepanties tussen de potenties van leerlingen en hun schoolresultaten. Daarbij treden twee problemen op: hoe, met welke instrumenten, kunnen we de potentie van een leerling bepalen; en wanneer spreken we van een discrepantie. Bij de instrumenten is het de vraag hoe we een zo scherp mogelijk beeld kunnen krijgen van zowel de werkelijke (cognitieve) mogelijkheden van leerlingen als van hun schoolprestaties. De in het cohortonderzoek gebruikte instrumenten vormen daarbij natuurlijk altijd een momentopname en kennen,

zoals alle instrumenten in sociaal-wetenschappelijk onderzoek, de nodige meetonnauwkeurigheid (in methodologische termen: *onbetrouwbaarheid* of *attenuatie*). Bij de bepaling van discrepanties tussen twee niet geheel nauwkeurig gemeten variabelen wordt deze onnauwkeurigheid nog vergroot. En voor het bepalen van wanneer er sprake is van een discrepantie is geen algemeen aanvaarde standaard voorhanden. Alleen al op basis van toeval zullen er altijd discrepanties zijn en het is niet op voorhand duidelijk wanneer deze groot genoeg zijn om te kunnen spreken van 'onbenut talent'. Het genoemde onderzoeksrapport eindigt dan ook met een waarschuwing: "De aard en omvang van onderbenutting is dus sterk afhankelijk van de gekozen definitie en instrumenten."(Mulder e.a., 2007, p. 83)
In dit hoofdstuk zullen we, met gegevens uit het recente COOL$^{5\text{-}18}$ onderzoek in groep 5 van het basisonderwijs, opnieuw proberen om in kaart te brengen of er sprake is van discrepanties tussen de cognitieve capaciteiten van leerlingen en hun schoolprestaties. En de eigenlijke onderzoeksvraag is vervolgens in hoeverre zulke discrepanties in verschillende mate voorkomen bij de in dit onderzoek onderscheiden leerlinggroepen (waaronder de doelgroepen van het achterstandenbeleid).
Vergeleken met het eerdere PRIMA-onderzoek heeft in COOL$^{5\text{-}18}$ een uitgebreidere meting van de cognitieve capaciteiten van leerlingen in groep 5 plaats gevonden. Hierdoor wordt het mogelijk de relatie tussen cognitieve capaciteiten en schoolprestaties te modelleren met behulp van latente variabelen. Daarmee kan het probleem van de onnauwkeurigheid van de onderzoeksinstrumenten verkleind worden.

Onderbenutting van leerpotentieel
De vraagstelling van het onderzoek betreft een kwestie die al lang de belangstelling heeft van zowel onderwijsbeleidsmakers als onderwijsonderzoekers, namelijk in hoeverre het onderwijs recht doet aan de ontwikkelingsmogelijkheden van een ieder. Onderzoek naar het zogenaamde leerpotentieel van leerlingen is eigenlijk ook op deze vraagstelling gebaseerd, hoewel daarin het accent sterk ligt op de individuele leerling en niet zozeer op de positie van maatschappelijke groepen. In dit type onderzoek wordt vaak gekeken naar de leerwinst die optreedt als gevolg van een training tussen een voormeting en een nameting van cognitieve capaciteit (*pretest – training – posttest design*, zie bijvoorbeeld Resing, 2001), of naar de hoeveelheid benodigde hulp tijdens een test (*train within test design;* (Meijer, Oostdam et al.

2002). In tegenstelling tot conventionele, statische tests worden dergelijke tests wel dynamische tests genoemd (Grigorenko and Sternberg 1998). Wanneer onderbenutting van talent of leerpotentieel gedetecteerd moet worden, zou een dergelijke aanpak erg voor de hand liggen. Er zou dan een vergelijking kunnen worden gemaakt tussen het zelfstandig vermogen van een leerling, zoals gemeten met conventionele tests, en het leerpotentieel van een leerling, gemeten met een zogenaamde dynamische test zoals hiervoor beschreven. Grote discrepanties tussen deze beide metingen zouden dan kunnen wijzen op onderprestaties of overprestaties van leerlingen. Een groot deel van het onderzoek naar leerpotentieel is gebaseerd op de theorie omtrent de zone van naaste ontwikkeling van (Vygotsky 1963).
In dit hoofdstuk zullen we de (ten opzichte van PRIMA) uitgebreidere meting van cognitieve capaciteiten van leerlingen benutten om verschillen in cognitieve reserves bij onderscheiden groepen leerlingen nader te analyseren. Met de tot nu toe beschikbare éénmalige data van deze uitgebreidere meting uit COOL$^{5\text{-}18}$ is de bovenstaande benadering echter niet uit te voeren. Bij longitudinale data zou een voormeting en een nameting kunnen worden gebruikt om de leerwinst als gevolg van genoten onderwijs in de tussenperiode vast te stellen. Aangenomen dat leerwinst aldus geoperationaliseerd als een soort dynamische meting kan worden opgevat, zou deze maat vervolgens vergeleken kunnen worden met een statische maat, zoals de in COOL$^{5\text{-}18}$ gebruikte gestandaardiseerde test voor het meten van de cognitieve capaciteit. We beschikken nu echter pas over een eerste meting onder leerlingen in groep 5 van het basisonderwijs.

Mulder e.a. (2007) definiëren onderbenutting als de discrepantie tussen verwachte en feitelijke prestaties. Als één van de indicatoren[21] voor verwachte prestaties hanteren deze auteurs de score op een (beknopte) intelligentietest van leerlingen en als indicator voor feitelijke prestaties gebruiken zij schoolprestaties. Onderbenutting wordt vervolgens opgevat als een overschatting van feitelijke schoolprestaties zoals voorspeld door de score op de intelligentietest van meer dan een standaarddeviatie (de werkelijke schoolprestaties zijn dus lager dan de voorspelde). Deze benadering nemen we ook voor dit hoofdstuk als uitgangspunt. Zoals eerder aangegeven kent deze benadering, methodologisch gezien, wel enkele problemen. In de eerste plaats

[21] Mulder e.a. gebruiken een hele reeks indicatoren voor onderbenutting. In dit onderzoek beperken wij ons tot de relatie tussen cognitieve capaciteiten en schoolprestaties.

is de gehanteerde grens voor de grootte van de standaarddeviatie arbitrair. Daarnaast is het mogelijk dat de aldus gevonden overschattingen het gevolg zijn van andere factoren, zoals:

1. Meetfouten van beide indicatoren (onbetrouwbaarheid van de metingen);
2. Verschil van structuur van beide indicatoren;
3. Persoonsgebonden onveranderlijke eigenschappen;
4. Persoonsgebonden veranderlijke eigenschappen en veranderlijke karakteristieken van de omgeving (Thorndike 1963).

Om werkelijk tot identificatie van onder- en overpresteren te kunnen komen, zou eigenlijk onderzoek nodig zijn waarin het verband tussen onderpresteren en deze vier factoren in kaart wordt gebracht. Dergelijk onderzoek vergt wel heel veel en is voor zover wij weten nog niet ergens op systematische wijze verricht.

In dit hoofdstuk zullen we proberen de benadering van Mulder e.a. aan te scherpen door het probleem van meetfouten en de structuur van de indicatoren expliciet in de analyses te betrekken en daarmee vervolgens tot een schatting van de mate van over- en onderschatting van prestaties in verschillende groepen te komen. Daarbij zullen we vooral kijken naar de verdeling van over- en onderschatting over de hier onderscheiden groepen (achterstands)leerlingen. We zijn dus niet in de eerste plaats geïnteresseerd in schattingen van het *absolute vóórkomen* van onderbenutting van talent in het Nederlandse basisonderwijs, maar vooral in de schattingen van de *verschillen tussen groepen* wat dit betreft.

5.2 Methode en data

In grote lijnen zullen we eenzelfde procedure volgen als Mulder e.a. in de zin dat we onderpresteren zullen opvatten als een negatieve discrepantie tussen verwachte en feitelijke prestaties. Of, in andere termen: als een overschatting van het feitelijke prestatieniveau (de cognitieve potentie doet hogere prestaties verwachten dan feitelijk worden gerealiseerd). Voor de verwachte prestatie worden subtests van een test voor cognitieve capaciteiten gebruikt; voor de feitelijke prestatie enkele gestandaardiseerde schoolvorderingentoetsen van het Cito. Eerst wordt via regressie-analyse het feitelijk bereikte prestatieniveau voorspeld door het verwachte prestatieniveau; vervolgens wordt met behulp van residu-analyse gekeken naar eventuele systematische verschillen tussen

gemiddelde residuscores van de tien in dit rapport onderscheiden groepen leerlingen. Om het probleem van de onbetrouwbaarheid van de metingen enigszins te ondervangen, wordt niet met geobserveerde scores gewerkt, maar met factorscores.
In de analyses zijn de volgende stappen ondernomen:

1. Eerst is met exploratieve factoranalyse nagegaan in hoeverre de gebruikte subtests en toetsen ondergebracht kunnen worden bij twee factoren, te weten intelligentie en schoolvordering.
2. Vervolgens is de passing van enkele modellen onderzocht met behulp van toetsende (confirmatieve) factoranalyse.
3. Aan de hand van het meest preferente model zijn factorscores berekend.
4. Vervolgens is een regressie-analyse uitgevoerd met de factorscore voor schoolvorderingen als afhankelijke en de factorscore voor intelligentie als onafhankelijke variabele, en is de residuscore (de geobserveerde score minus de voorspelde score) voor elke leerling bepaald.
5. Ten slotte is een variantie-analyse uitgevoerd op de residuscores met de tien onderscheiden (doelgroep)leerlingen als factor. Omdat er geen a priori verwachtingen zijn omtrent de richting van de verschillen tussen de groepen, zijn post hoc paarsgewijze vergelijkingen toegepast om de verschillen tussen de groepen op statistische significantie te kunnen toetsen, waarbij is gecorrigeerd voor kanskapitalisatie.

Instrumenten
In groep 5 van het COOL[5-18] cohortonderzoek zijn in het schooljaar 2007/2008 een aantal instrumenten afgenomen, die we in dit onderzoek gebruiken. Allereerst betreft dit een viertal standaardtoetsen die deel uitmaken van het leerlingvolgsysteem van het Cito:

1. Leeswoordenschat;
2. Drie Minuten Toets voor technisch lezen;
3. Begrijpend lezen;
4. Rekenen/Wiskunde.

Met behulp van deze toetsen zijn dus de schoolprestaties van de leerlingen gemeten.

Daarnaast zijn er een vijftal tests afgenomen die behoren tot de Niet-Schoolse Cognitieve Capaciteiten Test (NSCCT; Batenburg & Van der Werf, 2004):

1. Figuur samenstellen;
2. Exclusie;
3. Getalreeksen;
4. Categorieën;
5. Analogieën.

De eerste drie tests bestaan elk uit 15 items, de laatste twee uit elk twintig items. Deze versie van de NSCCT is speciaal voor het COOL-onderzoek bewerkt. Voor voorbeelditems verwijzen we hier kortheidshalve naar Driessen, Mulder et al., 2009. De NSCCT is geschikt voor een klassikale afname. De test is niet bedoeld om een individueel Intelligentie Quotiënt te bepalen, maar wel geschikt om de cognitieve capaciteiten van grotere groepen te schatten.
In de navolgende analyses beschikken we over gegevens van 8205 leerlingen uit groep 5 in de totale steekproef van COOL[5-18], waarbij al deze toetsen en (sub)tests zijn afgenomen. In Tabel 5.1 geven we de aantallen, onderscheiden naar de in dit onderzoek gehanteerde indeling in OAB-groepen naar gewicht, gebaseerd op het opleidingsniveau van de ouders, en de etnische herkomst.

Tabel 5.1 Aantallen leerlingen uitgesplitst naar OAB–indeling

Gewicht	Autochtoon	Allochtoon	Totaal
1.0, ho	1940	206	2146
1.0, mbo	1777	291	2068
ex 1.25 of ex 1.9	1470	375	1845
nu 0.3	982	213	1195
nu 1.2	189	762	951
Totaal	6358	1847	8205

5.3 Resultaten

5.3.1 De exploratieve factoranalyse

Het ligt voor de hand om de subtests van de NSCCT als indicatoren voor intelligentie te gebruiken en de Citotoetsen als metingen van het latente construct schoolvorderingen. Het is echter bekend dat intelligentie niet zonder meer als één homogene factor kan worden opgevat; met name het verschil

tussen *fluid* en *crystallized intelligence* is veelvuldig onderwerp van onderzoek geweest (Cattell 1971; Carroll 1993; Horn 1994). Met *fluïde* intelligentie wordt de capaciteit om nieuwe problemen op te lossen bedoeld, terwijl *crystallized* intelligentie veel sterker afhankelijk is van ervaring. Het oplossen van figurenreeksen of alfanumerieke reeksen doet bijvoorbeeld een sterk beroep op fluïde intelligentie, terwijl algemene ontwikkeling een typisch product is van gekristalliseerde intelligentie. Daarnaast is woordenschat één van de zuiverste schatters van hetgeen wel verbale intelligentie is genoemd (Meijer and Oostdam 2001) en doet rekenen vaak een beroep op het redeneervermogen of fluïde intelligentie.

Verder is er discussie over de vraag in hoeverre schoolvorderingen en intelligentie conceptueel goed van elkaar te onderscheiden zijn. De correlatie tussen beide is over het algemeen hoog, zoals ook uit analyses op de cohortdata al eens is gebleken (Driessen, 2005). Voor de analyse zoals we hier willen uitvoeren is het van belang dat de NSCCT iets anders meet dan de Citotoetsen.

Om deze redenen is eerst een exploratieve factoranalyse uitgevoerd op de negen subtesten en toetsen. Het aantal factoren is op twee gefixeerd en de extractiemethode was principale componenten. Er is VARIMAX rotatie toegepast om een eenvoudige structuur te forceren. Het is immers zeer onwaarschijnlijk dat beide factoren in werkelijkheid niet correleren. Tabel 5.2 geeft de factorladingen na rotatie[22].

Tabel 5.2 Exploratieve factoranalyse: factorladingen

Subtest of toets	Intelligentie	Schoolvorderingen
Exclusie	.74	.04
Figuren samenstellen	.72	.02
Getallenbewerking	.64	.20
Categorieën	.62	.19
Analogieën	.61	.34
Rekenen/Wiskunde	.54	.50
Drie Minuten Toets	-.14	.78
Begrijpend lezen	.38	.72
Leeswoordenschat	.29	.71

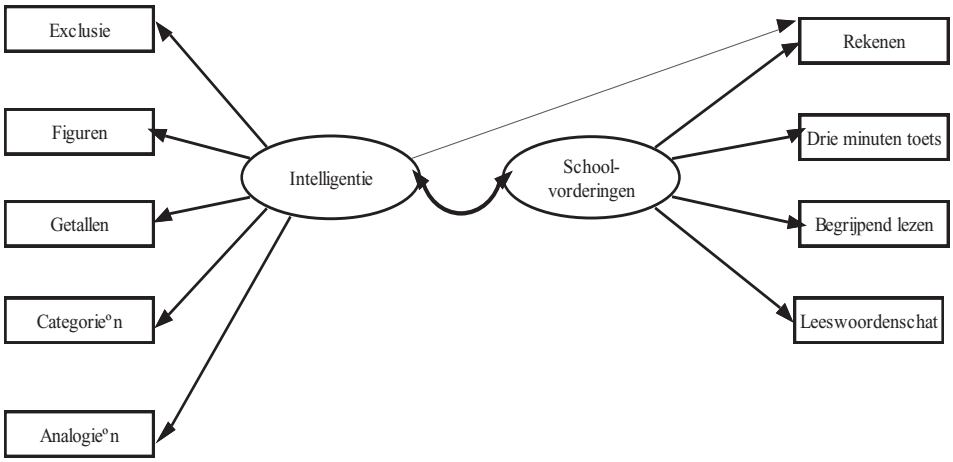
[22] De Kaiser-Meyer-Olkin measure of sampling adequacy bedraagt .87 en geeft aan dat de data zeer geschikt zijn voor factoranalyse.

We zien in deze tabel dat de subtests van de NSCCT alle substantieel laden op de intelligentiefactor. Van de subtests laadt Analogieën daarnaast nog het meest ook op de factor voor schoolvorderingen. De leestoetsen laden alle zeer substantieel op de factor voor schoolvorderingen. Daarnaast lijkt de rekentoets hybride te zijn, i.e., deze laadt op beide factoren.

5.3.2 De confirmatieve factoranalyse

Om te kunnen nagaan of de rekentoets werkelijk hybride is en of dit mogelijk ook geldt voor de Analogiën subtest, zijn verschillende modellen getoetst. In een eerste model is een toetsende (confirmatieve) factoranalyse uitgevoerd onder de aanname dat alle tests en toetsen slechts op één factor laden, dat wil zeggen alle NSCCT-tests alleen op de intelligentiefactor en alle Citotoetsen alleen op de factor voor schoolvorderingen. Dit model is weergegeven in Figuur 5.1.

Figuur 5.1 Model voor confirmatieve factoranalyse[23]



In een tweede model is de factorlading van rekenen op intelligentie vrijgelaten (de gestippelde lijn in Figuur 1). In een derde model is ook de factorlading van analogieën op schoolvorderingen vrijgelaten. Uit de passingindices voor de drie modellen (zie Bijlage A) blijkt dat het loslaten van restricties in principe beter

[23] De gestippelde lijn geeft rekenen als hybride factor weer, i.e. het tweede model.

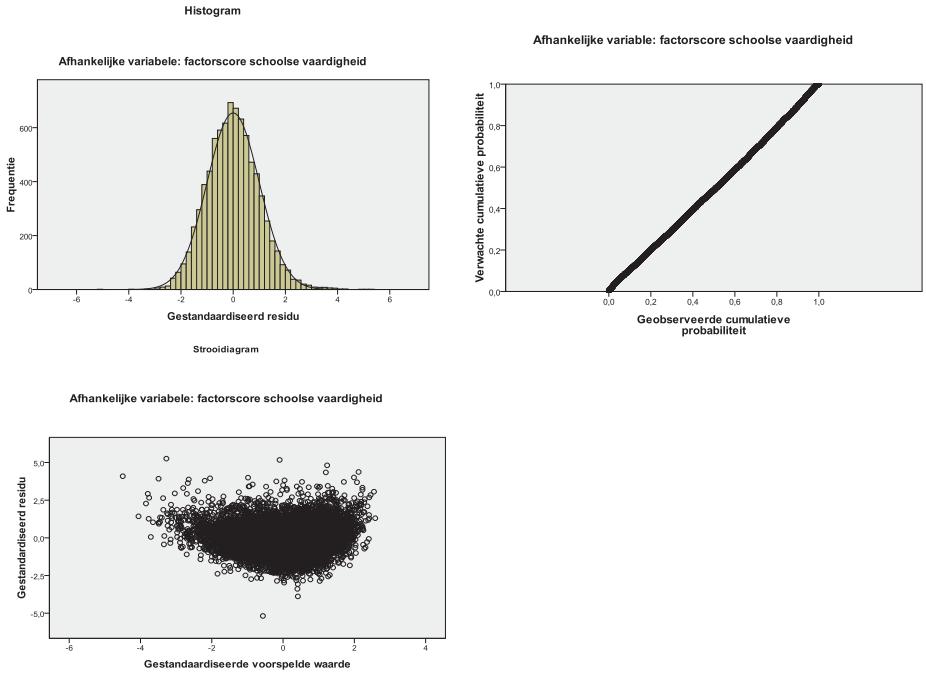
passende modellen oplevert. Omdat de factorlading van de subtest Analogieën op schoolvorderingen in het derde model kleiner is dan .30 (namelijk .24) en omdat het geen verbale analogieën maar figurale analogieën betreft is wegens spaarzaamheid besloten het berekenen van factorscores op het tweede model te baseren.
Daarmee zijn de factorscores voor intelligentie en schoolvorderingen dus berekend onder de assumpties dat:

1. Exclusie, Getallen bewerken, Figuren samenstellen, Categorieën en Analogieën alleen op intelligentie laden;
2. Drie Minuten Toets, Begrijpend lezen, en Leeswoordenschat alleen laden op schoolvorderingen;
3. Rekenen/Wiskunde zowel op intelligentie als schoolvorderingen laadt.

5.3.3 Regressie-analyse

Vervolgens is een regressie-analyse uitgevoerd met de factorscore voor schoolvorderingen als afhankelijke variabele en de intelligentie-factorscore als predictor. De correlatie tussen beide bedraagt .80; intelligentie en schoolvorderingen delen dus 64% van hun variantie. Dit is opmerkelijk hoger dan de gemiddelde correlatie tussen intelligentie en *academic achievement* in de Verenigde Staten, die volgens (Neisser, Boodoo et al. 1996) circa .50 bedraagt. De correlatie tussen factorscores is natuurlijk altijd hoger dan de correlatie tussen ruwe scores, want ze is gecorrigeerd voor de onbetrouwbaarheid in de metingen van die ruwe scores, maar de discrepantie is toch wel groot. Mogelijk speelt mee dat het hier om een relatief jonge groep gaat, namelijk leerlingen uit groep 5 van het basisonderwijs.
Figuur 5.2 geeft enkele karakteristieken van de residuscores, dat wil zeggen de voorspellingsfouten, gedefinieerd als de geobserveerde score minus de voorspelde score.

Figuur 5.2 Enkele karakteristieken van de verdelingen van voorspelde waarden en residuen



Te zien is dat de residuen normaal verdeeld zijn. In het strooidiagram van gestandaardiseerde residuen en voorspelde gestandaardiseerde waarden zijn wel enkele *outliers* te zien, maar de aanname van homoskedasticiteit lijkt niet geschonden (d.w.z. de grootte van de residuen hangt niet systematisch samen met de waarde van de voorspelde score).

5.3.4 Variantie–analyse

De op basis van de voorgaande analyses berekende residu-scores geven de afwijking tussen de voorspelde schoolvorderingen van leerlingen en hun feitelijk behaalde schoolresultaten. Een negatief residu geeft dus aan dat de feitelijke resultaten achterblijven bij wat op basis van de scores op de NSCCT verwacht kan worden en vormt daarmee een indicatie voor het onderbenutten van talent. Door middel van variantie-analyse is nagegaan of er op dit punt systematische verschillen bestaan tussen de in dit onderzoek onderscheiden tien groepen (doelgroep)leerlingen.

De variantie-analyse is uitgevoerd met de ongestandaardiseerde residuscores als afhankelijke variabele en de groepenindeling in tien categorieën als factor. De aanname van homogene varianties tussen groepen wordt niet geschonden (Levene's (9, 8195) = 1.77, n.s.) en het effect van de groepsindeling is statistisch significant ($F(9, 8195) = 31.42$, $p < .001$). De waarde van de effectgrootte bedraagt .032, hetgeen een klein tot medium effect impliceert (.01 is een klein effect, .06 een medium effect).
De volgende tabel geeft een overzicht van de gemiddelde residuscore en enkele andere statistieken per groep.

Tabel 5.3 Beschrijvende statistieken van de residuen

				95% cfi				
	gem	sd	se	onder	boven	min	max	n
auto 1.0, ho	1.58	6.09	.14	1.31	1.85	-16.43	28.87	1940
auto 1.0, mbo	.37	5.97	.14	.09	.65	-18.44	31.55	1777
auto 1.25-ex	-.42	5.88	.15	-.72	-.12	-17.19	31.03	1470
auto 1.25 (nu 0.3)	-.87	5.72	.18	-1.23	-.51	-31.09	20.51	982
auto 1.25-zwaar (nu 1.2)	-1.37	5.89	.43	-2.22	-.53	-13.11	23.26	189
allo 1.0, ho	.41	5.71	.40	-.38	1.19	-13.97	23.55	206
allo 1.0, mbo	-.20	6.18	.36	-.92	.51	-14.94	17.56	291
allo 1.9-ex	-1.00	6.02	.31	-1.61	-.38	-20.33	19.05	375
allo 1.9 (nu 0.3)	-.83	5.92	.41	-1.63	-.03	-17.15	19.82	213
allo 1.9-zwaar (nu 1.2)	-1.93	5.41	.20	-2.31	-1.54	-23.31	23.97	762

gem: gemiddelde; sd: standaardafwijking; se: standaardfout van gemiddelde; cfi: betrouwbaarheidsinterval

Negatieve residuscores duiden op overschattingen; dat wil zeggen dat er hogere toetsscores verwacht zouden kunnen worden op basis van de scores op de NSCCT-tests. Negatieve scores vormen daarmee, zoals gezegd, een indicatie voor onbenut talent. Bij positieve residuscores is er juist sprake van meer schoolvorderingen dan kon worden verwacht.
Er is in de tabel een vrij duidelijk patroon te zien. Hoe lager de opleiding van de ouders, des te lager het gemiddelde residu. Voorts lijken binnen elke opleidingcategorie de allochtone leerlingen in het algemeen een lager gemiddeld residu te hebben dan de autochtone leerlingen. Dat wijst daarmee op meer cognitieve reserve bij allochtone leerlingen.

Ofschoon er vastgesteld kan worden dat de variantie van de residuen systematisch is gebonden aan de (doel)groepindeling, blijkt uit Tabel 5.3 niet welke groepen significant van elkaar verschillen wat betreft gemiddelde residuscore. Om dat te onderzoeken zijn paarsgewijze vergelijkingen gedaan, met de Bonferroni-correctie om kanskapitalisatie te voorkomen. Tabel 5.4 geeft een overzicht van de uitkomsten.

Tabel 5.4 Paarsgewijze verschillen tussen gemiddelde residuscores

	Autochtoon					Allochtoon				
	1.0 ho	1.0 mbo	1.25 ex	1.25 (0.3)	1.25 (1.2)	1.0 ho	1.0 mbo	1.9 ex	1.9 (0.3)	1.9 (1.2)
auto 1.0, ho		+	+	+	+		+	+	+	+
auto 1.0, mbo	-		+	+	+			+		+
auto 1.25-ex	-	-								+
auto 1.25 (nu 0.3)	-	-								+
auto 1.25-zwaar (nu 1.2)	-	-								
allo 1.0, ho										+
allo 1.0, mbo	-									+
allo 1.9-ex	-	-								
allo 1.9 (nu 0.3)	-									
allo 1.9-zwaar (nu 1.2)	-	-	-	-		-	-			

Van links naar rechts: +: gemiddelde groep in rij hoger dan groep in kolom;
–: gemiddelde groep in rij lager dan groep in kolom.
(N.B. + en – zijn elkaars spiegelbeeld).

In de eerste rij van deze tabel valt te lezen dat autochtone leerlingen met ouders die hoger onderwijs hebben genoten, significant hogere residuals hebben dan alle andere groepen, met uitzondering van leerlingen van allochtone ouders met hoger onderwijs. Uit de laatste regel blijkt dat allochtone leerlingen met het huidige zware gewicht (1.2) significant lagere residuals hebben dan alle groepen autochtone leerlingen, uitgezonderd die met ook een zwaar gewicht. Verder is hun gemiddelde ook lager dan van allochtone leerlingen met middelbaar of hoger opgeleide ouders, maar verschilt het niet significant van de twee andere 1.9-groepen (de ex-groep en de 0.3-groep).

Deze paarsgewijze toetsing bevestigt daarmee het eerdere beeld dat leerlingen met ouders met lagere opleidingen doorgaans (significant) negatievere gemiddelde residuals hebben (dus lagere prestaties dan voorspeld) dan leerlingen met middelbaar of hoger opgeleide ouders.
Het lijkt merkwaardig dat de groep allochtone leerlingen met hoog opgeleide ouders alleen significant hogere residuscores heeft ten opzichte van de leerlingen met allochtone ouders die tot de zwaarste (1.2) gewichtscategorie behoorden. Dit komt echter door het relatief kleine aantal leerlingen in de groep leerlingen met allochtone ouders met hoger onderwijs, waardoor de power van de toets op verschillen kleiner is (zie ook de breedte van het betrouwbaarheidsinterval van het gemiddelde in Tabel 5.3).
Het lijkt er dus op dat leerlingen met (relatief) laag opgeleide ouders lagere prestaties hebben dan geschat, terwijl leerlingen met relatief hoger opgeleide ouders in het algemeen beter presteren dan verwacht. Dat zou dus betekenen dat onderbenutting groter zou zijn naarmate leerlingen uit sterker benadeelde sociaal-economische milieus afkomstig zijn. Dit is in overeenstemming met een gedeelte van de literatuur omtrent leerpotentieel, waar ook soms is gevonden dat kinderen met achterstanden meer leerpotentieel vertonen dan kinderen met een normale ontwikkeling (zie bijvoorbeeld (Badad and Budoff 1974).

5.3.5 Etniciteit of leerlinggewicht

Om na te gaan welke dimensie van de doelgroepindeling er het meest toe doet, is een tweewegvariantie-analyse uitgevoerd met etniciteit en gewicht, afgeleid uit de ouderlijke opleiding, als factoren en de residuscore als afhankelijke variabele (Tabel 5.5).

Tabel 5.5 Gemiddelde residu–score en (sd); naar gewichtklasse en etniciteit

Gewicht	Autochtone leerlingen		Allochtone leerlingen		Totaal	
1.0, ho	1.58	(6.09)	.41	(5.71)	1.47	(6.07)
1.0, mbo	.37	(5.97)	-.20	(6.18)	.29	(6.00)
1.25 en 1.9 ex	-.42	(5.88)	-1.00	(6.02)	-.54	(5.91)
nu 0.3	-.87	(5.71)	-.83	(5.92)	-.86	(5.75)
nu 1.2	-1.37	(5.89)	-1.93	(5.41)	-1.82	(5.51)
Totaal	.31	(6.02)	-1.08	(5.81)	0.00	(6.00)

Zowel het effect van etniciteit als het effect van gewichtsklasse is statistisch significant ($F(1, 8195) = 9.25$, $p = .002$ en $F(4, 8195) = 20.69$, $p < .001$, respectievelijk). Het effect van de interactie tussen beide factoren is niet significant. Statistische significantie is echter een slechte raadgever bij grote steekproeven zoals de onderhavige en daarom is ook gekeken naar de effectgroottes. Voor etniciteit bedraagt de partiële eta-kwadraat .001 en voor leerlingewicht .01. Het effect voor gewichtsklasse valt als klein te typeren en dat voor etniciteit als nog veel kleiner: in de orde tien maal zo klein als dat voor gewichtsklasse[24].

Wanneer we naar de randtotalen kijken dan zien we dat de gemiddelde residuscore van allochtone leerlingen significant lager is dan die van autochtone leerlingen (-1.08 versus .31). Wat betreft de gewichtsklassen geven post hoc vergelijkingen met Bonferroni-correctie aan dat alle gewichtsklassen significant van elkaar verschillen, met uitzondering van de groepen die voorheen een gewicht kregen toebedeeld (de ex-groepen; met een gemiddelde van -.54) en de groep die tegenwoordig een gewicht 0.3 krijgt (met een gemiddelde van -.86).

We kunnen concluderen dat er tussen de gewichtsklassen (gebaseerd op de opleidingen van de ouders) meer verschil zit dan tussen de groepen autochtoon versus allochtoon. En er is geen sprake van interactie: het verband tussen gewichtsklasse en gemiddelde residu-score is bij autochtone en allochtone leerlingen vergelijkbaar.

5.3.6 Onderbenutting bij een grens van één standaarddeviatie

Hoewel er dus systematische verschillen zijn tussen verschillende van de hier onderscheiden categorieën van (doelgroep)leerlingen in gemiddelde residu-scores, geven de effectgroottes aan dat er veel meer verschil zit tussen leerlingen *binnen* een zo'n categorie dan er systematische verschillen *tussen* de categorieën zijn. De vraag is wat dit kan betekenen voor een schatting van het aandeel 'onderbenutters', wanneer we daarvoor dezelfde operationele definitie gebruiken als bij Mulder e.a. (2007), namelijk schoolprestaties die één standaarddeviatie (of meer) lager liggen dan verwacht op basis van de cognitieve capaciteiten test.

Wanneer we het verschil nemen tussen de op basis van ons model berekende factorscores voor Intelligentie en schoolvorderingen, dan is er bij 4.7% van de

[24] Ofschoon er aan de effectgrootte η_p^2 interpretatieproblemen kleven, wijken de waarden van η^2 nauwelijks af van die van η_p^2.

leerlingen sprake van schoolvorderingen die tenminste één standaarddeviatie lager zijn dan de factorscore voor intelligentie. En dit percentage varieert niet sterk tussen de onderscheiden groepen (achterstands)leerlingen: van 6.3% bij de autochtone leerlingen met gewicht 1.2 tot 3.9% bij de autochtone leerlingen met hoog opgeleide ouders; deze kleine verschillen zijn statistisch niet significant.
Deze schatting van het aandeel onderbenutters is aanmerkelijk lager dan bij Mulder e.a. Daar werd het aandeel onderbenutters bij verschillende groepen en toetsen steeds op rond de 15% geschat; en bij sommige achterstandsgroepen nog duidelijk hoger (zie bijvoorbeeld Tabel 3 en 4 uit Mulder e.a., pp. 22-24). De hier gebruikte modellering, waarin expliciet rekening wordt gehouden met de onnauwkeurigheid (onbetrouwbaarheid) in de gebruikte instrumenten, leidt dus tot duidelijk lagere schattingen van discrepanties tussen cognitief vermogen en schoolprestaties[25].

5.4 Alternatieve verklaringen?

Er zijn mogelijke alternatieve verklaringen voor de in dit onderzoek gevonden verschillen in residuscores tussen groepen leerlingen. Ten eerste zouden de toetsen voor schoolvorderingen partijdig kunnen zijn in de zin dat zij met name leerlingen van allochtone ouders benadelen. Drie van de vier toetsen betreffen immers de Nederlandse taal. Dit zou op zichzelf geen verklaring vormen voor de overige verschillen die met leerlinggewicht van doen hebben, maar het is mogelijk dat ook hier taal een rol speelt in de zin dat kinderen uit hogere sociaal-economische milieus (met hoger opgeleide ouders) het Nederlands wellicht beter beheersen.
Er zijn ook nog andere mogelijkheden, die we met de beschikbare data nader kunnen verkennen. Zo is het mogelijk dat het model voor de factoranalyse, op grond waarvan de factorscores voor intelligentie en schoolvorderingen zijn berekend, niet gelijkelijk opgaat voor alle tien onderscheiden groepen. En verder is het mogelijk dat het regressiemodel niet hetzelfde is voor de tien groepen. Het is kortom goed mogelijk dat de samenhang tussen intelligentie en schoolvorderingen niet invariant is over groepen.

[25] De lagere schattingen komen door de modellering en niet doordat het hier om een andere dataset gaat dan bij Mulder e.a.; als we bij de hier gebruikte COOL-data van groep 5 uitgaan van de direct gemeten variabelen komen we ook op schattingen van rond de 15% onderbenutting.

5.4.1 Verschillende factorstructuren in verschillende groepen?

Om na te gaan in hoeverre de factorstructuren in de verschillende groepen aan elkaar gelijk gesteld kunnen worden, zijn opnieuw enkele toetsende factoranalyses uitgevoerd. De factorscores zijn berekend onder de aanname dat één en hetzelfde model geldt voor alle tien groepen en het is dus de vraag in hoeverre deze aanname houdbaar is. Er zijn vier modellen geanalyseerd, waarbij steeds meer gelijkheidsresticties zijn opgelegd.

1. De ongestandaardiseerde factorladingen zijn identiek in de groepen;
2. De ongestandaardiseerde factorladingen en de covariantie tussen beide factoren zijn identiek in de groepen;
3. De ongestandaardiseerde factorladingen, de covariantie tussen beide factoren en de varianties van beide factoren zijn identiek in de groepen;
4. De ongestandaardiseerde factorladingen, de covariantie tussen beide factoren, de varianties van beide factoren en de residuele varianties van de tests en toetsen zijn identiek in de groepen.

De details van de passing van deze modellen staan in Bijlage B.
De extra restricties van modellen 3 en 4 leveren significant slechter passende modellen op. Toch zijn de passingsmaten van deze modellen zeker niet onacceptabel te noemen. We mogen dus wel stellen dat de factorstructuren in de tien doelgroepen in voldoende mate gelijk zijn.

5.4.2 Past het regressiemodel in alle groepen?

Het is denkbaar dat de uitkomsten van de residuanalyse mede bepaald zouden kunnen worden door een verschillende sterkte van de samenhang tussen intelligentie en schoolvorderingen in de verschillende groepen. Er is in de residuanalyse gewerkt met één en hetzelfde regressiemodel voor alle groepen. Dus is de aanname dat dat model ook opgaat voor alle groepen.
Deze aanname is weer onderzocht door verschillende modellen te vergelijken. In het eerste model zijn de regressies van schoolvorderingen op intelligentie voor alle groepen apart geschat, dat wil zeggen voor alle groepen vrijgelaten. In het volgende model is de ongestandaardiseerde regressie van school-vorderingen op intelligentie gefixeerd over de groepen, dat wil zeggen er is gepostuleerd dat de samenhang tussen intelligentie en schoolvorderingen identiek is. In het daaropvolgende model is bovendien gesteld dat de residuele varianties van beide factorscores in alle groepen aan elkaar gelijk zijn, hetgeen

bijna volledige equivalentie van het regressiemodel over groepen heen impliceert. De uitkomsten van deze analyse staan in Bijlage C.
Wanneer de regressieparameters apart worden geschat voor elke groep (model 1) worden wel verschillen in samenhang tussen de factorscores voor intelligentie en factorscores voor schoolvorderingen gevonden. Het model met bijna volledige equivalentie van het regressiemodel lijkt echter toch ook redelijk te passen met de data. Maar bij een conservatieve toetsing moet de aanname van equivalentie worden verworpen. Al met al is het niet helemaal duidelijk of het regressiemodel in voldoende gelijke mate opgaat voor alle groepen.

5.5 Conclusies

De resultaten van de analyses laten zien dat er een systematische samenhang bestaat tussen de gemiddelde residuscore en de hier gehanteerde indeling van leerlingen in (doel)groepen. Deze samenhang houdt in dat leerlingen met lager opgeleide ouders tenderen naar sterker negatieve residuscores; ofwel: naar schoolprestaties die lager zijn dan voorspeld kan worden op basis van hun cognitieve capaciteiten. Hoe lager het opleidingsniveau des te meer zou er dus sprake zijn van onderpresteren of onderbenutting.
De trend in deze uitkomsten sluit daarmee aan bij het eerdere onderzoek van Mulder e.a. (2007). Tegelijk hebben we laten zien dat de modellering van intelligentie en schoolvorderingen met latente variabelen (factoren) leidt tot duidelijk kleinere verschillen tussen capaciteiten en prestaties dan de rechtstreekse metingen zoals die door Mulder e.a. zijn gebruikt. Het verdient daarom aanbeveling bij verder onderzoek naar onderbenutting expliciet rekening te houden met de meetonnauwkeurigheid van de gebruikte instrumenten.
De oorzaak van de hier gevonden onderbenutting zou kunnen zijn dat het Nederlandse onderwijs niet voldoende aansluit bij het potentieel van leerlingen met een sociaal-economische achterstand, zoals geoperationaliseerd door opleidingsniveau ouders en etniciteit.
Anderzijds is het echter mogelijk dat de modellen die zijn gebruikt voor het bepalen van de voorspellingsfouten ten onrechte voor elke groep identiek zijn gesteld. Dat geldt zowel voor het factormodel als het voorspellingsmodel. Voor beide modellen werd gevonden dat een model met minder gelijkheidsresticties beter bij de data past. Bij het factormodel kan een (vrijwel) identiek model in

alle groepen wel geaccepteerd worden. Bij het voorspellingsmodel is dit echter minder het geval.
Het is dus gerechtvaardigd te concluderen dat sociaal-economische achterstand samenhangt met onderbenutting, i.e., lagere prestaties dan verwacht. Maar in verband met de genoemde alternatieve verklaringen moet deze conclusie wel met de nodige voorzichtigheid worden benaderd. Het zou interessant zijn om in verder onderzoek naar onderbenutting ook eens gebruik te maken van één van de twee paradigma's die in het onderzoek naar leerpotentieel worden gebruikt. Gebruikmaking van een dynamische test met hetzij een *train within test* opzet, hetzij een *pretest – train – posttest* opzet, kan meer inzicht in de problematiek rond onderbenutting opleveren dan gebruikmaking van statische tests.

Bijlage 5A Uitkomsten confirmatieve factoranalyse

In een eerste model is een toetsende (confirmatieve) factoranalyse uitgevoerd onder de aanname dat alle tests en toetsen slechts op één factor laden, dat wil zeggen alle NSCCT-tests alleen op de intelligentiefactor en alle Citotoetsen alleen op de factor voor schoolvorderingen. In een tweede model is de factorlading van rekenen op intelligentie vrijgelaten (de gestippelde lijn in Figuur 5.1). In een derde model is ook de factorlading van analogieën op schoolvorderingen vrijgelaten. Onderstaande tabel geeft een overzicht van de passingsindices voor elk van de drie modellen.

Tabel A Passingsindices voor drie modellen voor toetsende factoranalyse

Model	chi^2	df	Δchi^2	CFI	RMSEA	90% cfi	p(RMSEA ≤.05)
1	1401	26	-	.93	.080	.077 - .084	<.001
2	950	25	451	.95	.067	.064 - .071	<.001
3	749	24	201	.96	.061	.057 - .064	<.001

Noot. CFI: comparative fit index, 90% cfi: 90% betrouwbaarheidsinterval van RMSEA. Overige verklaringen in de tekst.

Het is duidelijk dat het loslaten van beide restricties beter passende modellen oplevert. De respectieve afnamen van de chi^2-waarden (Δchi^2) zijn in beide gevallen significant. De waarde van de root mean square error of approximation (RMSEA) is voor het eerste model op het randje, en voor beide andere modellen niet optimaal, maar wel acceptabel (Browne and Cudeck 1993).

Bijlage 5B Toetsing van gelijke factorstructuren binnen groepen

Om na te gaan in hoeverre de factorstructuren in de verschillende groepen aan elkaar gelijk gesteld kunnen worden, zijn enkele toetsende factoranalyses uitgevoerd. Er zijn vier modellen geanalyseerd, waarbij steeds meer gelijkheidsrestricties zijn opgelegd:

1. De ongestandaardiseerde factorladingen zijn identiek in de groepen;
2. De ongestandaardiseerde factorladingen en de covariantie tussen beide factoren zijn identiek in de groepen;
3. De ongestandaardiseerde factorladingen, de covariantie tussen beide factoren en de varianties van beide factoren zijn identiek in de groepen;

4. De ongestandaardiseerde factorladingen, de covariantie tussen beide factoren, de varianties van beide factoren en de residuele varianties van de tests en toetsen zijn identiek in de groepen.

In Tabel B staat een overzicht van de passing van deze vier modellen.

Tabel B Passingsindices voor vier tien-groepsmodellen

Model	chi^2	df	Δchi^2	CFI	RMSEA	90% cfi
1	1261	322	-	.942	.060	.056 - .063
2	1276	331	15	.941	.059	.056 - .062
3	1360	349	84	.937	.059	.056 - .063
4	1866	430	506	.911	.064	.061 - .067

Noot. CFI: comparative fit index, 90% cfi: 90% betrouwbaarheidsinterval van RMSEA. Overige verklaringen in de tekst.

Alle extra restricties, behalve die van gelijke covariantie tussen intelligentie en schoolvorderingen in alle groepen, leveren significant slechter passende modellen op. Het verschil tussen de chi^2-waarden, waarmee modellen 1 en 2 zijn geassocieerd, is 15, terwijl de kritische chi^2-waarde bij 9 vrijheidsgraden ongeveer 18 is. Op grond van de toets met de verschillen in de waarde van chi^2 tussen modellen, zouden modellen 3 en 4 verworpen moeten worden en zou er alleen geconcludeerd kunnen worden dat de ongestandaardiseerde factorladingen en de covarianties tussen beide factoren over groepen heen constant zijn. Toch zijn de overige passingsmaten van modellen 3 en 4 zeker niet onacceptabel te noemen. Model 4 impliceert bijna volledige equivalentie van de factorstructuur over de groepen heen. Daardoor zijn de ook de verklaarde proporties variantie van de geobserveerde variabelen in de groepen aan elkaar gelijk. Weliswaar is de achteruitgang van de goodness of fit tussen model 3 en 4 het sterkst, maar de kans dat RMSEA groter is dan .08, is wel erg klein. De assumptie van min of meer gelijke factorstructuren in de tien doelgroepen kan dus niet met stelligheid worden verworpen.

Bijlage 5C Toetsing van gelijk regressiemodel binnen groepen

De aanname van een gelijk regressiemodel binnen all groepen is onderzocht door verschillende modellen te vergelijken. In het eerste model zijn de regressies van schoolvorderingen op intelligentie voor alle groepen apart geschat, dat wil zeggen voor alle groepen vrijgelaten. Dit is een tautologisch model omdat het altijd past: er zijn 10 groepen en er moeten 10 regressies (en

10 residuele varianties, die een functie zijn van de regressieparameters) worden geschat. Het aantal vrijheidsgraden van dit model is nul. In het volgende model is de ongestandaardiseerde regressie van schoolvorderingen op intelligentie gefixeerd over de groepen, dat wil zeggen er is gepostuleerd dat de samenhang tussen intelligentie en schoolvorderingen identiek is. In het daaropvolgende model is bovendien gesteld dat de residuele varianties van beide factorscores in alle groepen aan elkaar gelijk zijn, hetgeen bijna volledige equivalentie van het regressiemodel over groepen heen impliceert.
Tabel C geeft een overzicht van de passing van modellen 2 en 3. Voor het eerste model worden geen passingsmaten gegeven, omdat ze niet geëvalueerd kunnen worden.

Tabel C Passingsindices voor regressiemodellen in tien groepen

Model	chi^2	df	Δchi^2	CFI	RMSEA	90% cfi
2	101	9	-	.988	.112	.093 - .132
3	150	27	49	.984	.074	.063 - .086

Noot. CFI: comparative fit index, 90% cfi: 90% betrouwbaarheidsinterval van RMSEA. Overige verklaringen in de tekst.

Volgens de chi^2-verschiltoets past model 3 significant slechter dan model 2. De kritische waarde van chi^2 bij 18 vrijheidsgraden is ongeveer 28, terwijl het verschil in chi^2-waarden tussen beide modellen 49 is. Volgens het RMSEA-criterium is model 2 echter onacceptabel. Ook bij model 3 is er een niet te verwaarlozen kans dat RMSEA groter is dan .08, maar volgens dit criterium is model 3 beter dan model 2. In model 3 bedraagt het gestandaardiseerde beta-gewicht .77 in alle groepen, hetgeen impliceert dat bijna 60% van de schoolvorderingen van leerlingen verklaard kunnen worden door intelligentie. Wanneer de regressieparameters apart worden geschat voor elke groep (model 1) is het grootste beta-gewicht .83 (69% verklaarde variantie) in de groep leerlingen van allochtone afkomst met ouders die hoger onderwijs hebben genoten en het kleinste betaewicht .73 (54% verklaarde variantie) bij de ex-1.9 allochtone leerlingen. Dit lijkt toch wel een vrij substantieel verschil in verklaarde variantie. Mogelijk speelt ook hier de geringe omvang van de eerstgenoemde groep (allochtone leerlingen met hoger opgeleide ouders) een rol.

Al met al is het niet duidelijk of het regressiemodel in gelijke mate opgaat voor alle groepen. Bij conservatieve toetsing moet de aanname dat dit zo is worden verworpen.

6 Samenvatting

Guuske Ledoux

6.1 Achtergronden, vraagstellingen en opzet

Sinds de jaren zeventig van de vorige eeuw wordt er door de Nederlandse overheid beleid gevoerd ter bestrijding van onderwijsachterstanden die het gevolg zijn van sociale, economische en culturele factoren. Om vast te stellen of er verbetering optreedt in de onderwijskansen van leerlingen die behoren tot de doelgroepen van het achterstandenbeleid, wordt al vanaf het schooljaar 1988/89 in het basisonderwijs en vanaf het schooljaar 1989/90 in het voortgezet onderwijs cohortonderzoek uitgevoerd. Tot 2005 ging het om het LEO- en het PRIMA-cohortonderzoek in het basisonderwijs en om PRIMA-VO en VOCL in het voortgezet onderwijs. Daarna zijn deze cohorten geïntegreerd voortgezet onder de naam COOL$^{5\text{-}18}$. Dankzij deze onderzoeken zijn er uitgebreide databestanden opgebouwd, die informatie bevatten over meerdere leerlingcohorten in verschillende fasen van de schoolloopbaan in basis- en voortgezet onderwijs. Deze zijn al diverse malen benut voor onderzoek naar de ontwikkeling van de prestaties en loopbanen van de doelgroepen van het onderwijsachterstandenbeleid (OAB).
Dit rapport bevat over dit onderwerp nieuwe gegevens, op basis van nieuwe analyses op de cohortdata, waaronder de meest recente van COOL$^{5\text{-}18}$. Ditmaal is de periode 1994-2007 in beschouwing genomen. Nieuw is dat er gebruik is gemaakt van een meer gedifferentieerde indeling in groepen leerlingen. Zowel

binnen de groep doelgroepleerlingen als binnen de groep overige leerlingen zijn nieuwe subgroepen onderscheiden. Dit maakt het mogelijk nauwkeuriger uitspraken te doen over de relatieve positie van elk van deze groepen. Ook zijn nieuwe analysetechnieken toegepast, waarmee verschillen tussen groepen beter kunnen worden bepaald.

De onderzoeksvragen

De centrale vraag is hoe de prestaties en loopbanen van de doelgroepleerlingen zich verhouden tot die van de overige leerlingen. Daarbij horen verschillende deelvragen:

1. Wat zijn de actuele taal- en rekenprestaties van de verschillende groepen leerlingen in het basisonderwijs?
2. Welke positie nemen deze groepen leerlingen momenteel in in het voortgezet onderwijs?
3. Is er bij de doelgroepleerlingen sprake van een inhalen van onderwijsachterstanden in basis- en voortgezet onderwijs, wanneer cohorten worden vergeleken in de tijd en wanneer leerlingen longitudinaal worden gevolgd?

Daarnaast is nog een andere vraag gesteld, namelijk of en in hoeverre er bij leerlingen uit de doelgroepen sprake is van (verborgen) cognitief talent dat onvoldoende benut wordt. Ook hierbij horen enkele deelvragen:

4. Krijgen doelgroepleerlingen en overige leerlingen verschillende adviezen voor het voortgezet onderwijs? In hoeverre worden deze verklaard door cognitieve en niet-cognitieve competenties?
5. Krijgen sommige groepen leerlingen hogere of lagere adviezen dan verwacht?
6. Is er sprake van discrepantie tussen de cognitieve capaciteiten van leerlingen en hun schoolprestaties? Doet deze discrepantie zich meer/minder voor bij de verschillende groepen doelgroepleerlingen?

Welke groepen leerlingen worden onderscheiden?

In het basisonderwijs is de gewichtenregeling het meest centrale beleidsinstrument in de bestrijding van onderwijsachterstanden. Deze regeling zorgt ervoor dat scholen met kinderen uit de doelgroepen van het OAB over meer formatie (personeel) kunnen beschikken, omdat de doelgroepleerlingen zwaarder meewegen in de berekening van de formatie. De beleidsverwachting daarbij is dat meer formatie het mogelijk maakt om gerichte hulp en aandacht

te geven aan de doelgroepkinderen en de school daarmee dus meer mogelijkheden krijgt om te compenseren voor ongunstige omgevingsfactoren. De gewichtenregeling is in de loop der jaren enkele malen veranderd. Daarmee zijn ook de definities voor doelgroepleerling gewijzigd. De oorspronkelijke gewichtenregeling (ingevoerd in 1986) ging uit van vier categorieën leerlingen: leerlingen die niet tot een achterstandsgroep behoren (gewogen als 1.0-leerling, dus geen extra gewicht), autochtone achterstandsleerlingen (gewogen als 1.25), allochtone achterstandsleerlingen (gewogen als 1.9), schipperskinderen (gewogen als 1.4) en woonwagen- en zigeunerkinderen (gewogen als 1.7). Voor de 1.25- en de 1.9-leerlingen gold daarbij als criterium dat ten minste één van de ouders laag opgeleid moest zijn (niet hoger dan lager beroepsonderwijs), en/of dat de meeste verdienende ouder werkloos moest zijn dan wel een (ongeschoold) handarbeidersberoep moest uitoefenen. In de loop van de jaren negentig is voor de 1.25-leerlingen het criterium aangescherpt. Voortaan moesten beide ouders laag opgeleid zijn en werd niet meer naar de beroepssituatie gekeken. Voor de 1.9-leerlingen bleef de oude definitie gelden. In 2006 is de regeling opnieuw aangepast. Vanaf dat moment is het opleidingsniveau van beide ouders het enige criterium en wordt etnische herkomst buiten beschouwing gelaten. Er zijn nu nog maar drie categorieën, namelijk 0.3 voor kinderen van ouders met maximaal lbo/vbo-niveau, 1.2 voor kinderen van ouders met maximaal basisonderwijs en 0.0 voor kinderen die niet tot een achterstandsgroep behoren[26].
Omdat dit onderzoek een lange tijdsperiode bestrijkt, was het vanwege deze wijzigingen nodig om verschillende categorieën doelgroepleerlingen te onderscheiden, namelijk de ex-doelgroepleerlingen en de huidige doelgroepleerlingen. Daarnaast was het van belang om onderscheid te blijven maken tussen allochtone en autochtone leerlingen. Niet alleen om nog vergelijkingen te kunnen maken met de groep ex 1.9-leerlingen, maar ook omdat eerdere analyses lieten zien dat er nog steeds sprake was van ongelijkheid in prestaties naar etnische herkomst, ook na controle voor opleidingsniveau ouders. Verder is een differentiatie aangebracht bij de niet-doelgroepleerlingen. In die groep zijn kinderen van ouders die maximaal een mbo-opleiding hebben gevolgd onderscheiden van kinderen met ouders die hoger onderwijs hebben gevolgd. Dit heeft geleid tot de volgende indeling in tien groepen, die in dit rapport steeds is gehanteerd:

[26] Op een gegeven moment werd ook de manier van aanduiden anders: 1.0-leerlingen waren voortaan 0.0-leerlingen, het nieuwe gewicht 1.3 werd 0.3, etc.

Aanduiding	Criteria	Leerlinggewicht oud	na 1998	na 2006
autochtoon				
1.25 zwaar	beide ouders max. lbo; ten minste 1 max. lo	0.25	0.25	1.2
1.25	beide ouders max. lbo	0.25	0.25	0.3
1.25 ex	één ouder max. lbo, de ander meer	0.25	0	0
1.0 mbo	beide ouders meer dan lbo; maar geen ho	-	-	-
1.0 ho	beide ouders meer dan lbo; minstens één ho	-	-	-
allochtoon				
1.9 zwaar	beide ouders max. lbo; ten minste 1 max. lo	0.9	0.9	1.2
1.9	beide ouders max. lbo	0.9	0.9	0.3
1.9 ex	één ouder max. lbo, de ander meer	0.9	0.9	0
1.0 mbo	beide ouders meer dan lbo; maar geen ho	-	-	-
1.0 ho	beide ouders meer dan lbo; minstens één ho	-	-	-

6.2 Actuele taal- en rekenprestaties van de verschillende groepen in het basisonderwijs

De meeste recente gegevens over de prestaties van leerlingen in het basisonderwijs zijn afkomstig van de eerste meting van COOL$^{5\text{-}18}$, uit het schooljaar 2007/08. In COOL$^{5\text{-}18}$ worden onder meer toetsscores taal/lezen en rekenen verzameld in de groepen 2, 5 en 8. Voor dit onderzoek is gekeken naar de scores in groep 2 en groep 8, dus respectievelijk begin en eind basisonderwijs.
Het blijkt dat er nog steeds aanzienlijke prestatieverschillen in *taalvaardigheid* zijn tussen OAB-doelgroepleerlingen en overige leerlingen, maar ook binnen de doelgroepleerlingen en binnen de groep overige leerlingen. Hoe hoger de opleiding van de ouders, hoe beter de gemiddelde taalscores. 'Zware' achterstandskinderen (ouders zeer laag opgeleid) presteren slechter dan 'lichte' achterstandskinderen, beide groepen presteren slechter dan niet-

achterstandskinderen en binnen die laatste groep presteren kinderen van mbo-opgeleide ouders weer slechter dan kinderen van hbo/wo-opgeleide ouders. Dit geldt zowel voor groep 2 als voor groep 8, maar in groep 2 wel sterker dan in groep 8.
Daarnaast speelt ook etnische herkomst nog een rol: *binnen elke opleidingscategorie* doen autochtone leerlingen het beter dan allochtone leerlingen. Dit laatste is vooral zichtbaar in groep 2. In groep 2 hebben bijvoorbeeld kinderen van hoogopgeleide allochtone ouders (die dus niet tot de doelgroep van het beleid horen en ook niet in het verleden daartoe behoorden) slechtere taalprestaties dan de 'lichte' autochtone achterstandskinderen. In groep 8 is het verschil in taalprestaties minder groot. Kinderen van hoog opgeleide allochtone ouders presteren dan beter dan alle groepen autochtone achterstandskinderen, maar nog wel slechter dan kinderen van hoogopgeleide autochtone ouders.
Tussen allochtone leerlingen uit verschillende herkomstlanden zijn er ook nog verschillen in taalprestaties, maar die zijn vrij gering en er zit ook geen duidelijke lijn in.
De verschillen tussen de groepen in *rekenvaardigheid* zijn veel kleiner dan die in taalvaardigheid. Ook hiervoor geldt dat de verschillen in groep 8 weer wat minder groot zijn dan die in groep 2. De invloed van de thuissituatie is voor rekenen van minder belang dan voor taal en neemt kennelijk gedurende de onderwijsloopbaan in het basisonderwijs ook nog eens flink af. Maar ook voor rekenen geldt dat er in groep 8 nog steeds duidelijke verschillen bestaan tussen de OAB-doelgroepleerlingen en de niet-doelgroepleerlingen. Het effect van etnische herkomst doet zich bij rekenen echter minder voor dan bij taal. Binnen de opleidingscategorieën zijn er nog maar kleine verschillen tussen allochtone en autochtone kinderen, en in de zwaarste achterstandscategorie (1.2) is er helemaal geen verschil naar etnische herkomst.
De conclusie is dat er anno 2007/08 in het basisonderwijs nog steeds achterstanden bestaan bij de OAB-doelgroepleerlingen. Voor taal zijn die achterstanden groter dan voor rekenen, en in groep 2 groter dan in groep 8. Allochtone leerlingen blijven vooral bij taal meer achter dan autochtone leerlingen, ook als rekening gehouden wordt met het opleidingsniveau van de ouders.

6.3 Actuele positie van de verschillende groepen in het voortgezet onderwijs

COOL[5-18] bevat gegevens over de onderwijspositie van leerlingen in leerjaar 3 van het voortgezet onderwijs en over hun prestaties in Nederlands en wiskunde. Er blijken anno 2007/08 nog steeds zeer grote verschillen tussen de groepen te bestaan in *gerealiseerd onderwijsniveau* (type voortgezet onderwijs dat gevolgd wordt). De doelgroepleerlingen nemen een flink lagere positie in (volgen lagere schooltypen) dan de niet-doelgroepleerlingen. Opleidingsniveau van de ouders blijkt hier de belangrijkste voorspeller, zowel voor autochtone als voor allochtone leerlingen. Anders dan bij de prestaties in het basisonderwijs zien we hier dus geen effect van etnische herkomst.
Dat zien we echter wel weer als we kijken naar de *taal- en wiskundevaardigheid* in leerjaar 3 van het voortgezet onderwijs. In het algemeen presteren de allochtone leerlingen lager dan de autochtone leerlingen, en wel in elke opleidingscategorie. De enige uitzondering is er bij leerlingen met gewicht 1.2 (de leerlingen met de laagstopgeleide ouders): daar scoren juist de autochtone leerlingen lager dan de allochtone leerlingen.
Opvallend is verder dat in het voortgezet onderwijs de verschillen tussen allochtone en autochtone leerlingen voor wiskunde groter zijn dan voor taal. Kennelijk doet het wiskunde-onderwijs in het voortgezet onderwijs een (sterker) beroep op andere vaardigheden (zoals taal) dan het geval is bij rekenen in het basisonderwijs.
De conclusie is dat de OAB-doelgroepleerlingen in het voortgezet onderwijs nog steeds op forse achterstand staan. Opmerkelijk daarbij is dat de lagere prestaties die de allochtone leerlingen halen bij Nederlands en wiskunde (ook na controle voor opleidingsniveau ouders) zich niet vertalen in een *extra* lage onderwijspositie. Allochtone leerlingen bezoeken vergelijkbare schooltypen als autochtone leerlingen met dezelfde ouderlijke opleiding. Daarbinnen behalen ze echter weer wel lagere prestaties dan de autochtone leerlingen.

6.4 Halen doelgroepleerlingen achterstanden in?

Ook al hebben doelgroepleerlingen nog steeds achterstand, de relevante vraag voor het beleid is natuurlijk of er sprake is van *vermindering* van achterstanden. Daar is in dit onderzoek op twee manieren naar gekeken: via cross-sectionele analyses en via longitudinale analyses.

De *cross–sectionele analyses* zijn gebaseerd op een vergelijking tussen vier generaties: 1994, 1998, 2002, 2007. Voor de leerlingen die in deze jaren in groep 2 en groep 8 van het basisonderwijs en in leerjaar 3 van het voortgezet onderwijs zaten, is nagegaan of in de meer recente generaties sprake is van geringere verschillen tussen de groepen leerlingen (doelgroepleerlingen en overige leerlingen) dan in de oudere generaties. Het beeld dat daaruit naar voren komt is als volgt.

- In groep 2 van het basisonderwijs is er in de onderzochte periode met betrekking tot de taalvaardigheid nauwelijks sprake van ontwikkelingen. De taalachterstand van de OAB-doelgroepen aan het begin van het basisonderwijs is dus ongeveer even groot gebleven. Alleen bij de allochtone leerlingen is in groep 2 sprake van enige vermindering van taalachterstand tussen 2002 en 2007. Bij rekenen zien we eveneens dat er bij allochtone leerlingen in groep 2 in 2007 sprake is van minder achterstand, vooral bij de zwaarste achterstandscategorie. Bij de autochtone doelgroepleerlingen is in groep 2 geen sprake van minder rekenachterstand.
- In groep 8 zijn de allochtone leerlingen flink vooruit gegaan, niet alleen de allochtone doelgroepleerlingen, maar ook de allochtone niet-doelgroepleerlingen. Het geldt verder zowel voor taal als voor rekenen, maar voor rekenen het meest. Voor beide vakken is bij de allochtone leerlingen sprake van een continue stijging van de taal- en rekenprestaties tussen 1994 en 2007. De positie van de autochtone doelgroepleerlingen is in groep 8 bij taal niet verbeterd; tussen 2002 en 2007 is zelfs sprake van een lichte achteruitgang (hun taalachterstand neemt dus toe). Bij rekenen is bij hen wel sprake van inhalen van achterstand, dat geldt met name voor de autochtone 1.2-leerlingen (en minder voor de 0.3-leerlingen).
- Wanneer we alle doelgroepleerlingen samen vergelijken met alle leerlingen zonder gewicht, dan zijn de doelgroepen er in het

basisonderwijs ongeveer 20% op vooruitgegaan over de periode 1994 tot 2007. Wanneer we de totale doelgroep opsplitsen in de autochtone en de allochtone doelgroep, dan ontstaat echter een veel genuanceerder beeld. De allochtone doelgroep heeft een forse inhaalslag gemaakt: in de periode 1994-2007 zijn de achterstanden, zowel bij taal als bij rekenen, met rond de 40% afgenomen. Voor de autochtone doelgroep is de ontwikkeling echter lang niet zo rooskleurig. In de periode 1994-2007 is er qua taalvaardigheid sprake van stilstand of zelfs een lichte verslechtering. Met betrekking tot de rekenvaardigheid doet zich wel een positieve ontwikkeling voor: in de periode 1994-2007 zijn de achterstanden met rond de 10% gereduceerd.

- In het voortgezet onderwijs blijkt er weinig ontwikkeling te zijn in het onderwijsniveau dat de verschillende groepen halen. Alleen bij de allochtone 0.3-leerlingen is sprake van vooruitgang, dus van (enig) inhalen van achterstand. Bij de autochtone 1.2-leerlingen is echter juist sprake van een daling in onderwijsniveau. Voor de taal- en wiskundevaardigheid in leerjaar 3 van het voortgezet onderwijs kon alleen een vergelijking worden gemaakt tussen cohorten in 1999 en 2003. In deze periode deden zich nagenoeg geen ontwikkelingen voor.

De *longitudinale analyses* hebben wat betreft het basisonderwijs dit beeld nog wat verder aangescherpt. Bij longitudinale analyses wordt gekeken naar de leervorderingen die leerlingen uit eenzelfde generatie maken in de loop van hun schoolloopbaan. In dit geval is gekeken naar de vorderingen tussen groep 2 en groep 8. De opvallendste uitkomsten zijn de volgende.

- Alle doelgroepleerlingen (zowel volgens de oude als volgens de nieuwe regelingen) starten lager in het basisonderwijs dan de overige leerlingen, en dat geldt voor de allochtone leerlingen nog weer sterker dan voor de autochtone leerlingen.
- Allochtone doelgroepleerlingen (alle categorieën) halen in het basisonderwijs achterstand in; zij vorderen méér dan de autochtone niet-doelgroepleerlingen (de vergelijkingscategorie).
- Autochtone doelgroepleerlingen halen geen achterstand in, hun achterstand neemt ten opzichte van de vergelijkingscategorie juist toe.

Dit beeld ligt in lijn met bevindingen uit eerder onderzoek.

6.5 Krijgen doelgroepleerlingen en niet-doelgroepleerlingen verschillende adviezen voor het voortgezet onderwijs?

Naar de advisering voor het voortgezet onderwijs is al veel onderzoek gedaan. Dat geldt onder meer voor de vraag of er sprake is van overadvisering dan wel onderadvisering van bepaalde groepen. Van overadvisering is sprake als leerlingen hogere adviezen voor het voortgezet onderwijs krijgen dan op basis van hun leerprestaties in groep 8 verwacht zou worden. Van onderadvisering is sprake bij lagere adviezen dan verwacht. In dit onderzoek is de advisering opnieuw onderzocht, nu met gebruikmaking van de nieuwe doelgroepindeling en van enkele nieuwe gegevens over niet-cognitieve leerlingkenmerken uit COOL[5-18].

Eerst is nagegaan hoe over het geheel genomen (zonder uitsplitsing naar doelgroepen) de adviezen samenhangen met leerlingkenmerken. Zoals verwacht blijkt er een sterke samenhang te zijn tussen adviezen en leerprestaties in groep 8. Dat geldt echter niet voor alle leerdomeinen even sterk. De sterkste samenhang is er met de totaalscore op de Cito Eindtoets, daarna met de scores op toetsen rekenen en begrijpend lezen. De samenhang met score voor technisch lezen is relatief zwak. Het niveau technisch lezen is dus een minder goede voorspeller van de hoogte van het advies. Verder zien we samenhangen tussen het advies en de werkhouding van de leerling (zoals beoordeeld door de leerkracht) en tussen het advies en de ouderbetrokkenheid (idem), al zijn deze samenhangen niet heel erg sterk. Leerkrachten houden bij het opstellen van het advies dus tot op zekere hoogte ook rekening met deze twee kenmerken. Ook blijkt er een matige samenhang te zijn met twee persoonlijkheidskenmerken van leerlingen: ordelijkheid en autonomie (onafhankelijkheid). Leerlingen die volgens hun leerkracht hoog scoren op deze eigenschappen krijgen ook wat hogere adviezen.

Vervolgens is nagegaan welke verschillen er zijn tussen de doelgroepleerlingen en overige leerlingen in advies en in met advies samenhangende leerlingkenmerken. Zoals verwacht blijken er bij advieshoogte duidelijke verschillen tussen doelgroepleerlingen en niet-doelgroepleerlingen. Autochtone én allochtone doelgroepleerlingen krijgen vaker lage adviezen en minder vaak hoge adviezen dan niet-doelgroepleerlingen. Dat geldt voor de huidige doelgroepleerlingen (0.3 en 1.2) nog weer wat sterker dan voor de voormalige doelgroepleerlingen (1.25 en 1.9). Er blijken ook duidelijke verschillen te bestaan naar etnische herkomst, maar alleen bij de *niet*-doelgroepleerlingen. Zo

krijgt van de allochtone leerlingen met middelbaar opgeleide ouders 15% een havo/vwo-advies, tegen 26% van de autochtone leerlingen met middelbaar opgeleide ouders. Bij leerlingen met hoger opgeleide ouders is er een vergelijkbaar groot verschil: daar krijgt 33% van de allochtone leerlingen een havo/vwo-advies tegen 46% van de autochtone leerlingen.
De verschillen tussen de groepen leerlingen zijn tamelijk gering als het om niet-cognitieve leerlingkenmerken gaat die relevant kunnen zijn voor het advies. Eén kenmerk blijkt echter wel van sterk belang: ouderbetrokkenheid. Leerkrachten vinden dat er bij (alle categorieën) doelgroepleerlingen sprake is van lagere ouderbetrokkenheid dan bij niet-doelgroepleerlingen en bij allochtone leerlingen bovendien nog weer lager dan bij autochtone leerlingen. Het verschil allochtoon-autochtoon doet zich ook voor bij niet-doelgroepleerlingen.
Tot slot is nagegaan of er nog verschillen in advieshoogte blijven bestaan nadat is gecontroleerd voor de verschillen in leerlingkenmerken. Daarbij is ook de vraag over over- en onderadvisering betrokken.
Het blijkt dat, wanneer rekening wordt gehouden met de cognitieve competenties zoals gemeten met de Eindtoets, met name de huidige 0.3-categorie, en daarbinnen de autochtone leerlingen meer nog dan de allochtone leerlingen, in lichte mate wordt ondergeadviseerd. Zij krijgen dus wat lagere adviezen dan op basis van hun prestaties zou worden verwacht. Van een lichte overadvisering is alleen sprake bij de autochtone leerlingen met hoger opgeleide ouders. Zij krijgen dus wat hogere adviezen dan hun prestaties doen verwachten. Bij dit alles dient overigens te worden aangetekend dat het in statistische zin weliswaar om significante effecten gaat, maar dat deze empirisch gezien maar zeer beperkt zijn.
De cognitieve competenties van de leerlingen verklaren het grootste deel van de verschillen in advieshoogte weg. De niet-cognitieve competenties voegen daar nog maar heel weinig aan toe. Maar ze zijn niet geheel zonder betekenis: de werkhouding van de leerling en de ouderbetrokkenheid verklaren nog iets van de advieshoogte, ook na controle voor de cognitieve competenties van de leerlingen.

6.6 Cognitieve reserves bij doelgroepleerlingen?

Een achterliggende veronderstelling bij het onderwijsachterstandenbeleid is dat de doelgroepen van dat beleid, door allerlei oorzaken, minder goed presteren in het onderwijs dan eigenlijk zou kunnen. Door de inzet van extra middelen zouden de scholen in staat moeten zijn om meer uit deze leerlingen te halen dan er tot dan toe uit is gekomen. Het doel van het onderwijsachterstandenbeleid is dus in essentie het tegengaan van onderpresteren/van het onderbenutten van capaciteiten. Of met een oudere term, het doel is te komen tot volledige benutting van 'verborgen talent'.
Naar het vóórkomen van onderbenutting/onderpresteren in het basisonderwijs is al eerder onderzoek gedaan. Een lastig probleem daarbij is het bepalen van de grens: wanneer is sprake van onderbenutting? In dit onderzoek zijn nieuwe analyses gedaan, waarbij rekening is gehouden met enkele methodologische kwesties, zoals de meetonnauwkeurigheid van onderzoeksinstrumenten. Ook is gebruik gemaakt van nieuwe meetinstrumenten, met name van een uitgebreidere meting van cognitieve capaciteiten die met COOL$^{5\text{-}18}$ beschikbaar is gekomen.
De leidende vraag daarbij was óf onderbenutting voorkomt en vooral of dat in verschillende mate het geval is bij de in dit onderzoek onderscheiden leerlinggroepen. Onderbenutting is daarbij gedefinieerd als een discrepantie tussen verwachte en feitelijke prestaties. Er is gebruik gemaakt van de meest recente gegevens uit COOL$^{5\text{-}18}$. Verwachte prestaties zijn de prestaties die verwacht kunnen worden op basis van een cognitieve capaciteitentest die in COOL$^{5\text{-}18}$ wordt afgenomen in groep 5. Feitelijke prestaties zijn de toetsscores op de in groep 5 afgenomen toetsen voor lezen, rekenen en woordenschat.
De resultaten van de analyses laten zien dat leerlingen met lager opgeleide ouders (dus de leerlingen uit de OAB-doelgroepen) tenderen naar schoolprestaties die lager zijn dan voorspeld kan worden op basis van hun cognitieve capaciteiten. Hoe lager het opleidingsniveau, des te meer is er sprake van onderpresteren of onderbenutting. De nieuwe analyses maken echter duidelijk dat de verschillen tussen groepen leerlingen aanmerkelijk kleiner zijn dan op grond van eerdere analyses werd gedacht. Ook is zichtbaar geworden dat de gebruikte modellen niet voor alle onderscheiden groepen als precies gelijk verondersteld mogen worden. Vandaar de voorzichtige formulering van de conclusie ('tenderen naar'…). Het verdient in ieder geval aanbeveling bij verder onderzoek naar onderbenutting expliciet rekening te

houden met de meetonnauwkeurigheid van de gebruikte instrumenten. Ook zou het nuttig zijn om in de toekomst onderzoek naar onderbenutting te doen met gebruik van een specifieke leerpotentieel-test.

6.7 Resumerend

Leerlingen uit de doelgroepen van het onderwijsachterstandenbeleid hebben nog altijd achterstand in het onderwijs. Dat geldt bij aanvang van het basisonderwijs, aan het eind van het basisonderwijs en in het voortgezet onderwijs. Wel is er, bezien over een langere periode (1994-2007), sprake van inhalen van achterstanden door doelgroepleerlingen. Dat geldt echter vooral voor allochtone doelgroepleerlingen. Autochtone doelgroepleerlingen raken – relatief gezien - eerder verder achterop; dat geldt in het bijzonder voor de (kleine) groep autochtone leerlingen met zeer laag opgeleide ouders.
Sinds 2006 speelt etnische herkomst geen rol meer in de definitie van doelgroepleerling. Alleen de opleiding van de ouders bepaalt nog of een leerling meetelt als doelgroepleerling. Etnische herkomst blijkt echter nog wel degelijk van belang als het gaat om leerprestaties en loopbanen. Binnen elke opleidingscategorie (dus ook bij kinderen met hoger opgeleide ouders) doen autochtone leerlingen het namelijk beter dan allochtone leerlingen bij toetsen taal en rekenen (basisonderwijs) en Nederlands en wiskunde (voortgezet onderwijs). Ook krijgen vooral in de hogere opleidingscategorieën autochtone leerlingen hogere adviezen dan allochtone leerlingen.
Doelgroepleerlingen krijgen aan het eind van het basisonderwijs in lichte mate lagere adviezen dan verwacht op grond van hun prestaties (onderadvisering). Kinderen van hoogopgeleide ouders krijgen juist in lichte mate hogere adviezen dan verwacht (overadvisering). Deze afwijkingen zijn echter statistisch gezien wel significant, maar stellen empirisch gezien nauwelijks iets voor. Verder geldt, onder enig voorbehoud, dat naarmate de ouderlijke opleiding lager is, er vaker sprake is van onderbenutting van talent.

Referenties

Babeliowsky, M., & Boer, R. den (2007). *Voortgezet onderwijs in beeld. De leerlingen en hun resultaten in het Amsterdamse VO. Schooljaar 2005/2006.* Almere: Babeliowsky Onderwijsonderzoek.

Badad, E. Y., & Budoff, M. (1974). Sensitivity of learning potential measurement in three levels of ability. *Journal of Educational Psychology, 3*(66), 439-447.

Batenburg, T. v., & Werf, M. v. d. (2004). NSCCT: Verantwoording, normering en handleiding. Groningen: GION.

Bergh, L. van den, Denessen, E., Hornstra, L., Voeten, M., & Holland, R. (2010).The implicit prejudiced attitudes of teachers relations to teacher expectations and the ethnic achievement gap. *American Educational Research Journal, 47(* 2), 497-527.

Boer, H. de (2009). *Schoolsucces van Friese leerlingen in het voortgezet onderwijs.* Groningen: GION.

Boer, H. de, Bosker, R., & Werf, M. van der (2007). De gevolgen van onder- en overadvisering. In Inspectie van het Onderwijs, *Onderadvisering in beeld* (pp. 83-92). Utrecht: Inspectie van het Onderwijs.

Browne, M. W., & Cudeck, R. (1993). Alternative Ways of Assessing Model Fit. In K. A. Bollen & J. S. Long (Eds.), Testing Structural Equation Models. London: Sage.

Carroll, J. B. (1993). *Human cognitive abilities. A survey of factor–analytic studie*s. New York: Cambridge University Press.

Cattell, R. B. (1971). Abilities. Their structure, growth and action. Boston: Houghton Mifflin.

Cito (2008). *Terugblik en resultaten Eindtoets Basisonderwijs 2008 Versie 29 februari 2008. Eindtoets Basisonderwijs. Groep 8.* Arnhem: Cito.

Claassen, A. (i.v.) *Effecten van wijzigingen in de gewichtenregeling en andere oab–faciliteiten* Nijmegen: ITS.

Claassen, A., & Mulder, L. (2003). *Leerlingen na de overstap. Een vergelijking van vier cohorten leerlingen na de overgang van basisonderwijs naar voortgezet onderwijs met nadruk op de positie van doelgroepleerlingen van het onderwijsachterstandenbeleid.* Nijmegen: ITS.

Claassen, A., & Mulder, L. (2006). *Na vier jaar voortgezet onderwijs. Afsluitende rapportage over Prima cohort 00–08 in het vierde jaar en de eindexamens van vmbo–leerlingen.* Nijmegen: ITS.

Claassen, A., Mulder, L., & Hulsen, M. (2005). *In het derde jaar voortgezet onderwijs. Een vergelijking van leerlingen vóór en na de invoering van het vmbo met nadruk op cognitieve en sociale competenties.* Nijmegen: ITS.

Coe, R. (2002). *It's the effect size, stupid! What effect size is and why it is important.* Paper presented at the British Educational Research Association annual conference, Exeter, 12-14 September, 2002.

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

Dagevos, J., & Gijsberts, M. (red.) (2007). *Jaarrapport integratie 2007.* WODC/SCO/CBS. Den Haag: SCP.

Dagevos, J., & Gijsberts, M. (red.) (2009). *Jaarrapport integratie 2009.* WODC/SCO/CBS. Den Haag: SCP.

Dagevos, J., Gijsberts, M., & Praag, C. van (2003). *Rapportage minderheden 2003. Onderwijs, arbeid en sociaal–culturele integratie.* Den Haag: SCP.

Davis, J. (1966). The campus as a frog pond: An application of the theory of relative deprivation to career decisions of college men. *American Journal of Sociology, 72*, 17-31.

Doolaard, S. & Leseman, P. (2008). *Versterking van het fundament. Integrerende studie n.a.v. de opbrengsten van de onderzoekslijn Sociale en institutionele context van scholen uit het Onderzoeksprogramma beleidsgericht onderzoek primair onderwijs 2005–2008.* Groningen: GION.

Doolaard, S., & Mulder, L. (2009). *Programma Beleidsgericht Onderzoek Primair Onderwijs (BOPO) 2009–2012.* Den Haag: NWO.

Driessen, G. (1991). Discrepanties tussen toetsresultaten en doorstroomniveau. Positieve discriminatie bij de overgang basisonderwijs - voortgezet onderwijs? *Pedagogische Studiën*, *68*(1), 27-35.

Driessen, G. (2005). De totstandkoming van de adviezen voortgezet onderwijs: Invloeden van thuis en school. *Pedagogiek*, *25*(4), 279-298.

Driessen, G. (2005) *De cross–sectionele en longitudinale ontwikkeling van de intelligentietestscores in PRIMA.* Nijmegen: ITS

Driessen, G. (2006). Het advies voortgezet onderwijs: Is de overadvisering over? *Mens & Maatschappij*, *81*(1), 5-23.

Driessen, G. (2006). Ontwikkelingen in het gebruik van streektalen en dialecten in de periode 1995-2003. *Toegepaste Taalwetenschap in Artikelen*, *75*(1), 103-113, 118.

Driessen, G. (2007). '*Peer group' effecten op onderwijsprestaties. Een internationaal review van effecten, verklaringen en theoretische en methodologische aspecten.* Nijmegen: ITS.

Driessen, G. (2009). *Prestaties, gedrag en houding van basisschoolleerlingen. Stand van zaken in 2008 en ontwikkelingen sinds 2001*. Nijmegen: ITS.

Driessen, G. (2010). Structurele en culturele integratie, generaties en onderwijsprestaties. *Migrantenstudies, 26*(3), 262-279.

Driessen, G., & Dekkers, H. (2008). Dutch policies on socio-economic and ethnic inequality in education. *International Social Science Journal*, *59*(193/194), 449-464.

Driessen, G., & Doesborgh, J. (2005a). Relaties tussen achtergrondkenmerken en competenties van leerlingen en hun advies voor voortgezet onderwijs. In G. Driessen, J. Doesborgh, G. Ledoux, M. Overmaat, J. Roeleveld & I. van der Veen, *Van basis– naar voorgezet onderwijs. Voorbereiding, advisering en effecten* (pp. 39-70). Nijmegen/Amsterdam: ITS/SCO-Kohnstamm Instituut.

Driessen, G., & Doesborgh, J. (2005b). De rol van ouders en scholen bij de totstandkoming van de adviezen voortgezet onderwijs. In G. Driessen, J. Doesborgh, G. Ledoux, M. Overmaat, J. Roeleveld & I. van der Veen, *Van basis– naar voorgezet onderwijs. Voorbereiding, advisering en effecten* (pp. 11-38). Nijmegen/Amsterdam: ITS/SCO-Kohnstamm Instituut.

Driessen, G., Doesborgh, J., Ledoux, G., Overmaat, M., Roeleveld, J., & Veen, I. van der (2005). *Van basis– naar voorgezet onderwijs. Voorbereiding, advisering en effecten*. Nijmegen/Amsterdam: ITS/SCO-Kohnstamm Instituut.

Driessen, G., & Langen, A. van (2010). *De onderwijsachterstand van jongens. Omvang, oorzaken en interventies.* Nijmegen: ITS.

Driessen, G., Langen, A. van, & Vierke, H. (2000). *Basisonderwijs: Veldwerkverslag, leerlinggegevens en oudervragenlijsten. Basisrapportage PRIMA–cohortonderzoek. Derde meting 1998/99.* Nijmegen: ITS.

Driessen, G., Langen, A. van, & Vierke, H. (2004). *Basisonderwijs: Veldwerkverslag, leerlinggegevens en oudervragenlijsten. Basisrapportage PRIMA–cohortonderzoek. Vijfde meting 2002–2003*. Nijmegen: ITS.

Driessen, G., & Merry, M. (2011).The effects of the integration and generation of immigrants on language and numeracy achievement. *Educational Studies, 37*(4/5).

Driessen, G., Mooij, T., & Doesborgh, J. (2007). *Hoogbegaafdheid van leerlingen in het primair onderwijs. Ontwikkelingen en samenhangen met kenmerken van thuis, de groep en de school*. Nijmegen: ITS.

Driessen, G., Mulder, L., Ledoux, G., Roeleveld, J., & Veen, I. van der (2009). *Cohortonderzoek COOL$^{5-18}$. Technisch rapport basisonderwijs, eerste meting 2007/08*. Nijmegen: ITS/ Amsterdam: SCO-Kohnstamm Instituut.

Driessen, G., Mooij, T., & Doesborgh, J. (2007). *Hoogbegaafdheid van leerlingen in het primair onderwijs. Ontwikkelingen en samenhangen met kenmerken van thuis, de groep en de school.* Nijmegen: ITS.

Driessen, G., Sleegers, P., & Smit, F. (2008). The transition from primary to secondary education: Meritocracy and ethnicity. *European Sociological Review, 24*(4), 527-542.

Driessen, G., & Smeets, E. (2007). De relatie tussen prestaties en advies: onder- of overadvisering bij de overgang van basis- naar voortgezet onderwijs? In Inspectie van het Onderwijs, *Onderadvisering in beeld* (pp. 59-81). Utrecht: Inspectie van het Onderwijs.

Driessen, G., & Smeets, E. (2011). The use of teacher recommendations and achievement tests as valid instruments for selecting ethnic minority students into secondary school tracks. In L. Madsen (Ed.), *Achievement tests: Types, interpretations and uses.* Hauppauge, NY: Nova Science Publishers.

Driessen, G., & Withagen, V. (1999). Language varieties and educational achievement of indigenous primary school pupils. *Language, Culture and Curriculum*, *12*(1), 1-22.

Dronkers, J., Erp, M. van, Robijns, M., & Roeleveld, J. (1998). Krijgen leerlingen in de grote steden en met name in Amsterdam te hoge adviezen? *Tijdschrift voor Onderwijsresearch*, *23*, 17-30.

Field, A. (2009). Discovering statistics using SPSS (3rd ed.). London: Sage.

Gemeente Amsterdam (2007a). *Basisschooladviezen en etniciteit. Onderzoeksverslag, 29 januari 2007*. Amsterdam: Gemeente Amsterdam.

Gemeente Amsterdam (2007b). *Brief aan de leden van de commissie WIJ, 20 februari 2007*. Amsterdam: Gemeente Amsterdam.

Gemeente Amsterdam (2009). *Amsterdamse resultaten primair onderwijs 2009. Stedelijke rapportage basisschooladviezen en cito–resultaten*. Amsterdam: Gemeente Amsterdam.

Gemeente Amsterdam (2010). *Welke factoren spelen een rol bij onderadvisering in het basisonderwijs?* Amsterdam: Gemeente Amsterdam. Op 08/07/2010 gedownload van http://www.os.amsterdam nl/pdf/2010_factsheets_3.pdf.

Grigorenko, E. L., & Sternberg, R. J. (1998). Dynamic Testing. *Psychological Bulletin, 124*(1), 75-111.

Guldemond, H. & Bosker, R. (2006). *Onderwijsachterstanden en ontwikkelingen in leerprestaties*. Groningen: GION

Hanushek, E., Kain, J., Markman, J., & Rivkin, S. (2003). Does peer ability affect student achievement? *Journal of Applied Economics*, 18, 527-544.

Harker, R., & Tymms, P. (2004). *Symposium on 'Compositional effects'.* AERA Annual Meeting, San Diego, USA, 13 April 2004.

Horn, J. L. (1994). Theory of fluid and crystallized intelligence. In R. J. Sternberg (Ed.), Encyclopedia of human intelligence (pp. 443-451). New York: Prentice Hall.

Hox, J. J. (1998). Er is nieuws onder de zon: nieuwe oplossingen voor oude problemen. *Kwantitatieve Methoden* 19: 95-118.

Hustinx, P. (2002). School careers of pupils of ethnic minority background after the transition to secondary education: Is the ethnic factor always negative? *Educational Research and Evaluation, 8*, 169-195.

Inspectie van het Onderwijs (2001). *Onderwijsverslag over het jaar 2000.* Utrecht: Inspectie van het Onderwijs.

Inspectie van het Onderwijs (2007). *Onderadvisering in beeld*. Utrecht: Inspectie van het Onderwijs.

Jong, M.J. de (1987). *Herkomst, kennis en kansen. Allochtone en autochtone leerlingen tijdens de overgang van basis– naar voortgezet onderwijs.* Lisse: Swets & Zeitlinger.

Jungbluth, P., Langen, A. van, Peetsma, Th. & Vierke, H. (1996) *Leerlinggegevens basisonderwijs en speciaal basisonderwijs*. Nijmegen/Amsterdam: ITS/SCO-Kohnstamm Instituut

Jungbluth, P., Peetsma, & Roeleveld, J. (1996). *Leerlingprestaties en leerlinggedrag in het primair onderwijs.* Ubbergen: Tandem Felix.

Kamphuis, F., Mulder, L, Vierke, H, Overmaat, M. & Koopman, P. (1998) *De relatie tussen PRIMA–toetsen en toetsen uit het CITO–leerlingvolgsysteem*. CITO, Arnhem; ITS Nijmegen; SCO-Kohnstamm Instituut, Amsterdam.

Kerkhoff, A. (1988). *Taalvaardigheid en schoolsucces. De relatie tussen taalvaardigheid Nederlands en schoolsucces van allochtone en autochtone leerlingen aan het einde van de basisschool*. Amsterdam/Lisse: Swets & Zeitlinger.

Koeslag, M., & Dronkers, J. (1994). Overadvisering en de schoolloopbanen van migrantenleerlingen en autochtone leerlingen in het voortgezet onderwijs. *Tijdschrift voor Onderwijsresearch, 19*, 240-258.

Koopman, P. (2002) *Cum grano salis: de omrekening van de PRIMA–rekenvaardigheidsscore naar de LVS–rekenvaardigheidsscore*. (Interne notitie; SCO-Kohnstamm Instituut)

Kuyper, H., & Werf, M. van der (2005). *VOCL'99–3. Prestaties en opvattingen van leerlingen in de derde klas van het voortgezet onderwijs.* Groningen: GION.

Kuyper, H., Suhre, C., Jansen, G., & Pijl, Y. (2000). *Geïntegreerd rendementsonderzoek in het voortgezet onderwijs.* Groningen: GION.

Langen, A. van, & Driessen, G. (2006). *Sekseverschillen in onderwijsloopbanen. Een internationaal comparatieve trendstudie.* Nijmegen: ITS.

Langen, A. van, & Suhre, C. (2001). *Ontwikkelingen in de schoolloopbanen van achterstandsleerlingen. Vergelijkende analyses van een aantal leerlingcohorten in basis– en voortgezet onderwijs.* Nijmegen: ITS.

Langen, A. van, Vierke, H. & Robijns, M. (1996) *Veldwerkverslag basisonderwijs en speciaal onderwijs.* Nijmegen/Amsterdam: ITS/SCO-Kohnstamm Instituut

Ledoux, G., & Veen, A. (2009). *Beleidsdoorlichting onderwijsachterstandenbeleid. Periode 2002–2008.* Amsterdam: SCO-Kohnstamm Instituut.

Ledoux, G. (2003). Ontwikkelingen in het basisonderwijs. In: W. Meijnen (red.), Onderwijsachterstanden in basisscholen. Antwerpen-Apeldoorn: Garant.

Ledoux, G., Overmaat, M., Boogaard, M., Felix, C., & Triesscheijn, B. (2005). *Onderwijskansen bekeken. De stand van zaken in het onderwijskansenbeleid*. Amsterdam: SCO-Kohnstamm Instituut.

Lubbe, M. van der, Verhelst, N., Heuvelmans, T., & Staphorsius, G. (2005). *Verslag van een onderzoek naar de toelating van leerlingen in het voortgezet onderwijs.* Arnhem: Cito.

Luyten, H. (2004). Succes in het voortgezet onderwijs: Capaciteiten, inzet of achtergrond? *Pedagogische Studiën*, *81*(1), 151-1166.

Luyten, H., & Bosker, R. (2004). Hoe meritocratisch zijn schooladviezen? *Pedagogische Studiën*, *81*(1), 89-103.

Luyten, H., Cremers-Van Wees, L., & Bosker, R. (2001). *Mattheus–effecten voor taal, rekenen en non–verbaal IQ. Verschillen tussen scholen, lichtingen en leerlingen.* Enschede: Twente University Press.

Meijer, J., & Oostdam, R. J. (2001). The factor structure and criterion validity of a test for assessment of cognitive abilities. Comparing students in special

and regular education. *European journal of psychology of education*, 16(4), 531-547.

Meijer, J., Oostdam, R., & Sluis, S. v. d. (2002). A feasibility study into incremental predictive validity of dynamic assessment of verbal-reasoning ability. In G. M. v. d. Aalsvoort, W. C. M. Resing & A. J. J. M. Ruijssenaars (Eds.), *Learning Potential Assessment and Cognitive Training: actual research and perspectives in theory building and methodology* (Vol. 7, pp. 273-291). Oxford: Elsevier.

Meijnen, W. (red.) (2003). *Onderwijsachterstanden in basisscholen*. Antwerpen-Apeldoorn: Garant.

Meijnen, G.W. (2004). Het concept meritocratie en het voortgezet onderwijs. *Pedagogische Studiën*, *81*, 79-88.

MinOCW (2010). *Rijksbegroting 2010. VIII Onderwijs, Cultuur en Wetenschap.* Den Haag: Ministerie van OCW.

MinOCW (2010). *Vraag en antwoord. Mag de school mijn kind uitsluiten van deelname aan de Citotoets of een andere eindtoets?* Op 14/07/10 gedownload van http://www.rijksoverheid.nl/documenten-en-publicaties/vragen-en-antwoorden/mag-de-school-mijn-kind-uitsluiten-van-deelname-aan-de-citotoets-of-een-andere-eindtoets html.

Mulder, L. (1993). De invloed van het advies op de schoolloopbaan van de OVB-doelgroepleerlingen in het voortgezet onderwijs. *Pedagogische Studiën, 70*, 242-251.

Mulder, L. (1996). *Meer voorrang, minder achterstand? Het Onderwijsvoorrangsbeleid getoetst*. Nijmegen: ITS.

Mulder, L., Roeleveld, J., Veen, I. van der & Vierke, H. (2005) *Onderwijsachterstanden tussen 1988 en 2002: ontwikkelingen in basis– en voortgezet onderwijs*. Nijmegen/Amsterdam: ITS/SCO-Kohnstamm Instituut

Mulder, L., Roeleveld, J., & Vierke, H. (2007). *Onderbenutting van capaciteiten in basis– en voortgezet onderwijs.* Den Haag: Onderwijsraad.

Mulder, L., & Suhre, C. (1995). *OVB–doelgroepleerlingen in het voortgezet onderwijs. Deel 1: De overgang van basis– naar voortgezet onderwijs in schooljaar 1992/1993.* Nijmegen: ITS.

Mulder, L., & Tesser, P. (1992). *De schoolkeuze van allochtone leerlingen.* Nijmegen: ITS.

Mulder, L., Van der Hoeven, A., Vierke, H., Ledoux, G., Van der Veen, I., Oud, W. & Roeleveld, J. (2008). *Inrichting en effecten van schakelklassen. Resultaten*

van het evaluatieonderzoek schakelklassen in het schooljaar 2006/2007. Nijmegen/Amsterdam: ITS/SCO-Kohnstamm Instituut

Mulder, L., Van der Hoeven, A., Vierke, H., Van der Veen, I., & Elshof, D. (2009). *Inrichting en effecten van schakelklassen. Resultaten van het evaluatieonderzoek schakelklassen in het schooljaar 2007/2008*. Nijmegen/Amsterdam: ITS/SCO-Kohnstamm Instituut

Muthén, L.K. & Muthén, B.O. (2004) *Mplus. Statistical Analysis With Latent Variables. User's Guide*. Los Angeles, CA: Muthén & Muthén

Neisser, U., Boodoo, G., Bouchard, T. J., Boykin, A. W., Brody, N., Ceci, S. J., e.a. (1996). Intelligence: Knowns and Unknowns. *American Psychologist*, 51(2), 77-101.

OECD (2008). *Measuring Improvements in Learning Outcomes. Best Practices to Assess the Value–Added of Schools.*

Onderwijsraad (2001a). *Wat 't zwaarst weegt… Een nieuwe aanpak voor het onderwijsachterstandenbeleid*. Den Haag: Onderwijsraad.

Onderwijsraad (2001b). *Met 't oog op onderwijsachterstanden. Studies bij het advies Wat 't zwaarst weegt*. Den Haag: Onderwijsraad.

Peschar, J.L. (2007) Over leerwinst als stelselindicator. Haren

Pijl, Y. (1994). De overschatte overadvisering van allochtonen. *Tijdschrift voor Onderwijswetenschappen*, *23*(6), 240-245.

Rasbash, J., Steele, F., Browne, F., & Goldstein, H. (2009). *A user's guide to MLwiN.* Bristol: University of Bristol.

Resing, W. C. M. (2001). Beyond Binet. *Issues in Education*, 7(2), 225-236.

Roeleveld, J. & R. Portengen (1998). *Uitval en instroom bij het PRIMA–cohortonderzoek.* Amsterdam/Nijmegen: SCO-Kohnstamm Instituut / ITS

Roeleveld, J. & Vierke, H. (2003). *Uitval en instroom bij de derde meting van het PRIMA–cohortonderzoek.* Intern onderzoeksrapport PRIMA. Amsterdam/Nijmegen: SCO-Kohnstamm Instituut/ITS

Roeleveld, J., Veen, I. van der, & Ledoux, G. (2009). *Verkenning leerwinst als indicator voor onderwijskwaliteit.* Amsterdam: SCO-Kohnstamm Instituut.

Ruijven, E. van (2003). V*oorsprong of achterstand? Onderzoek naar het onderwijsniveau van de Friese leerlingen in het basisonderwijs en het voortgezet onderwijs.* Ljouwert: Fryske Akademy.

Smit, F., Driessen, G., & Doesborgh, J. (2005). *Opvattingen van allochtone ouders over onderwijs: tussen wens en realiteit. Een inventarisatie van de verwachtingen en wensen van ouders ten aanzien van de basisschool en educatieve activiteiten in Rotterdam.* Nijmegen: ITS.

Snijders, T., & Bosker, R. (1994). Modeled variance in two-level models. *Sociological Methods Research, 22,* 342-363.

Stoel, R. (2003) *Issues in growth curve modeling.* Proefschrift Universiteit van Amsterdam

Tesser, P., & Iedema, J. (2001). *Rapportage minderheden 2001. Vorderingen op school.* Den Haag: SCP.

Tesser, P., Dagevos, J., & Iedema, J. (2003). *Rapportage minderheden 2001*. Den Haag: SCP.

Thompson, B. (1998). *Five methodology errors in educational research: The pantheon of statistical significance and other faux pas.* Invited address AERA annual meeting, San Diego, US, April 1998.

Thorndike, R. L. (1963). *The concepts of over– and underachievement*. New York: Teachers College Columbia University.

Tweede Kamer (2010). *Doorstromen en stapelen in het onderwijs.* Op 16/07/10 gedownload van
http://www.tweedekamer nl/kamerstukken/dossiers/doorstromen_en_stapelen_in_het_onderwijs.
jsp.

Veen, I. van der (2001). *Successful Turkish and Moroccan students in the Netherlands.* Leuven/Apeldoorn: Garant.

Vierke, H. (1995). *De PRIMA–toetsen gekalibreerd. De ontwikkeling van vaardigheidsscores over de leerjaren heen op basis van de jaargroeptoetsen in het cohort primair onderwijs (PRIMA).* Nijmegen: ITS.

Vogels, R., & Bronnemans-Helmers, R. (2003). *Autochtone achterstandsleerlingen: een vergeten groep.* Den Haag: SCP.

Vygotsky, L. S. (1963). Learning and Mental Development at School Age. In B. Simon & J. Simon (Eds.), *Educational Psychology in the U.S.S.R.* (pp. 21-34). London: Routledge & Kegan Paul.

Werf, G. van der, & Kuyper, H. (2004). Brugklasdocenten gaan niet op de hurken. *Didaktief, 34*, 22-24.

Zijsling, D., Keuning, J., Kuyper, H., Batenburg, T. van, & Hemker, B. (i.v.). *Technisch rapport. Eerste meting van COOL5–18 in het derde leerjaar van het voortgezet onderwijs.* Groningen/Arnhem: RUG/Cito.

Recent uitgegeven Kohnstamm Instituut rapporten

853 Veen, van der I., Derriks, M., & Elshof, D.P., m.m.v. Hoeven, van der A.
Amsterdamse schakelklassen in het schooljaar 2009/2010

851 Derriks M., Vergeer M.M.
Van havo naar hbo

850 Koopman, P.N.J., Derriks, M., Voncken, E.
PrO-loopbaan Vervolgd

849 Meijer, J., Ledoux, G., m.m.v. Elshof, D.P.
Gebruikersvriendelijke leerlingvolgsystemen in het primair onderwijs

848 Derriks M., Vergeer M.M.
Doorstroom in het ROC

847 Sligte, H., Breetvelt, I.S., Heemskerk, I.M.C.C.
Technologische ontwikkelingen voor leerlingen met een visuele beperking of dyslexie

846 Emmelot, Y., Veen, A.M., Heurter, A.M.H., Bongers, C., Roode de, N., Vegt van der, A.L.
De samenwerking tussen voorschoolse voorzieningen in het kader van de harmonisatie

845 Veen, I. van der, Derriks, M., m.m.v. van der Hoeven A.
Amsterdamse schakelklassen in het schooljaar 2008/2009

844 Stams, G.J.J.M., Top Eem van der, M., Limburg, S., Vugt van E.S., Laan van der, P.H
Implementatie en doelmatigheid van de Deltamethode Gezinsvoogdij

Deze rapporten zijn te bestellen via: secr@kohnstamm.uva nl
Voor meer informatie, zie; http://www kohnstamminstituut.uva nl

Kohnstamm Instituut UVA bv
Postbus 94208
1090 GE Amsterdam
T 020 5251226
www.kohnstamminstituut.uva.nl