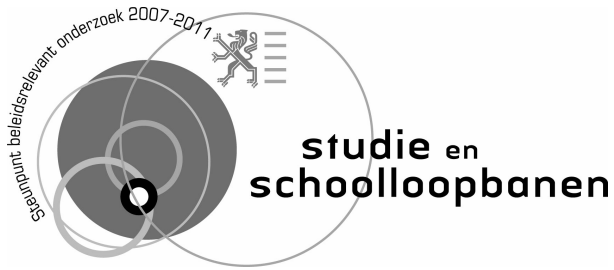




Leerwinst en toegevoegde waarde voor wiskunde, technisch lezen en spelling in eerste en tweede leerjaar

Jean Pierre Verhaeghe & Jan Van Damme

Onderzoek gefinancierd door de Vlaamse Regering in het kader van het
programma 'Steunpunten voor Beleidsrelevant Onderzoek'



Longitudinaal onderzoek in het basisonderwijs

Leerwinst en toegevoegde waarde voor wiskunde, technisch lezen en spelling in eerste en tweede leerjaar

Jean Pierre Verhaeghe & Jan Van Damme

Promotoren: J. Van Damme, B. De Fraine, P. Ghesquière, R. Janssen,
F. Laevers, I. Nicaise, P. Onghena, R. Rymenans, M. Valcke,
P. Van Petegem, L. Verschaffel & K. Verschueren

Onderzoek in opdracht van de Vlaamse minister van Onderwijs en Vorming, in het kader van het programma 'Steunpunten voor Beleidsrelevant Onderzoek'

2007

SSL-rapport nr. OD1/05
datum oplevering eerste versie: 04.03.2008
datum publicatie: 18.04.2008

Voor meer informatie over deze publicatie:

Steunpunt SSL, onderzoeksdomein “Onderwijsloopbanen”

Auteurs: Jean Pierre Verhaeghe & Jan Van Damme

Adres: Dekenstraat 2, 3000 Leuven

Tel.: +32 16 32 57 58 of +32 16 32 57 47

Fax: +32 16 32 58 59

E-mail: Petra.Dewaele@ped.kuleuven.be

Website: <http://www.steunpuntloopbanen.be>

Copyright (2007) Steunpunt SSL
p/a Parkstraat 47, 3000 Leuven

Niets uit deze uitgave mag worden verveelvoudigd en/of openbaar gemaakt zonder uitdrukkelijk te verwijzen naar de bron.

No part of this material may be made public without an explicit reference to the source.

De verantwoordelijkheid voor dit rapport berust volledig bij de auteurs en vertolkt niet noodzakelijk de officiële visie van de Vlaamse Overheid.

Woord vooraf

Door het Steunpunt *Loopbanen doorheen Onderwijs naar Arbeidsmarkt* (ondertussen: Steunpunt *Studie- en Schoolloopbanen*) wordt sedert het schooljaar 2002-2003 een longitudinaal onderzoek bij leerlingen basisonderwijs uitgevoerd. In dat onderzoek worden de ontwikkeling en de schoolloopbanen van 6000 leerlingen gevolgd vanaf het begin van de derde kleuterklas tot het begin van het secundair onderwijs. De bedoeling is die ontwikkeling te beschrijven en verschillen erin te verklaren aan de hand van factoren in het kind zelf, factoren in de thuisomgeving en factoren in de klas en op school. Behalve een representatieve referentiesteekproef van 4000 leerlingen uit 120 basisscholen werden in het onderzoek enkele *oversamplings* opgenomen, onder meer om beter zicht te krijgen op de ontwikkelingen in scholen met veel kansarme leerlingen.

In voorliggend rapport worden de resultaten gepresenteerd van een reeks analyses over de gegevens verzameld in het eerste en het tweede leerjaar (schooljaren 2003-2004 en 2004-2005) in de referentiesteekproef, de oversampling GOK-scholen en de oversampling Gentse scholen. Op basis van de toetsen taalvaardigheid (begin eerste leerjaar), wiskunde (begin en einde eerste leerjaar, einde tweede leerjaar), technisch lezen en spelling (einde eerste en tweede leerjaar) worden groeicurves geschetst en wordt nagegaan welke leerwinst de leerlingen in de loop van het eerste en het tweede leerjaar voor de leergebieden wiskunde, technisch lezen en spelling geboekt hebben. Door die leerwinst in verband te brengen met de sociale achtergrond van de leerlingen, wordt voor de betrokken scholen ook de toegevoegde waarde bepaald. Tenslotte wordt ook nagegaan in hoeverre de samenstelling van de gevolgde leerlingencohorte binnen elke school een effect heeft op de gemiddelde leerwinst van de leerlingen.

Qua aanpak bouwen de analyses waarover hier gerapporteerd wordt, voort op de werkwijze die voor LOA-rapporten 43 (Van de gaer e.a., 2006) en 44 (Verhaeghe e.a., 2006) gehanteerd werd. D.w.z. dat de sociale en etnisch-culturele achtergrond van de leerlingen in rekening gebracht werd aan de hand van zogenaamde GOK-categorieën. Die werkwijze sluit aan bij de indeling die de Vlaamse overheid in het kader van het Gelijke Onderwijskansenbeleid hanteert en biedt als methodologisch voordeel dat een aantal complexe verbanden op eenvoudiger en eenduidiger wijze te interpreteren zijn (Verhaeghe e.a., 2006). Ook het begrippenkader dat in dit rapport gehanteerd wordt, sluit aan bij de LOA-rapporten 43 en 44. De belangrijkste elementen daaruit en de belangrijkste elementen uit de gevolgde onderzoeks- en analysemethode worden toegelicht in een inleidend hoofdstuk. Voor meer details verwijzen we naar voornoemde LOA-rapporten.

Een basiskarakter van de data-set waarmee gewerkt wordt, is het longitudinale karakter ervan. Dat stelt een aantal bijzondere problemen die samenhangen met de in-, door- en uitstroom van leerlingen bij elke overgang naar een volgend leerjaar (en toetsmoment). In het eerste hoofdstuk gaan we daarom even in op die in-, door- en uitstroom.

In hoofdstuk 2 bespreken we de instroomkenmerken. Behalve de verdeling bekijken we ook de samenhang met de toetsresultaten.

In hoofdstukken 3, 4 en 5 ten slotte komen voor elk leergebied afzonderlijk de toetsresultaten aan de orde. Na een korte bespreking van de inhoud van de toetsen, komen achtereenvolgens de algemene evolutie van de leerlingen voor het betreffende leergebied, de evolutie per GOK-categorie, en de evolutie van de zittenblijvers aan bod. We eindigen de bespreking per leergebied met een onderzoek naar groepscompositie-effecten. In het kader van het debat rond gelijke onderwijskansen is dat een niet onbelangrijk thema. We gaan na in hoeverre behalve de eigen individuele achtergrond ook de achtergrond van de medeleerlingen een rol speelt in de leerwinst die leerlingen maken. Is het van belang met welke andere leerlingen men samen op school zit? We kijken daarbij verder dan wat de analyses aan groepscompositie-effecten over de hele steekproef heen reveleren. We gaan met name na in welke mate die groepscompositie-effecten variëren van school tot school en brengen dat in verband met de toegevoegde waarde die scholen realiseren. Op die manier proberen we tot meer genuanceerde uitspraken te komen. De analyses inzake groepscompositie-effecten worden meer diepgaand gerapporteerd voor het leergebied wiskunde. Voor technisch lezen en spelling volgen we dezelfde analysemethode maar houden we de rapportage beknopter.

We eindigen het rapport met een algemeen besluit en een beknopte vergelijking over de drie leergebieden heen, waarna enkele meer technische bijlagen volgen.

Bij het voltooien van dit rapport wensen al diegenen te danken die dit onderzoek mogelijk hebben gemaakt. Dat is in de eerste plaats uiteraard de Vlaamse Overheid die via het Steunpunt Studie- en Schoolloopbanen voor de nodige financiële middelen zorgde. We danken ook de scholen die aan het SiBO-onderzoek deelnemen, de directies en de betrokken leraren, de leerlingen en hun ouders, voor hun bijzonder gewaardeerde medewerking. Ten slotte wensen wij de collega's van het SiBO-team en de mede-promotoren van onderzoeksdomein 1 binnen het Steunpunt te danken voor hun bijdragen in de verzameling en de verwerking van de gegevens en hun kritische en ondersteunende commentaren bij de totstandkoming van dit rapport.

J.P. Verhaeghe

J. Van Damme

Inhoud

Woord vooraf	3
Inhoud	5
Inleiding.....	7
<i>Leerwinst</i>	7
<i>Toegevoegde waarde</i>	8
<i>Onderzoeksvragen</i>	11
<i>Onderzoeksmethode</i>	11
1. In-, door- en uitstroom in het eerste en het tweede leerjaar	15
2. Instroomkenmerken.....	19
2.1 Kindfactoren.....	19
2.2 Sociale en etnisch-culturele achtergrond	31
2.3 Aanvankelijke Nederlandse taalvaardigheid (begin eerste leerjaar)	35
2.4 Wat met de toets Rekenbegrip begin eerste leerjaar?	38
2.5 Besluit	39
3. Leerwinst en toegevoegde waarde voor wiskunde	41
3.1 De toetsen wiskunde	41
3.2 Algemene evolutie van begin L1 naar einde L2	43
3.3 Evolutie voor de verschillende GOK-categorieën	48
3.4 Evolutie van de zittenblijvers	53
3.5 Maakt de samenstelling van de leerlingengroep een verschil uit?	55
3.6 Besluit	80
4. Leerwinst en toegevoegde waarde voor technisch lezen.....	85
4.1 De Drie-Minuten-Toets voor Vlaanderen (DMT).....	85
4.2 Algemene evolutie van begin L1 naar einde L2	86
4.3 Evolutie voor de verschillende GOK-categorieën	89
4.4 Evolutie van de zittenblijvers	92

4.5 Maakt de samenstelling van de leerlingengroep een verschil uit?	92
4.6 Besluit.....	99
5. Leerwinst en toegevoegde waarde voor spelling.....	103
5.1 De Schaal Vorderingen Spellingvaardigheid voor Vlaanderen (SVS)	103
5.2 Algemene evolutie van begin L1 naar einde L2.....	104
5.3 Evolutie voor de verschillende GOK-categorieën	108
5.4 Evolutie van de zittenblijvers	110
5.5 Maakt de samenstelling van de leerlingengroep een verschil uit?	111
5.6 Besluit.....	118
6. Algemeen besluit en vergelijking over de leergebieden.....	121
7. Referenties.....	127
 Bijlage 1. GCE voor wiskunde met en zonder correctie voor aanvankelijk Nederlands taalvaardigheidsniveau	 129
Bijlage 2. Tabellen met de resultaten van de multiniveau-analyses: verdeling onverklaarde variantie en aandeel verklaarde variantie	131

Inleiding

Qua begrippenkader en analysemethode bouwt voorliggend rapport voort op LOA-rapporten 43 (Van de gaer e.a., 2006) en 44 (Verhaeghe e.a., 2006). In dit inleidend hoofdstuk bespreken op beknopte wijze de belangrijkste elementen uit dat begrippenkader en die analysemethode. Voor een meer gedetailleerde toelichting verwijzen we naar de genoemde LOA-rapporten.

Centraal in onze analyses staan de begrippen *leerwinst* en *toegevoegde waarde*.

Leerwinst

Onder leerwinst verstaan we de vooruitgang die een leerling tijdens een bepaalde periode voor een bepaald leergebied boekt. De leerwinst kan bepaald worden door het verschil te maken tussen de toetsscore die een leerling op een bepaald moment (het einde van een periode) behaalde en de toetsscore die die leerling voor hetzelfde leergebied op een vorig moment (het begin van die periode) behaalde. Aan die simpele “aftrekmethod” zijn wel enkele voorwaarden verbonden.

- De toetsscores van de verschillende momenten moeten onderling vergelijkbaar zijn. Dat kan bijvoorbeeld door meerdere keren dezelfde toets af te nemen. Maar voor vaardigheden waarin kinderen snel evolueren, ligt dat natuurlijk moeilijk. Een toets rekenen/ wiskunde die geschikt is voor leerlingen in het zesde leerjaar kan natuurlijk niet gebruikt worden in het eerste leerjaar en omgekeerd. Er bestaat gelukkig ook de mogelijkheid om zulke toetsen, die éénzelfde achterliggende vaardigheid meten maar op een verschillend (lager - hoger) beheersingsniveau mikken, onderling te *calibreren* aan de hand van een statistische techniek die gebaseerd is op de zgn. *Item Response Theory (IRT)*. Op die manier worden de scores van verschillende toetsen (b.v. één in elk leerjaar) op éénzelfde meetschaal geplaatst en daardoor onderling vergelijkbaar gemaakt. De vaardigheidsscore op zo’n meetschaal is heel eenvoudig te interpreteren. Wie bijvoorbeeld bij het begin van het eerste leerjaar een vaardigheidsscore van 20 haalt, een jaar later 40 en nog eens jaar later 60, is in de tweede periode evenveel vooruitgegaan als in de eerste. Eén van de voorwaarden voor het gebruik van IRT-modellering is dat de verschillende toetsen in wezen éénzelfde achterliggende vaardigheid meten.
- Uit de eerste voorwaarde volgt dat leerwinst slechts per leergebied (b.v. wiskunde) of soms zelfs per subdomein (technisch lezen, spelling, begrijpend lezen, ...) bepaald kan worden. Het is goed mogelijk dat de leerwinst die een leerling voor het ene leergebied of subdomein boekt, verschilt van de leerwinst die hij of zij voor een ander leergebied of subdomein haalt. Ook voor een school kan de gemiddelde leerwinst voor het ene leerdomein (b.v. wiskunde) verschillen

van de gemiddelde leerwinst voor een ander domein. Het heeft weinig zin het gemiddelde te maken over verschillende leergebieden.

- Om de leerwinst van een leerling correct te kunnen bepalen is het van belang dat op elk moment het vaardigheidsniveau van die leerling zo accuraat mogelijk bepaald wordt. Dat veronderstelt dat de toetsen niet te gemakkelijk noch te moeilijk zijn én dat plafond- en vloereffecten vermeden worden. Wanneer leerlingen het maximum halen spreken we van een plafondeffect; wanneer leerlingen nul halen spreken we van een vloereffect. In beide gevallen kan het vaardigheidsniveau eigenlijk niet echt bepaald worden. Van leerlingen die het maximum halen, weten we niet hoeveel die nog beter hadden gekund als de toets moeilijker of uitgebreider was geweest. Van leerlingen die nul halen weten we niet hoeveel ze gehaald zouden hebben, als de toets makkelijker was geweest.

In geval de leerlingen onderling erg van niveau verschillen, kan het aangewezen zijn dat toetsen van verschillende moeilijkheidsgraad gebruikt worden, zodat elke leerling die toets aflegt die het best bij zijn/haar niveau aansluit. Dat is niet alleen aangenamer (en minder demotiverend of vervelend) voor de leerling, het levert ook een betrouwbaarder schatting van het vaardigheidsniveau op. Toetsen voor hetzelfde meetmoment die qua moeilijkheidsgraad verschillen, kunnen net als toetsen die voor verschillende meetmomenten bedoeld zijn, door middel van IRT-modellering op éénzelfde schaal geplaatst worden.

Toegevoegde waarde

Principieel verwijst het begrip *toegevoegde waarde* naar de mate waarin een school bijdraagt aan de ontwikkeling (en dus aan de leerwinst) van zijn leerlingen. In de onderzoekspraktijk (en bij schoolfeedback) wordt de toegevoegde waarde uitgedrukt door een relatief getal dat aangeeft hoeveel groter of kleiner de toegevoegde waarde van een school is in vergelijking met de gemiddelde toegevoegde waarde in een referentiegroep.

Waarom is de toegevoegde waarde van scholen van belang?

De reden om in schoolfeedback naast ruwe gemiddelden voor leerpeil of leerwinst ook te informeren over de toegevoegde waarde van een school, steunt op de volgende overwegingen:

1. Behalve de school zijn ook andere factoren van invloed op het leren van kinderen.
2. Met betrekking tot die factoren zijn er verschillen tussen kinderen: sommigen hebben een achtergrond die bevorderend is voor het leren, anderen een achtergrond die minder bevorderend of zelfs remmend is.
3. Scholen verschillen in de samenstelling van hun schoolbevolking: sommige scholen hebben meer kinderen met een gunstige achtergrond, andere meer kinderen met een ongunstige achtergrond.

4. Als gevolg daarvan kunnen we op basis van “ruwe” schoolgemiddelden niet bepalen in welke mate een school meer of minder dan gemiddeld bijdraagt tot het leren van zijn leerlingen. Dat geldt niet alleen voor de ruwe gemiddelden van het leerpeil op een bepaald moment, maar ook voor het ruwe gemiddelde van de leerwinst (al is in die leerwinst het aanvankelijke prestatieniveau al verrekend).

Behalve op school steken kinderen ook binnen het gezin en in de omgang met leeftijdsgenoten buiten de school heel wat op dat rechtstreeks of onrechtstreeks hun schoolse prestaties ten goede kan komen. We moeten daarbij niet alleen aan formele instructie of hulp bij huiswerk e.d. denken. Het gezinsklimaat, de wijze waarop ouders met hun kinderen omgaan, de aard van de activiteiten die ze met hun kinderen samen doen, ... kunnen meer of minder bevorderend werken op de ontwikkeling van de kinderen. Ook zijn er factoren binnen het kind zelf (intelligentie, sekse, geboortemaand ...) die maken dat sommige kinderen een hoger dan wel een lager niveau halen of zich sneller c.q. trager ontwikkelen. Kortom, kinderen verschillen, zowel wat hun persoonlijke eigenschappen betreft als wat hun sociale achtergrond betreft. En die verschillen hebben een invloed op hun schools presteren en de leerwinst die ze boeken.

Zoals reeds gezegd, verschillen scholen inzake de aard van hun schoolbevolking. Sommige scholen tellen bijvoorbeeld meer kinderen met een sociale achtergrond die over het algemeen bevorderlijk is voor het leren op school, andere scholen hebben meer kansarme kinderen. Scholen verschillen m.a.w. met betrekking tot hun *instroomkenmerken*.

Willen we weten in welke mate verschillen tussen scholen echt te maken hebben met verschillen in de werking van de scholen, met verschillen in de kwaliteit van het verstrekte onderwijs, dan moeten we corrigeren voor de impact van de instroomkenmerken op de schoolprestaties. Dat is precies wat gebeurt als we de toegevoegde waarde van scholen bepalen.

Hoe wordt de toegevoegde waarde van een school bepaald?

Het komt erop aan de leerprestaties (en ook de leerwinst) die binnen scholen gemiddeld geboekt worden, uit te zuiveren van de effecten van de instroomkenmerken. Hoe dat *uitzuiveren* gebeurt, wordt verderop kort besproken. De essentie bestaat erin dat we de prestaties (of de leerwinst) uit elkaar trekken in twee delen: een deel dat verklaard kan worden door de samenhangen met de instroomkenmerken die we in rekening brengen (sekse, geboortemaand, geboortemaand, sociale achtergrond en thuistaal, aanvankelijke Nederlandse taalvaardigheid) en een niet-verklaarde *rest of residu*. Het gemiddelde van die rest per school (in het jargon *schoolresidu* genoemd) levert een maat (getal) voor de *toegevoegde waarde* van de school.

Hoe moet het getal dat de toegevoegde waarde van een school weergeeft, geïnterpreteerd worden?

De maat voor toegevoegde waarde (= het schoolresidu) is een relatief getal. Een waarde 0 betekent dat de toegevoegde waarde van de betreffende school even groot is als de gemiddelde toegevoegde

waarde van alle scholen in Vlaanderen. Een positieve waarde betekent dat de toegevoegde waarde van de school in kwestie groter is dan gemiddeld in Vlaamse scholen. Een negatieve waarde betekent dat de toegevoegde waarde kleiner is dan gemiddeld in Vlaamse scholen.

Een negatief getal voor de toegevoegde waarde betekent dus NIET dat er in die school geen leerwinst werd geboekt en al helemaal niet dat de leerlingen op het eind minder kunnen dan aan het begin. Het betekent enkel dat de bijdrage van die school aan de ontwikkeling van de leerlingen kleiner is dan in de gemiddelde Vlaamse school. Als gevolg daarvan ligt de gerealiseerde leerwinst of het behaalde niveau in zo'n school gemiddeld lager dan op grond van de instroomkenmerken van de leerlingen verwacht mocht worden. Omgekeerd wijst een positief *toegevoegde waarde*-getal erop dat de leerwinst c.q. het niveau in de school gemiddeld hoger ligt dan op grond van de instroomkenmerken verwacht kon worden.

DE toegevoegde waarde van een school? Eén cijfer per school?

We spreken steeds van de toegevoegde waarde van een *school*, maar het is van belang te noteren dat een concreet *toegevoegde waarde*-getal steeds betrekking heeft op een bepaalde groep leerlingen binnen een school, voor een bepaald leergebied en over een bepaalde periode. We hebben het dus bijvoorbeeld over de toegevoegde waarde van school X voor *wiskunde in het eerste leerjaar voor de groep leerlingen die in 2003-2004 in het eerste leerjaar zaten*. Of voor dezelfde groep en hetzelfde leerjaar maar dan voor spelling. Of voor technisch lezen. Of voor het tweede leerjaar. Of voor de periode van begin eerste tot einde tweede leerjaar. En daarbij kunnen we steeds kijken naar het prestatieniveau op het einde van een periode of naar de leerwinst tijdens die periode.

Zoals ook voor leerwinst het geval is, kunnen we de toegevoegde waarde van een school over één of meerdere leerjaren bepalen. Het heeft echter weinig zin om het gemiddelde over verschillende leergebieden te maken. De toegevoegde waarde die een school realiseert, kan verschillen van leergebied tot leergebied en van leerjaar tot leerjaar. Het is voor een school (en de instanties die de school begeleiden) informatiever om die verschillen in toegevoegde waarde te zien en veel minder interessant om slechts één globaal cijfer te zien.

Hoe stabiel is de toegevoegde waarde van een school?

Uit onderzoek weten we dat de toegevoegde waarde van schooljaar tot schooljaar kan schommelen, ook voor hetzelfde leergebied en hetzelfde leerjaar of dezelfde periode van meerdere leerjaren. Dat is vooral voor kleine scholen met slechts weinig leerlingen per leerjaar het geval. Het gaat meestal om "natuurlijke" schommelingen die niets te maken hebben met wijzingen in het gevoerde beleid of de aanpak in de klas. Voor kleinere scholen wordt in de literatuur daarom aanbevolen de toegevoegde waarde over minstens drie opeenvolgende schooljaren te bekijken. Dat levert een betrouwbaarder beeld op.

Onderzoeksvragen

De analyses werden gestuurd door de volgende onderzoeksvragen:

1. Hoe sterk verschillen Vlaamse basisscholen onderling qua gemiddeld niveau voor Nederlandse taalvaardigheid en wiskunde bij het begin van het eerste leerjaar?
2. In hoeverre kunnen verschillen tussen leerlingen in niveau en leerwinst voor wiskunde, technisch lezen en spelling in het eerste en het tweede leerjaar verklaard worden door kind- en thuisfactoren als sekse, geboortjaar en geboortemaand, thuistaal en sociale en etnisch-culturele achtergrondkenmerken?
3. Hoe sterk verschillen Vlaamse basisscholen onderling qua gemiddelde leerwinst in het eerste en het tweede leerjaar voor de leergebieden wiskunde, technisch lezen en spelling? Hoe groot zijn de verschillen in toegevoegde waarde die in dat verband gerealiseerd wordt?
4. Is de toegevoegde waarde die scholen realiseren voor wiskunde, technisch lezen en spelling in het eerste en het tweede leerjaar verschillend voor verschillende categorieën van leerlingen, onderscheiden naar sociale achtergrond resp. schoolloopbaan?
5. In hoeverre is er, bovenop de samenhang met individuele achtergrondkenmerken, voor het beginniveau, de leerwinst en het niveau op het einde van het eerste resp. tweede leerjaar, voor wiskunde, technisch lezen en spelling, ook nog een samenhang met de samenstelling van de schoolbevolking (met name de leerlingencohorte die binnen het SiBO-onderzoek gevolgd wordt)? Zijn er met betrekking tot die samenhangen verschillen tussen scholen?

Onderzoeksmethode

De analyses: basisprincipe

Voor het bepalen van de toegevoegde waarde per school werd gebruik gemaakt van een statistische techniek die bekend staat als *regressie-analyse*. Bij een regressie-analyse wordt gepoogd de scores voor een bepaalde zgn. “afhankelijke” variabele (b.v. de scores op een toets wiskunde) te schrijven als de gewogen som van de scores op een reeks voorspellende variabelen. Dat waren in ons geval: de sekse van de leerling, het geboortjaar en de geboortemaand, of een leerling het eerste, resp. het tweede leerjaar al dan niet overdeed, de GOK-categorie (zie verder) en het taalvaardigheidsniveau bij het begin van het eerste leerjaar.

Ruwe score - verwachte score - residu - gecorrigeerde score

De gewogen som van de waarden voor de voorspellende variabelen geeft aan welke toetsscore men voor een leerling op grond van zijn persoons- en achtergrondkenmerken (de voorspellende variabelen) kon verwachten. Ze wordt daarom de *verwachte score* genoemd. Die komt niet steeds overeen met de *ruwe toetsscore*. Het verschil tussen ruwe toetsscore en verwachte score wordt de rest of het

residu genoemd. Dat residu vormt de basis voor het bepalen van de toegevoegde waarde van de scholen (zie hoger). Technisch gezien is het getal dat we gebruiken om de toegevoegde waarde van een school uit te drukken (het schoolresidu) niets anders dan het verschil tussen de gemiddelde ruwe score van een school en zijn gemiddelde verwachte score. Daarom zegt de toegevoegde waarde van een school ons hoeveel beter (of slechter) de gemiddelde score van die school is dan wat op grond van de instroomkenmerken van de leerlingen verwacht mocht worden. Die gemiddelde “score” kan zowel het gemiddelde prestatieniveau op één bepaald moment als de gemiddelde leerwinst tijdens een periode betreffen.

In eerdere rapporten spraken we ook van *gecorrigeerde (leerwinst)scores*. Een gecorrigeerde score is de score die een leerling in een bepaalde school gehaald zouden hebben, mochten zijn/haar instroomkenmerken precies

dezelfde zijn als die in de *Vlaamse referentiegroep*. *Gecorrigeerde gemiddelden* laten toe scholen op een faire wijze met elkaar te vergelijken. De verschillen tussen de scholen (in prestatieniveau of in leerwinst) worden namelijk gecorrigeerd voor de invloed van leerlingkenmerken (persoons- en achtergrondkenmerken). De enige verschillen die nog overblijven zijn de verschillen die samenhangen met de toegevoegde waarde van elke school. Het kan wiskundig eenvoudig aangetoond worden dat het getal dat we hanteren om de toegevoegde waarde van een school uit te drukken (het schoolresidu) niets anders is dan het verschil tussen het gecorrigeerde gemiddelde van die school en het gemiddelde van de *referentiegroep*.

Leerlingen - klassen - scholen

Bij onze analyses hebben we rekening gehouden met het gegeven dat leerlingen gegroepeerd zitten in scholen. Dat gebeurde door een bijzondere vorm van regressie-analyse toe te passen, bekend als *multiniveau-analyse*. Bij de multiniveau-analyse zoals wij die toepasten, wordt nagegaan in welke mate verschillen tussen leerlingen verband houden met (1) verschillen tussen scholen en (2) verschillen tussen leerlingen *binnen* scholen. In plaats van de toetsscores van leerlingen rechtstreeks met het “Vlaamse gemiddelde” te vergelijken, worden leerlingen eerst met het gemiddelde van de eigen

Wie is de Vlaamse referentiegroep?

Voor de Vlaamse referentiegroep wordt uitgegaan van de ong. 4000 leerlingen uit de referentiesteekproef van 120 scholen die zo getrokken is dat ze een representatief beeld geeft van het Vlaamse basis-onderwijs. Wanneer in een regressie-analyse variabelen als sekse, geboortjaar, geboortemaand, GOK-categorie e.d. worden ingevoerd, moet telkens een referentiecategorie gekozen worden. Daardoor wordt de referentiegroep ingeperkt tot: Nederlandstalige, niet-kansarme meisjes uit de referentiesteekproef, geboren in juni 1997, zonder enige schoolse achterstand en met een gemiddelde Nederlandse taalvaardigheid bij begin L1.

Wat is de relatie met het “Vlaamse gemiddelde”?

Soms verwijzen we naar het *Vlaamse gemiddelde* als vergelijkingsbasis. In feite bedoelen we daar steeds het gemiddelde van de Vlaamse referentiegroep mee. In een regressie-analyse met alleen maar continue variabelen als voorspellers zou het *Vlaamse gemiddelde* en het *gemiddelde van de Vlaamse referentiegroep* trouwens samenvallen.

school vergeleken en vervolgens worden die schoolgemiddelden vergeleken met het “Vlaamse gemiddelde”. Op die manier leveren multiniveau-analyses rechtstreeks residuen op schoolniveau op, die als maat voor de toegevoegde waarde van scholen gebruikt worden.

Hoewel we gegevens hebben over de klasgroep waartoe de leerlingen behoorden, werd *klas* niet als niveau opgenomen. Dat komt omdat we de evolutie over twee leerjaren bekijken en in veel scholen met meer dan één klas per leerjaar de leerlingen bij de overgang van het eerste naar het tweede leerjaar van klasgroep veranderen. Als we in de analyses ook een niveau *klas* zouden onderscheiden, zouden we ook die wisseling van klasgroep expliciet in het analysemodel moeten opnemen. Dat zou heel complexe modellen opgeleverd hebben. Vanwege de praktische haalbaarheid werd daarom besloten geen rekening te houden met mogelijke verschillen tussen klassen binnen scholen.

Evolutie over drie meetmomenten: leerwinst voor eerste en leerwinst voor tweede leerjaar

In onze analyse brachten we de gegevens uit drie meetmomenten samen: begin eerste leerjaar, einde eerste leerjaar en einde tweede leerjaar. We hebben daarbij nagegaan hoe leerlingen van het ene naar het andere meetmoment evolueerden, hoe verschillend die evoluties zijn tussen de leerlingen én hoe verschillend ze zijn tussen scholen. Om dat te kunnen doen, hebben we bij onze multiniveau-analyses de zgn. *herhaalde-metingen-techniek (repeated measures design)* toegepast. De specifieke herhaalde-metingen-techniek die wij toepasten, levert behalve een coëfficiënt die aangeeft hoe hoog het Vlaamse gemiddelde beginniveau is, ook een coëfficiënt op die aangeeft hoe groot de gemiddelde leerwinst voor het eerste leerjaar in Vlaanderen is en een coëfficiënt die zegt hoe groot de Vlaamse gemiddelde leerwinst is voor het tweede leerjaar. Bij elk van die leerwinst-coëfficiënten hoort een schoolresidu dat voor elke school aangeeft hoe groot de toegevoegde waarde voor het betrokken leerjaar is. De som van beide schoolresiduen geeft weer hoe groot de toegevoegde waarde over beide leerjaren samen is.

Verschillen tussen categorieën van leerlingen

In de analyses gingen we na of er inzake beginniveau en/of leerwinst significante verschillen zijn tussen categorieën van leerlingen, onderscheiden naar sociale en etnisch-culturele achtergrond. De categorisering van leerlingen is gebaseerd op de criteria die in het Vlaamse Gelijke Onderwijskansenbeleid gehanteerd worden (zie hoofdstuk 2). We gingen ook na of de gemiddelde aanvankelijke achterstand van sommige categorieën van leerlingen en de gemiddelde leerwinst van die leerlingen groter of kleiner is naargelang van de school.

Oversampling GOK-scholen en Gents stedelijk onderwijs

De gegevens van de oversampling GOK-scholen en van het Gents stedelijk onderwijs (dat tegen betaling aan het SiBO-onderzoek participeert) werden samen met de gegevens van de referentie-steekproef geanalyseerd. Door het grotere aandeel GOK-leerlingen in de scholen uit beide

toegevoegde deelsteekproeven, kregen we een veel steviger basis voor analyses waarin de GOK-categorieën een rol spelen. Op die manier konden we bijvoorbeeld een zinvol onderscheid maken tussen taalgroepen binnen de categorie *anderstalige GOK-leerlingen*. Ook de analyses met betrekking tot groepscompositie-effecten krijgt een steviger basis door het grotere aantal scholen met hoge aandelen kansarme en allochtone leerlingen.

Door die toevoeging wordt echter de representativiteit van de steekproef verstoord. Om dat te herstellen werd voor de oversamplings een dummie-variabele in de analyses ingevoerd, samen met de nodige interacties met de overige voorspellende variabelen.

Zittenblijvers

In onze analyses namen we ook de gegevens op van de leerlingen die in het schooljaar 2003-2004 het eerste leerjaar overdeden of die in het schooljaar 2004-2005 het tweede leerjaar overzaten. Dat zijn leerlingen die door zittenblijven in dezelfde klassen terechtkwamen als de hoofdgroep van de SiBO-leerlingen. Omdat die leerlingen deel uit maken van de klas- en schoolrealiteit werd besloten ze niet te weren uit de analyses waarin de gemiddelde leerwinst en de toegevoegde waarde van de scholen geschat werden. Door twee voorspellende variabelen in de analyses op te nemen (“zittenblijver L1” resp. “zittenblijver L2”) konden twee vliegen in één klap geslagen worden. Enerzijds geven de algemene resultaten (coëfficiënten en residuen) het gemiddelde beginniveau en de gemiddelde leerwinst voor de normaalvorderende leerlingen weer en krijgen we geen “vertekening” door het opnemen van de zittenblijvers; anderzijds geven de coëfficiënten voor “zittenblijver L1” resp. “zittenblijver L2” aan hoe verschillend de scores van die twee categorieën van zittenblijvers zijn in vergelijking met die van de normaalvorderende leerlingen. We konden ook nagaan of die verschillen groter of kleiner zijn naargelang van de school.

Voor de “zittenblijvers L2 ('04-'05)” ontbreekt uiteraard een beginmeting. Door het gebruik van de zgn. *herhaalde-metingen-techniek* binnen een *multiniveaubenadering* (zie hoger) vormt dat evenwel geen bezwaar. (Met die techniek kunnen leerlingen in de analyses opgenomen worden, van zodra ze voor minstens één toetsmoment een score hebben.) Wel is voor die “zittenblijvers L2 ('04-'05)” enkel de schatting van het gecorrigeerd prestatieniveau einde L2 betekenisvol (en niet de schatting van de leerwinst). Voor de “zittenblijvers L1 ('03-'04)” kan - net als voor de normaalvorderende leerlingen - zowel het gecorrigeerde beginniveau als gecorrigeerde leerwinstscores geschat worden.

De leerlingen die door zittenblijven uit de SiBO-hoofdgroep uitstroonden (i.c. de leerlingen die in het schooljaar 2004-2005 het eerste leerjaar overdeden), vormen binnen het SiBO-onderzoek het voorwerp van aparte analyses (net als de leerlingen die uitstromen wegens overgang naar het buitengewoon onderwijs). Van hen werden in de analyses voor het voorliggend rapport enkel de gegevens voor het schooljaar 2003-2004 opgenomen. Voor het schooljaar 2004-2005 worden zij als uitstromers beschouwd en buiten beschouwing gelaten.

1. In-, door- en uitstroom en het eerste en het tweede leerjaar

Longitudinale analyse brengen een meer genuanceerd beeld van de prestaties in een school. Zo blijkt bijvoorbeeld dat een hoger dan gemiddelde leerwinst in het ene leerjaar nogal eens gevolgd wordt door een minder dan gemiddelde leerwinst in het volgende leerjaar en omgekeerd. Maar zo'n longitudinale analyse stelt ook bijzondere problemen. Die hebben onder meer te maken met de in- en uitstroom bij de overgang van het ene leerjaar naar het andere. Doordat we voor het eerste leerjaar met twee meetmomenten werkten (begin en einde), konden we bovendien ook de in- en uitstroom in de loop van het eerste leerjaar bekijken.

Welke leerlingen werden in de analyses opgenomen?

In onze analyses hebben we steeds alle leerlingen - van zodra ze op één van de meetmomenten een toets aflegden - in de analyses opgenomen. Op die manier konden we voor elk leerjaar voor de gemiddelde leerwinst van de leerlingen en voor de toegevoegde waarde van de school een schatting maken die het best met de realiteit overeenkomt. Het weglaten van leerlingen die op één of meer meetmomenten geen toetsscores hebben, kan immers tot gevolg hebben dat de overblijvende gegevens geen representatief beeld meer vormen van de schoolrealiteit.

Waarom is de analyse van de in- en uitstroom van belang?

De analyse van de in-, uit- en doorstroom is dan wel om meer dan één reden van belang.

- 1) Wanneer we de ruwe leerwinstcijfers bekijken, moeten we ons realiseren dat een grote in- of uitstroom voor een ernstige vertekening kan zorgen. Als bijvoorbeeld vooral zwakkere leerlingen uitstromen en sterkere instromen, bestaat de mogelijkheid dat de gemiddelde leerwinst van de school in de volgende leerjaren in positieve zin vertekend wordt. Men krijgt dan een *selectie-effect*.
- 2) Omdat we op elk meetmoment corrigeren voor de persoons- en achtergrondkenmerken van leerlingen, leveren in- en uitstroom in principe geen problemen op voor de interpretatie van de toegevoegde waarde en van de gecorrigeerde gemiddelden (voor leerwinst of prestatieniveau). We zeggen "in principe" omdat er technisch gezien zich enkele uitzonderingen kunnen voordoen. Wat we echter niet kunnen controleren, is in hoeverre instromers het jaar voordien gemiddeld hetzelfde vaardigheidsniveau hadden (en dus gemiddeld dezelfde leerwinst realiseerden) als normaal doorstromende leerlingen met dezelfde achtergrondkenmerken. In de analyses wordt daar wel van uitgegaan. Mocht die aanname niet kloppen, dan levert dat bij een klein aantal instromers nog geen problemen op voor de interpretatie van de gemiddelde gecorrigeerde leerwinst of voor de inter-

pretatie van de toegevoegde waarde. De vertekening zal immers zeer klein zijn. Bij een groot aantal instromers zou die vertekening wel eens groter kunnen zijn. Een analoge onzekerheid is er in het geval van een groot aantal uitstromers.

3) Voor de betrokken scholen kunnen de in-, uit- en doorstroomcijfers op zich een bron van reflectie vormen. Als de in- en uitstroomcijfers bijvoorbeeld veel groter zijn dan gemiddeld in Vlaanderen het geval is, kan het nuttig zijn dat het schoolteam zich buigt over mogelijke oorzaken. Moet de oorzaak in de context gezocht worden, b.v. dat de school zich in een wijk bevindt die gekenmerkt wordt door een groot verloop van de bevolking, waardoor ook de kinderen vaak van school veranderen? Of hangt die in- en uitstroom samen met het imago van de school en de positie t.o.v. andere scholen? Spelen aspecten van het eigen schoolbeleid (b.v. beleid i.v.m. zittenblijven) een rol?

Die laatste overweging is uiteraard vooral van belang in zoverre men de resultaten van afzonderlijke scholen gaat bekijken, b.v. wanneer ze aan de betrokken scholen zelf ter beschikking gesteld worden als hulpmiddel bij zelfevaluatie.

Figuur 1.1 geeft een overzicht van de gemiddelde in-, uit- en doorstroomcijfers in het eerste en tweede leerjaar in scholen die tot de Vlaamse referentiesteekproef behoren. We zien dat bij de overgang van het eerste naar het tweede leerjaar een uit- en instroom van 14,5% resp. 10%.

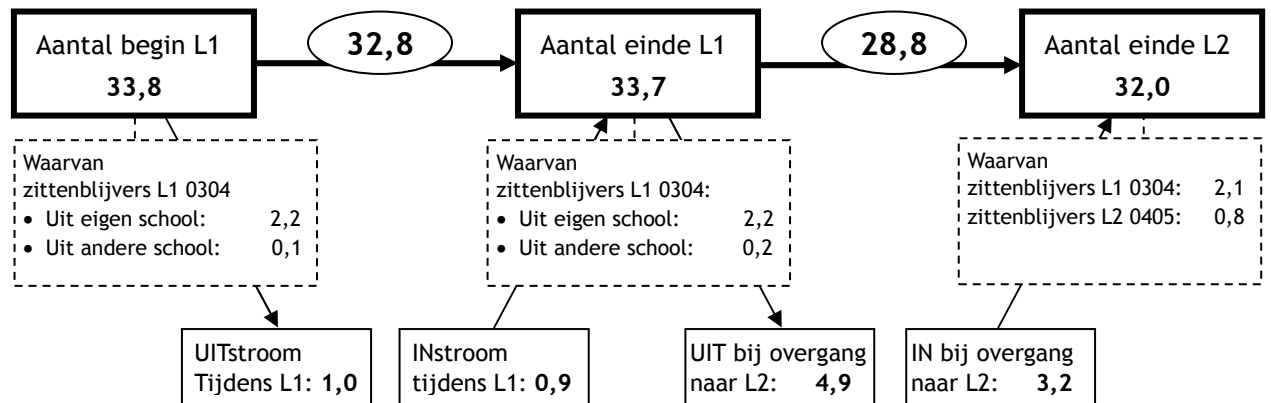
Tot de categorie leerlingen die in figuren 1.1 en 1.2 bij de overgang van L1 naar L2 bij de uitstroom gerekend worden, horen ook de kinderen die in het schooljaar 2004-2005 het eerste leerjaar overdoen in de eigen school. Zij worden bij de uitstromers gerekend omdat zij net als andere uitstromers verdwijnen uit de hoofdgroep, die in 2004-2005 in het tweede leerjaar zat.

Zittenblijvers

De leerlingen die in 2003-2004 voor de tweede keer het eerste leerjaar deden, werden - zoals gezegd - samen met de hoofdgroep van SiBO-leerlingen in de analyses opgenomen. En dat was ook het geval voor hen die in 2004-2005 het tweede leerjaar overdeden. Die laatsten worden bij de instromers gerekend omdat ze nieuw instromen in de klassen waar de hoofdgroep op dat moment zit. In de huidige analyses werden op die manier twee categorieën van zittenblijvers meegeteld, met name de twee categorieën leerlingen die een leerjaar (hetzij het eerste, hetzij het tweede) overdoen in het schooljaar dat de hoofdgroep uit het SiBO-onderzoek er aankomt (zie hoger). Figuur 1.1 geeft voor elk van beide categorieën van zittenblijvers aan om hoeveel leerlingen het in de “gemiddelde Vlaamse school” gaat. Voor wat de zittenblijvers L1 betreft, maken we een onderscheid tussen zittenblijvers die het eerste leerjaar eerder al in de eigen school deden en zittenblijvers die het eerste leerjaar het jaar voordien in een andere school volgden. Voor het tweede leerjaar bleken de aantallen te klein om zo’n onderscheid te maken.

Beïnvloeden de scores van de zittenblijvers het beginniveau en dus ook de leerwinst en de schatting van de toegevoegde waarde van de scholen niet? Neen. Zittenblijvers hebben inderdaad wel een ander beginniveau en ook een afwijkend groeipatroon, maar door ze in de analyses als “zitten-

blijver” te markeren, konden die effecten “uitgevlakt” worden. De algemene groeicurves en de per GOK-categorie opgesplitste groeicurves in hoofdstukken 3 tot 5 weerspiegelen bijgevolg steeds de gemiddelde vorderingen van normaalvorderende leerlingen (in de betreffende categorie). Door de zittenblijvers als aparte categorie op te nemen, konden wel ook aparte groeicurves getekend worden voor zittenblijvers enerzijds en voor normaalvorderende leerlingen anderzijds (zie hoofdstuk 2).



Figuur 1.1 In-, uit- en doorstroom in L1 en L2 in de gemiddelde Vlaamse school

2. Instroomkenmerken

Met het oog op het schatten van de toegevoegde waarde van scholen, worden de gemiddelde leerwinstscores gecorrigeerd voor de effecten van de instroomkenmerken. Die instroomkenmerken hebben we geordend in drie categorieën: kindfactoren, sociale achtergrondkenmerken en scores voor aanvankelijke Nederlandse taalvaardigheid. Bij elk van de instroomkenmerken geven we aan hoe groot de samenhang met de toetscores wiskunde, technisch lezen en spelling is.

2.1 Kindfactoren

Onder de noemer kindfactoren rekenen we drie voorspellende variabelen: de sekse van de leerlingen, hun geboortjaar en hun geboortemaand. We moeten hier opmerken dat in ons onderwijssysteem geboortjaar strikt genomen niet zonder meer als een instroomkenmerk beschouwd kan worden. Het gros van de verschillen in geboortjaar zijn immers het gevolg van beslissingen of adviezen vanuit de school, vooral in verband met zittenblijven. Maar omdat kinderen met schoolvertraging (soms) een afwijkend groeipatroon vertonen, is het wel relevant een onderscheid te maken tussen normaalvorderende leerlingen en leerlingen met een atypische schoolloopbaan.

2.2.1 Sekse van de leerlingen

Voor elk meetmoment is de verdeling naar sekse gemiddeld ongeveer 50-50 (zie tabel 2.1). Negentig procent van de scholen uit de referentiesteekproef heeft in de SiBO-leerlingencohorte minstens één derde jongens en minstens één derde meisjes. Het kleinste aandeel jongens is 25%; het kleinste aandeel meisjes 16%.

Tabel 2.1 Verdeling naar sekse in de gemiddelde Vlaamse school

	Begin eerste leerjaar			Einde eerste leerjaar			Einde tweede leerjaar		
	Totaal	jongen	meisje	Totaal	jongen	meisje	Totaal	jongen	meisje
Gemidd.	33,8	17,1	16,7	33,7	17,0	16,6	32,0	16,1	15,9
Pct.	100%	50,5%	49,5%	100%	50,5%	49,5%	100%	50,3%	49,7%

Waarom sekse van belang is: verschillen tussen jongens en meisjes op de toetsscores

Voor een correcte schatting van de toegevoegde waarde van scholen is een afwijkende verdeling jongens - meisjes maar van belang in de mate dat jongens en meisjes gemiddeld verschillen in de scores die ze op de toetsen halen. Tabel 2.2 geeft voor de drie toetsen het verschil tussen jongens en meisjes voor het beginniveau (indien van toepassing), de leerwinst in het eerste leerjaar en de leerwinst in het tweede leerjaar. De verschillen die in tabel 2.2 worden gerapporteerd, zijn

gebaseerd op de resultaten van multiniveau regressie-analyses (zie inleiding). Daardoor zijn ze gecorrigeerd voor de effecten van de overige voorspellende variabelen. In tabel 2.2 geven we telkens aan of het verschil significant is. Om de grootte van het verschil beter te kunnen inschatten, geven we ook weer wat de gemiddelde leerwinst is die door de *Vlaamse referentiegroep* tijdens het eerste resp. het tweede leerjaar gemaakt wordt.¹

Tabel 2.2 Verschillen tussen jongens en meisjes op de drie toetsen, na correctie voor overige voorspellende variabelen (Vlaanderen)

Toets	Gemiddeld verschil ⁽¹⁾ jongens - meisjes voor...			Gemiddelde leerwinst (LW) in Vlaamse referentiegroep	
	Beginniveau	Leerwinst L1	Leerwinst L2	Eerste leerjaar	Tweede leerjaar
Wiskunde	0,83	1,83	0,86	11,48	13,43
Sign.	p < 0,001	p < 0,001	p < 0,001		
% t.o.v. LW	7,2 %	15,9 %	6,4 %		
Technisch lezen		1,73	-0,29	78,30 ⁽²⁾	83,19
Sign.		n.s.	n.s.		
% t.o.v. LW		2,2 %	-0,4 %		
Spelling		-1,66	-0,52	80,71 ⁽²⁾	14,46
Sign.		p < 0,001	p < 0,05		
% t.o.v. LW		-2,1 %	-3,6 %		

¹Een positief verschil betekent dat de jongens het beter doen.

²Deze getallen betreffen het gemiddelde prestatieniveau einde eerste leerjaar.

Na correctie voor andere voorspellende variabelen bedraagt het verschil tussen jongens en meisjes op de aanvangstoets voor **wiskunde** (begin L1) in Vlaanderen gemiddeld 0,83 punten in het voordeel van de jongens. Dat oogt op het eerste zicht misschien niet zoveel, maar komt toch overeen met 7,2 % van de leerwinst die leerlingen uit de *Vlaamse referentiegroep* gemiddeld in de loop van het eerste leerjaar maken. Men zou kunnen zeggen dat - alle andere kenmerken gelijkgeschakeld - jongens voor wiskunde het eerste leerjaar beginnen met een voorsprong van gemiddeld 7,2 % op meisjes. Dat komt overeen met een voorsprong van goed twee en een halve week.

Het is belangrijk voor ogen te houden dat dit aanvankelijke verschil voor wiskunde een statistische schatting betreft die gemaakt wordt wanneer alle andere kenmerken gelijkgeschakeld worden, met andere woorden: als

Ter herinnering...

Wie is de Vlaamse referentiegroep?

Na opname in de regressie-analyse van voorspellers als sekse, geboortejahr, geboortemaand, GOK-categorie e.d. kan de referentiegroep omschreven worden als: Nederlandstalige, niet-kansarme meisjes uit de referentiesteekproef, geboren in juni 1997, zonder enige schoolse achterstand en met een gemiddelde Nederlandse taalvaardigheid bij het begin van het eerste leerjaar (zie p. 12).

¹ Voor technisch lezen en spelling hebben we op eind van L1 nog maar een eerste prestatie meting. Als we aannemen dat leerlingen uit de referentiegroep bij het begin van L1 gemiddeld nog helemaal niet kunnen lezen noch schrijven, kan het niveau dat op het eind van L1 bereikt is, als leerwinst beschouwd worden.

geboortjaar en geboortemaand dezelfde zijn, als de sociale achtergrond dezelfde is én als de aanvankelijke Nederlandse taalvaardigheid even hoog is. Voor geboortemaand en sociale achtergrond zijn tussen jongens en meisjes nu niet direct systematische verschillen te verwachten. Maar systematische, niet-toevallige verschillen zijn er wel met betrekking tot de Nederlandse taalvaardigheid bij het begin van het eerste leerjaar. Daar scoren meisjes - bij een gelijke sociale achtergrond en een gelijk geboortjaar en -maand - gemiddeld beter voor dan jongens (zie verder). Bovendien blijken de scores voor wiskunde significant samen te hangen met die aanvankelijke Nederlandse taalvaardigheid. Hoe hoger de scores voor aanvankelijke Nederlandse taalvaardigheid, hoe hoger de aanvangsscores voor wiskunde ($r_{(3935)} = 0,65$; $p < 0,001$).² Als we de verschillen in taalvaardigheid gewoon laten doorspelen en er niet voor corrigeren, blijken de jongens hun voorsprong voor wiskunde kwijt te zijn. Het verschil tussen jongens en meisjes bedraagt dan gemiddeld slechts 0,04 punten (nu in het voordeel van de meisjes) en dat is niet significant. De intrinsieke voorsprong die jongens gemiddeld voor wiskunde lijken te hebben, wordt - althans voor de toets die binnen het SiBO-onderzoek bij het begin van het eerste leerjaar werd afgenomen - teniet gedaan door hun gemiddeld geringere Nederlandse taalvaardigheid.

Des te opvallender is daarom de vaststelling dat jongens voor wiskunde gemiddeld een hogere *leerwinst* halen dan meisjes, en vooral de vaststelling dat die verschillen in leerwinst nauwelijks veranderen of we nu wel of niet corrigeren voor de invloed van andere voorspellende variabelen. In het eerste leerjaar halen jongens voor wiskunde gemiddeld bijna 16 % meer leerwinst dan meisjes, in het tweede leerjaar 6,4 % meer. Als we niet corrigeren voor de invloed van de verschillen in aanvankelijke Nederlandse taalvaardigheid, krijgen we nagenoeg dezelfde cijfers. In de loop van het eerste leerjaar nemen jongens m.a.w. voor wiskunde gemiddeld een voorsprong van zes schoolweken. In de loop van het tweede leerjaar scherpen ze die voorsprong met nog eens 2,4 weken aan.

Hoe komen we aan die percentages?

Als we de gemiddelde leerwinst van de Vlaamse referentiegroep (zie tabel 2.2) gelijkstellen aan 100%, dan zeggen de percentages in tabel 2.2 hoeveel % leerwinst jongens méér of minder halen dan meisjes. In de analyses namen we meisjes immers als referentiecategorie voor de variabele sekse.

Voor wiskunde doen zich nog een aantal *interactie-effecten* voor. Zo zijn de verschillen tussen jongens en meisjes inzake leerwinst in L1 kleiner bij de anderstalige leerlingen dan bij de Nederlandstaligen.

Een toets *technisch lezen* werd pas vanaf het einde van het eerste leerjaar afgenomen. Er is geen toets technisch lezen begin eerste leerjaar. In tabel 2.2 zijn er bijgevolg ook geen waarden opgenomen voor *beginniveau*. Inzoverre we mogen aannemen dat geen enkele (normaalvorderende) leerling bij het begin van het eerste leerjaar al kan lezen, mogen we de toetscore einde L1 inter-

² Correlatie voor de referentiesteekproef.

preteren als een weergave van wat een leerling in de loop van het eerste leerjaar op het vlak van technisch lezen heeft bijgeleerd, als *leerwinst* dus. De verschillen tussen jongens en meisjes zijn klein en niet-significant. Ook als we niet corrigeren voor aanvankelijke Nederlandse taalvaardigheid, zijn de verschillen jongens - meisjes niet-significant, alhoewel meisjes significant beter scoren voor aanvankelijke Nederlandse taalvaardigheid en aanvankelijke Nederlandse taalvaardigheid significant positief correleert met technisch lezen, zowel op het einde van het eerste leerjaar ($r_{(3717)} = 0,37$; $p < .001$) als op het einde van het tweede leerjaar ($r_{(3285)} = 0,21$; $p < 0,001$).

Ook voor *spelling* werd slechts vanaf het einde van het eerste leerjaar een toets afgenomen. Na correctie voor de invloed van andere predictoren, blijken meisjes in het eerste leerjaar beter te scoren en in het tweede leerjaar ook meer leerwinst te halen dan jongens. De verschillen zijn klein, maar wel significant. Corrigeren we niet voor de effecten van aanvankelijke Nederlandse taalvaardigheid, dan ligt het verschil in het eerste leerjaar anderhalve keer zo hoog. Voor het tweede leerjaar verandert er weinig. Een deel van de gemiddeld hogere leerwinst die meisjes in het eerste leerjaar voor spelling halen, kan bijgevolg toegeschreven worden aan hun gemiddeld hogere Nederlandse taalvaardigheid bij het begin van het eerste leerjaar.

2.2.2 Geboortear en geboortemaand van de leerlingen

Het geboortear van de leerlingen is een goede indicator voor de schoolse achterstand. In het SiBO-onderzoek zijn leerlingen geboren in 1997 *op leeftijd*. Leerlingen die vroeger geboren zijn, hebben meestal een jaar overgedaan. De zittenblijvers van wie sprake in hoofdstuk 1 (diegenen die in 2003-2004 het eerste leerjaar overdoen en op die manier in de klasgroepen van de SiBO-hoofdgroep terecht kwamen) zijn allen in 1996 of vroeger geboren. Maar niet alle leerlingen geboren in 1996 of vroeger deden in 2003-2004 het eerste leerjaar al voor de tweede keer. Sommigen van hen hebben de derde kleuterklas overgedaan, sommigen hebben tussen de derde kleuterklas en het eerste leerjaar een brugklas gedaan, enz.

Geboortemaand is een potentieel belangrijke voorspellende variabele vanwege het zgn. *geboortemaandeffect*. Dat geboortemaandeffect zou erin bestaan dat kinderen die later op het jaar geboren zijn, in de beginjaren van de lagere school lagere schoolprestaties neerzetten. Zij zijn in maanden gerekend jonger - en mogelijk dus minder rijp - op het moment dat het formele onderwijs in het eerste leerjaar begint. En dat zou tot uiting komen in lagere schoolresultaten. Of er een geboortemaandeffect is, kan uiteraard maar zuiver nagegaan worden als ook het geboortear in rekening gebracht. Daarom werden in de multiniveau regressie-analyses behalve *geboortear* en *geboortemaand* ook het interactie-effect *geboortear x geboortemaand* als voorspeller ingevoerd.

Tabel 2.3 geeft voor de referentiesteekproef de gemiddelde verdeling over de geboortearen. In die tabel werden de geboortearen 1996, 1995 en 1994 samengenomen. De enkele kinderen voor wie de geboortedatum onbekend is en zij die een jaar vóór zitten werden niet in de tabel opgenomen.

Tabel 2.3 Verdeling naar geboortejaar in de gemiddelde Vlaamse school

	Begin eerste leerjaar			Einde eerste leerjaar			Einde tweede leerjaar		
	Totaal	1997	≤1996	Totaal	1997	≤1996	Totaal	1997	≤1996
Gemidd.	33,8	29,6	3,9	33,7	29,4	4,0	32,0	26,8	4,9
Pct.	100%	87,5%	11,5%	100%	87,2%	11,9%	100%	83,7%	15,4%

Er zijn tussen scholen nogal wat verschillen in aandeel leerlingen met schoolse vertraging. De tien procent scholen met het kleinste aandeel leerlingen met vertraging telt minder dan 2% leerlingen met één of meer jaar schoolse vertraging. Aan de andere kant van het spectrum vinden we tien procent scholen met aan het begin van L1 25% of meer leerlingen met een vertraging van één of meer schooljaren en op het einde van L2 zelfs 35% of meer leerlingen met vertraging.

Het effect van geboortejaar en geboortemaand op de leerprestaties en de leerwinst

Er is een complex samenspel tussen de effecten van geboortejaar, geboortemaand, zittenblijven én initiële taalvaardigheid op de leerwinst voor **wiskunde** in het eerste en het tweede leerjaar.

Figuur 2.1 toont de effecten van geboortejaar, geboortemaand en zittenblijven ná correctie voor de effecten van initiële taalvaardigheid, sekse en sociale achtergrond. Op de horizontale as staat de leeftijd van de leerling uitgedrukt in maanden op 1 september 2003 (d.i. de start van het eerste leerjaar in het SiBO-onderzoek). De verticale as geeft het vaardigheidsniveau voor wiskunde aan. De gebroken, trapvormige lijnen in de grafiek tonen de samenhang tussen vaardigheidsniveau en leeftijd op drie meetmomenten: begin L1 (onderste twee lijnen), einde L1 (middelste twee lijnen) en einde L2 (hoogste drie lijnen). Het verschil tussen de onderste lijnen en de overeenkomstige middelste weerspiegelt de leerwinst voor wiskunde in L1, het verschil tussen de middelste en de overeenkomstige bovenste lijnen toont de leerwinst in L2.

Leerlingen met een leeftijd tussen 69 en 80 maanden (op 1 sept. 2003) zijn leerlingen geboren in 1997. Dat zijn de leerlingen die in het eerste leerjaar “op leeftijd” waren. In de leeftijdscategorie tussen 81 tot 92 maanden vinden we de leerlingen geboren in 1996. Zij hadden dus op 1 september 2003 bij de intrede in het eerste leerjaar één jaar achterstand. Het volgende interval (93 tot 104 maanden) toont de leerlingen geboren in 1995, met twee jaar achterstand op 1 september 2003, enz. Links in de grafiek, in het leeftijdsinterval van 57 tot 68 maanden, vindt men de leerlingen geboren in 1998. Dat zijn leerlingen die het eerste leerjaar met één jaar voorsprong begonnen.

Voor elk van de drie meetmomenten is een globaal neerwaartse trend te zien. Bij de leerlingen die het eerste leerjaar met één of twee jaar achterstand begonnen, zien we gemiddeld een lagere aanvangsscore voor wiskunde én een kleinere leerwinst in het eerste en in het tweede leerjaar. (Men maakt best abstractie van de leerlingen met drie of vier jaar achterstand op 1 sept. 2003; het gaat in die categorieën om zeer beperkte aantallen.)

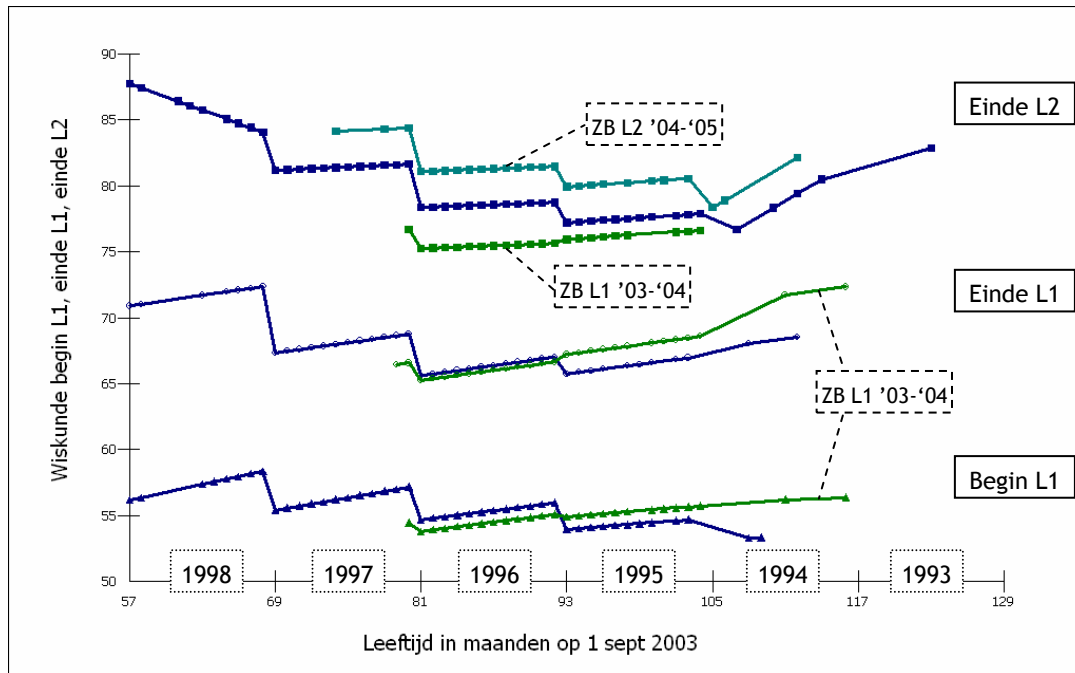


Fig. 2.1 Samenhang van geboortjaar, geboortemaand en zittenblijven in L1 en L2 met vaardigheidsniveau voor wiskunde op drie meetmomenten na correctie voor sekse, sociale achtergrond en initieel taalvaardigheidsniveau

De neerwaartse trend is sterker op het einde van L2. Ook al kennen leerlingen na een eerder opgelopen achterstand, in het eerste en het tweede leerjaar verder een “normale” schoolloopbaan, de kloof in wiskundevaardigheid met de echt normaalvorderende leerlingen wordt steeds groter. Wie het eerste leerjaar met twee jaar achterstand begon, maakt wel wat meer leerwinst in het eerste leerjaar, maar raakt vervolgens in het tweede leerjaar weer meer achterop. Het beperkte aantal leerlingen dat een jaar “te vroeg” aan het eerste leerjaar begon, heeft gemiddeld niet alleen een hogere beginscore, maar maakt ook nog eens meer leerwinst in het eerste en het tweede leerjaar.

Die algemene samenhang tussen geboortjaar en vaardigheidsniveau wordt doorkruist door het *geboortemaandeffect*. Wie eerder op het jaar geboren is, start het eerste leerjaar vanuit een hoger vaardigheidsniveau voor wiskunde. Dit tegengestelde effect (binnen elk geboortjaar zijn het nu de oudere leerlingen die een voordeel hebben) zorgt voor de “getrapte vorm” van de lijnen. Men kan zien dat de hellingsgraad van die trapjes afneemt (1) naarmate het kind in een vroeger jaar geboren is, en (2) bij elk volgend meetmoment. Het geboortemaandeffect neemt dus met de tijd af en geldt ook minder voor leerlingen die het eerste leerjaar met achterstand begonnen.

Die twee effecten worden nog eens doorkruist door het kenmerk “zittenblijven”. Sommige leerlingen liepen hun achterstand in het kleuteronderwijs op. Anderen deden in 2003-2004 het eerste leerjaar al voor de tweede maal. Die laatsten (de “zittenblijvers L1 '03-'04”) zijn “nieuw” in onze steekproef (zie hoofdstuk 1). Opvallend genoeg beginnen de “zittenblijvers L1 '03-'04” die slechts één jaar vertraging hebben, het eerste leerjaar (dat ze al eens deden!) met een wiskundeniveau dat - welis-

waar na correctie voor verschillen in taalvaardigheid - nog iets lager ligt dan bij hun leeftijdsgenoten (die ook één jaar vertraging hebben, maar die vertraging eerder opliepen). Ze maken in L1 wel wat meer leerwinst dan de andere leerlingen met een jaar achterstand, maar verliezen in de loop van het tweede leerjaar dan weer drastisch terrein.

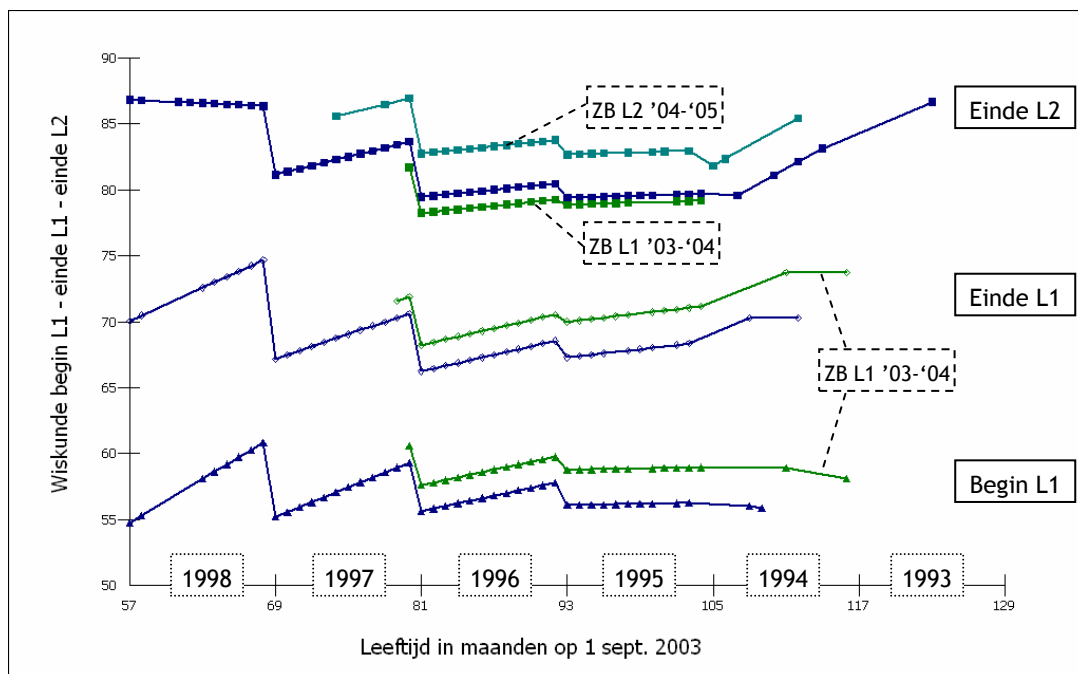


Fig. 2.2 Samenhang van geboortemaand en zittenblijven in L1 en L2 met wiskundevaardigheid op drie meetmomenten na correctie voor sekse en sociale achtergrond maar niet voor initieel taalvaardigheidsniveau (Vlaanderen)

Bij de intrede in L2 komen we opnieuw enkele voor ons “nieuwe leerlingen” tegen: zij die het jaar voordien het tweede leerjaar al eens deden (de “zittenblijvers L2 '04-'05”). Die leerlingen blijken op het einde van het tweede leerjaar (schooljaar '04-'05) gemiddeld op eenzelfde niveau te komen als de normaalvorderende leerlingen van het geboortemaand 1997.

Figuur 2.1 geeft de samenhangen weer na correctie voor de effecten van onder meer aanvankelijke Nederlandse taalvaardigheid. Wanneer we de verschillen in initiële Nederlandse taalvaardigheid laten meespelen (er niet voor corrigeren), dan zien we een scherper geboortemaandeffect (zie figuur 2.2). Ook ziet de situatie van de zittenblijvers L1 '03-'04 er iets rooskleuriger uit: zij starten en eindigen het eerste leerjaar op een gelijk niveau als de normaalvorderende leerlingen uit geboortemaand 1997. In de loop van het tweede leerjaar verliezen ze toch terug terrein, al is de extra achterstand t.o.v. leeftijdsgenoten die eerder een schoolse achterstand opliepen niet zo groot.

Voor technisch lezen en spelling vonden we vrij gelijklopende samenhangen. In figuur 2.3 worden de samenhangen met *technisch lezen* afgebeeld.

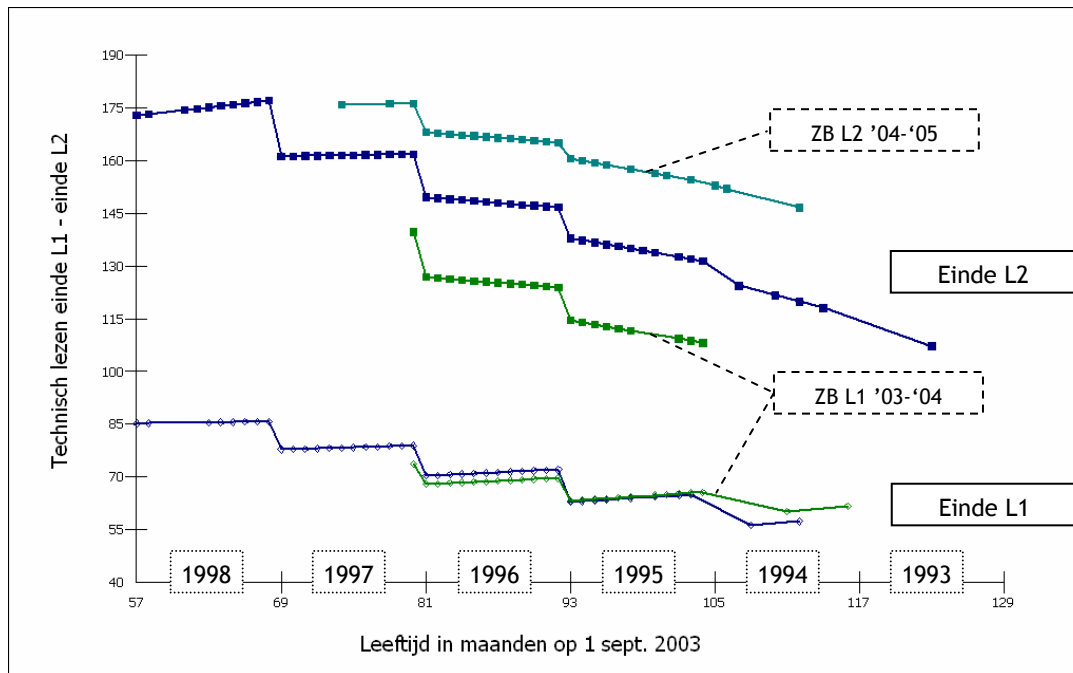


Fig. 2.3 Samenhang van geboortemaand, geboortemaand en zittenblijven in L1 en L2 met vaardigheidsscore voor technisch lezen op twee meetmomenten na correctie voor sekse, sociale achtergrond en initiële taalvaardigheid

We zien op de twee meetmomenten een duidelijk effect van het geboortemaand. Kinderen uit vroegere geboortemaanden scoren lager. De zittenblijvers L1 '03-'04 scoren op het einde van het eerste leerjaar niet of nauwelijks lager dan de andere leerlingen met een even grote schoolse achterstand. Maar één jaar verder hebben de zittenblijvers L1 weer duidelijk een extra achterstand opgelopen. Het verschil bedraagt ruim 20 punten *bovenop de achterstand die samenhangt met het geboortemaand* en is significant. De leerlingen die in 2004-2005 het tweede leerjaar overdeden, hebben op het einde van dat tweede leerjaar wel (nog) een voorsprong. Die voorsprong is behoorlijk groot (gemiddeld 19 punten) wanneer met het eigen geboortemaand vergeleken wordt, maar klein wanneer met de normaalvorderende leerlingen (geboortemaand 1997) vergeleken wordt. Uit figuur 2.3 is duidelijk dat er voor technisch lezen van een geboortemaandeffect nauwelijks sprake is.

Corrigeren we niet voor de effecten van aanvankelijke Nederlandse taalvaardigheid, dan komt er voor technisch lezen wel een (licht) geboortemaandeffect tevoorschijn (zie figuur 2.4). Aan de verschillen tussen de geboortemaanden verandert weinig. Wel scoren zittenblijvers L1 op het einde van het eerste leerjaar nu iets hoger dan leeftijdsgenoten die hun achterstand eerder opliepen, maar in de loop van het tweede leerjaar lopen de zittenblijvers L1 weer een flinke extra achterstand op.

Figuur 2.5 toont de samenhangen (op Vlaams niveau) tussen geboortemaand, geboortemaand, zittenblijven en *spelling*vaardigheid einde L1 en einde L2. Het beeld is zeer gelijkend aan dat voor technisch lezen, al is er nu wel een geboortemaandeffect zichtbaar, zij het een "omgekeerd" geboortemaandeffect bij de leerlingen die een jaar voorop zijn en zij die twee jaar achterop zijn.

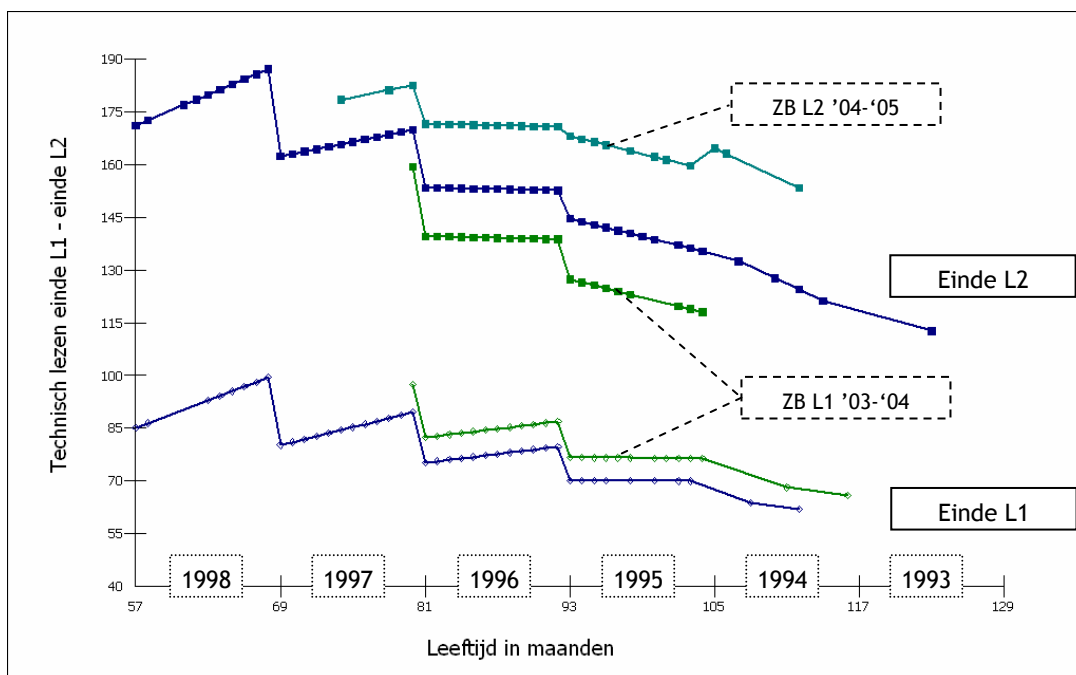


Fig. 2.4 Samenhang van geboortemaand en zittenblijven in L1 en L2 met vaardigheid in technisch lezen op twee meetmomenten na correctie voor sekse en sociale achtergrond, maar niet voor initiële taalvaardigheid

Enkel bij de normaalvorderende leerlingen met geboortemaand 1997 doet zich een “gewoon” geboortemaandeffect voor. Dat effect is klein (-0,1 punt per maand later geboren), maar significant ($p < 0,01$), althans op het einde van L1. Een jaar later is het geboortemaandeffect voor de normaalvorderende

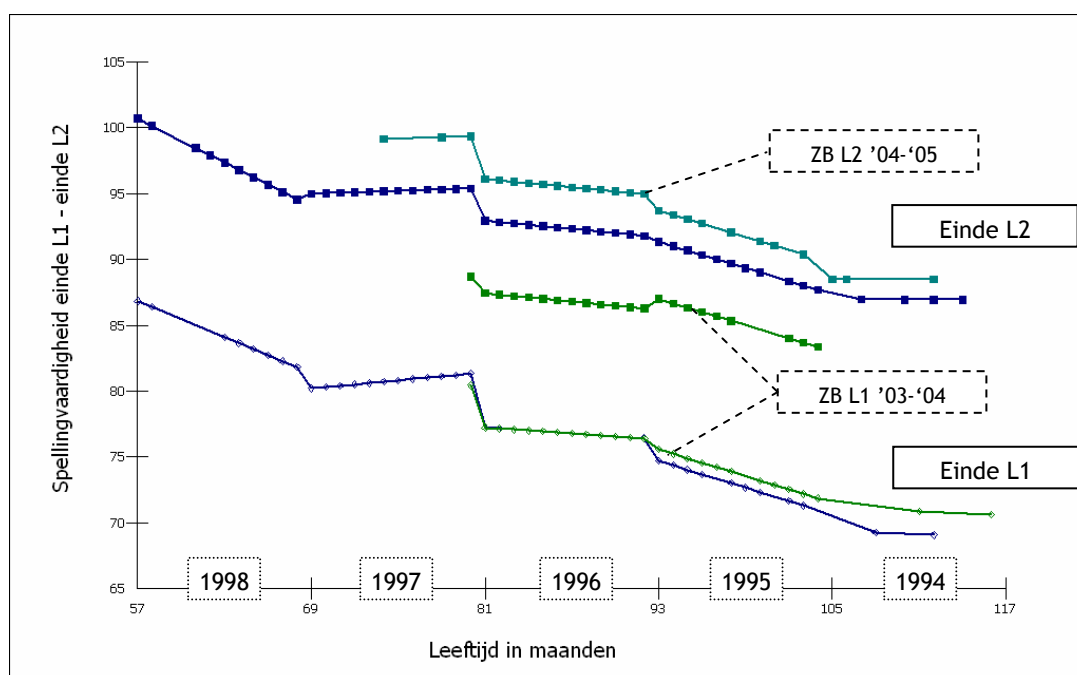


Fig. 2.5 Samenhang van geboortemaand en zittenblijven in L1 en L2 met spellingvaardigheid op twee meetmomenten na correctie voor sekse, sociale achtergrond en aanvankelijk taalvaardigheidsniveau

leerlingen uit 1997 nog kleiner en niet langer significant. Leerlingen die in 2003-2004 het eerste leerjaar overdeden (zittenblijvers L1 '03-'04) halen op het einde van L1 een significant lagere score dan normaalvorderende leerlingen (gemiddeld 3,8 punten minder, $p < 0,001$ voor wie in 1996 geboren is), maar scoren daarmee wel even hoog (voor geboortjaar 1996) als of zelfs iets hoger (voor geboortjaar 1995) dan de andere leerlingen met een even grote schoolse achterstand. Of leerlingen hun achterstand (ook) in het eerste leerjaar opliepen of (enkel) eerder, maakt voor de score op spelling einde eerste leerjaar bijgevolg geen enkel verschil uit. Maar tegen het einde van het tweede leerjaar bouwen die zittenblijvers L1 '03-'04 wel een significante extra achterstand van gemiddeld 5,5 punten op ($p < 0,05$), waardoor hun achterstand t.o.v. de normaalvorderende leerlingen tegen het einde van L2 gemiddeld ruim 8 punten bedraagt.

Niet corrigeren voor aanvankelijke Nederlandse taalvaardigheid laat voor de normaalvorderende leerlingen wel een scherper geboortemaandeffect zien, maar verandert nauwelijks iets aan de verschillen tussen geboortejaren (zie fig. 2.6). Rekenen we de verschillen in initiële Nederlandse taalvaardigheid gewoon door (= niet corrigeren), dan blijken zittenblijvers L1 '03-'04 op het einde van L1 gemiddeld wel een voorsprong te hebben van gemiddeld 2,7 punten. Maar dat voordeel - dat ze dus aan hun gemiddeld hogere Nederlandse taalvaardigheid te danken hebben - verliezen ze in de loop van het tweede leerjaar. Ondanks hun grotere Nederlandse taalvaardigheid bij het begin van hun bisjaar in L1 bouwen zittenblijvers L1 in vergelijking met leeftijdsgenoten die eerder vertraging opliepen, in de loop van het tweede leerjaar voor spelling een behoorlijke extra achterstand op, waardoor hun totale achterstand t.o.v. normaalvorderende leerlingen tegen einde L2 nog vergroot.

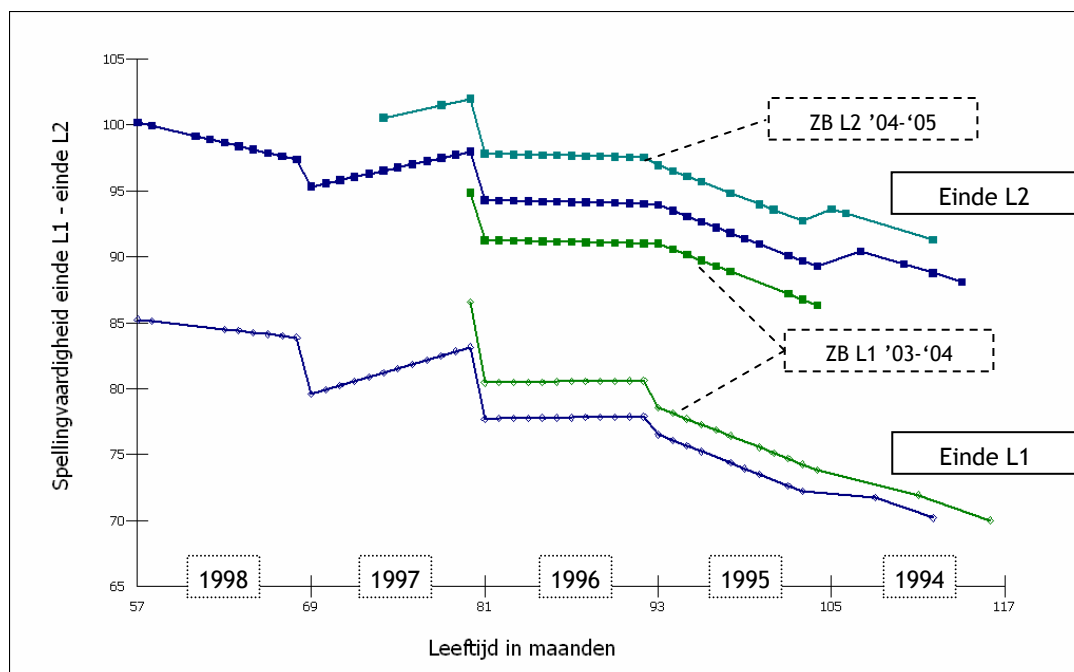


Fig. 2.6 Samenhang van geboortjaar, geboortemaand en zittenblijven in L1 en L2 met spellingvaardigheid op twee meetmomenten na correctie voor sekse en sociale achtergrond, maar niet voor aanvankelijk taalvaardigheidsniveau

Samenvattend voor de drie leerdomeinen kunnen we besluiten:

- Er is slechts in eerder beperkte mate sprake van een *geboortemaandefect*.
 - Dat geboortemaandefect is groter voor wiskunde en voor spelling dan voor technisch lezen.
 - Het is duidelijk geassocieerd met verschillen in aanvankelijke Nederlandse taalvaardigheid. Van zodra we corrigeren voor de effecten van die initiële Nederlandse taalvaardigheid, verkleint het geboortemaandefect (wiskunde, spelling) of verdwijnt het zelfs volledig (technisch lezen). Dat kinderen die eerder op het jaar geboren zijn, iets beter presteren, komt dus (groten)deels door hun grotere aanvankelijke Nederlandse taalvaardigheid.
 - Het geboortemaandefect is meer uitgesproken voor normaalvorderende leerlingen (met geboortjaar 1997) en aan het begin of het einde van het eerste leerjaar. Het is veel kleiner of onbestaande voor leerlingen die met vertraging in het eerste leerjaar aankomen of het eerste leerjaar overdoen. Het wordt ook kleiner naarmate leerlingen vorderen in hun schoolloopbaan. Voor wiskunde en technisch lezen is er ook bij de leerlingen die een jaar voorop zijn een uitgesproken geboortemaandefect. Dat is ook voor spelling het geval, maar dan wel in de omgekeerde zin: wie een jaar voorop, scoort voor spelling nog beter naarmate men later op het jaar geboren is.
- Er zijn duidelijke verschillen in leerprestaties naargelang van het geboortjaar van de kinderen, d.w.z. naargelang ze al of niet schoolse achterstand opgelopen hebben. Halen kinderen met vertraging - dankzij hun naar verhouding betere Nederlandse taalvaardigheid - in het eerste leerjaar nog een niveau dat vergelijkbaar is met of maar weinig verschilt van het niveau van normaalvorderende leerlingen, tegen het eind van het tweede leerjaar vertonen ze een onmiskenbare achterstand.
- Leerlingen die in het *eerste leerjaar* een achterstand oplopen (zittenblijvers L1) vormen een speciale categorie binnen de groep leerlingen met schoolse achterstand.
 - Door hun hogere initiële Nederlandse taalvaardigheid hebben die leerlingen in het eerste leerjaar een voetje voor op leerlingen die eerder achterstand opliepen. Ze scoren in L1 gemiddeld ongeveer even hoog als normaalvorderende leerlingen. Corrigeren we voor dat taalvaardigheidseffect, dan scoren ze in L1 niet hoger dan wie eerder vertraging opliep.
 - Leerlingen die het eerste leerjaar overgedaan hebben, boeken in het tweede leerjaar minder leerwinst, zowel in vergelijking met normaalvorderende leerlingen als in vergelijking met hun leeftijdsgenoten die vóór het eerste leerjaar vertraging opliepen. Daardoor wordt hun achterstand tegen het einde van L2 groter dan bij die leeftijdsgenoten. Zonder hun gemiddeld hogere aanvankelijke Nederlandse taalvaardigheid zou de achterstand van zittenblijvers L1 trouwens gemiddeld nóg groter zijn. De extra achterstand voor zittenblijvers L1 is het meest uitgesproken bij technisch lezen en spelling, maar zien we toch ook bij wiskunde.

- o Leerlingen die het tweede leerjaar overdoen presteren op het einde van L2 gemiddeld even goed als of (bij technisch lezen en spelling) zelfs iets beter dan normaalvorderende leerlingen. Van enig effect van aanvankelijke Nederlandse taalvaardigheid is hier geen sprake.

Tabel 2.4 geeft een vereenvoudigd overzicht van de verschillende leeftijdsgebonden effecten op de prestaties en leerwinst voor wiskunde, technisch lezen en spelling in het eerste en het tweede leerjaar. Vanwege de wellicht grotere relevantie voor de praktijk beperken we ons in tabel 2.4 tot de effecten *niet* gecorrigeerd voor verschillen in aanvankelijke Nederlandse taalvaardigheid, maar wel gecorrigeerd voor sekse en sociale achtergrond van de leerlingen.

Tabel 2.4 Effecten van geboortemaand en van eerder verworven schoolse vertraging op beginniveau en leerwinst in het eerste en het tweede leerjaar na correctie voor sekse en sociale achtergrond van de leerling maar niet voor initiële Nederlandse taalvaardigheid

Toets	Effect	Gemiddeld effect op ... ¹			Gemiddelde leerwinst (LW) in Vlaamse referentiegroep	
		Beginniveau	Leerwinst L1	Leerwinst L2	Eerste leerjaar	Tweede leerjaar
Wiskunde	Geboortemaand ³	-3,2 % *	+0,5 % ***	+0,7 % **	11,48	13,43
	Geboortejaar ⁴	-3,9 %	-8,5 % ***	-6,8 % **		
	Zittenblijven L1 ⁵	+10,8 %	+0,2 %	-23,8 % ***		
Technisch lezen	Geboortemaand ³		-1,0 % ***	+0,2 %	78,30 ⁽²⁾	83,19
	Geboortejaar ⁴		-8,4 % ***	-6,8 % **		
	Zittenblijven L1 ⁵		+8,3 % **	-26,1 % ***		
Spelling	Geboortemaand ³		-0,4 % ***	+0,5 % *	80,71 ⁽²⁾	14,46
	Geboortejaar ⁴		-4,3 % ***	+7,1 % *		
	Zittenblijven L1 ⁵		+3,4 % ***	-37,4 % ***		

¹ Effect op beginniveau en leerwinst L1 in % t.o.v. gemiddelde leerwinst L1 bij de Vlaamse referentiegroep. Effect op leerwinst L2 in % t.o.v. gemiddelde leerwinst L2 bij de Vlaamse referentiegroep. + wijst op voorsprong/inhaalbeweging, - op (extra) achterstand. Significantieniveaus: * p < 0,05 (significant) ** p < 0,01 *** p < 0,001 (heel significant).

² Deze getallen betreffen in feite het gemiddelde prestatieniveau einde eerste leerjaar, maar mogen voor leerlingen uit de Vlaamse referentiegroep als leerwinst voor het eerste leerjaar opgevat worden gezien leerlingen bij het begin van L1 in Vlaanderen niet geacht worden al te kunnen lezen of schrijven.

³ Extra voor- (+) of nadeel (-) per maand later geboren voor normaalvorderende leerlingen (° 1997).

⁴ Effect voor 1 jaar schoolse vertraging. Dit effect is niet lineair, maar neemt toe of af naarmate de vertraging meer jaren omvat, afhankelijk van leergebied en meetmoment. Zie figuren 2.2, 2.4 en 2.6.

⁵ Gemiddeld extra effect *bovenop* effect per jaar schoolse vertraging voor leerlingen die vertraging oplopen door zittenblijven in het eerste leerjaar. Om de gemiddelde achterstand van een zittenblijver geboren in 1996 te schatten, moet men het effect voor zittenblijver L1 en het effect voor geboortejaar bij elkaar optellen.

Het geboortemaandeffect dat in tabel 2.4 wordt weergegeven, is het effect voor normaalvorderende leerlingen (geboren in 1997). Voor andere geboortejaren is de omvang van het geboortemaandeffect verschillend (zie figuren 2.1 tot 2.6). Het lijkt op het eerste zicht veel kleiner dan de effecten van geboortejaar en zittenblijven, maar om bijvoorbeeld het verschil in beginniveau te schatten van twee kinderen die met zes maanden verschil geboren zijn, moet het betreffende percentage in tabel 2.4 met 6 vermenigvuldigd worden. Het negatieve percentage voor beginniveau wiskunde (-3,2 %) duidt erop dat kinderen die later op het jaar geboren een lager beginniveau hebben. De positieve

percentages bij leerwinst L1 (+0,5 %) en bij leerwinst L2 (+0,7 %) wijzen erop dat het negatieve geboortemaandeffect *afzwakt*. Op het einde van L1 rest er nog een *negatief* geboortemaandeffect ten belope van 2,7% van de gemiddelde leerwinst in L1 ($-3,2 + 0,5 = -2,7$). Tegen het einde van L2 is het negatieve geboortemaandeffect voor wiskunde nog verder afgezwakt. Iets gelijkaardigs geldt voor technisch lezen en spelling. Er is eerst (op het einde van L1) een negatief geboortemaandeffect, dat vervolgens (tegen einde L2) afgezwakt wordt. De noemers waarop de percentages berekend zijn, vindt men rechts in de tabel, onder “gemiddelde leerwinst in de Vlaamse referentiegroep”.

De effecten voor zittenblijven in tabel 2.4 moeten bekeken worden in samenhang met de effecten voor geboortjaar. Wie in 1996 geboren is (en die vertraging van een jaar niet in het eerste leerjaar opgelopen heeft), begint het eerste leerjaar met een achterstand voor wiskunde ten belope van 3,9% van de gemiddelde leerwinst in L1. Maar voor zittenblijvers L1 geboren in 1996 wordt dat overschaduwed door een tegengesteld effect dat overeenkomt met 10,8% van de gemiddelde leerwinst in L1. Zij beginnen L1 dus per saldo met een voorsprong voor wiskunde ten belope van 6,9 % van de gemiddelde leerwinst voor wiskunde in L1 ($-3,9 + 10,8 = +6,9$). Wat voor elk van de leergebieden opvalt, is het grote negatieve effect van zittenblijven op de leerwinst in L2: -23,8% voor wiskunde, -26,1% voor technisch lezen en -37,4% voor spelling. En daar moet dan telkens nog eens het geboorteaareffect op de leerwinst in L2 bijgeteld worden.

Spelling is het enige leergebied waarvoor een positief geboorteaareffect genoteerd wordt t.a.v. de leerwinst in L2. Dat wil zeggen dat leerlingen die vóór het eerste leerjaar al een vertraging van één jaar opgelopen hadden (en dat geldt in nog iets grotere mate ook voor wie in 1995 geboren is en dus twee jaar vertraging heeft) in het tweede leerjaar méér leerwinst boeken dan normaalvorderende leerlingen. Dat die leerlingen tegen het einde van L2 per saldo toch nog op achterstand staan, heeft te maken met het feit dat die grotere leerwinst in L2 niet opweegt tegen de (in absolute termen) veel grotere achterstand op het einde van L1 (4,3 % van 80,71 > 7,1 % van 14,46).

2.2 Sociale en etnisch-culturele achtergrond

In dit rapport brengen we de sociale en etnisch-culturele achtergrond van de leerlingen in rekening door de leerlingen op te delen in verschillende categorieën. We gaan daarbij uit van de categorieën die in het Gelijke Onderwijskansenbeleid (GOK-beleid) van de Vlaamse overheid gehanteerd worden. In het kader van het ondersteuningsluik van het GOK-decreet spreekt de Vlaamse overheid van een GOK-leerling als de leerling beantwoordt aan één of meer van de volgende voorwaarden:

- de moeder van de leerling is laag opgeleid (geen diploma secundair onderwijs),
- het gezin leeft van een vervangingsinkomen,
- het gezin behoort tot de zgn. “trekkende bevolking”,
- het kind wordt buiten het eigen gezin opgevoed (pleeggezin, geplaatst...).

Als in combinatie met één of meer van die vier criteria bij een leerling thuis geen Nederlands gesproken wordt, wordt een extra gewicht toegekend. Maar een leerling die enkel “thuis taal niet-Nederlands” als kenmerk heeft en geen van de vier andere kenmerken, komt niet voor extra subsidiëring via de GOK-lestijden in aanmerking.

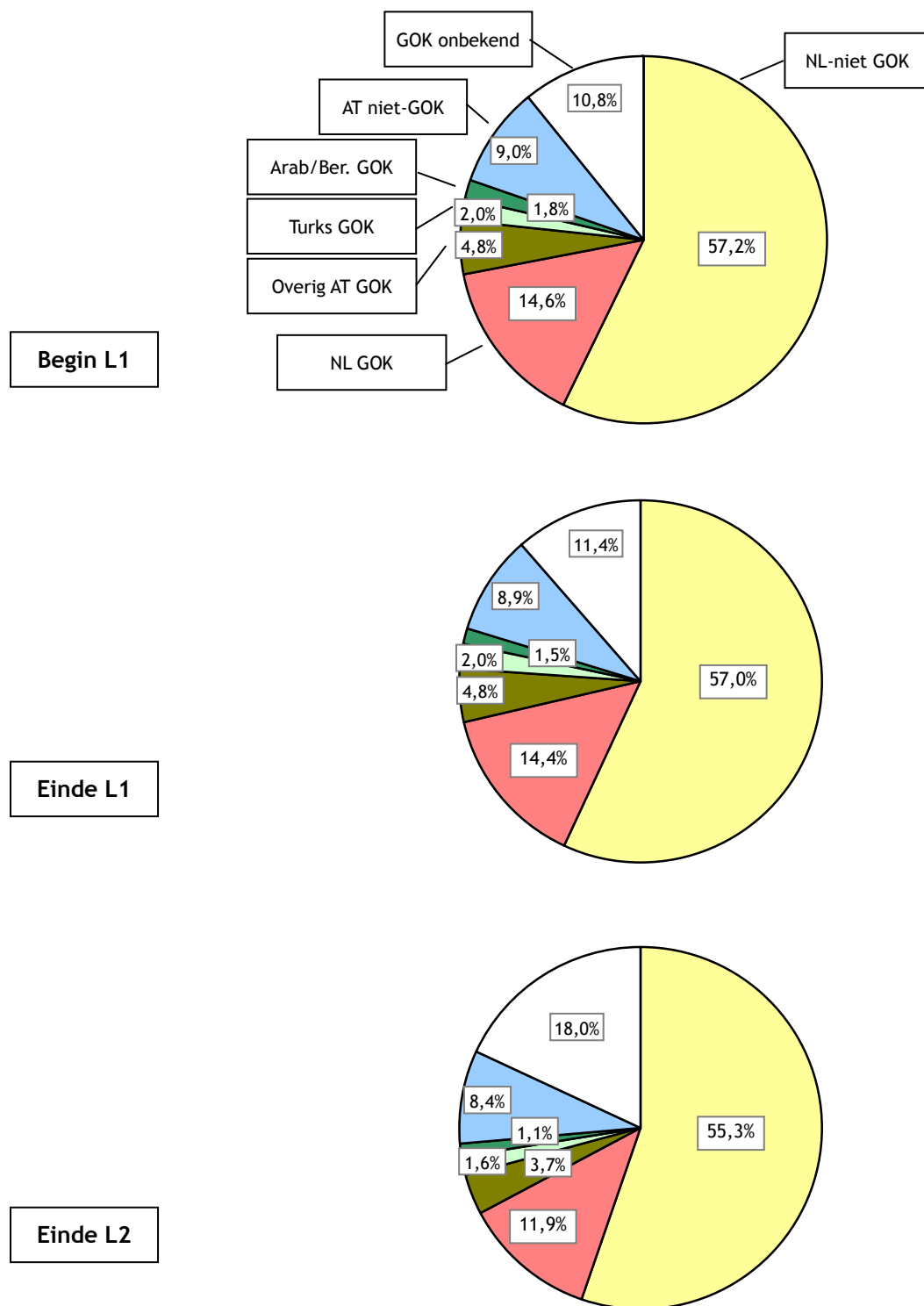
Op basis van de GOK-criteria kan een onderscheid gemaakt worden tussen:

- Nederlandstalige niet-GOK-leerlingen (die we verder ook modale Vlaamse leerlingen noemen),
- Nederlandstalige GOK-leerlingen,
- anderstalige GOK-leerlingen,
- anderstalige niet-GOK-leerlingen.

Tot 2007 werden alle nodige GOK-gegevens door de scholen op leerlingniveau verzameld, maar naar naar de Vlaamse overheid werden alleen samenvattende tabellen op school- of vestigingsplaatsniveau doorgestuurd. Binnen het SiBO-onderzoek werden de oorspronkelijke gegevens op leerlingniveau *gereconstrueerd* op basis van gegevens uit een oudervragenlijst, afgenomen toen de leerlingen in het eerste leerjaar zaten. De SiBO-gegevens laten een meer verfijnde indeling naar taalgroep toe. Omdat gebleken is dat binnen de categorie anderstalige GOK-leerlingen het beginniveau en de leerwinst voor Turkse en Arabisch of Berber sprekende leerlingen sterk afwijkt van die van overige anderstalige GOK-leerlingen werd de categorie anderstalige GOK-leerlingen nog verder opgesplitst in drie subcategorieën: (1) Turkse GOK-leerlingen, (2) Arabisch of Berber sprekende GOK-leerlingen en (3) overige anderstalige GOK-leerlingen.

Binnen het SiBO-onderzoek beschikken we voor sommige leerlingen dan weer niet over de nodige gegevens om hun “GOK-status” te bepalen. Om die reden voegden we nog een categorie “GOK-status onbekend” toe. Dat levert een indeling met zeven categorieën op. Figuur 2.7 toont de samenstelling naar GOK-categorie voor “de gemiddelde Vlaamse school”.

De in- en uitstroom in de loop van het eerste leerjaar verandert nauwelijks iets aan de samenstelling van de schoolbevolking. De instroom bij de overgang naar het tweede leerjaar heeft wel tot gevolg dat het aandeel leerlingen van wie de GOK-status onbekend is met de helft toeneemt. Als we aannemen dat in- en uitstroom vooral de zwakkere categorieën van leerlingen betreft en dat de populatie als geheel bij het opschuiven van L1 naar L2 niet veel van samenstelling verandert, kunnen we op basis van de vaststelling dat bij de overgang naar L2 het aandeel modale Vlaamse leerlingen en het aandeel anderstalige niet-GOK-leerlingen zo goed als gelijk blijven, vermoeden dat die toename van “GOK-status onbekend” vooral GOK-leerlingen betreft.



Figuur 2.7 Verdeling van de leerlingen naar GOK-categorieën in de Vlaamse referentie-steekproef

Waarom is het aandeel leerlingen met "GOK-status onbekend" zoveel groter in het tweede leerjaar?

Een deel van de leerlingen stroomde nieuw in bij de overgang van het eerste naar het tweede leerjaar. Voor die leerlingen beschikken we nog niet over voldoende achtergrondgegevens om te bepalen tot welke GOK-categorie ze behoren. In de loop van 2007 wordt de oudervragenlijst opnieuw verspreid en dat zal ons helpen de ontbrekende gegevens aan te vullen.

Worden leerlingen met ontbrekende gegevens voor instroomkenmerken mee in de analyses opgenomen?

Ja. Voor zulke leerlingen werd een aparte categorie "gegeven ontbreekt" gemaakt, zoals hier geïllustreerd wordt met de GOK-categorieën. Dat gebeurde - waar nodig - ook voor de andere instroomkenmerken. Op die manier konden we ook die leerlingen in de analyses opnemen en de uitval beperken tot leerlingen die helemaal geen toetscores hebben afgelegd.

Wat met leerlingen uit taalgemengde gezinnen?

In de SiBO-gegevens wordt een afzonderlijke categorie gehanteerd voor leerlingen bij wie behalve een andere taal thuis ook Nederlands gesproken wordt. Volgens een uitspraak van de Commissie Leerlingenrechten kunnen zulke leerlingen niet beschouwd worden als leerlingen die beantwoorden aan de indicator "de thuistaal is niet het Nederlands". We hebben de leerlingen uit deze zgn. "taalgemengde gezinnen" toch bij de niet-Nederlandstalige leerlingen ingedeeld. Een aantal overwegingen speelde hierin een rol.

De meest doorslaggevende was dat bekeken vanuit de samenhang met de toetsresultaten begin en einde eerste leerjaar en einde tweede leerjaar er een grotere gelijkheid bleek tussen leerlingen uit "taalgemengde" gezinnen en leerlingen uit niet-Nederlandstalige gezinnen dan tussen leerlingen uit "taalgemengde" gezinnen en leerlingen uit Nederlandstalige gezinnen. Een andere overweging heeft betrekking op de twijfel die vanuit de scholen vaak geuit wordt omtrent de geldigheid van de bewering van vele allochtone ouders dat ze thuis "ook Nederlands" spreken. Een laatste overweging ten slotte was dat de afwijkingen van de referentiesteekproefverdeling t.o.v. de cijfers voor de Vlaamse populatie groter zouden zijn - meer bepaald wat de aandelen Nederlandstalige en niet-Nederlandstalige GOK-leerlingen betreft - als de kinderen uit "taalgemengde" gezinnen bij de Nederlandstaligen gerekend zouden worden.

Door die indeling tellen de categorieën "Turks" en "Arabisch-Berbers" ook de leerlingen voor wie uit de SiBO-data blijkt dat ze thuis Turks én Nederlands, resp. Arabisch of Berbers én Nederlands spreken. Uit een vergelijking van de samenhangen tussen thuistaal, resultaten op de toets voor aanvankelijke Nederlandse taalvaardigheid en de resultaten op de toetsen voor wiskunde, technisch lezen en spelling blijkt overigens dat de thuistaal van een leerling vooral als een indicator van zijn of haar etnisch-culturele achtergrond gezien moet worden en niet als een indicator van minder Nederlandse taalvaardigheid.

Waarom de verdeling naar GOK-categorieën belangrijk is...

Tussen de verschillende GOK-categorieën zijn er grote verschillen met betrekking tot het aanvangsniveau en de leerwinst voor wiskunde, technisch lezen of spelling in het eerste en het tweede leerjaar. Scholen met verschillende verdelingen van hun leerlingen over de GOK-categorieën zijn daarom onderling moeilijk te vergelijken, tenzij men corrigeert voor de effecten van het behoren tot deze of gene GOK-categorie. Of tenzij men de groeicurves uitsplitst per GOK-categorie. De grote verschillen tussen scholen met betrekking tot de verdeling naar GOK-categorie roept bovendien ook de vraag op of het geconcentreerd samenzitten van leerlingen uit bepaalde GOK-categorieën nog een bijkomend effect heeft (= groepscompositie-effect), bovenop de effecten van de eigen achtergrond van elke individuele leerling.

De samenhang tussen GOK-categorie en beginniveau / leerwinst worden besproken worden in de hoofdstukken waarin de resultaten per leerdomein aan bod komen. Ook de resultaten van de analyses met betrekking tot groepscompositie-effecten op basis van de GOK-categorieën worden in die hoofdstukken besproken.

2.3 Aanvankelijke Nederlandse taalvaardigheid (begin eerste leerjaar)

Bij de aanvang van het eerste leerjaar (begin september) werd van de leerlingen een Nederlandse taalvaardigheidstoets afgenomen. Deze toets is gebaseerd op de CiTO-toets *Taal voor Kleuters* en peilt enerzijds naar een aantal “leesvoorwaarden” (visuele en auditieve discriminatie, auditieve synthese, ...) en anderzijds naar de vaardigheid in kritisch luisteren. Een meer gedetailleerd beschrijving vindt men in Verachtert e.a. (2005). De score op de Nederlandse taalvaardigheidstoets begin L1 werd zowel in de analyses voor wiskunde als in de analyses voor spelling en technisch lezen als voorspellende variabele opgenomen.

Daarnaast werd ook een multiniveau-analyse uitgevoerd met aanvankelijke taalvaardigheid als afhankelijke variabele. Een opmerkelijke vaststelling bij die analyse is dat de achterstand in initiële Nederlandse taalvaardigheid voor anderstalige en Nederlandstalige GOK-leerlingen, jongens, zittenblijvers L1 en andere leerlingen met schoolse vertraging significant blijkt te verschillen tussen scholen. Overigens wordt slechts 30 % van de variantie op schoolniveau inzake initiële Nederlandse taalvaardigheid door verschillen in instroomkenmerken verklaard. Een mogelijke bron voor de resterende verschillen zou in het kleuteronderwijs kunnen liggen. Maar ook andere gemeenschappelijke factoren of ervaringen van de leerlingen (b.v. de buurt waarin ze wonen) kunnen voor die verschillen een verklaring bieden.

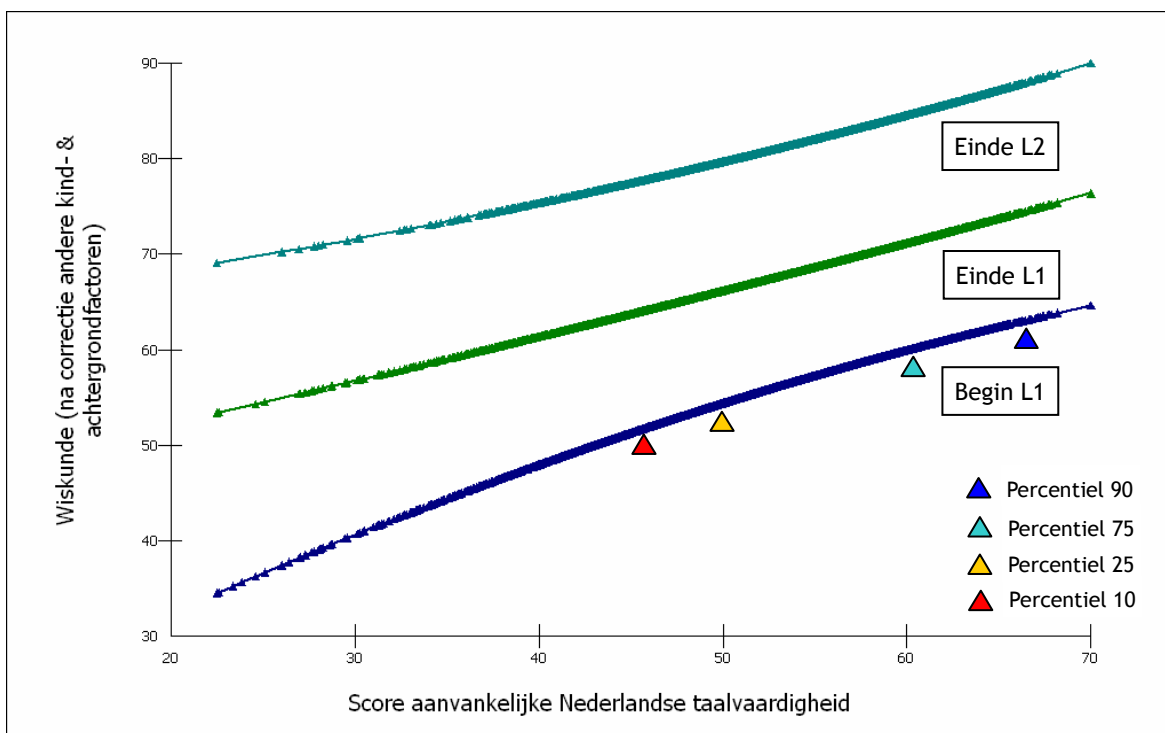
Waarom verschillen in aanvankelijke Nederlandse taalvaardigheid belangrijk zijn...

De initiële Nederlandse taalvaardigheid van de leerlingen werd niet alleen voor technisch lezen en spelling, maar ook voor wiskunde als voorspellende variabele opgenomen. Voor elk van die drie

leerdomeinen bleek de score voor aanvankelijke Nederlandse taalvaardigheid een significante factor te zijn, bovenop de invloed van de andere kindfactoren en van de achtergrondfactoren.

Figuur 2.8 toont de samenhang met de toetscores voor *wiskunde* op de drie meetmomenten. Bij het begin van L1 blijkt er sprake te zijn van een curvilineaire samenhang. De samenhang is sterker bij lagere scores voor aanvankelijke Nederlandse taalvaardigheid en zwakt geleidelijk aan iets af. In figuur 2.8 is duidelijk te zien dat vooral bij de allerlaagste scores (< percentiel 10) de samenhang het sterkst is. Tegen het eind van het eerste leerjaar is de samenhang nagenoeg lineair geworden. Dat betekent dat het de leerlingen met een zeer lage aanvankelijke Nederlandse taalvaardigheid zijn, die in het eerste leerjaar gemiddeld de grootste vooruitgang voor wiskunde maken. Die extra groei van kinderen met een heel lage aanvankelijke Nederlandse taalvaardigheid zet zich nog een beetje door in het tweede leerjaar, waardoor de curve zelfs iets in de andere richting ombuigt. Maar in het rechterdeel van de grafiek verandert de hellingsgraad van de curves niet van het ene naar het andere meetmoment. Aan de rechterkant lopen de drie curves zo goed als evenwijdig. Dat kan op twee wijzen geïnterpreteerd worden: (1) de invloed van de aanvankelijke Nederlandse taalvaardigheid op de scores voor wiskunde blijft doorspelen tot minstens het einde van het tweede leerjaar, óf (2) er is een derde, achterliggende factor in het spel die zowel op de scores voor aanvankelijke taalvaardigheid als op de scores voor wiskunde een invloed heeft en die ook in de verdere ontwikkeling van de kinderen een invloed blijft uitoefenen op (onder meer) de wiskundescores.

De samenhang tussen aanvankelijke Nederlandse taalvaardigheid en wiskundevaardigheid is behalve

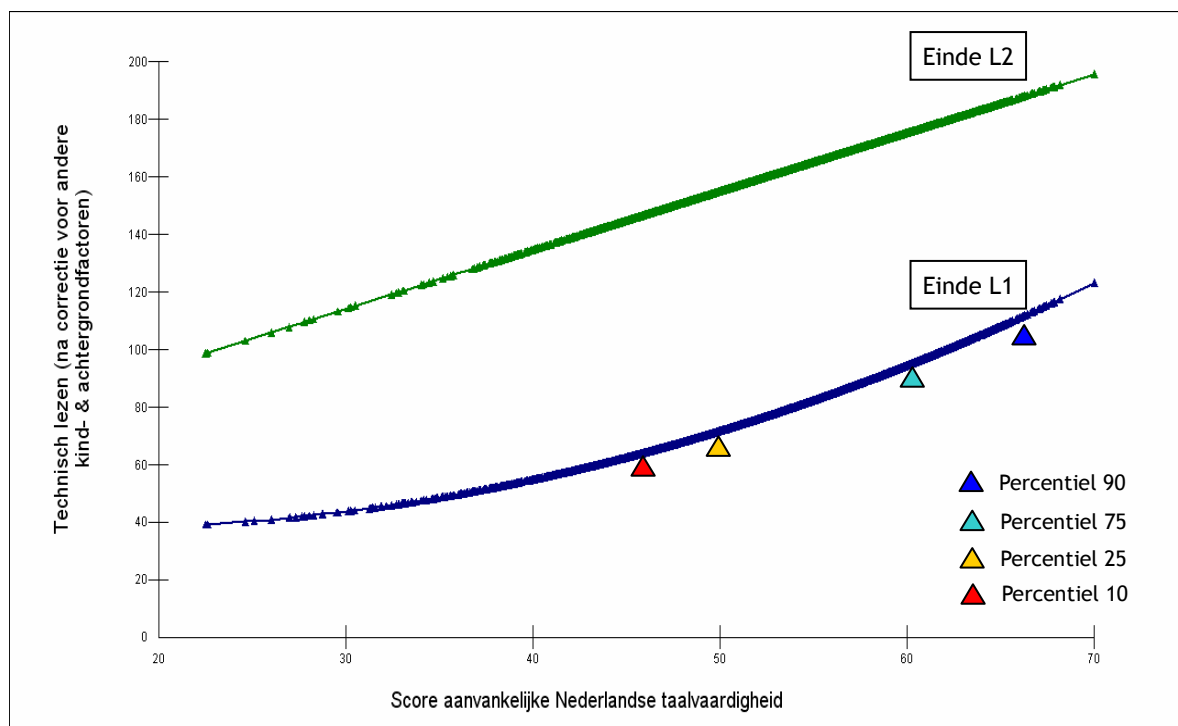


Figuur 2.8 Samenhang tussen de wiskundescores en de scores voor aanvankelijke Nederlandse taalvaardigheid op drie meetmomenten

significant ook behoorlijk groot ($ES = 0,57$). Het verschil in wiskundescore begin L1 van een leerling die voor initiële Nederlandse taalvaardigheid op het 25^e percentiel zit (▲) en van een leerling die voor taalvaardigheid het 75^e percentiel haalt (▲), komt overeen met ruim de helft van de leerwinst die in het eerste leerjaar door modale Vlaamse leerlingen gemiddeld voor wiskunde geboekt wordt. Voor leerlingen op resp. het 10^e (▲) en het 90^e (▲) percentiel voor aanvankelijke Nederlandse taalvaardigheid komt het verschil in wiskundescore begin L1 overeen met ruim de volledige leerwinst die in het eerste leerjaar door modale Vlaamse leerlingen gemiddeld voor wiskunde geboekt wordt.

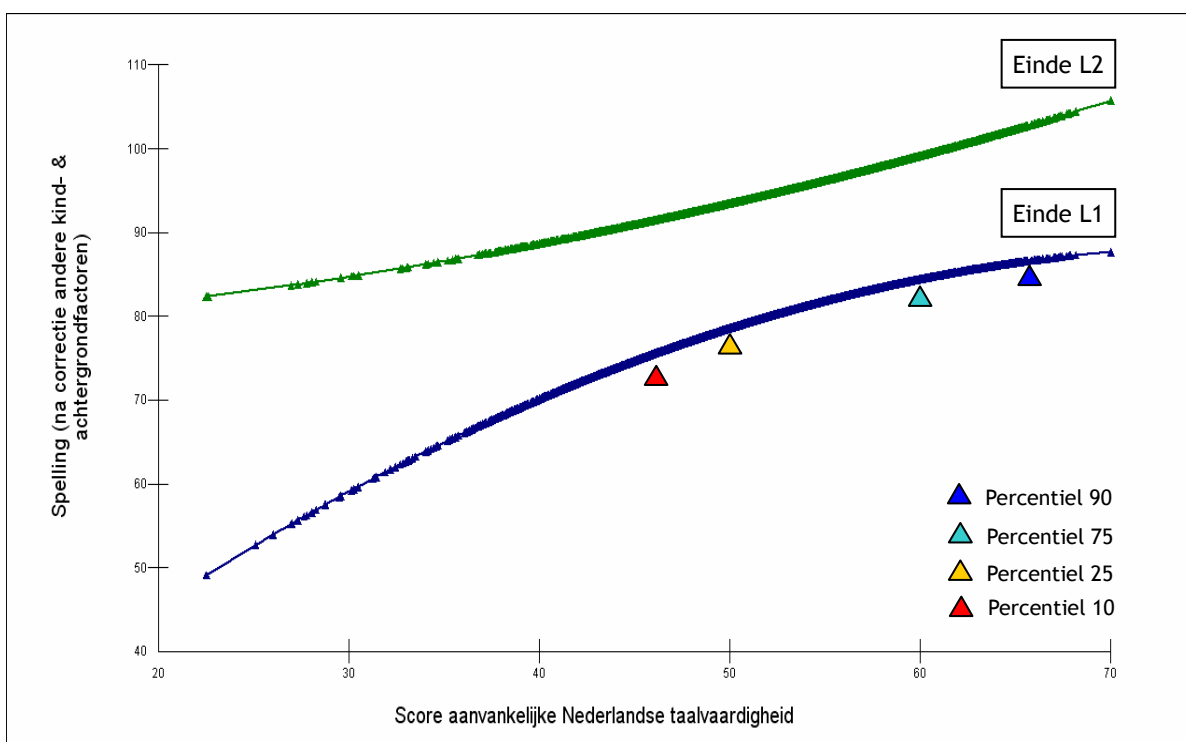
Voor *technisch lezen* vinden we op het einde van het eerste leerjaar een curvilineaire samenhang met de score voor aanvankelijke Nederlandse taalvaardigheid. Hier is sprake van een samenhang die sterker wordt naarmate de taalvaardigheidsscore hoger is (zie figuur 2.9). Een jaar later is de samenhang globaal nog altijd even groot, maar nu lineair. Het effect is minder groot dan bij wiskunde ($ES = 0,37$). Het verschil in score voor technisch lezen einde L1 tussen een leerling met een aanvankelijke Nederlandse taalvaardigheidsscore die overeenkomt met het 25^e percentiel en een leerling op het 75^e percentiel, komt overeen met ruim een kwart van de leerwinst voor technisch lezen die modale Vlaamse leerlingen gemiddeld in het tweede leerjaar boeken. Het P10-P90 interval voor aanvankelijke Nederlandse taalvaardigheid komt voor Vlaanderen overeen met goed de helft van die gemiddelde leerwinst in L2.

De samenhang van aanvankelijke Nederlandse taalvaardigheid met *spelling* is iets sterker dan de samenhang met technisch lezen en iets minder sterk dan de samenhang met wiskunde ($ES = 0,45$). Op



Figuur 2.9 Samenhang tussen de scores voor technisch lezen en de scores voor aanvankelijke Nederlandse taalvaardigheid op twee meetmomenten

het einde van het eerste leerjaar neemt de samenhang af naarmate de aanvankelijke Nederlandse taalvaardigheid hoger ligt (zie fig. 2.10). In de loop van het tweede leerjaar maken de taalzwakste leerlingen ook voor spelling een inhaalbeweging, waardoor tegen het einde van het tweede leerjaar de curve zelfs iets in de andere richting doorbuigt. Maar de globale sterkte van de samenhang verandert nauwelijks. Met het 25^e - 75^e percentiel-interval in aanvankelijke Nederlandse taalvaardigheid hangt op het einde van L1 een verschil in spellingvaardigheid samen dat overeenkomt met 45 % van de leerwinst in spellingvaardigheid die door modale Vlaamse leerlingen gemiddeld in L2 geboekt wordt; het P10-P90 interval is geassocieerd met een verschil in spellingvaardigheid dat gelijkstaat met ongeveer 90 % van die leerwinst. Dat zijn cijfers van dezelfde grootte-orde als bij wiskunde.



Figuur 2.10 Samenhang tussen de scores voor spelling en de scores voor aanvankelijke Nederlandse taalvaardigheid op twee meetmomenten (Vlaanderen)

2.4 Wat met de toets Rekenbegrip begin eerste leerjaar?

Naast de taalvaardigheidstoets werd bij het begin van het eerste leerjaar nog een toets aanvankelijk rekenbegrip afgenomen die peilt naar enkele voorbereidende rekenvaardigheden (b.v. vergelijken van hoeveelheden, rangordening, ...). De scores op de toets “Rekenbegrip begin eerste leerjaar” konden op één meetschaal gebracht worden met de scores op de toetsen wiskunde einde eerste leerjaar en einde tweede leerjaar. De resultaten van die toets worden daarom opgenomen in de groeicurves voor wiskunde die in hoofdstuk 3 gerapporteerd worden.

2.5 Besluit

Bij de analyses op de scores voor wiskunde, technisch lezen en spelling, waarover in de volgende hoofdstukken gerapporteerd zal worden, werd gecorrigeerd voor de effecten van de volgende instroomkenmerken: sekse van de leerlingen, geboortjaar en geboortemaand, het al dan niet overdoen van het eerste leerjaar in 2003-2004 of van het tweede leerjaar in 2004-2005, sociale en etnisch-culturele achtergrond van de leerlingen en hun aanvankelijke Nederlandse taalvaardigheid. Het gaat om individuele leerlingkenmerken die van invloed blijken te zijn op de prestaties en leerwinst voor wiskunde, technisch lezen en spelling. Om een faire en correcte vergelijking tussen scholen mogelijk te maken, is het daarom van belang voor deze instroomkenmerken te corrigeren.

Bij de bespreking van de effecten van de individuele leerlingkenmerken op de leerprestaties, hebben we aan twee elementen bijzondere aandacht besteed.

Het eerste element betreft de onderlinge verwevenheid van de instroomkenmerken, in het bijzonder het gegeven dat de initiële Nederlandse taalvaardigheid significant blijkt te verschillen tussen jongens en meisjes, tussen zittenblijvers en normaalvorderenden, tussen Nederlandstaligen en anderstaligen en tussen kansarme en niet-kansarme leerlingen. Daarom leek het relevant om de samenhangen van de leerresultaten met elk van de instroomkenmerken te bekijken mét én zónder correctie voor de score voor aanvankelijke Nederlandse taalvaardigheid.

Verder bleek ook dat de samenhangen tussen aanvankelijke Nederlandse taalvaardigheid (begin eerste leerjaar) en de overige instroomkenmerken significant verschillen van school tot school. Hier spelen mogelijks effecten van het kleuteronderwijs. Behalve effecten van proceskenmerken als zodanig, kan het ook om verschillen in participatiegraad gaan. Ook andere, nog niet in kaart gebrachte gemeenschappelijke kenmerken die leerlingen van de ene school zouden onderscheiden van de andere (en die b.v. te maken kunnen hebben met de buurt waarin de kinderen wonen), kunnen hierin een rol spelen.

3. Leerwinst en toegevoegde waarde voor wiskunde

De leerlingen in het SiBO-onderzoek kregen bij het begin van het eerste leerjaar en daarna telkens op het einde van elk leerjaar een toets wiskunde voorgelegd. De toetsen zijn speciaal voor het SiBO-onderzoek ontwikkeld. We bespreken eerst kort de inhoud van de toetsen en de ontwikkeling van de meetschaal voor wiskundevaardigheid. Daarna bespreken we de eigenlijke resultaten.

3.1 De toetsen wiskunde

De toets *Rekenbegrip begin eerste leerjaar* telt 40 opgaven die in vier rubrieken kunnen worden onderverdeeld: vergelijken van hoeveelheden (15 opgaven), begrippen van plaats en rang (10 opgaven), meten en rekentaal (10 opgaven) en tellen (ter bepaling van een hoeveelheid, 5 opgaven). De eerste rubriek peilt naar de beheersing van de begrippen '(hoeveel) meer', '(hoeveel) minder' en 'evenveel'. In de tweede rubriek gaat het om begrippen als 'voorlaatste', 'middelste' en 'juist na' evenals rangtelwoorden. De inhoud van *meten en rekentaal* is vrij verscheiden. Er wordt gepeild naar de beheersing van begrippen en uitdrukkingen als 'halfvol', 'even ver van', 'breedst', enzovoort. De opgaven uit de vierde rubriek tenslotte vragen de kinderen een bepaalde hoeveelheid te tellen. Analyse van de metrische kenmerken wijzen op een goede betrouwbaarheid en validiteit, één achterliggende algemene wiskundige vaardigheid en een ruime spreiding van de scores (Verachtert e.a., 2005).

De toets *Wiskunde einde eerste leerjaar* telt 70 opgaven, verdeeld over twee, apart af te nemen delen. Voor het eerste deel (40 opgaven) is geen tijdslimiet voorzien, voor het tweede deel (30 opgaven) wel. Inhoudelijk kunnen zeven rubrieken onderscheiden worden (zie tabel 3.1). Bij de opgaven 'getallenkennis' wordt gepeild naar inzicht in de getallenstructuur waarbij een beroep gedaan wordt op begrippen als 'meer dan', 'de helft', e.d. en van plaats- en rangordebegrippen (b.v. 'even ver van', 'twee plaatsen voor', enz.). De vraagstukken betreffen concrete probleemsituaties waarbij gerekend moet worden om tot de oplossing te komen. In 'toepassingsituaties metend rekenen' wordt gepeild naar de kennis van maateenheden en van de dagen van de week. De eenvoudige optel- en aftreksommen liggen binnen het getalbereik tot 20. Voor elk van beide reeksen is er een tijdslimiet van één minuut. Analyse van de metrische kenmerken wijzen op een hoge betrouwbaarheid en validiteit, één achterliggende algemene wiskundige vaardigheid, een ruime spreiding van de scores en een eerder hoge moeilijkheidsgraad (Verachtert e.a., 2005).

De toets *Wiskunde einde tweede leerjaar* telt 80 opgaven verdeeld over, afzonderlijk af te nemen delen met respectievelijk 50 en 30 opgaven. Inhoudelijk kunnen negen rubrieken onderscheiden worden (zie tabel 3.1). De meeste rubrieken bouwen voort op de overeenkomstige rubrieken uit de toets einde L1, zij het (aangevuld) met nieuwe, moeilijker begrippen en bewerkingen. 'Metend

rekenen' bijvoorbeeld bevat nu ook opgaven waarin gepeild wordt naar de kennis van de maanden en waarin met Euro gerekend moet worden. De twee reeksen eenvoudige optel- resp. aftreksommen liggen binnen het getalbereik tot 100; voor elke reeks is een tijdslimiet van anderhalve minuut voorzien. Nieuw zijn de rubrieken 'combinatie-opgaven'. Het gaat om hoofdrekenopgaven bijvoorbeeld $(100 : 2) - 20 = \dots$. Nieuw is ook de rubriek 'Meetkunde' met opgaven over ruimtelijke oriëntatie en vormleer. Analyses van de metrische kenmerken wijzen op een hoge betrouwbaarheid en goede validiteit, de mogelijkheid tot het berekenen van een score voor een ééndimensionele achterliggende algemene wiskundige vaardigheid en een ruime, zij het licht scheve spreiding van de scores (Hendriks e.a., 2005).

Tabel 3.1 Inhoud van de toetsen wiskunde einde L1 en einde L2

Toetsinhoud	Aantal opgaven einde L1	Aantal opgaven einde L2
Getallenkennis	10	10
Vraagstukken	10	10
Combinatie-opgaven		5
Structureren of splitsen van getallen	10	5
Getallenreeksen	5	5
Toepassingsituaties metend rekenen	5	10
Meetkunde		5
Totaal zonder tijdsdruk	40	50
Eenvoudige optelsommen met tijdslimiet	15	15
Eenvoudige aftreksommen met tijdslimiet	15	15
Totaal met tijdsdruk	30	30

IRT-calibratie

Het bleek mogelijk de scores van de verschillende toetsen door middel van IRT-calibratie op één meetschaal te plaatsen. Dat bevestigt dat er achter de kennis en vaardigheden die nodig zijn om de verschillende soorten opgaven correct te kunnen oplossen één algemene wiskundevaardigheid schuilt, waarvan de ontwikkeling over meerdere leerjaren door middel van de toetsen wiskunde gemeten kan worden.

In verband met de sommen die onder tijdsdruk werden afgenomen, stelt zich een probleem. Vanuit testtheoretische overwegingen kunnen die niet zomaar samen met de overige opgaven meegecalibreerd worden. Reeksen onder tijdsdruk peilen immers eerder naar de automatisatiegraad dan naar het beheersen op zich van de specifieke vaardigheid waar de opgave een beroep op doet. Met het oog op de validiteit van de scores werden de opgaven uit de tijdsdrukreeksen niet opgenomen in de calibratiescores die in de voorliggende analyses werden gebruikt.

Om tot een goede *fit* te komen, dienden bovendien uit de resterende groep opgaven nog een aantal opgaven verwijderd te worden. Het gaat om respectievelijk drie (begin L1), acht (einde L1) en één

opgave (einde L2). De niet mee verrekende opgaven zijn vrij evenwichtig over de verschillende rubrieken gespreid, zodat de inhoudelijke structuur van de toetsen niet in het gedrang kwam.

De doosdiagrammen in figuur 3.1 tonen (voor de referentiesteekproef) hoe de gecalibreerde toetsen elkaar qua bereik op de wiskundemeetschaal overlappen. De hoogte van elk “doosje” geeft de interkwartielafstand weer; het platte streepje onder en boven de verticale staafjes geven de grenzen aan waarbinnen 95% van de scores liggen. Een leerling die hoog scoort op de toets begin L1 komt even hoog op de meetschaal als een leerling die eerder laag scoort op de toets einde L1 en zelfs even hoog als een leerling die heel laag scoort op de toets einde L2. Tegelijk is uit de positie van de *doosjes* duidelijk dat de drie toetsen elkaar netjes *opvolgen* op de meetschaal, zoals verwacht kan worden van toetsen die geacht worden op verschillende, opeenvolgende meetmomenten een steeds verder ontwikkelende vaardigheid in kaart te brengen.

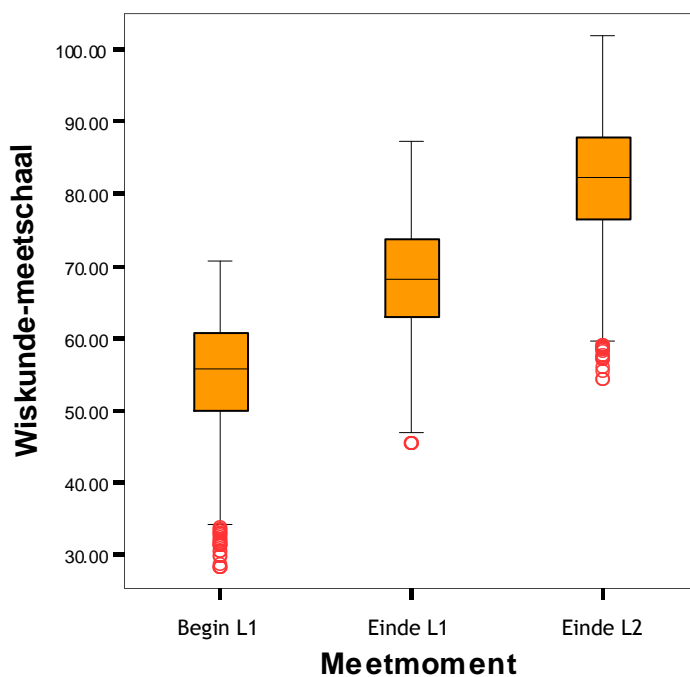
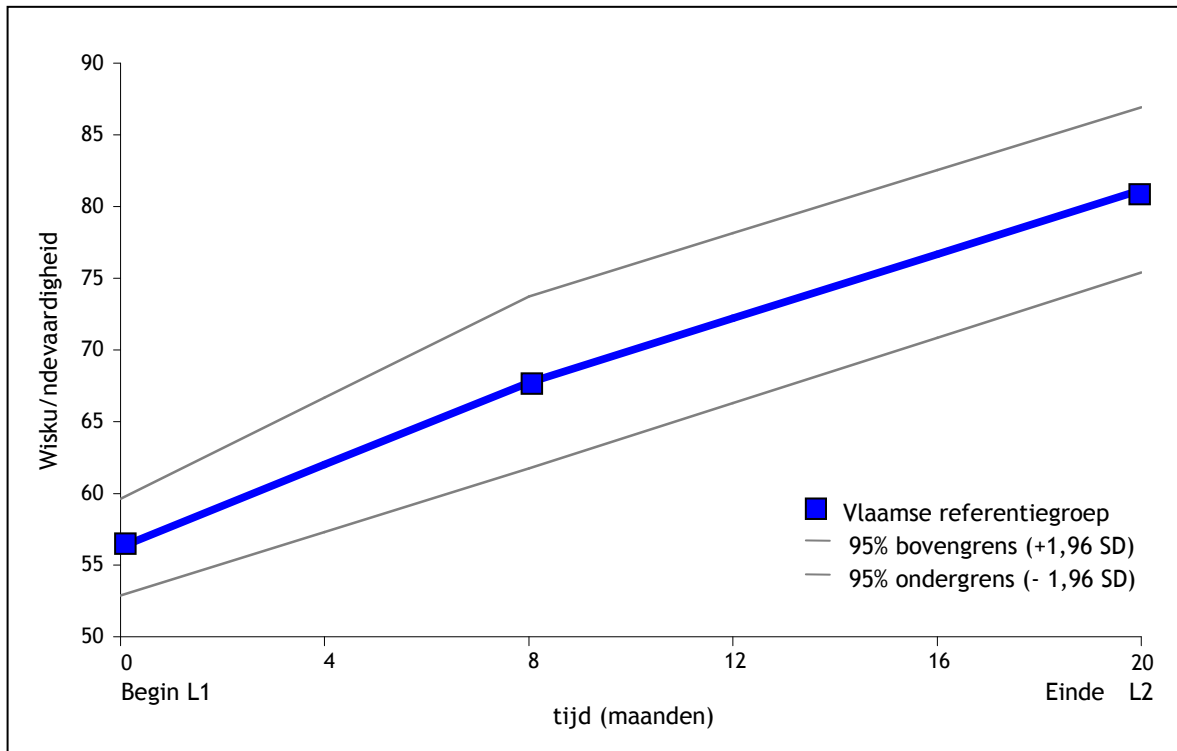


Fig. 3.1 Spreiding van de calibratiescores wiskunde op de meetschaal voor drie meetmomenten (referentiestkpf.)

3.2 Algemene evolutie van begin L1 naar einde L2

Met *algemene evolutie* bedoelen we: de gemiddelde evolutie van de *hele groep* leerlingen die aan het SiBO-onderzoek deelneemt, zonder onderscheid te maken naar subcategorieën van leerlingen. Figuur 3.2 toont de *gemiddelde groeicurve van de Vlaamse referentiegroep* (zie p. 13). Dat is de groeicurve waarmee de gecorrigeerde gemiddelde groeicurves van afzonderlijke scholen vergeleken kunnen worden.

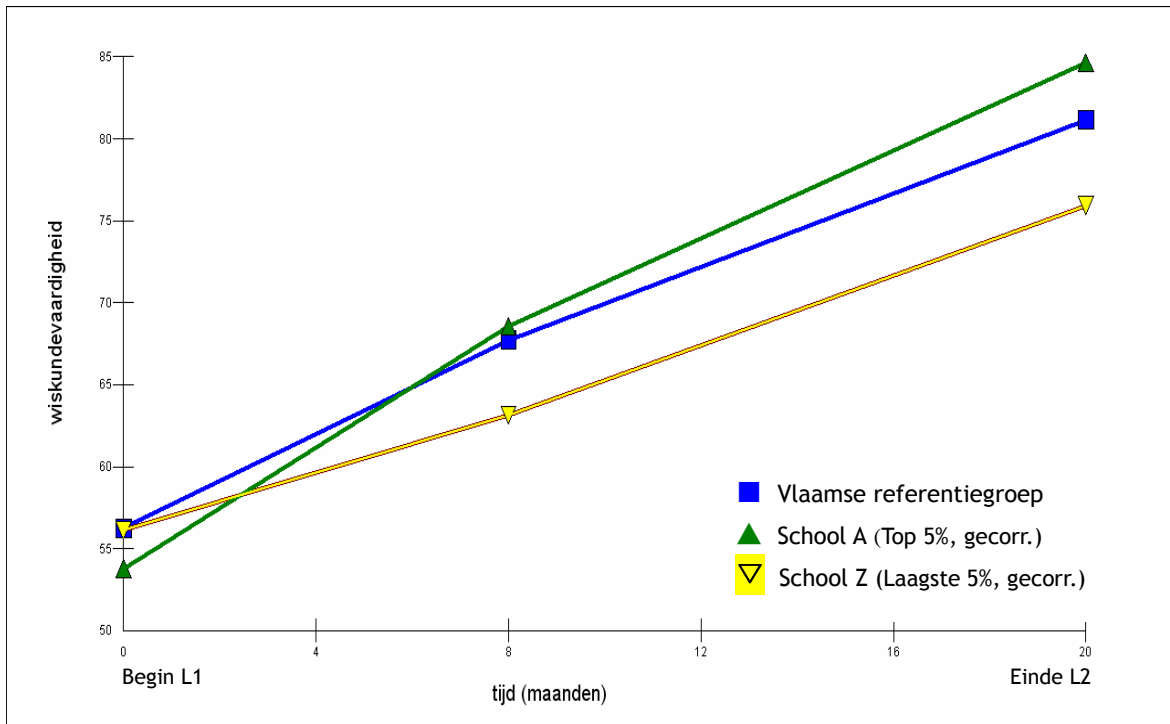


Figuur 3.2 Gemiddelde evolutie voor wiskunde van begin eerste leerjaar naar einde tweede leerjaar in de Vlaamse referentiegroep (+/- 1,96 x standaarddeviatie op schoolniveau)

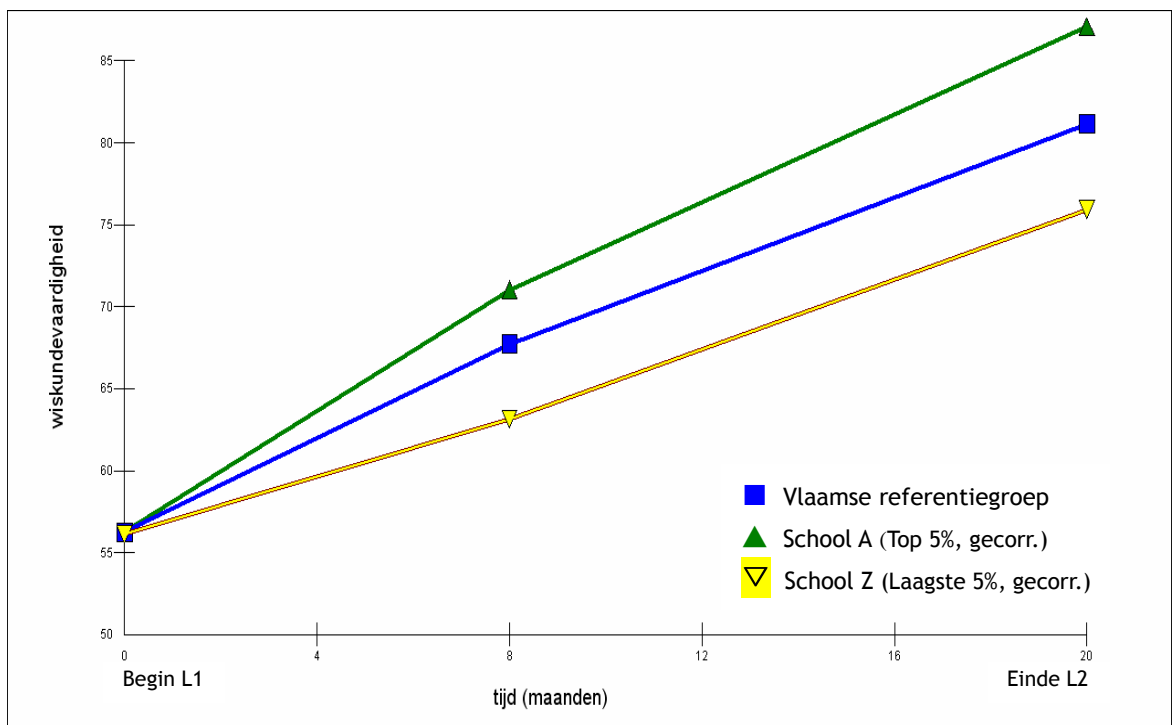
In figuur 3.2 geeft de horizontale as het verloop van de tijd weer, uitgedrukt in maanden. Tellend vanaf het begin van het eerste leerjaar (september = maand 0), geeft maand 8 het toetsmoment op het einde van het eerste leerjaar (eind mei) aan en maand 20 het toetsmoment op het einde van het tweede leerjaar. Tussen het eerste en het tweede toetsmoment (einde L1) lagen gemiddeld acht en een halve maand, tussen het tweede en het derde een jaar. Op de verticale as wordt het vaardigheidsniveau voor wiskunde afgebeeld.

Figuur 3.2 laat een gestage gemiddelde groei zien van begin L1 tot einde L2, met een lichte vertraging in het tweede leerjaar. De gemiddelde leerwinst voor wiskunde in de Vlaamse referentiegroep bedraagt 11,5 punten in het eerste leerjaar (1,35 per maand) en 13,4 in het tweede (1,12 per maand). De twee dunnere lijnen boven de gemiddelde groeicurve geven de grenzen aan waarbinnen zich (uitgaande van een normale verdeling) 90% van de schoolgemiddelden bevinden. Die boven- en ondergrens zijn afgeleid van de variantie op schoolniveau op elk van de toetsmomenten en de onderlinge covarianties. In de loop van het eerste leerjaar is er sprake van een significante toename van de variantie tussen scholen. De lichte afname van de variantie tussen scholen in de loop van het tweede leerjaar is niet significant.

In figuur 3.3 geven de gecorrigeerde gemiddelde groeicurves van een school die qua toegevoegde waarde (TW) tot de top 5 behoort (school A) resp. een school die qua TW tot de laagste vijf behoort (school Z) op meer concrete wijze een idee van de variatie binnen het Vlaamse basisonderwijs.



Figuur 3.3 Gemiddelde evolutie voor wiskunde van begin L1 naar einde L2 in de Vlaamse referentiegroep, in een school met een hoge en in een school met een lage toegevoegde waarde



Figuur 3.4 Gemiddelde evolutie voor wiskunde van begin eerste leerjaar naar einde tweede leerjaar in de Vlaamse referentiegroep, in een school met een hoge en in een school met een lage toegevoegde waarde, na correctie voor verschillen in beginniveau

Waar zien we de leerwinst?

Hoe groot de gemiddelde leerwinst voor het eerste of voor het tweede leerjaar is, kan men in de grafiek aflezen door op de verticale as het verschil tussen eindpunt en het beginpunt van dat leerjaar te nemen.

Waar zien we de toegevoegde waarde (TW)?

Uit het verschil tussen de gecorrigeerde gemiddelde groeicurve van een school en de groeicurve van de Vlaamse referentiegroep kan men de TW van die school afleiden. Is de hellingsgraad van de groeicurve van de school steiler, dan is de TW van die school groter dan gemiddeld in Vlaanderen. Is de hellingsgraad minder steil, dan is de TW van die school lager dan gemiddeld. Lopen beide curves parallel, dan is de TW van die school even groot als de gemiddelde Vlaamse toegevoegde waarde.

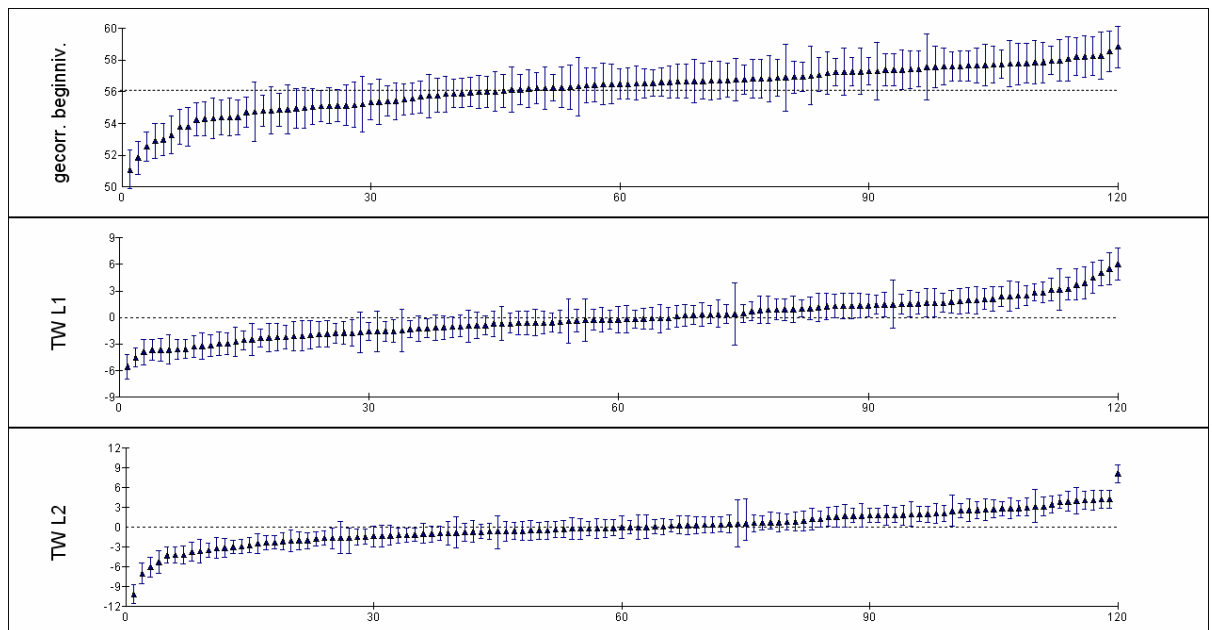
Om in de grafieken het verschil in gemiddelde gecorrigeerde leerwinst en toegevoegde waarde te zien, kijken we naar de hellingsgraad van de curves. Dat gaat makkelijker als we de beginpunten van de curves laten samenvallen (zie fig. 3.4). Die beginpunten laten samenvallen komt eigenlijk neer op het corrigeren voor verschillen in beginniveau.

Tabel 3.2 *Beginniveau, gemiddelde leerwinst en toegevoegde waarde voor eerste en tweede leerjaar*

	Beginniveau (begin L1) na correctie voor instroomkenm.	Eerste leerjaar		Tweede leerjaar	
		Gecorrig. gemidd. leerwinst	Toegevoegde waarde	Gecorrig. gemidd. leerwinst	Toegevoegde waarde
Vlaamse ref.gr.	56,35	11,53	0,00	13,42	0,00
School A	53,79	14,78	3,25	16,06	2,64
School Z	56,15	7,00*	-4,53*	12,78	-0,64

Hoe groot de verschillen qua beginniveau en leerwinst in L1 en L2 precies zijn, ziet men in tabel 3.2. De getallen die met * aangeduid zijn, zijn significant verschillend van het gemiddelde van de Vlaamse referentiegroep.

Opvallend in tabel 3.2 is dat enkel de gecorrigeerde gemiddelde leerwinst en de toegevoegde waarde voor L1 bij school Z significant afwijken van het Vlaamse gemiddelde. Zelfs de top 5-school verschilt - na correctie voor instroomkenmerken - niet significant van het Vlaamse gemiddelde. Toch is het verschil tussen Vlaamse scholen niet echt klein te noemen. Corrigerend voor de effecten van instroomkenmerken én voor het verschil in beginniveau, bedraagt het verschil tussen school A en Z in toegevoegde waarde op het einde van het tweede leerjaar 11,06 punten. Dat komt overeen met bijna 90% van de leerwinst die de Vlaamse referentiegroep jaarlijks gemiddeld voor wiskunde boekt. Dat betekent dat qua vooruitgang in wiskundevaardigheid school A in de loop van amper twee schooljaren bijna een vol schooljaar uitgelopen is op school Z. In figuur 3.3 kan men zien dat in het geval van school A dat maar in zeer beperkte mate door een lager beginniveau verklaard kan worden.



Figuur 3.5 Verschillen tussen scholen (referentiesteekproef) voor gecorrigeerd beginniveau wiskunde en toegevoegde waarde i.v.m. de leerwinst in L1 resp. leerwinst in L2

Figuur 3.5 toont de scholen uit de referentiesteekproef geordend naar hetzij het gecorrigeerde beginniveau voor wiskunde (boven), de toegevoegde waarde met betrekking tot de leerwinst in het eerste leerjaar (midden) of de toegevoegde waarde met betrekking tot de leerwinst in het tweede leerjaar (onderaan). De verticale staafjes geven het betrouwbaarheidsinterval weer. Dat is zo gekozen ($CI = 1,39 \times SE$) dat voor elke twee scholen waarvan de staafjes elkaar niet overlappen, geldt dat ze significant van elkaar verschillen. Voor elk van de drie elementen (gecorrigeerd beginniveau, TW L1 en TW L2) zijn er behoorlijk veel scholen die significant verschillen van andere scholen.

De samenhang tussen beginniveau en de toegevoegde waarde voor L1 en L2

Er is over het algemeen een significant, negatief verband tussen de toegevoegde waarde met betrekking tot de leerwinst in L1 en de toegevoegde waarde met betrekking tot de leerwinst in L2 ($r_{(120)} = -0,54$; $p < 0,001$). Scholen die een hogere toegevoegde waarde realiseren in het eerste leerjaar, neigen er eerder toe een lagere toegevoegde waarde te realiseren in het tweede leerjaar en vice versa. Maar die samenhang is zeker niet bij alle scholen even uitgesproken.

Duidelijk geen systematisch verband is er tussen de TW met betrekking tot de leerwinst in L1 en het beginniveau voor wiskunde, gecorrigeerd voor instroomkenmerken ($r_{(120)} = 0,06$; ns). Een zwak en maar nipt significant, negatief verband is er tussen de TW met betrekking tot de leerwinst in L2 en het gecorrigeerde beginniveau ($r_{(120)} = -0,18$; $p < 0,05$).

3.3 Evolutie voor de verschillende GOK-categorieën

Kansarme en/of anderstalige leerlingen vertonen vanaf het begin van het eerste leerjaar een zekere achterstand voor wiskunde en evolueren soms ook anders in de loop van L1 en L2 dan modale (niet-kansarme) Vlaamse leerlingen. Voor de analyse van de verschillen tussen de *GOK-categorieën* (zie hoofdstuk 2), werden de toetscores gecorrigeerd voor de effecten van de sekse van de leerlingen en van de leeftijdsgebonden voorspellende variabelen, maar niet voor de effecten van aanvankelijke Nederlandse taalvaardigheid. De reden daarvoor is dat een lagere aanvankelijke Nederlandse taalvaardigheid één van de belangrijkste factoren in de achterstand van kansarme en/of anderstalige leerlingen is. Afhankelijk van de GOK-categorie komt 30% tot ruim 40% van de initiële achterstand in wiskunde voor rekening van een minder goede aanvankelijke Nederlandse taalvaardigheid (zie tabel 3.3).

Wegfilteren van dat deel van de achterstand dat samenhangt met (en wellicht het gevolg is van) een minder goede aanvankelijke Nederlandse taalvaardigheid zou een vertekend beeld geven van de werkelijke achterstand van die GOK-leerlingen. Dat ligt anders voor sekse en leeftijdsgebonden voorspellende variabelen. Die vertonen ten eerste maar een heel beperkte samenhang met de achterstand van kansarme of anderstalige leerlingen. Bovendien komen in elk van de GOK-categorieën in principe evenveel jongens als meisjes voor en mag binnen elke categorie ook een gelijke spreiding naar geboortemaand verwacht worden. Afwijkingen daarvan zijn als toevallig en niet-systematisch te beschouwen.

Tabel 3.3 Deel van de initiële achterstand voor wiskunde (begin L1), dat samenhangt met een minder goede aanvankelijke Nederlandse taalvaardigheid, per GOK-categorie

GOK-categorie	Achterstand verklaard door TvK-score
Nederlandstalig GOK	44,2%
Turks GOK	41,2%
Arabisch/Berber GOK	31,4%
Overige anderstalig GOK	37,8%
Anderstalig niet-GOK	28,6%
GOK-status onbekend	38,5%

Hoe komen we aan de percentages in tabel 3.3?

Nemen we het voorbeeld van de Nederlandstalige GOK-leerlingen. Als we *niet* corrigeren voor de effecten van aanvankelijke Nederlandse taalvaardigheid (en verschillen in taalvaardigheid dus laten doorspelen) bedraagt hun achterstand op modale Vlaamse leerlingen bij het begin van L1 gemiddeld 3,6 punten. Na correctie voor de effecten van aanvankelijke Nederlandse taalvaardigheid, bedraagt die achterstand nog slechts 2 punten. Het verschil tussen beide ($3,6 - 2 = 1,6$) is het aandeel in achterstand voor wiskunde waarvan aangenomen mag worden dat het voor rekening van een lagere aanvankelijke Nederlandse taalvaardigheid komt. En dat verschil (1,6) komt overeen met 44% van 3,6.

Zoals besproken in hoofdstuk 2 blijkt een minder goede aanvankelijke Nederlandse taalvaardigheid een vrij hardnekkige factor in de achterstand voor wiskunde, die - hetzij rechtstreeks hetzij via een gemeenschappelijke achterliggende factor - tot minstens het einde van het tweede leerjaar geassocieerd is met de wiskundeprestaties. Afhankelijk van de GOK-categorie wordt slechts 5% (anderstalige niet-GOK-leerlingen) tot 15% (Turkse GOK-leerlingen) van dat deel van de initiële achterstand in wiskunde dat voor rekening komt van een minder goede aanvankelijke Nederlandse taalvaardigheid, in de loop van L1 ingelopen. In de loop van L2 wordt nog eens 1% tot 4% ingelopen. Hoe dan ook, het grootste deel (bijna 80% tot bijna 95%) van de initiële achterstand voor wiskunde die samenhangt met een minder goede aanvankelijke Nederlandse taalvaardigheid, blijven de kinderen met zich meedragen tot (minstens) het einde van het tweede leerjaar.

Tabel 3.4 toont de achterstand van de verschillende GOK-categorieën t.o.v. modale Vlaamse leerlingen op de drie toetsmomenten (na correctie voor de effecten van sekse, leeftijd en zitten-blijven). De achterstand wordt uitgedrukt in procent t.o.v. de gemiddelde leerwinst in de de Vlaamse referentiegroep. De achterstand bij het begin van L1 en op het einde van L1 wordt uitgedrukt t.o.v. de leerwinst in L1; de achterstand einde L2 t.o.v. de leerwinst die, L1 en L2 samengenomen, jaarlijks gemiddeld gemaakt wordt.

Tabel 3.4 Achterstand voor wiskunde van de verschillende GOK-categorieën t.o.v. modale Vlaamse leerlingen

GOK-categorie	Toetsmoment		
	Begin L1 ¹	Einde L1 ¹	Einde L2 ²
Nederlandstalig GOK	-30,8%	-32,3%	-33,3%
Turks GOK	-91,5%	-56,8%	-52,2%
Arabisch/Berber GOK	-78,2%	-57,0%	-64,5%
Overige anderstalig GOK	-72,4%	-40,8%	-44,7%
Anderstalig niet-GOK	-38,9%	-17,6%	-19,1%
GOK-status onbekend	-56,7%	-35,9%	-32,6%

¹Achterstand uitgedrukt als percentage van de gemiddelde leerwinst van de Vlaamse referentiegroep voor wiskunde in L1

²Achterstand als percentage van $(LW1 + LW2)/2$ met LW1 en LW2 zijnde de gemiddelde leerwinst van de Vlaamse referentiegroep voor wiskunde in L1 resp. L2

Initiële achterstand voor wiskunde

De achterstand voor wiskunde die Nederlandstalige GOK-leerlingen bij het begin van het eerste leerjaar gemiddeld t.a.v. Vlaamse niet-kansarme leerlingen hebben, komt overeen met bijna één trimester. De initiële achterstand van Turkse GOK-leerlingen in Vlaanderen is gemiddeld drie keer zo groot. De achterstand voor de andere categorieën ligt daar tussenin.

Evolutie tijdens het eerste en het tweede leerjaar

Opvallend is dat de achterstand van Nederlandstalige GOK-leerlingen tijdens L1 en L2 zo goed als gelijk blijft. Ze neemt een klein beetje toe, maar die toename is niet significant. Bij de andere

categorieën zien we tijdens het eerste leerjaar een significante *afname* van de achterstand. Die leerlingen maken met andere woorden een *inhaalbeweging*. Turkse GOK-leerlingen bijvoorbeeld hebben op het einde van L1 nog een achterstand voor wiskunde die overeenkomt met ruim de helft (56,8%) van de gemiddelde leerwinst van modale Vlaamse leerlingen in L1. Ze hebben tegen het einde van het eerste leerjaar dus ruim vier tiende van hun oorspronkelijke achterstand (91,5%) voor wiskunde goedgemaakt.

In het tweede leerjaar verandert er nog maar weinig aan de achterstandspositie die tegen het einde van het eerste leerjaar ingenomen werd. Voor sommige categorieën gaat de eerder ingezette inhaalbeweging nog een beetje door (b.v. de Turkse GOK-leerlingen), voor andere neemt de achterstand weer een klein beetje toe (b.v. de Arabisch of Berber sprekende GOK-leerlingen).

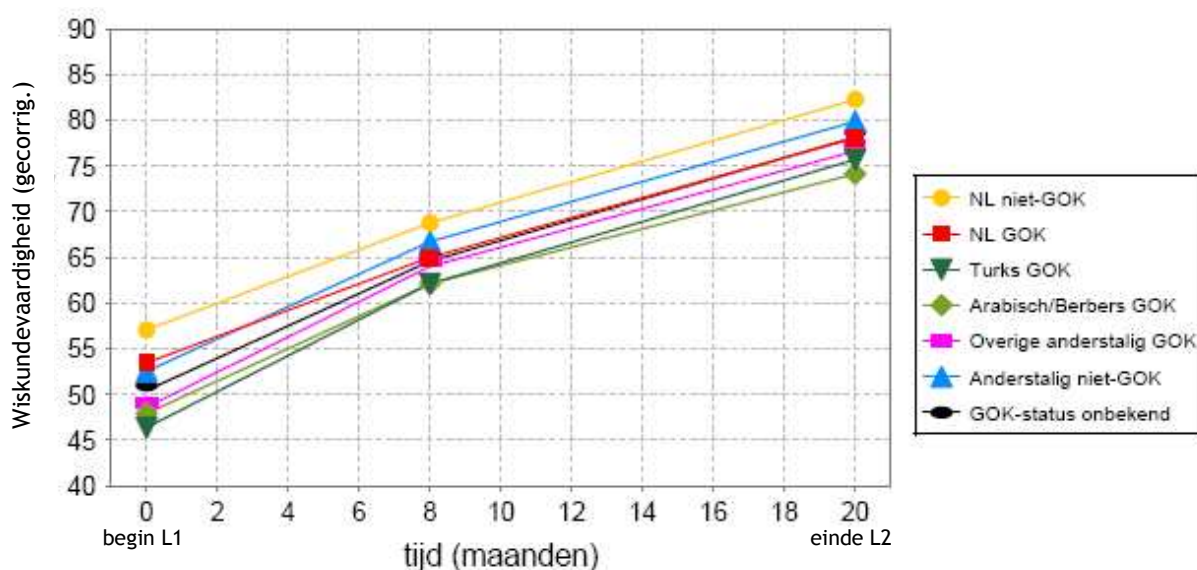
Hoe moeten de percentages in tabel 3.5 geïnterpreteerd worden?

De leerwinst die de modale Vlaamse (niet-kansarme) leerlingen gemiddeld tijdens het eerste leerjaar voor wiskunde maken, bedraagt 11,7 punten. De achterstand van Nederlandstalige GOK-leerlingen bij het begin van L1 belooft gemiddeld 3,6 punten. $3,6 = 30,8\%$ van 11,7.

Doordat we de gemiddelde jaarlijkse leerwinst van modale Vlaamse leerlingen als eenheidsmaat nemen, kunnen we de achterstand ook uitdrukken in de tijd die een modale Vlaamse leerling zou nodig hebben om ze in te halen. Als we aannemen dat leerlingen van begin september tot eind juni gestadig groeien, dan komt een achterstand van 30,8% overeen met ruim 11 weken of twee en een halve maand.

Figuur 3.6 toont de onderscheiden groeicurves voor elk van de GOK-categorieën. Opvallend is dat de curves voor de verschillende categorieën van anderstalige GOK-leerlingen dichter bij de curve voor de Nederlandstalige GOK-leerlingen komen, maar er op geen enkel moment bovenuit stijgen. De enige curves die daar wel bovenuit stijgen, zijn de curves voor de twee categorieën van *niet*-GOK-leerlingen. Dat wijst erop dat een wezenlijk deel van de achterstand van GOK-leerlingen, of het nu om Nederlandstalige dan wel om niet-Nederlandstalige GOK-leerlingen gaat, te maken heeft met de zgn. *sociaal-economische* criteria (waarvan “opleidingsniveau van de moeder” en “gezin leeft van een vervangingsinkomen” de twee belangrijkste indicatoren zijn). Het gegeven dat de achterstand van Vlaamse GOK-leerlingen constant blijft over de twee leerjaren lijkt erop te wijzen dat die sociaal-economische factor een heel hardnekkige factor is in de achterstand van leerlingen (zie ook Verhaeghe e.a., 2006).

Tabel 3.5 geeft voor elk van de GOK-categorieën het gemiddelde beginniveau en de gemiddelde gecorrigeerde leerwinst in L1 en in L2, telkens voor Vlaanderen (V) en voor een exemplarische school (S) met hoge aandelen GOK-leerlingen.



Figuur 3.6 Gemiddelde evolutie voor wiskunde van begin L1 naar einde L2 voor verschillende GOK-categorieën, na correctie voor overige instroomkenmerken maar niet voor aanvankelijke Nederlandse taalvaardigheid

Tabel 3.5 Gemiddeld beginniveau en gemiddelde gecorrigeerde leerwinst voor elk van de GOK-categorieën voor Vlaanderen en voor een school met veel GOK-leerlingen

GOK-categorie	Begin L1		Einde eerste leerjaar			Einde tweede leerjaar		
	Aantal begin	Gemidd. gecorr. niveau	Aantal einde	Gemidd. gecorr. leerwinst	Toegevoegde waarde	Aantal	Gemidd. gecorr. leerwinst	Toegevoegde waarde
NL-niet-GOK	V	2279	57,06	2245	11,70	2084	13,50	1,01
	S	4	54,96	4	12,88			
NL-GOK	V	576	53,46	568	11,53	449	13,06	
	S	3		3				
Turks-GOK	V	79	46,38	77	15,74	60	13,55	0,18
	S	14	43,11	13	21,85 *			
Arab/Ber GOK	V	59	47,93	54	14,17	40	12,02	2,22
	S	6	45,54	6	18,33 *			
Overig AT GOK	V	192	48,60	175	15,39	140	12,63	
	S	9	47,80	9	15,08			
AT niet-GOK	V	352	52,52	348	14,19	316	13,15	
	S	6	50,42	6	15,36			
GOK onbek.	V	411	50,43	409	14,14	679	13,57	
	S	7	48,33	6	15,31			

V = Vlaanderen, S = exemplarische school NL-niet-GOK = modale Vlaamse leerlingen

De leerlingenaantallen die in tabel 3.5 worden vermeld voor Vlaanderen zijn de aantallen voor de referentiesteekproef. Zoals in de inleiding vermeld, gebeurden de analyses op een uitgebreide steekproef, maar door het analyseprocédé dat gehanteerd werd, gelden de schattingen van de

gemiddelden voor de referentiesteekproef. Daarom vermelden we voor de betreffende GOK-categorieën de leerlingenaantallen die voor de referentiesteekproef gelden.

Hoe wordt de toegevoegde waarde bepaald voor subcategorieën van leerlingen?

De algemene principes zijn uiteraard dezelfde als voor de globale toegevoegde waarde. De TW is gelijk aan het verschil tussen de gemiddelde gecorrigeerde leerwinst van de school en het "Vlaamse gemiddelde". Maar in plaats van het gemiddelde van de Vlaamse referentiegroep, nemen we nu het gecorrigeerde Vlaamse gemiddelde van de betreffende subcategorie als vergelijkingsbasis.

Voorbeeld: de TW voor Turkse GOK-leerlingen in het Gentse stedelijke basisonderwijs voor wiskunde in L1 is gelijk aan het verschil tussen:

- de gemiddelde gecorrigeerde leerwinst van Turkse GOK-leerlingen in L1 in het Gentse stedelijke basisonderwijs min
- de gemiddelde gecorrigeerde leerwinst van Turkse GOK-leerlingen in L1 in de Vlaamse referentiesteekproef.

Voor de exemplarisch vermelde school vermeldt tabel 3.5 voor elke categorie van leerlingen afzonderlijk ook de toegevoegde waarde (behalve voor de categorie "NL-GOK" omdat het om zeer kleine aantallen gaat). Men kan zien dat voor de categorieën *AT-niet-GOK* en *GOK-onbekend* de TW gelijk is aan de TW voor modale Vlaamse leerlingen (en dat zou ook voor de categorie *NL-GOK* zo geweest zijn). Dat komt omdat voor die categorieën geen differentiële schooleffecten vastgesteld werden. De achterstand voor wiskunde van die categorieën van leerlingen t.a.v. modale Vlaamse leerlingen is in alle scholen gelijk. Daarmee samenhangend is de TW die scholen voor die categorieën van leerlingen realiseren, niet verschillend van de TW die zij voor de referentiecategorie van modale Vlaamse leerlingen realiseren.

Voor de anderstalige GOK-leerlingen is de situatie anders. Voor die drie categorieën verschilt de achterstand in sommige scholen van het algemene beeld in Vlaanderen (differentiële schooleffecten). Daarmee samenhangend is de TW voor die categorieën van leerlingen in die scholen verschillend van de TW die ze voor modale Vlaamse leerlingen realiseren. We spreken daarom ook van "differentiële TW". De in tabel 3.5 exemplarisch voorgestelde school bijvoorbeeld realiseert in het eerste leerjaar voor haar Turkse GOK-leerlingen een TW (6,11) die ruim vijf keer zo groot is als de TW voor haar modale Vlaamse leerlingen (1,18). De gecorrigeerde leerwinst tijdens L1 van de Turkse GOK-leerlingen in die school (21,85) ligt namelijk een flink stuk hoger dan de gemiddelde gecorrigeerde leerwinst van Turkse GOK-leerlingen in Vlaanderen (15,74). De inhaalbeweging die zij in school S maken is daardoor een flink stuk scherper dan in Vlaanderen gemiddeld het geval is. Een vergelijking van het gemiddelde gecorrigeerde beginniveau van Turkse GOK-leerlingen in school S met het gemiddelde gecorrigeerde beginniveau van Turkse GOK-leerlingen in Vlaanderen, laat zien dat die

extra groei van Turkse GOK-leerlingen in school S niet louter een correctie t.g.v. een grotere aanvankelijke achterstand is. Voor de categorie Arabisch of Berber sprekende GOK-leerlingen in school S gelden gelijkaardige vaststellingen. Ook voor de categorie overige anderstalige GOK-leerlingen is sprake van een differentiële TW, maar het verschil met de categorie modale Vlaamse leerlingen is in school S niet significant.

3.4 Evolutie van de zittenblijvers

In vergelijking met andere leerlingen die in hetzelfde jaar geboren zijn en in jaren uitgedrukt dus evenveel schoolse achterstand hebben, vertonen zittenblijvers L1 een significant afwijkend ontwikkelingspatroon. Bij het overdoen van het eerste leerjaar beginnen ze én eindigen ze gemiddeld met een relatieve voorsprong, maar een jaar later (einde L2) hebben ze gemiddeld alweer een extra achterstand opgelopen (zie pp. 23-25). Uit de analyses blijkt verder dat er wat die achterstand tegen einde L2 betreft, tussen scholen significante verschillen zijn. (We spreken hier van *relatieve* voorsprong en *extra* achterstand omdat we vergelijken met de andere leerlingen met evenveel vertraging en niet met de Vlaamse referentiegroep.) Voeg daarbij de grote verschillen tussen scholen in aantal en aandeel zittenblijvers (zie hoofdstuk 1) en dan wordt het duidelijk dat het van belang is in de analyses en de rapportering zittenblijvers als aparte categorie te beschouwen.

Net als voor de GOK-categorieën het geval is (zie pp. 48-49), zou wegfilteren van het deel van de achterstand dat voor rekening komt van een geringere taalvaardigheid, een vertekend beeld geven van de positie van de zittenblijvers (zie pp. 23-25). Daarom wordt bij de voorstelling van het beginniveau en de vooruitgang van de zittenblijvers niet gecorrigeerd voor de effecten van aanvankelijke Nederlandse taalvaardigheid.

Uit de vergelijking van figuur 2.2 met figuur 2.1 (pp. 24-25) kan men afleiden dat de relatieve voorsprong van zittenblijvers L1 in 2003-2004 op hun leeftijdsgenoten die eerder in hun schoolloopbaan een jaar achterstand opliepen, grotendeels voor rekening komt van hun hogere Nederlandse taalvaardigheid in sept. 2003 (moment waarop ze L1 voor de tweede maal aanvatten). Dat voordeel blijven ze behouden tot het einde van L1 (juni 2004), maar het is tegen het einde van L2 (juni 2005) helemaal weggesmolten. Ze eindigen L2 zelfs met een wiskundescore die gemiddeld iets onder die van hun leeftijdsgenoten ligt.

Tabel 3.6 toont het gemiddelde gecorrigeerde beginniveau en de gemiddelde gecorrigeerde leerwinst in het eerste resp. het tweede leerjaar voor normaalvorderende leerlingen (N) en voor zittenblijvers L1 (Z) in Vlaanderen gemiddeld (V) en in een exemplarische school (S). De gemiddelden voor de zittenblijvers L1 in tabel 3.6 zijn de gemiddelden die geschat werden voor zittenblijvers L1 met niet meer dan één jaar schoolse vertraging op het ogenblik van de toetsafname begin L1. In de leerlingenaantallen die vermeld worden, zijn alle zittenblijvers L1 inbegrepen.

Tabel 3.6 Gemiddeld gecorrigeerd beginniveau, gemiddelde gecorrigeerde leerwinst en toegevoegde waarde voor de normaalvorderende leerlingen en voor de zittenblijvers L1 in Vlaanderen en in een exemplarische school S

Leerlingen-categorie		Begin L1		Einde eerste leerjaar			Tweede leerjaar		
		Aantal begin	Gemidd. gecorrig. niveau	Aantal einde	Gemidd. gecorrig. leerwinst	Toegevoegde waarde	Aantal	Gemidd. gecorrig. leerwinst	Toegevoegde waarde
Normaalvord.	V	3684	57,06	3603	11,70	0,00	3428	13,50	0,00
	S	18	49,13 *	14	13,84	2,14	6	16,75	3,25
ZB L1	V	264	58,57	273	10,73	0,00	244	9,36	0,00
	S	6	50,64 *	5	12,87	2,14	4	15,54 *	6,17 *

ZB L1 = Zittenblijvers L1 V = Vlaanderen (met aantallen voor referentiesteekproef) S = exemplarische school

* sign. verschillend van Vlaams gemiddelde voor die categorie

Een significant differentieel schooleffect werd vastgesteld met betrekking tot de gemiddelde leerwinst in het tweede leerjaar. In het statistisch model werd daarom de mogelijkheid opgenomen dat met betrekking tot de leerwinst in het tweede leerjaar de toegevoegde waarde die een school realiseert, kan verschillen tussen normaalvorderende leerlingen enerzijds en zittenblijvers L1 anderzijds. In enkele scholen (zoals bijvoorbeeld school S) is de TW die in de loop van L2 t.a.v. ex-zittenblijvers L1 gerealiseerd wordt, dubbel zo groot als de TW t.a.v. normaalvorderende leerlingen. In sommige andere scholen ligt het verschil juist in de andere richting. In de meeste scholen zijn de verschillen in TW voor L2 tussen ex-zittenblijvers L1 enerzijds en normaalvorderende leerlingen anderzijds evenwel niet erg groot. Met betrekking tot de leerwinst in L1 werd t.a.v. zittenblijvers L1 geen differentiële schooleffecten vastgesteld.

Waarom zijn de waarden voor beginniveau, gemiddelde gecorrigeerde leerwinst en TW voor normaalvorderende leerlingen (tabel 3.6) dezelfde als voor modale Vlaamse leerlingen (tabel 3.5)?

Dat komt omdat in beide gevallen gecorrigeerd werd voor de effecten van alle overige predictoren (sekse, geboortjaar en geboortemaand, ...). In feite gaat het dus steeds om de waarden voor de referentiecategorie: normaalvorderende, Nederlandstalige niet-kansarme meisjes, geboren in juni 1997. (We corrigeerden niet voor het effect van aanvankelijke Nederlandse taalvaardigheid.) Omdat we in tabel 3.5 focussen op het verschil tussen GOK-leerlingen en modale Vlaamse leerlingen, spreken we daar van "modale Vlaamse leerlingen". In tabel 3.6 spreken we van "normaalvorderende leerlingen" omdat we ons daar richten op het verschillen tussen zittenblijvers L1 en normaalvorderende leerlingen.

Omdat dit voor de lezer meer informatief is, geven we wel steeds de aantallen weer van de categorieën waarop we focussen. In tabel 3.6 wordt dus wel steeds het totale aantal normaalvorderende leerlingen c.q. zittenblijvers L1 aangegeven (inclusief de jongens, niet in juni 1997 geboren, enz.) In tabel 3.5 gaat het om de totale aantallen modale Vlaamse leerlingen c.q. diverse categorieën GOK-leerlingen.

3.5 Maakt de samenstelling van de leerlingengroep een verschil uit?

Er wordt vaak geopperd dat behalve de eigen achtergrond van een leerling ook de samenstelling van de groep waarin hij/zij terechtkomt, een invloed heeft op de leerprestaties / leerwinst. Om na te gaan in hoeverre er inderdaad sprake is van zo'n *groepscompositie-effect* hebben we ook analyses uitgevoerd waarin het percentage Nederlandstalige GOK-leerlingen (in L1 resp. L2), het percentage Turkse GOK-leerlingen, het percentage Arabisch of Berber sprekende GOK-leerlingen, het percentage overige anderstalige GOK-leerlingen, het percentage anderstalige niet-GOK-leerlingen en het percentage leerlingen met een onbekende GOK-status opgenomen werden. Eenmaal die percentages opgenomen, was het niet meer nodig ook nog het percentage Nederlandstalige niet-GOK-leerlingen (modale Vlaamse leerlingen) op te nemen. Dat percentage valt immers perfect af te leiden uit al die andere. Gezien het klasniveau niet in de analyses werd opgenomen, gaat het over de percentages in de gehele leerlingengroep van een bepaald leerjaar (L1 resp. L2) binnen elke school.

De analyse van bovenvermelde groepscompositie-effecten is in het bijzonder van belang in het licht van de inschrijvingsproblematiek (zie b.v. de voorrangspriodes) en de vraag in hoeverre het (ook) vanuit onderwijseffectiviteitspunt van belang is te streven naar een meer gelijkmatige spreiding van kansarme (en in het bijzonder allochtone kansarme) leerlingen. Sommige steden en regio's tellen een belangrijk aantal zgn. *concentratiescholen* en dat wordt door velen als een problematisch gegeven ervaren.

Voor een samenvatting van de bevindingen, zie pp. 77-79.

Vier methodologische opmerkingen zijn hier van belang:

- 1) Om tot een zuivere inschatting van het groepscompositie-effect te komen, is het nodig tegelijk minstens alle overeenkomstige variabelen op leerlingniveau in de analyse op te nemen. Het groepscompositie-effect geeft immers aan in welke mate de groepscompositie een invloed heeft bovenop de invloed van de individuele achtergrondkenmerken. Niet-opname van de overeenkomstige variabelen op leerlingniveau is een belangrijke oorzaak van allerlei statistische artefacten.
- 2) In overeenstemming met de redeneerwijze die we volgden voor de analyses waarin we de verschillen tussen de GOK-categorieën belichten, hebben we niet gecorrigeerd voor het aanvankelijke Nederlandse taalvaardigheidsniveau van de leerlingen. Gezien verschillen in initiële Nederlandse taalvaardigheid een groot deel van de verschillen tussen GOK-categorieën verklaren, vonden we dat correctie voor die factor mogelijk ook tot een onderschatting van de "werkelijke" groepscompositie-effecten zou leiden. In de discussie rond de vraag of concentratiescholen nu al dan niet nadelig zijn voor de ontwikkeling van leerlingen, heeft men immers geen boodschap aan de bevinding dat er terzake geen noemenswaardige groepscompositie-effecten zijn, als die bevinding onmiddellijk daarna genuanceerd moet worden met de bijkomende bevinding dat in concentratiescholen het gemiddeld lagere niveau van aanvankelijke taalvaardigheid vervolgens wél een belangrijke negatieve invloed op de leerwinst van leerlingen heeft.

Er is ook een methodologische reden om niet te corrigeren voor de effecten van aanvankelijke taalvaardigheid. Vanwege de complexe samenhangen met sociale achtergrond (GOK-categorieën) zou zo'n correctie aanleiding kunnen geven tot een omkering van de groepscompositie-effecten (b.v. dat de gemiddelde leerwinst groter is naarmate er meer GOK-leerlingen in de groep zijn). Door die complexe samenhangen zou niet meer duidelijk zijn hoe zo'n omgekeerd groepscompositieeffect geïnterpreteerd moet worden.

Niet-opname van een belangrijke verklarende leerlingfactor als aanvankelijke Nederlandse taalvaardigheid is geen vanzelfsprekende manier van doen. In de onderzoeksliteratuur wordt immers betoogd dat wanneer variabelen op groepsniveau in de analyse opgenomen worden, tegelijk alle relevante verklarende variabelen op leerlingniveau in de analyse meegenomen moet worden, ook de "niet-overeenkomstige" - zoniet dreigen toch nog statistische artefacten op te duiken. Ter controle en ter vergelijking hebben we daarom ook analyses uitgevoerd waarin ook voor de effecten van initiële taalvaardigheid gecontroleerd wordt.

- 3) Het klassieke onderzoek naar groepscompositie-effecten vertrekt vaak van het gemiddelde prestatieniveau van leerlingen op een voorgaand meetmoment. In de analyses die we hier rapporteren, hebben we dat aanvankelijk niet gedaan. Ook hier was de redenering dat GOK-leerlingen zich precies door een lager beginniveau sterk onderscheiden van niet-GOK-leerlingen. Opname van het gemiddelde beginniveau voor wiskunde (of analoog: voor aanvankelijke taalvaardigheid) zou het werkelijke effect van hogere concentraties van GOK-leerlingen kunnen versluieren. Vanwege de tegengestelde samenhang met leerprestaties en leerwinst, zou ook opname van het gemiddelde beginniveau voor wiskunde aanleiding kunnen geven tot niet-interpreteerbare omgekeerde groepscompositie-effecten. Ter exploratie hebben we naderhand toch nog eens nagegaan in hoeverre het een verschil uitmaakt als we toch ook het gemiddelde beginniveau als verklarende factor in de analyses opnemen.
- 4) De analyses weerspiegelen het longitudinale karakter van het onderzoek met - in dit geval - drie meetmomenten. Groepscompositie-effecten kunnen worden gedefinieerd voor elk van de meetmomenten. Voor het tweede en het derde meetmoment stelt dat geen enkel interpretatieprobleem. We spreken dan over de invloed van de groepssamenstelling op de gemiddelde leerwinst in de loop van het eerste resp. tweede leerjaar. Eventuele groepscompositie-effecten voor het eerste meetmoment (begin L1) zijn moeilijker interpreteerbaar. Gaat om effecten uit het kleuteronderwijs? En zo ja, wat schuilt er achter die effecten?

Bij groepscompositie-effecten denkt men spontaan aan de effecten van groepsdynamische processen op individuele leerprocessen of aan de effecten van de groep op de (didactische) aanpak in de klas (b.v. aanpassing aan het gemiddelde beginniveau), die vervolgens weer een effect hebben op de individuele prestaties of vorderingen van de leerlingen. Maar eigenlijk volstaat

het dat leerlingen uit éénzelfde categorie een gemeenschappelijk kenmerk hebben (een kenmerk dat niet als verklarende variabele op leerlingniveau in de analyse opgenomen is, maar wel een samenhang met leerprestaties of vorderingen vertoont) opdat er een statistisch “groepscompositie-effect” gevonden kan worden. Er zou bijvoorbeeld voor het aandeel kansarme kinderen in de leergroep (ook al komen ze uit verschillende kleuterscholen) bij het begin van het eerste leerjaar een groepscompositie-effect naar voor kunnen komen als die categorie van leerlingen zich van andere categorieën onderscheidt door een gemiddeld veel lagere participatie- of aanwezigheidsgraad in het kleuteronderwijs. Ook dat is een “erfenis uit het kleuteronderwijs”, maar dan één die niet (of toch niet rechtstreeks) met processen in de klas te maken heeft. Dat is precies de reden waarom groepscompositie-effecten in de literatuur met argwaan bekeken worden. Ze kunnen het puur statistische effect zijn van niet geëxpliciteerde leerlingkenmerken.

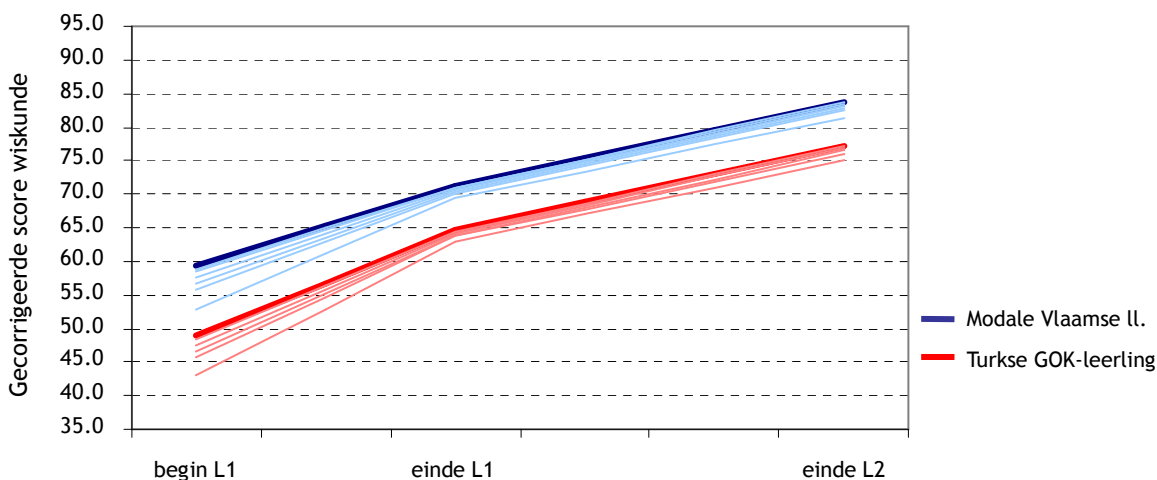
Puur statistisch bekeken ligt de essentie van een groepscompositie-effect namelijk in een gemeenschappelijkheid (binnen een groep van leerlingen) die niet op een lager niveau (leerlingniveau) gedetecteerd werd. Zo’n gemeenschappelijkheid kan het gevolg zijn van reëel voorgekomen groepsprocessen (die de effecten van de individuele kenmerken versterken of milderden) of voortvloeien uit de aggregatie van niet gedetecteerde kenmerken bij categorieën van leerlingen die in wisselende proporties voorkomen. Puur statistisch valt tussen beide geen onderscheid te maken. Is het onderscheid praktisch relevant? Niet als het erop aankomt na te gaan of bij hoge concentraties van bepaalde categorieën van kinderen nog bepaalde, niet nader geïdentificeerde achterstandsfactoren aan het licht gebracht kunnen worden en of er op dat vlak verschillen tussen scholen zijn. Wel als men wil identificeren wat die achterstandsfactoren precies zijn om ze vervolgens met een gepast beleid te kunnen aanpakken.

Met die opmerkingen in gedachte, bekijken we nu de resultaten van de speurtocht naar groepscompositie-effecten.

De afzonderlijke groepscompositie-effecten

Voor het percentage Nederlandstalige GOK-leerlingen, het percentage Turkse GOK-leerlingen, het percentage Arabisch of Berber sprekende GOK-leerlingen, het percentage overige anderstalige GOK-leerlingen en het percentage leerlingen voor wie de GOK-status niet bepaald kon worden, werden significante effecten gevonden op het wiskundeniveau bij het begin van het eerste leerjaar. Al die groepscompositie-effecten liggen in de verwachte richting: hoe hoger het betreffende aandeel GOK-leerlingen hoe lager het beginniveau voor wiskunde. T.a.v. de leerwinst in L1 en L2 samen werd slechts één significant groepscompositie-effect gevonden, nl. voor het percentage Turkse GOK-leerlingen op de gemiddelde leerwinst in het eerste leerjaar.

Figuur 3.7 geeft een idee van de impact en de richting van de groepscompositie-effecten die gevonden werden voor het percentage Turkse GOK-leerlingen in de groep. De dikkere donkerblauwe lijn geeft de gemiddelde groeicurve weer van modale Vlaamse leerlingen in scholen met (voor die cohorte) alleen maar modale Vlaamse leerlingen, de dikkere rode lijn de gemiddelde groeicurve van Turkse GOK-leerlingen die in hun leerjaar op hun school de enige niet-modale Vlaamse leerling zijn. De dunneren en bleke blauwe lijntjes onder de dikkere donkerblauwe lijn geven de gemiddelde groeicurves weer van modale Vlaamse leerlingen in groepen met resp. 5%, 10%, 20%, 30%, 40% en tenslotte 70% Turkse GOK-leerlingen en voor het overige modale Vlaamse leerlingen. Op analoge wijze geven de dunneren, roze lijntjes onder de dikkere rode lijn de gemiddelde groeicurves weer van Turkse GOK-leerlingen in groepen met resp. 10%, 20%, 30%, 40% en 70% Turkse GOK-leerlingen en voor het overige alleen maar modale Vlaamse leerlingen. De percentages zijn gekozen in overeenstemming met reëel voorkomende aandelen Turkse GOK-leerlingen. De lichtblauwe resp. roze band die zo ontstaat onder de blauwe resp. rode lijn geeft een idee van het groepscompositie-effect. Hoe meer Turkse GOK-leerlingen in de groep, hoe lager men in die band gesitueerd is. In figuur 3.7 is de wiskundescore gecorrigeerd voor de effecten van sekse, leeftijd, geboortjaar en zittenblijven, maar niet voor de effecten van aanvankelijke Nederlandse taalvaardigheid en GOK-status.



Figuur 3.7 Groepscompositie-effect m.b.t. aandeel Turkse GOK-leerlingen (wiskunde)

In figuur 3.7 vallen enkele zaken op:

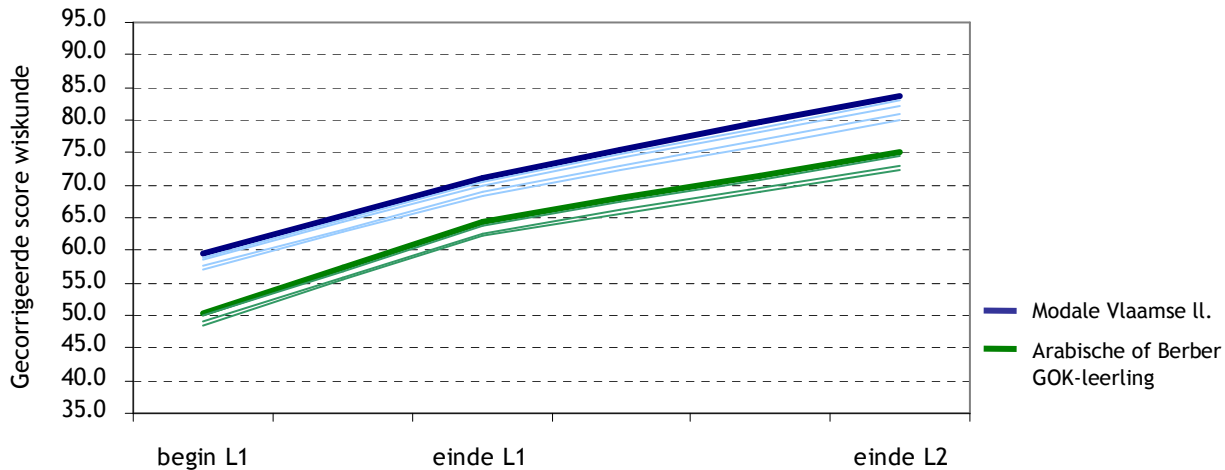
1. De verschillen binnen de lichtblauwe resp. roze band zijn kleiner dan de verschillen tussen de blauwe en de rode lijn. Dat betekent dat het groepscompositie-effect een kleinere impact heeft dan de individuele achtergrond (althans waar het om Turkse GOK-leerlingen versus modale Vlaamse leerlingen gaat).
2. Het groepscompositie-effect is het grootst bij het begin van het eerste leerjaar. Het gaat blijkbaar vooral om een “erfenis” uit het kleuteronderwijs (of om het effect van niet-gedetecteerde

leerlingkenmerken die meer bij Turkse GOK-leerlingen vóórkomen en die van invloed zijn op het beginniveau voor wiskunde, zie hoger).

3. Van begin L1 naar einde L1 neemt de band die het groepscompositie-effect voorstelt *af* in plaats van toe. Dat wijst op een omgekeerd groepscompositie-effect: hoe groter het aandeel Turkse GOK-leerlingen, hoe *hoger* de gemiddelde leerwinst. Dit is een effect dat parallel loopt met de inhaalbeweging die Turkse GOK-leerlingen voor wiskunde maken in het eerste leerjaar. Het zou erop kunnen wijzen dat leerkrachten in het eerste leerjaar met veel Turkse GOK-leerlingen in de klas een aanpak hanteren die de achterstand helpt weg te werken. Maar het is evengoed mogelijk dat het puur om een weerspiegeling gaat van de inhaalbeweging die Turkse GOK-leerlingen maken en daarmee eigenlijk om een soort van statistisch artifact.
4. In het tweede leerjaar ligt het groepscompositie-effect dan weer in de verwachte richting: de band wordt weer iets breder, zij het in heel minieme mate. Dat duidt erop dat er in L2 nauwelijks een groepscompositie-effect met betrekking tot het aandeel Turkse GOK-leerlingen is.
5. De dunne roze lijntjes lopen parallel met de lichtblauwe lijntjes. Dat komt omdat in het model dat hier gebruikt werd geen interactie-effecten tussen groepscompositie en individuele achtergrondkenmerken opgenomen werden. Doen we dat wel dan blijkt dat het effect van het aandeel Turkse GOK-leerlingen op de gemiddelde leerwinst in L1 verschilt naargelang van de individuele achtergrond van de leerlingen. Voor Turkse GOK-leerlingen en Arabisch of Berber sprekende GOK-leerlingen is er inderdaad sprake van een positief groepscompositie-effect. Hoe meer Turkse GOK-leerlingen in de groep, hoe meer leerwinst Turkse GOK-leerlingen en Arabisch of Berber sprekende GOK-leerlingen gemiddeld boeken in het eerste leerjaar. Voor de overige anderstalige leerlingen is het effect juist omgekeerd. Voor anderstalige niet-GOK-leerlingen, Nederlandstalige GOK-leerlingen en modale Vlaamse leerlingen speelt het aandeel Turkse GOK-leerlingen eigenlijk geen rol: het effect is niet significant. Dat betekent dat het negatieve groepscompositie-effect dat bij het begin van L1 speelt, voor o.m. modale Vlaamse leerlingen minder sterk gecompenseerd wordt in de loop van het eerste leerjaar. (Eigenlijk hoort de lichtblauwe band in figuur 3.7 dus niet zo sterk te versmallen tussen begin en einde L1.)

T.a.v. de leerwinst in L2 werden geen interactie-effecten gevonden tussen het aandeel Turkse GOK-leerlingen in de groep en de individuele achtergrond van de leerlingen. Zo'n interactie-effect was er evermin bij het begin van L1.

Figuur 3.8 toont het groepscompositie-effect voor het aandeel Arabisch of Berber sprekende GOK-leerlingen in de groep. Ook figuur 3.8 is gebaseerd op een model zonder interacties tussen het groepscompositie-effect en de individuele achtergrond van leerlingen. De meeste vaststellingen die voor het effect van het aandeel Turkse GOK-leerlingen gedaan werden, gelden ook voor het aandeel Arabisch of Berber sprekende GOK-leerlingen. Verschillen zijn:



Figuur 3.8 Groepscompositie-effect m.b.t. aandeel Arabische of Berber GOK-leerlingen voor het leergebied wiskunde

1. Het effect van het aandeel Arabisch of Berber sprekende GOK-leerlingen is kleiner dan het effect van het aandeel Turkse GOK-leerlingen. Het oogt in figuur 3.8 des te kleiner omdat het maximale aandeel Arabische of Berber sprekende leerlingen ook nog eens kleiner is (max 25%).
2. Het effect op de leerwinst is miniem en voor de twee leerjaren afzonderlijk op zich niet-significant. Het loopt wel twee keer in de verwachte richting (hoe meer Arabische of Berber sprekende GOK-leerlingen hoe minder leerwinst in L1 en in L2). Doordat ze twee keer in dezelfde richting gaan, cumuleren die twee effecten (op de leerwinst in L1 en op de leerwinst in L2).
3. Voor de effecten op de beginsituatie werden significante interacties gevonden met de individuele achtergrond. Hoe meer Arabisch of Berber sprekende GOK-leerlingen in de groep, hoe lager de beginscore voor wiskunde bij Nederlandstalige GOK-leerlingen en anderstalige niet-GOK-leerlingen. Voor modale Vlaamse leerlingen en voor anderstalige GOK-leerlingen (Turkse, Arabische of Berber en overige anderstaligen) blijkt de beginscore voor wiskunde niet significant samen te hangen met het aandeel Arabisch of Berber sprekende GOK-leerlingen in de groep.

Voor de andere groepscompositie-effecten (percentage Nederlandstalige GOK-leerlingen, percentage overige anderstalige GOK-leerlingen, enz.) werden geen interactie-effecten met de individuele achtergrond van leerlingen gevonden.

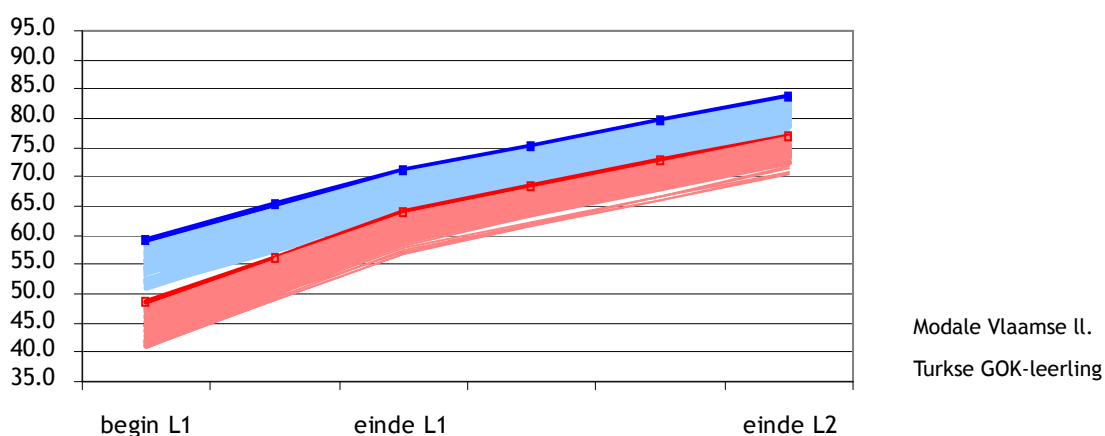
Een opvallende bevinding is dat het opnemen van die groepscompositie-effecten in de analyse nagenoeg niets wijzigt aan de coëfficiënten voor de overeenkomstige effecten op leerlingniveau. De achterstand of inhaalbeweging waartoe elk van de groepscompositie-effecten aanleiding geeft, betreft dus een pure *extra* achterstand respectievelijk *extra* inhaalbeweging bovenop de achterstand c.q. inhaalbeweging die met de individuele achtergrondkenmerken samenhangt. Dat geldt zowel voor de effecten op het beginniveau als voor de effecten op de leerwinst in L1 en de leerwinst in L2.

Bijkomend corrigeren voor de effecten van aanvankelijke Nederlandse taalvaardigheid verandert in essentie niets aan de hoger beschreven patronen van groepscompositie-effecten. Sommige groepscompositie-effecten (b.v. effect van percentage Turkse GOK-leerlingen op beginniveau voor wiskunde) worden iets minder uitgesproken, andere iets méér uitgesproken, hetzij in positieve zin (b.v. effect van pct. Turkse GOK-leerlingen op de leerwinst in L1), hetzij in negatieve zin (b.v. effect van pct. Turkse GOK-leerlingen op de leerwinst in L2). Maar de ene wijziging blijkt de andere te neutraliseren, zodat - wanneer men het geheel bekijkt - er weinig verandert.

Bijkomend controleren voor het gemiddelde beginniveau in wiskunde verandert nog minder aan de hoger beschreven groepscompositie-effecten. Dat gemiddelde beginniveau heeft geen significant effect op de leerwinst in L1, maar wel op de leerwinst in L2. Dat laatste effect is negatief als tegelijk ook gecontroleerd wordt voor de effecten van aanvankelijke Nederlandse taalvaardigheid, maar positief wanneer daar niet voor gecontroleerd wordt. De instabiliteit van dat effect houdt verband met de hoge correlaties tussen beide variabelen en vormt een reden om het niet verder in de analyses op te nemen. We werken daarom alleen verder met de hoger beschreven groepscompositie-effecten op basis van het aandeel leerlingen in elk van de GOK-categorieën.

Gecumuleerde groepscompositie-effecten

De figuren 3.7 en 3.8 presenteren elk een bepaald groepscompositie-effect afzonderlijk, alsof in scholen met b.v. een zeker aandeel Turkse GOK-leerlingen verder alleen maar modale Vlaamse leerlingen zitten. In werkelijkheid tellen sommige scholen behalve Turkse GOK-leerlingen ook nog Arabische of Berber GOK-leerlingen, overige anderstalige GOK-leerlingen, Nederlandstalige GOK-leerlingen, enz. Dat betekent dat in sommige scholen verschillende groepscompositie-effecten samen voorkomen, in wisselende sterkte, afhankelijk van het aandeel van elk van die categorieën leerlingen. Figuur 3.9 toont voor twee categorieën van leerlingen (modale Vlaamse leerlingen en Turkse GOK-leerlingen) in hoeverre de *gecumuleerde groepscompositie-effecten* leiden tot scores



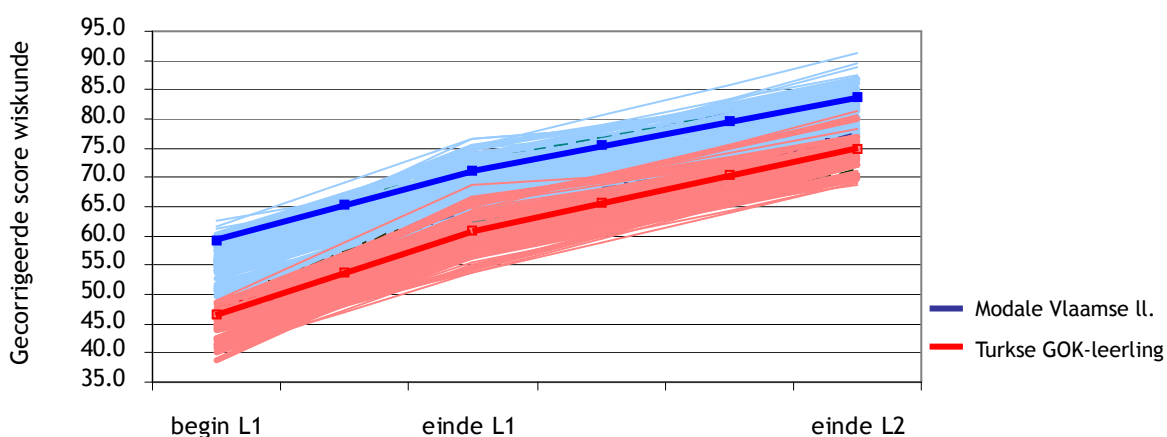
Figuur 3.9 Gecumuleerde groepscompositie-effecten bij twee categorieën van leerlingen voor het leergebied wiskunde

die lager zijn dan wanneer die modale Vlaamse leerling of die Turkse GOK-leerling in een school zou zitten met in hun leerjaar verder alleen maar modale Vlaamse leerlingen. In figuur 3.9 werden alle in de steekproef reëel voorkomende groepscomposities opgenomen. Net als figuren 3.7 en 3.8 gaat figuur 3.9 uit van een model zonder interacties met individuele achtergrondkenmerken.

De lichtblauwe en roze band onder de blauwe en rode groeicurves vallen nu een stuk breder uit. De hele zone tussen de blauwe en de rode groeicurve is gevuld. In de meest ongunstige gevallen (die echter slechts heel beperkt voorkomen) blijkt het gecumuleerde groepscompositie-effect even sterk te zijn als het verschil tussen modale Vlaamse en Turkse GOK-leerlingen.

De compenserende werking in sommige scholen

Behalve het feit dat er geen rekening gehouden wordt met de hoger vermelde interacties tussen groepscompositie en individuele achtergrond, vertoont figuur 3.9 nog een belangrijke afwijking van de realiteit. Er wordt geen rekening gehouden met verschillen tussen scholen in *toegevoegde waarde*. Dat is een belangrijk tekort omdat men bij het beantwoorden van de vraag of concentratiescholen nu al dan niet nadelig zijn voor de ontwikkeling van kinderen, niet enkel naar de algemene effecten van groepscompositie mag kijken. We moeten ook bekijken in hoeverre de geschatte effecten zich effectief ook in elke school voordoen. Dat geldt in het bijzonder voor de problematiek van concentratiescholen omdat groepscompositie-effecten geschat worden op basis van toch wel erg onregelmatige verdelingen van aandelen leerlingen uit verschillende categorieën (voor sommige categorieën komen hoge percentages maar zelden voor). Voor concentratiescholen berust de schatting van de groepscompositie-effecten op vrij zware extrapolaties van de algemeen vastgestelde samenhangen. Het is goed mogelijk dat in scholen met zeer gelijkaardige leerlingpopulaties leerlingen gemiddeld toch een sterk verschillend prestatieniveau of een sterk verschillende leerwinst halen. Figuur 3.10 toont wat er verandert aan figuur 3.9 als bij de gecumuleerde groepscompositie-effecten ook nog eens de toegevoegde waarde van elke school geteld wordt. Het beeld wordt



Figuur 3.10 Gecumuleerd effect van groepscompositie en toegevoegde waarde van de school voor het leergebied wiskunde voor twee categorieën van leerlingen

flink wat onregelmatiger. Men kan zien dat er nu scholen zijn die boven de dikke blauwe en de dikke rode lijn uitstijgen. Dat zijn blijkbaar scholen die erin slagen de effecten van groepscompositie te neutraliseren of te overstijgen.

Groepscompositie-effecten in vier types van scholen

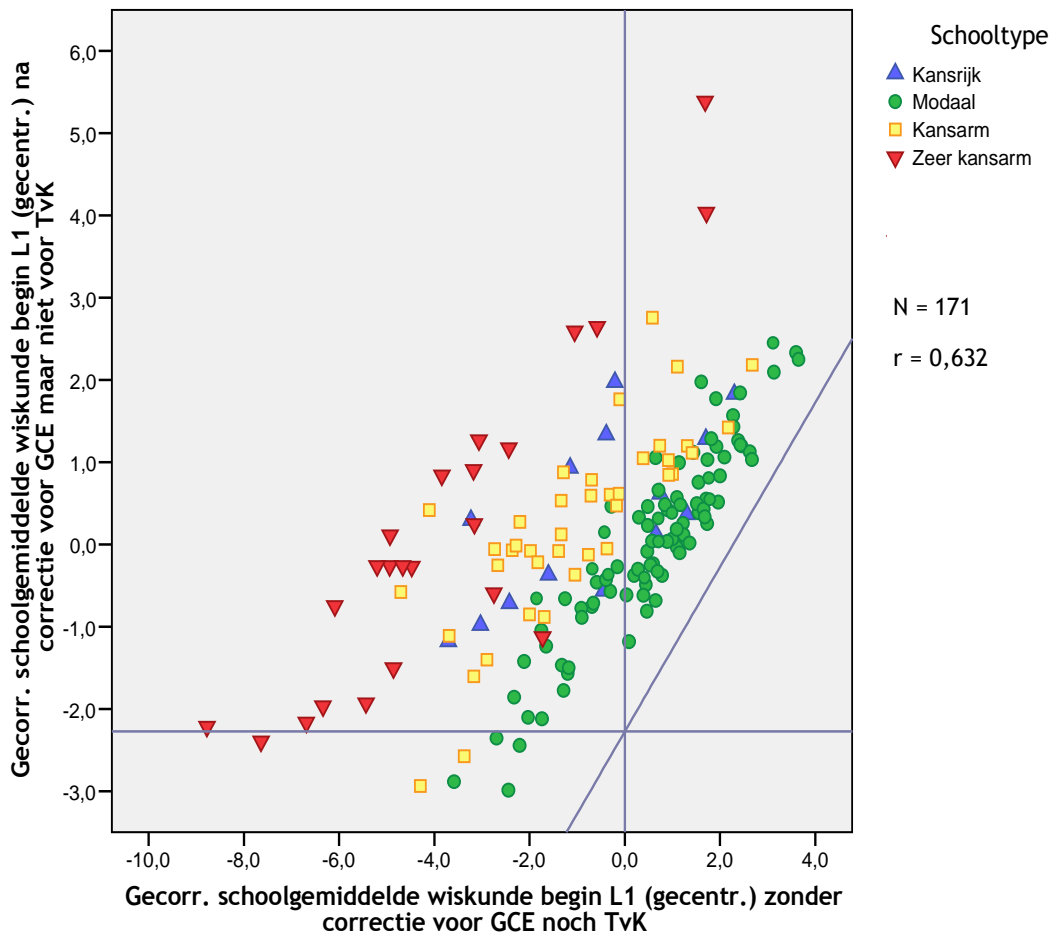
In figuur 3.10 is helemaal niet meer duidelijk welke scholen boven- of onderaan liggen. De toegevoegde waarde verschilt zo van school tot school en ook nog eens van leerjaar tot leerjaar dat de volgorde van de scholen totaal dooreen gehaald kan zijn. Figuren 3.11 en volgende bieden een beter beeld van waar scholen met verschillende groepscomposities zich situeren.

In figuren 3.11 e.v. zijn de scholen ingedeeld in vier categorieën op basis van het profiel waar de samenstelling van hun leerlingenpopulatie het dichtst bij aansluit. De grootste groep van scholen (70% van de referentiesteekproef) behoort tot de categorie van de zgn. *modale* scholen. Dat zijn scholen met 50% tot 100% (gemiddeld 70%) modale Vlaamse leerlingen. De tweede grootste groep scholen (20% van de referentiesteekproef) wordt gevormd door de categorie van de *kansarme* scholen. Bij hen bedraagt het aandeel modale Vlaamse leerlingen gemiddeld slechts 36% (minimum 0% en maximum 52%) en is het aandeel Nederlandstalige GOK-leerlingen verhoudingsgewijs het grootst (gemiddeld 24%). Ze tellen daarnaast gemiddeld ook 16% leerlingen voor wie de GOK-status niet bepaald kon worden. De categorie *zeer kansarme* scholen (4% van de referentiesteekproef, 13% van de uitgebreide steekproef) telt nauwelijks modale Vlaamse leerlingen (gemiddeld 6%) maar wel veel anderstalige GOK-leerlingen (gemiddeld 44%) en ook nog relatief veel anderstalige niet-GOK-leerlingen (gemiddeld 15%) en leerlingen voor wie de GOK-status niet bepaald kon worden (gemiddeld 28%). De categorie van de *kansrijke* scholen tenslotte (6% van de referentiesteekproef) telt vooral veel anderstalige niet-GOK-leerlingen (tussen 25% en 80%, gemiddeld 52%) en daarnaast gemiddeld ook nog 17% anderstalige GOK-leerlingen en 17% “onbekend”. We noemen deze scholen *kansrijk* omdat uit eerdere analyses (Verachtert e.a., 2006) bleek dat in deze scholen het gemiddelde SES beduidend hoger ligt dan in de modale scholen. De anderstalige niet-GOK-leerlingen in die scholen blijken vooral Franstalig te zijn. In figuren 3.11 e.v. wordt elk van de scholen door een stip voorgesteld. Kansrijke scholen zijn blauw gekleurd, modale scholen groen, kansarme scholen oranje-geel en zeer kansarme scholen rood.

Groepscompositie-effect op de wiskundescores begin eerste leerjaar

In figuur 3.11 geeft de horizontale as aan wat het gecorrigeerde gemiddelde van de school voor wiskunde bij het begin van het eerste leerjaar is wanneer wél gecorrigeerd wordt voor individuele achtergrondkenmerken (behalve initiële Nederlandse taalvaardigheid - TvK) maar niét voor groepscompositie-effecten (GCE). De verticale as geeft het gecorrigeerde schoolgemiddelde bij het begin van L1 aan wanneer óók voor het geheel van de groepscompositie-effecten gecorrigeerd wordt. Op

beide assen werd het gecorrigeerde schoolgemiddelde gecentreerd rond het Vlaamse gemiddelde. Daarom lopen die gecorrigeerde schoolgemiddelden van -10 over 0 (het nieuwe Vlaamse gemiddelde) tot 4 voor de horizontale as en van resp. -3 over 0 tot resp. 6 voor de verticale as.



Figuur 3.11 Gecorrigeerde schoolgemiddelden (gecentreerd) voor wiskunde begin L1 mét en zonder correctie voor groepscompositie voor vier types van scholen

De verticale lijn verdeelt de scholen in twee ongeveer gelijke helften: rechts liggen de scholen die zonder correctie voor groepscompositie-effecten maar wel voor individuele achtergrond beter scoren dan gemiddeld, links zij die minder scoren dan gemiddeld. Opvallend in dat opzicht is het verschil in positie (van links naar rechts) tussen rood, geel en groen. Van onder naar boven bekeken is er tussen de categorieën van scholen geen verschil in positie. Eenmaal gecorrigeerd voor groepscompositie-effecten, verschilt het gemiddelde beginniveau voor wiskunde van zeer kansarme scholen, kansarme scholen, modale scholen of kansrijke scholen niet meer van elkaar.

De horizontale lijn loopt niet door het nulpunt van de verticale as maar 2,275 punten lager. Door correctie voor groepscompositie-effecten is het Vlaamse gemiddelde namelijk met 2,275 punten gestegen. Door de horizontale lijn evenveel punten lager te leggen, neutraliseren we die verschuiving en kunnen we beter zien welke veranderingen correctie voor groepscompositie teweegbrengt.

De positie en afstand ten opzichte van de schuine lijn in figuur 3.11 geven aan in hoeverre scholen erbij “winnen” als we ook corrigeren voor groepscompositie-effecten. Alle scholen liggen links boven de schuine lijn. Dat komt omdat alle scholen wel enkele GOK-leerlingen of anderstalige leerlingen tellen. De zeer kansarme scholen liggen duidelijk het verst links boven de schuine lijn, dan komen de kansarme. Dat wijst erop dat de groepscompositie bij die twee categorieën van scholen een grotere rol speelt dan bij de modale scholen. Die modale scholen liggen het dichtst bij de schuine lijn. Voor hen verandert correctie voor de groepscompositie-effecten weinig aan het gecorrigeerde gemiddelde begin L1. Opvallend is dat de zgn. *kansrijke* scholen ook een middenpositie links boven de schuine lijn innemen. Dat komt omdat de anderstaligheid die de meeste leerlingen in die scholen kenmerkt, ervoor zorgt dat die scholen gemiddeld lager scoren dan de modale Vlaamse scholen.

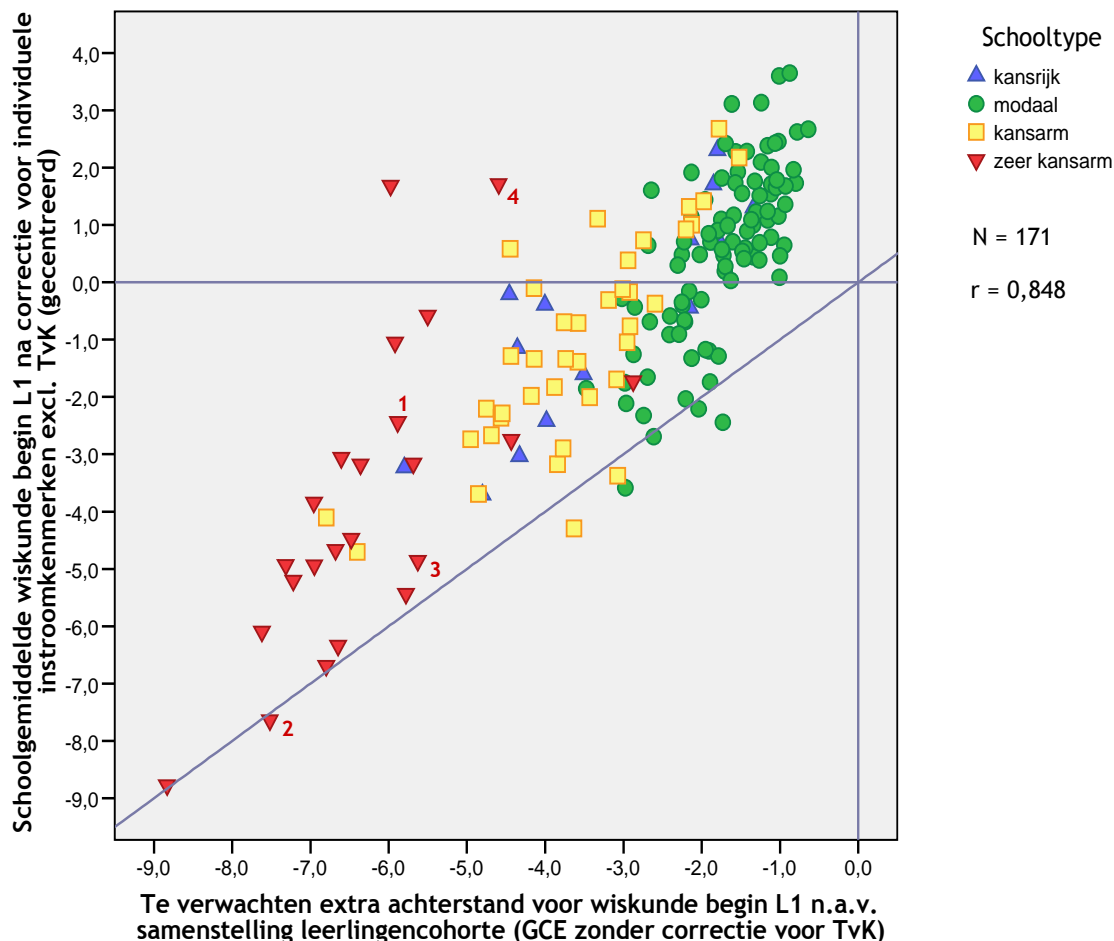
De correlatiecoëfficiënt ($r = 0,632$) wijst op een redelijk sterke samenhang tussen beide gecorrigeerde schoolgemiddelden in fig. 3.11. Het verschil tussen beide scores (-2,275 ter correctie van de verschuiving van het Vlaamse gemiddelde), levert een goede schatting van de grootte van het gecumuleerde groepscompositie-effect op de wiskundescores begin L1.

Op de horizontale as in figuur 3.12 worden de scholen naar dat verschil gerangschikt. Hoe meer naar links een stip ligt, hoe meer bij de kinderen in (de SiBO-cohorte in) die school bij begin L1 voor wiskunde gemiddeld een *extra* achterstand verwacht mag worden die samenhangt met de samenstelling van de SiBO-cohorte en die voor alle leerlingen in die groep in principe tot een lagere aanvangsscore leidt dan op grond van hun individuele achtergrond verwacht kon worden.

Men kan zien dat het gecumuleerde groepscompositie-effect op de beginscore voor wiskunde bij alle scholen negatief is, maar bij zeer kansarme en kansarme scholen duidelijk in veel sterkere mate. Voor de meeste zeer kansarme scholen is het verschil met de meeste modale scholen statistisch significant. Leerlingen in zeer kansarme scholen blijken bij het begin van L1 voor wiskunde gemiddeld 6,3 punten minder te halen dan op grond van hun individuele achtergrond verwacht kon worden. Gezien voor wiskunde de gemiddelde leerwinst van de Vlaamse referentiegroep in L1 11,7 punten bedraagt, komt dat neer op een extra achterstand van ruim een half schooljaar. Voor het gros van de kansarme scholen gaat het om een extra achterstand van twee tot zes maanden (gemiddeld 3 maanden). Die “theoretische” extra achterstand moet wel vergeleken worden met de extra achterstand die ook leerlingen in modale scholen t.g.v. een beperkt groepscompositie-effect ondervinden (gemiddeld 1,5 maand).

Bij de beoordeling van de reële betekenis van dat gecumuleerde groepscompositie-effect dient men te bedenken dat het om schattingen gaat op basis van *algemene* samenhangen tussen de aanvaankelijke wiskundeprestaties van leerlingen en de samenstelling van de cohorte leerlingen binnen elke school die aan het SiBO-onderzoek deelneemt. Met andere woorden: voor elk van de scholen is de “extra achterstand” waarvan sprake, een gemiddelde extra achterstand die men op basis van de

groepssamenstelling bij de aanvang van het eerste leerjaar zou mogen *verwachten*. Of kinderen in een concrete school met een gecumuleerd groepscompositie-effect ten belope van bijvoorbeeld 6 punten ook effectief het eerste leerjaar gemiddeld met een half jaar extra achterstand voor wiskunde aanvangen, valt nog te bezien. Waar we in dat verband naar moeten kijken, is de verticale as in figuur 3.12.



Figuur 3.12 Verschillen tussen scholen en schooltypes in omvang van gecumuleerd groepscompositie-effect op wiskundescores begin eerste leerjaar (zonder correctie voor effecten van aanvankelijke Nederlandse taalvaardigheid)

In figuur 3.12 geeft de positie t.o.v. de verticale as het schoolgemiddelde voor wiskunde begin L1 weer na correctie voor individuele instroomkenmerken (excl. aanvankelijke Nederlandse taalvaardigheid), maar niet voor het GCE. In dat gecorrigeerde gemiddelde speelt het gecumuleerde groepscompositie-effect op de beginscore voor wiskunde dus wel nog mee. De correlatiecoëfficiënt ($r = 0,848$) wijst op een sterk algemeen verband. Toch kan men zien dat scholen met eenzelfde gecumuleerd groepscompositie-effect vaak verschillende gecorrigeerde gemiddelden hebben. Voor school 1 en school 3 bijvoorbeeld kon op grond van de groepssamenstelling een gemiddelde extra achterstand van ongeveer 6 punten verwacht worden. In werkelijkheid bedraagt de gemiddelde (niet

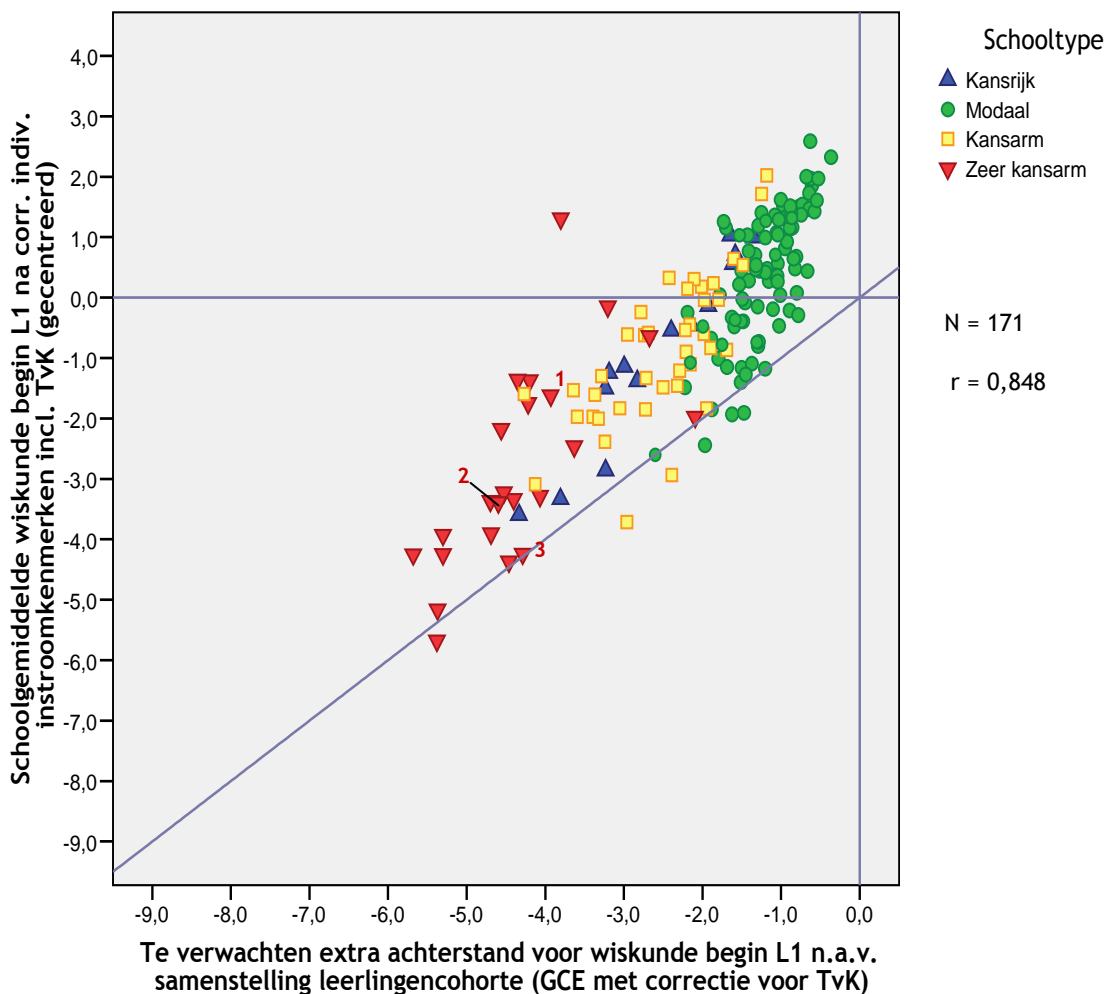
door individuele achtergrondkenmerken verklaarde) *extra achterstand* in school 1 slechts iets minder dan 2,5 punten. In school 3 daarentegen gaat het om 5,5 punten. In school 1 wordt het verwachte groepscompositie-effect dus in belangrijke mate gemilderd; in school 3 nauwelijks. In school 2 is van enige mildering helemaal geen sprake. De verwachte extra achterstand ten gevolge van het GCE bedraagt 7,5 punten en dat is ook de extra achterstand waarmee de leerlingen het eerste leerjaar effectief begonnen.

In figuur 3.12 valt op dat de spreiding ten opzichte van de verticale as veel groter is voor de zeer kansarme scholen dan voor de andere categorieën. Dat houdt verband met een ander fenomeen: er liggen nauwelijks scholen onder de schuine lijn, die in feite de diagonaal is. En voor de enkele scholen die er onder liggen, is de afstand tot de schuine lijn zeer klein. Scholen op de schuine lijn zijn scholen met een effectieve gemiddelde extra achterstand voor wiskunde begin L1 die precies overeenkomt met het gecumuleerde groepscompositie-effect. Scholen die boven de schuine lijn liggen, zijn scholen waarin dat gecumuleerde groepscompositie-effect reeds van bij de aanvang van L1 in zekere mate gemilderd is. Scholen onder de schuine lijn zijn scholen waarin de extra achterstand die samenhangt met het gecumuleerde groepscompositie-effect nog door andere negatieve factoren is aangedikt. Dat er geen scholen zijn die op een betekenisvolle afstand onder de schuine lijn liggen, wijst erop dat het gecumuleerde groepscompositie-effect (zo goed als) de enige factor is die nog voor extra achterstand zorgt, bovenop de individuele instroomkenmerken.

Waarom in de ene school het gecumuleerde groepscompositie-effect op de beginscore voor wiskunde ten volle doorspeelt terwijl het in andere scholen in meer of minder belangrijke mate gemilderd wordt, kan op grond van de voorliggende analyses niet bepaald worden. Net als voor de groepscompositie-effecten zelf kunnen de verklaringen zowel gezocht worden in “erfenissen uit het kleuteronderwijs” als in “niet-gedetecteerde gemeenschappelijke kenmerken” die leerlingen uit de ene school onderscheiden van leerlingen uit de andere school. Gezien de toetsen helemaal bij het begin van het schooljaar afgenomen werden, is het weinig waarschijnlijk dat die verschillen tussen scholen te maken hebben met schoolkenmerken of kenmerken van het didactisch handelen in L1. Zoals hoger al geargumenteed werd, kan een grotere of kleinere participatie aan het kleuteronderwijs ook als een “erfenis uit het kleuteronderwijs” gezien worden, ook al is de school er misschien niet direct verantwoordelijk voor. In de mate dat de leerlingen uit dezelfde wijken afkomstig zijn, kunnen ook wijk- of buurtkenmerken tot die “niet-gedetecteerde gemeenschappelijke kenmerken” behoren. Wat we in elk geval wel al met zekerheid kunnen stellen is dat de samenhang tussen de samenstelling van de leerlingengroep in het eerste leerjaar en de beginscore voor wiskunde niet in alle scholen even sterk is. In sommige kansarme en zeer kansarme scholen is van zo’n samenhang zelfs niets of bijna niets merkbaar. Dat is op zich een belangrijke vaststelling. Het wijst erop dat op basis van de samenstelling van de schoolbevolking geen uitspraken gedaan kunnen worden over de effectiviteit van een school.

De rol van aanvankelijke Nederlandse taalvaardigheid in groepscompositie-effecten begin L1

Zoals we hoger uitlegden, hebben we in de analyses van de groepscompositie-effecten niet gecorrigeerd voor de samenhang tussen aanvankelijke Nederlandse taalvaardigheid en de beginscores voor wiskunde. Hoe zou het plaatje uit figuur 3.12 er uitzien als we dat wel deden? Dat toont figuur 3.12b. Die figuur toont in essentie hetzelfde als figuur 3.12, maar mét correctie voor de “effecten” van het aanvankelijke Nederlandse taalvaardigheidsniveau van de leerlingen (TvK) en niet alleen van andere instroomkenmerken. Om de vergelijking goed te kunnen maken, hebben we voor beide assen in figuur 3.12b de schaalverdeling uit figuur 3.12 behouden. Wat opvalt, is dat (1) het gecumuleerde groepscompositie-effect op de beginscores voor wiskunde nu een stuk kleiner is (maximaal nog -6 i.p.v. -9 punten) en dat (2) ook t.o.v. de verticale as de scholen dichter bij elkaar liggen.



Figuur 3.12b Verschillen tussen scholen en schooltypes in omvang van gecumuleerd groepscompositie-effect op wiskundescores begin eerste leerjaar (na correctie voor effecten van aanvankelijke Nederlandse taalvaardigheid)

Dat betekent dat zowel het gecumuleerde groepscompositie-effect als de gemiddelde beginscore voor wiskunde van een school in belangrijke mate door het aanvankelijke Nederlandse taalvaardigheidsniveau van de leerlingen bepaald wordt. Maar ook na correctie voor verschillen in initiële Nederlandse taalvaardigheid is er nog wel degelijk een groepscompositie-effect. En het is - zeker voor de zeer kansarme scholen - nog altijd aanzienlijk te noemen. De verhoudingen tussen de verschillende categorieën van scholen (zeer kansarm - kansarm - modaal - kansrijk) zijn grosso modo ongewijzigd

Een nadere vergelijking van de figuren 3.12 en 3.12b leert dat de verschuivingen niet voor alle scholen even groot zijn. Voor scholen 1 en 3 bijvoorbeeld mag, na correctie voor TvK, een iets kleiner gecumuleerd groepscompositie-effect verwacht worden. Hun gecorrigeerd aanvangsgemiddelde voor wiskunde verhoogt in iets mindere mate. Zij liggen nu net iets dichterbij de diagonaal. Bij school 2 resulteert correctie voor de samenhang met de initiële Nederlandse taalvaardigheidsscore in een positie die nu niet meer op maar boven de diagonaal uitkomt. Dat wijst erop dat de lagere positie van school 2 in figuur 3.12 in belangrijke mate te wijten was aan de gemiddeld relatief lage initiële Nederlandse taalvaardigheidsscores van haar leerlingen (lager dan op grond van de overige instroomkenmerken verwacht kon worden). Bij school 4 zien we het omgekeerde: in figuur 3.12b ligt die school net onder de horizontale nul-lijn, in figuur 3.12 ongeveer 2 punten hoger. Die hoge positie is te danken aan de relatief hoge taalvaardigheid van de leerlingen in school 4 bij begin L1 (hoger dan op grond van de overige instroomkenmerken verwacht kon worden).

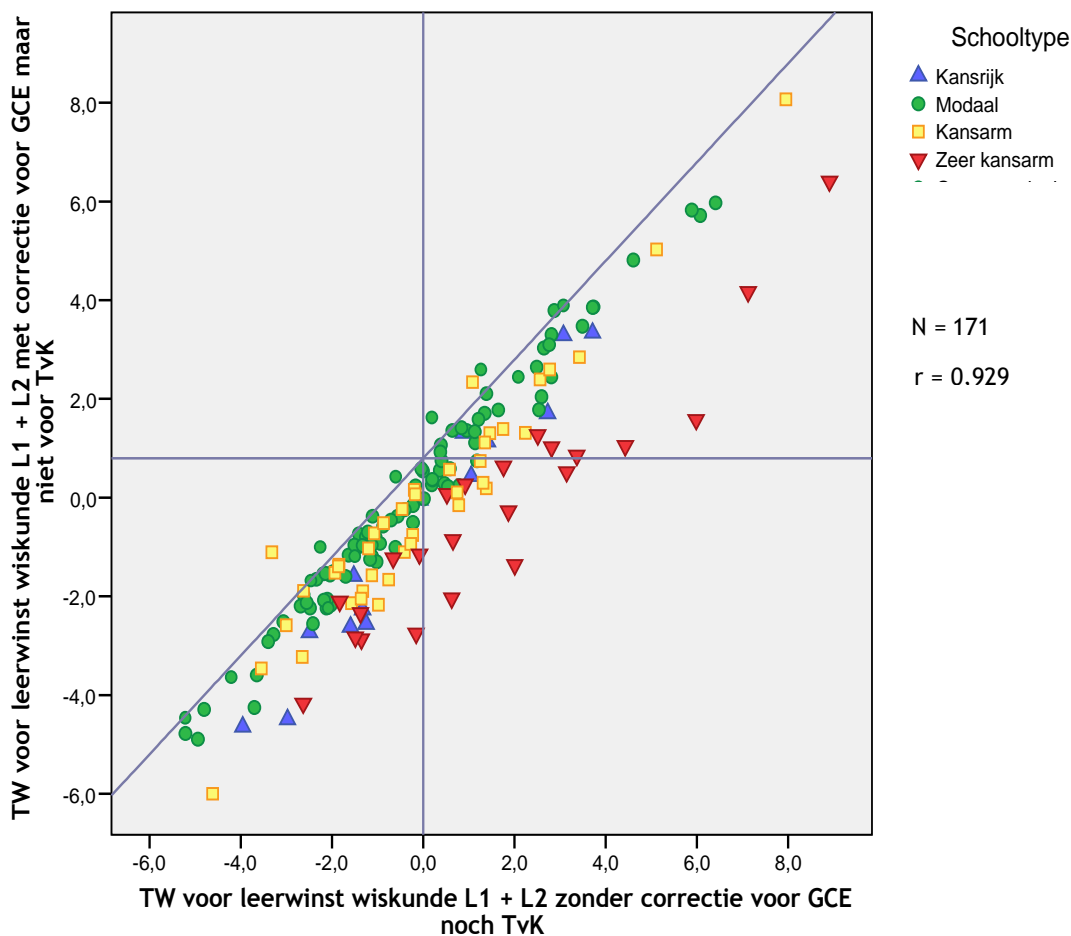
Een verdere vergelijking van het gecumuleerde groepscompositie-effect met en zonder correctie voor de samenhang met het aanvankelijke Nederlandse taalvaardigheidsniveau vindt men in bijlage 1. In wat volgt, keren we terug naar de analyses waarin niet gecorrigeerd werd voor de samenhang met aanvankelijke Nederlandse taalvaardigheid.

Groepscompositie-effecten op de leerwinst in eerste en tweede leerjaar

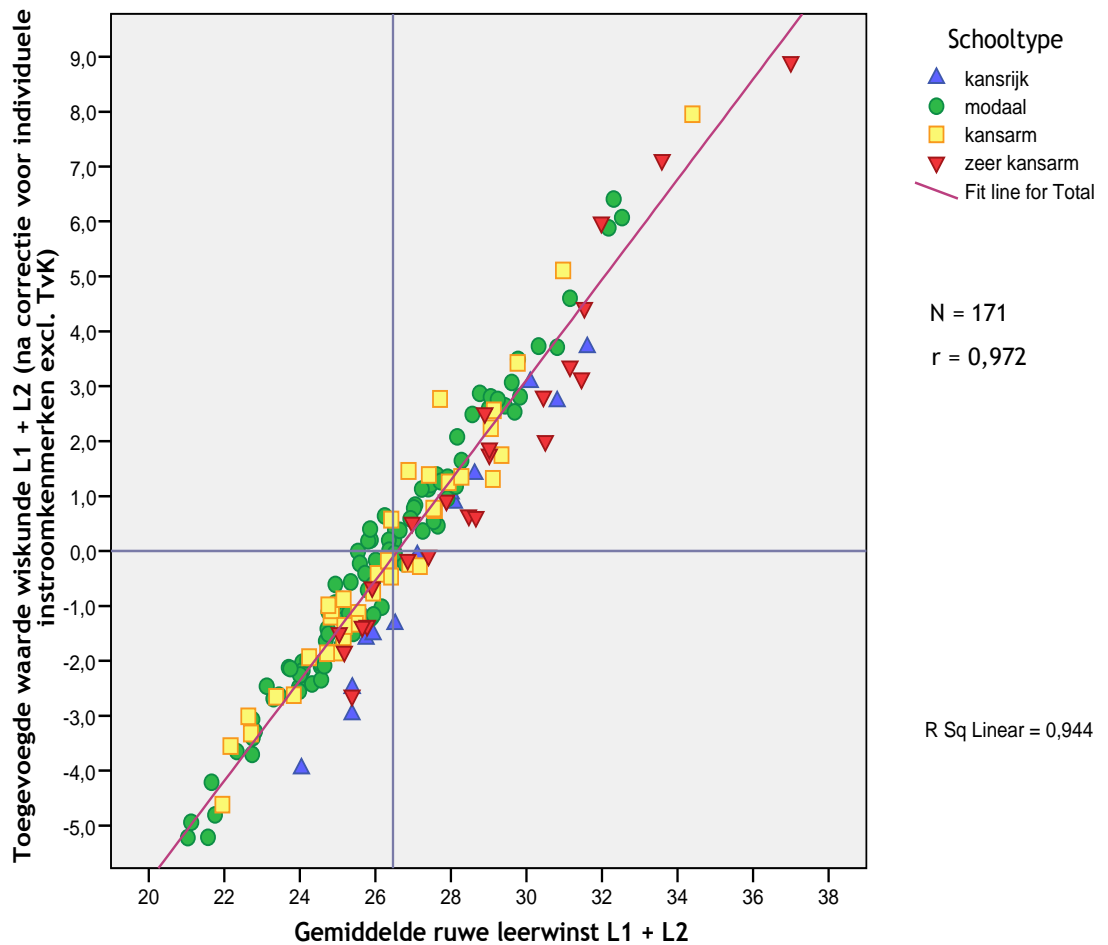
Figuur 3.13 toont per school en per schooltype welke invloed het gecumuleerde groepscompositie-effect heeft op de totale leerwinst in eerste en tweede leerjaar samen. De horizontale as geeft de toegevoegde waarde (TW) inzake de leerwinst in L1 en L2 aan zónder correctie voor de groepscompositie-effecten (GCE); de verticale as mét correctie voor de groepscompositie-effecten. Net als in figuur 3.11 werd wel steeds gecorrigeerd voor individuele achtergrondkenmerken behalve voor aanvankelijke Nederlandse taalvaardigheid (TvK). Door meteen de leerwinst voor L1 en L2 samen te tellen, neutraliseren we schommelingen in gemiddelde leerwinst tussen eerste en tweede leerjaar. De correlatiecoëfficiënt ($r = 0,929$) geeft aan dat de twee TW's heel sterk samenhangen.

Figuur 3.13 laat een heel ander beeld zien dan figuur 3.11. Een eerste punt van verschil is dat alle scholen heel wat dichterbij de schuine lijn liggen. Dat wijst erop dat de groepscompositie-effecten

heel wat kleiner zijn als het om de vooruitgang (de leerwinst) gaat die leerlingen in het eerste en het tweede leerjaar boeken. Een tweede punt van verschil is dat de rode stippen en driehoekjes nu rechts ónder de schuine lijn liggen. Dat wijst op een “omgekeerd” groepscompositie-effect. Zo’n omgekeerd effect zagen we al in figuur 3.11. Hoewel de afzonderlijke groepscompositie-effecten wisselen (nu eens in de verwachte richting liggen en dan weer omgekeerd zijn, zie ook figuur 3.12), blijkt de som van alle groepscompositie-effecten (het gecumuleerde GCE) over L1 en L2 samen omgekeerd te zijn: hoe meer GOK-leerlingen (Nederlandstalige en anderstalige samen) hoe *groter* de leerwinst voor wiskunde tussen begin L1 en einde L2 gemiddeld is. Voor de *zeer kansarme* scholen en de meeste *kansarme* scholen brengt dat met zich dat correctie voor de groepscompositie-effecten leidt tot *lagere* gemiddelde gecorrigeerde leerwinstscores. In overeenstemming met dat omgekeerde GCE is de vaststelling dat na correctie voor groepscompositie-effecten het Vlaamse gemiddelde nu 0,8 punten lager ligt. Om beter te zien tot welke verschuivingen de correctie voor het groepscompositie-effect leidt, werd de middellijn t.a.v. de verticale as in figuur 3.13 niet door het nulpunt maar 0,8 punten hoger getrokken.



Figuur 3.13 Toegevoegde waarde voor de leerwinst wiskunde tussen begin L1 en einde L2 mét en zonder correctie voor groepscompositie voor vier types van scholen



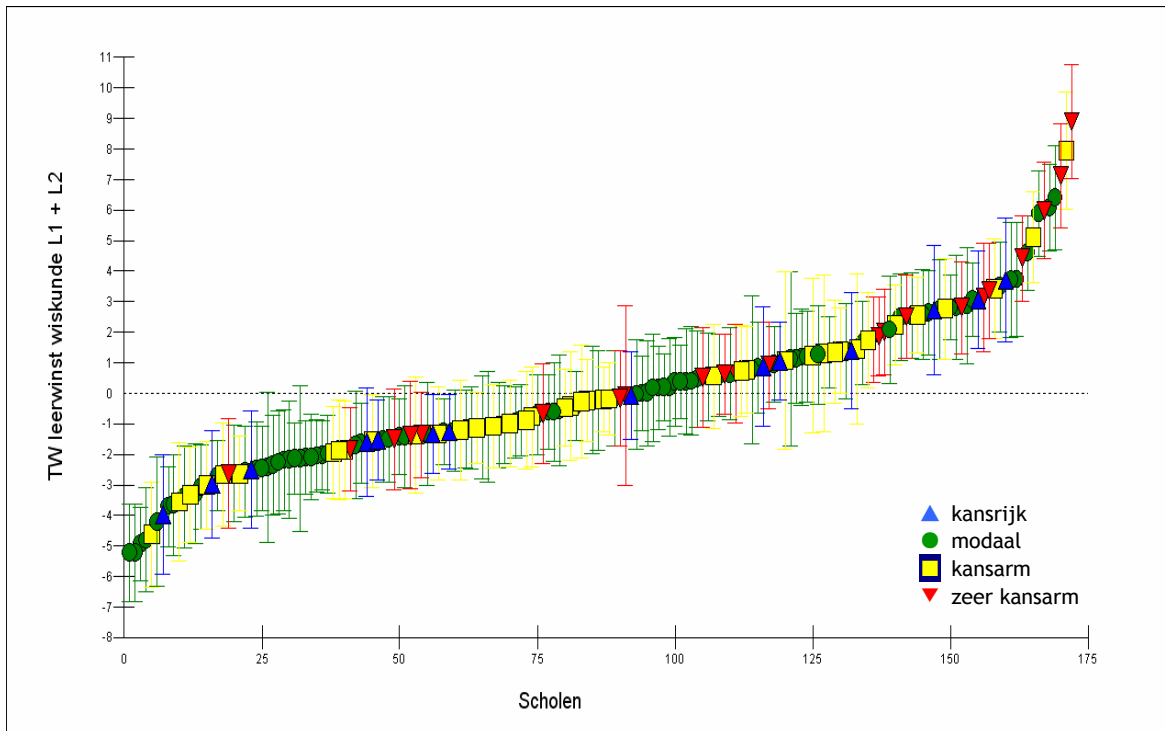
Figuur 3.14 Schoolgemiddelden voor de ruwe leerwinst voor wiskunde tussen begin L1 en einde L2 en toegevoegde waarde zonder correctie voor groepscompositie

Het meest opmerkelijke verschil zien we echter als we kijken naar de positie die scholen t.o.v. de *horizontale* as innemen. Vijftien van de 23 zeer kansarme scholen bevinden zich rechts van de verticale middellijn. Ook een groot deel van de kansarme scholen bevindt zich aan de rechterkant. Ook wanneer *niet* gecorrigeerd wordt voor groepscompositie-effecten, blijkt dus dat zich onder de kansarme en zeer kansarme scholen net zo goed effectieve en heel effectieve scholen bevinden als onder de modale scholen. Die conclusie houdt stand wanneer naar de ruwe leerwinst gekeken wordt. In figuur 3.14 geeft de horizontale as de gemiddelde ruwe leerwinst weer.

Uit figuur 3.14 is duidelijk dat de verschillen tussen scholen in gemiddelde leerwinst en toegevoegde waarde behoorlijk kunnen oplopen. Tussen de school met de laagste en die met de hoogste TW bedraagt het verschil ruim 14 punten (voor de gemiddelde ruwe leerwinst bijna 16 punten). Ter vergelijking: tussen begin L1 en einde L2 gaan modale Vlaamse leerlingen gemiddeld 25 punten vooruit.

In hoeverre de verschillen tussen scholen significant zijn, kan men zien in figuur 3.15. In die figuur staan de scholen gerangschikt van lage naar hoge toegevoegde waarde met betrekking tot de leerwinst in L1 + L2. De verticale lijntjes geven aan hoe betrouwbaar de schatting van de toe-

gevoegde waarde is ($CI = 1,39 \times SE$). Hoe korter het verticale lijntje, hoe betrouwbaarder. Twee scholen waarvan de verticale lijntjes elkaar niet overlappen, verschillen significant van elkaar. Aan de rechterkant van de grafiek vindt men scholen (waaronder ook kansarme en zeer kansarme) die significant hoger scoren dan scholen (waaronder ook modale) aan de linkerkant van de grafiek.

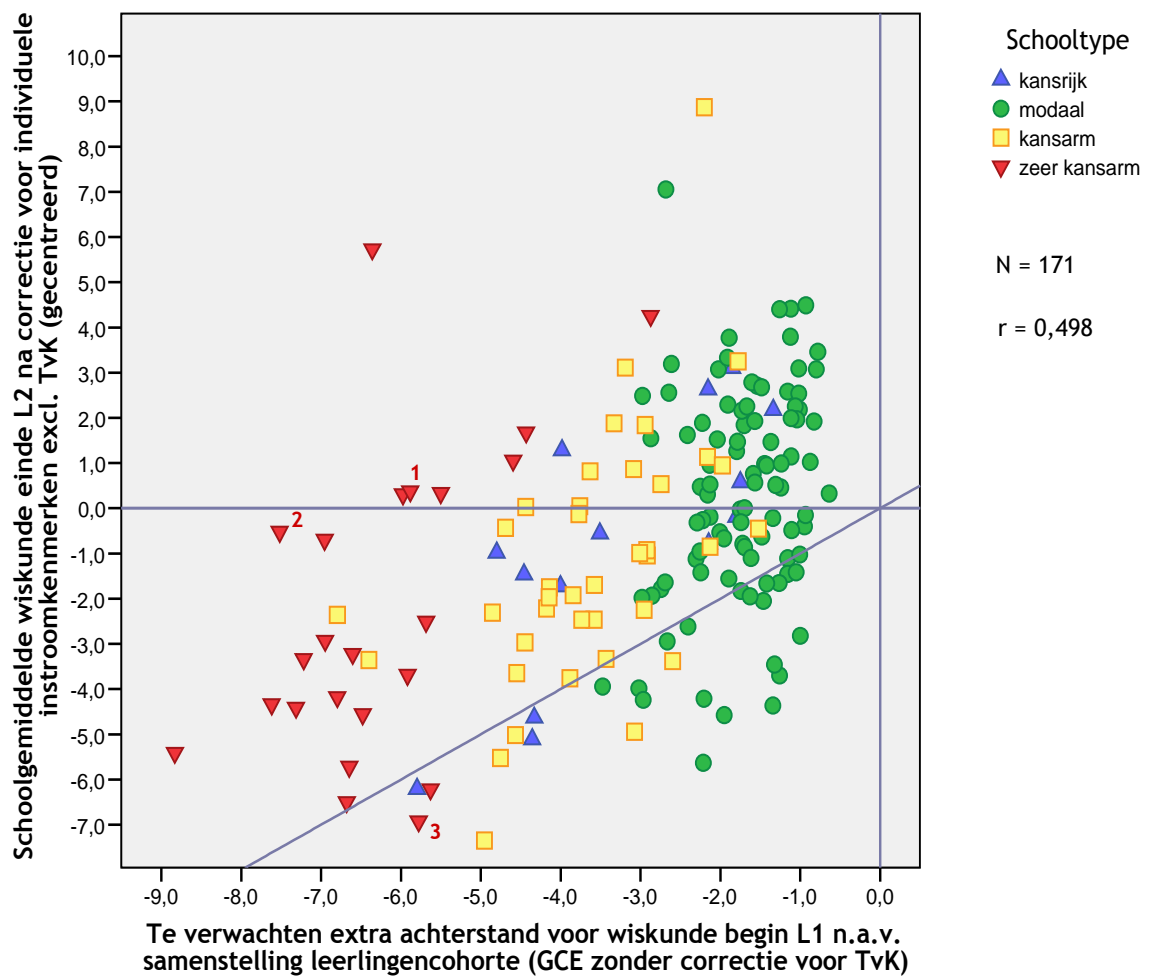


Figuur 3.15 Verschillen tussen scholen ($N=172$) en schooltypes in toevoegde waarde m.b.t. de leerwinst voor wiskunde tussen begin L1 en einde L2 (zonder correctie voor GCE)

Neutraliseert de leerwinst in L1 en L2 de aanvankelijke groepscompositie-effecten?

In hoeverre dragen de gunstige gemiddelde leerwinstcijfers die heel wat kansarme en zeer kansarme scholen laten optekenen, ertoe bij de negatieve effecten van de groepscompositie op de aanvangscore voor wiskunde (voor zover effectief aanwezig) te neutraliseren? Om daar beter zicht op te krijgen, hebben we in figuur 3.16 twee afgeleide scores tegenover elkaar afgezet.

De horizontale as in figuur 3.16 geeft de omvang van het gecumuleerde groepscompositie-effect bij het begin van L1 weer (zoals in figuur 3.12). Hoe meer naar links een stip ligt, hoe groter dat (te verwachten) negatieve effect. De verticale as in figuur 3.16 geeft per school het gemiddelde wiskundeniveau einde L2 weer, na correctie voor individuele instroomkenmerken (exclusief aanvankelijke Nederlandse taalvaardigheid), maar zonder correctie voor groepscompositie-effecten. Voor zover er nog groepscompositie-effecten uit begin L1 zijn blijven doorwerken tot op einde L2, liggen die dus in die gecorrigeerde gemiddelden einde L2 vervat. In figuur 3.16 is de gecorrigeerde wiskundescore einde L2 gecentreerd rond het gemiddelde van de Vlaamse referentiegroep.



Figuur 3.16 Mate waarin scholen er voor het leergebied wiskunde in slagen tegen einde L2 het groepscompositie-effect van bij begin L1 te neutraliseren, opgesplitst naar schooltype

De verdere interpretatie van figuur 3.16 is analoog de interpretatie van figuur 3.12. Scholen die op of boven de horizontale lijn liggen, zijn erin geslaagd het negatieve gecumuleerde groepscompositie-effect van bij het begin van L1 te overwinnen, hetzij doordat het effect van meet af aan al geheel of gedeeltelijk gemilderd was, hetzij door de toegevoegde waarde die de scholen in de loop van het eerste en het tweede leerjaar realiseerden, of door een combinatie van beide. De precieze positie in figuur 3.16 (hoger of lager) geeft aan hoeveel beter (of slechter) kinderen met een gegeven reeks instroomkenmerken op het einde van L2 gemiddeld scoren dan leerlingen met eenzelfde achtergrond en overige instroomkenmerken in andere scholen.

Scholen die onder de horizontale lijn liggen maar nog boven de diagonaal zijn scholen waarin nog maar gedeeltelijk gecompenseerd werd voor het negatieve groepscompositie-effect van bij begin L1. Tegen het einde van het tweede leerjaar vertonen hun leerlingen gemiddeld nog een zekere achter-

stand t.o.v. leerlingen met dezelfde instroomkenmerken in hoger gesitueerde scholen, maar die achterstand is minder groot dan op grond van het GCE begin L1 verwacht had kunnen worden.

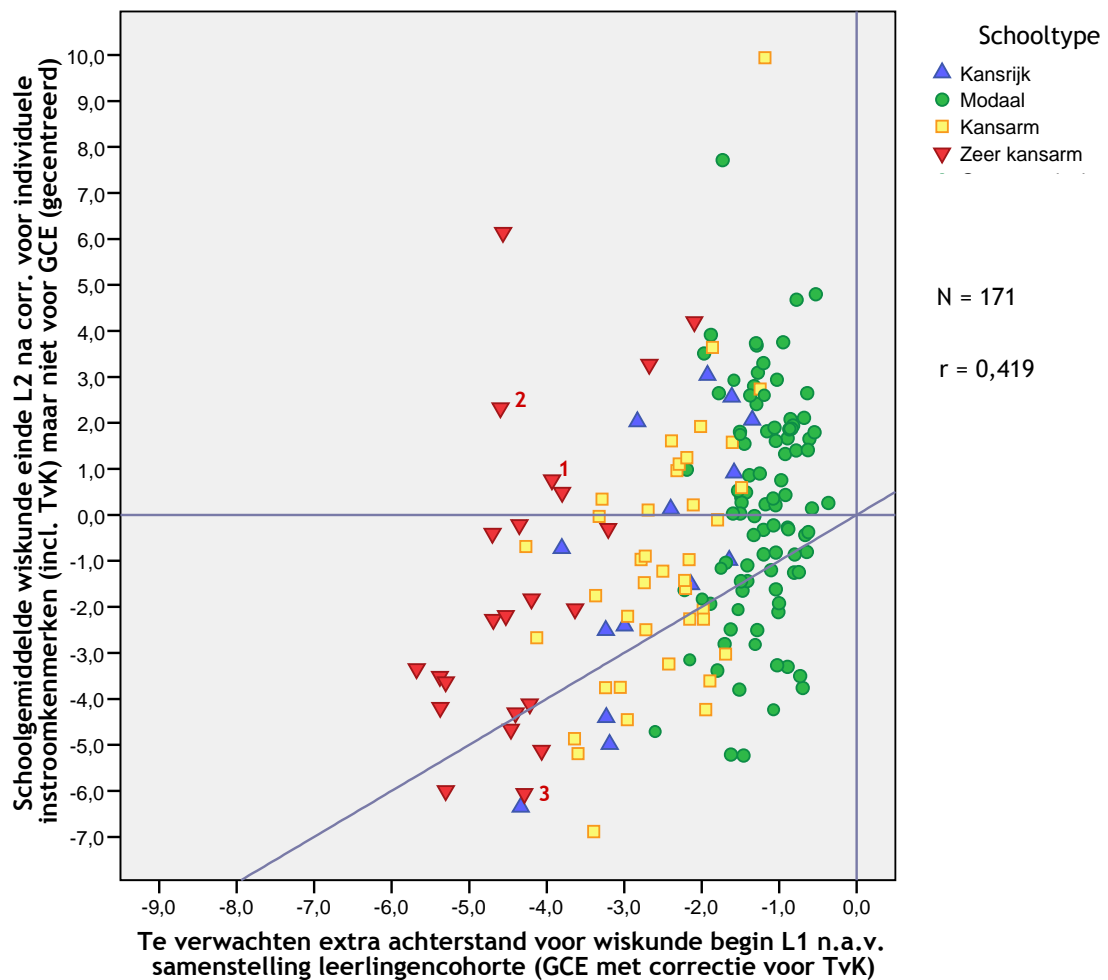
Scholen op de diagonaal zijn er niet in geslaagd het negatieve groepscompositie-effect van bij het begin van het eerste leerjaar te neutraliseren. In vergelijking met leerlingen met dezelfde instroomkenmerken in andere scholen vertonen hun leerlingen op het einde van L2 gemiddeld een achterstand die precies overeenkomt met wat op grond van het groepscompositie-effect bij het begin van L1 verwacht kon worden.

Scholen onder de diagonaal realiseerden tussen begin L1 en einde L2 een toegevoegde waarde die zo beperkt is dat tegen einde L2 hun leerlingen ten opzichte van leerlingen met dezelfde instroomkenmerken in andere scholen gemiddeld een achterstand hebben opgelopen die nog groter is dan op grond van het negatieve groepscompositie-effect van bij begin L1 verwacht kon worden.

Vergelijking met figuur 3.12 laat twee tegengestelde bewegingen zien: aan de ene kant meer zeer kansarme scholen die tegen het eind van L2 erin slagen te compenseren voor het (relatief grote) negatieve gecumuleerde groepscompositie-effect van begin L1 (zie b.v. scholen 1 en 2); aan de andere zijde van het spectrum ook een aantal modale scholen die tot rond of onder de diagonaal zakken en hun leerlingen een extra achterstand zien opbouwen. Het omgekeerde groepscompositie-effect dat met betrekking tot de leerwinst in L1 en L2 vastgesteld werd, speelt hier een rol in. Toch zijn er ook scholen die tegen die algemene trend ingaan (b.v. school 3). Het resultaat van dat alles is dat de samenhang tussen het gecumuleerde groepscompositie-effect begin L1 en het schoolgemiddelde (gecorrigeerd voor individuele instroomkenmerken) zwakker wordt. De correlatie daalt van 0,836 bij het begin van L1 naar 0,498 op het einde van L2. De impact van de samenstelling van de groep wordt m.a.w. geleidelijk aan een heel stuk kleiner.

Er blijven wel nog belangrijke verschillen tussen scholen. Negen van de zeer kansarme scholen slagen er tegen einde L2 (zo goed als) volledig in te compenseren voor het negatieve groepscompositie-effect van bij begin L1; twee van die scholen doen het zelfs een heel stuk beter dan dat. Een tiental zeer kansarme scholen realiseert een gedeeltelijke compensatie. In de overige vier vertonen de leerlingen (tegen de algemene trend in) op het einde van L2 grosso modo nog steeds de extra achterstand die op grond van het groepscompositie-effect bij het begin van L1 te verwachten viel; bij één van die vier (school 3) wordt die extra achterstand zelfs nog iets aangedikt.

Een vrij analoog beeld zien we voor de kansarme scholen. Een deel van de kansarme scholen slaagt er tegen einde L2 in (meer dan) te compenseren voor de extra achterstand ten gevolge van het gecumuleerde groepscompositie-effect bij het begin van L1, een ander deel slaagt daar gedeeltelijk in en nog een ander, kleiner deel helemaal niet. Bij de modale en de kansrijke scholen valt vooral op dat de onderlinge verscheidenheid op het einde van L2 groter is dan bij begin L1.



Figuur 3.16b Mate waarin scholen er voor het leergebied wiskunde in slagen tegen einde L2 het groepscompositie-effect van bij begin L1 te neutraliseren in analyse mét correctie voor aanvankelijke Nederlandse taalvaardigheid

De rol van aanvankelijke Nederlandse taalvaardigheid in de compensatie van groepscompositie-effecten in L1 en L2

Figuur 3.16b toont - analoog aan figuur 3.12b - wat er gebeurt als we in de analyses ook corrigeren voor de effecten van het aanvankelijke Nederlandse taalvaardigheidsniveau (TvK). De horizontale as is gelijk aan de horizontale as in figuur 3.16 en geeft de omvang van het gecumuleerde groepscompositie-effect bij het begin van L1 aan. De positie t.a.v. de verticale as in figuur 3.16b geeft het schoolgemiddelde einde L2 weer na correctie voor de effecten van alle individuele instroomkenmerken, inclusief TvK. Dat schoolgemiddelde is gecentreerd rond het gemiddelde van de Vlaamse referentiegroep (idem als in figuur 3.16).

Vergelijking van figuur 3.16b met figuur 3.16 laat weinig verschillen zien in de positie van scholen t.o.v. de verticale as. Sommige scholen nemen een iets hogere positie in, andere een iets lagere.

Voor 80% van de scholen bedraagt die verschuiving niet méér dan 1 punt, voor 50% zelfs niet meer dan een half punt. In de mate waarin scholen er tegen het einde van L2 in slagen te compenseren voor het negatieve GCE van bij het begin van L1, speelt de aanvankelijke Nederlandse taalvaardigheid nog wel een rol, maar die rol is bescheiden. Een lagere positie in figuur 3.16b dan in figuur 3.16 wijst op een stimulerend effect van een initiële Nederlandse taalvaardigheid die hoger uitviel dan verwacht. Een hogere positie wijst op een remmend effect van een initiële taalvaardigheid die lager uitviel dan verwacht kon worden op basis van de overige instroomkenmerken.

Voor een meer gedetailleerde analyse van de samenhang tussen de gecumuleerde groepscompositie-effecten op de leerwinst in L1 en L2 in een analyse zonder en een analyse met correctie voor de effecten van aanvankelijke Nederlandse taalvaardigheid, zie bijlage 1.

Vier voorbeelden

Bij wijze van illustratie presenteren we in tabel 3.7 de belangrijkste waarden voor vier scholen: twee zeer kansarme en twee modale. (De nummers in de kolom “school” verwijzen *niet* naar de schoolnummers in de voorgaande figuren.)

Tabel 3.7 Groepscompositie-effecten en feitelijke extra achterstand of voorsprong voor wiskunde in vier exemplarische basisscholen (zonder corr. voor initiële Nederlandse taalvaardigheid)

School	Schooltype	GCE begin L1 (verwachte extra achter- stand)	Feitelijke extra achterstand / voorsprong begin L1	TW leerwinst L1+L2	Feitelijke extra achterstand / voorsprong einde L2	N begin L1	N einde L2
1	Zeër kansarm	-7,52	-7,64 *	7,12 *	-0,52	24	13
2	Zeër kansarm	-5,78	-5,43 *	-1,49	-6,93 *	26	23
3	Modaal	-2,86	-0,43	-1,50	-1,93	48	45
4	Modaal	-1,32	1,76	-5,22 *	-3,46 *	30	28
Vlaanderen		-2,29	0,00	0,00	0,00		

* Significant verschillend van het gemiddelde van de Vlaamse referentiegroep

Voor elke school kan de graad van kansarmoede afgelezen worden in de derde kolom, die het gecumuleerde groepscompositie-effect begin L1 weergeeft, d.i. de extra achterstand die op grond van de samenstelling van de SiBO-cohorte in de betreffende school te verwachten is. Die extra achterstand is duidelijk veel groter in beide zeer kansarme scholen. Merk op dat het gemiddelde voor Vlaanderen hier niet nul is. Het GCE begin L1 is maar nul voor scholen zonder enige GOK-leerling en geen enkele anderstalige niet-GOK-leerling. Dat is voor Vlaanderen gemiddeld zeker niet het geval. In beide zeer kansarme scholen komt de feitelijke gemiddelde extra achterstand begin L1 (vierde

kolom) sterk overeen met wat op grond van de samenstelling van de SiBO-cohorte in die scholen te verwachten viel. Voor beide modale scholen is de feitelijke extra achterstand begin L1 (kolom 4) minder groot dan het GCE in kolom 3 aangeeft. Tussen beide zeer kansarme, maar ook tussen beide modale scholen zijn er grote verschillen in TW met betrekking tot de gecorrigeerde leerwinst die over de hele periode van begin L1 tot einde L2 voor wiskunde geboekt wordt. Tussen beide zeer kansarme scholen bedraagt het verschil 8,61 punten (wat erop neerkomt dat de ene school in de loop van twee schooljaren 25 weken uitloopt op de andere). Tussen beide modale scholen is het verschil in TW met betrekking tot de leerwinst 3,72 punten (= 11 schoolweken verschil).

Door de grote TW die school 1 realiseert, komt het gemiddelde niveau voor wiskunde tegen het einde van het tweede leerjaar uit op het niveau dat op grond van de *individuele instroomkenmerken* te verwachten viel (feitelijke extra achterstand in kolom 6 zo goed als gelijk aan nul). Van enige extra achterstand ten gevolge van de samenstelling van de leerlingengroep is in school 1 op het einde van L2 geen sprake meer. De leerlingen in school 2 daarentegen eindigen het tweede leerjaar op een niveau dat gemiddeld bijna zeven punten onder het niveau ligt dat op grond van hun individuele instroomkenmerken verwacht had mogen worden. Dat is zelfs nog iets lager dan op grond van de samenstelling van de leerlingengroep te voorzien was. Tussen de twee modale scholen zien we gelijkaardige, zij het iets minder uitgesproken verschillen.

De grote uitstroom in school 1 (van 24 leerlingen begin L1 naar 13 einde L2) dient uiteraard tot voorzichtigheid aan te manen. De gebruikte analysetechniek (repeated measures multiniveau-analyse) corrigeert weliswaar voor wijzigingen in instroomkenmerken t.g.v. in- en uitstroom, maar die correcties berusten op de samenhangen die in de steekproef aanwezig zijn. Dat betekent dat de schatting van de TW van een school niet beïnvloed wordt door uitstroom, in zoverre we mogen aannemen dat uitstromers na het verlaten van de school op dezelfde wijze verder evolueren als doorstromers met dezelfde kind- en achtergrondkenmerken en met gelijkaardige vroegere prestaties in dezelfde school (of in gelijkaardige scholen). Een analoge redenering geldt voor instroom. Afwijkingen van die assumptie leveren ook nog geen grote vertekening van de TW op zolang de in- en uitstroom beperkt is. Maar bij scholen met b.v. grote uitstroomcijfers rijst de vraag in hoeverre de uitstromende leerlingen echt vergelijkbaar zijn met op het eerste zicht qua achtergrond en vroegere prestaties gelijkaardige, maar doorstromende leerlingen in de betreffende of gelijkaardige scholen.

Samengevat

De algemene resultaten in verband met groepscompositie-effecten in het leergebied wiskunde kunnen als volgt worden samengevat:

1. Het beginniveau voor wiskunde van een leerling blijkt over het algemeen significant samen te hangen met de samenstelling van de leerlingengroep in het eerste leerjaar. Naarmate die groep meer (Nederlandstalige of anderstalige) GOK-leerlingen telt, mag een lager begin-

niveau (en dus een extra achterstand) verwacht worden dan voor leerlingen met gelijkaardige instroomkenmerken in scholen met minder GOK-leerlingen in het eerste leerjaar. Dat is mogelijk een “erfenis” uit het kleuteronderwijs maar kan ook wijzen op de grotere aanwezigheid bij (sommige categorieën) GOK-leerlingen van relevante, niet gedetecteerde leerlingkenmerken die van invloed zijn op het beginniveau voor wiskunde. Verschillen in aanvankelijke Nederlandse taalvaardigheid - op zich een belangrijke factor in de verklaring van verschillen tussen GOK-categorieën - blijken slechts ten dele een verklaring te kunnen vormen voor deze groepscompositie-effecten.

2. Die groepscompositie-effecten t.a.v. het beginniveau voor wiskunde zijn in grote mate algemene effecten in deze zin dat ze in ongeveer gelijke mate voor alle leerlingen gelden, wat ook hun individuele achtergrond is. Een uitzondering werd vastgesteld voor het effect van het percentage Arabische of Berber GOK-leerlingen in de klas. Dat percentage leerlingen vertoont wel een significante negatieve samenhang met het beginniveau voor wiskunde voor sommige categorieën van leerlingen (Nederlandstalige GOK-leerlingen, anderstalige niet-GOK-leerlingen), maar niet voor andere (anderstalige GOK-leerlingen).
3. Voor verschillende GOK-categorieën werd een groepscompositie-effect t.a.v. het beginniveau voor wiskunde gevonden. Die effecten verschillen op zich in grootte, tellen van school tot school meer of minder mee naargelang er van de betrokken categorie meer of minder leerlingen zijn en komen in verschillende combinaties met elkaar voor. De gewogen som van al die effecten vormt het gecumuleerde groepscompositie-effect. Dat geeft voor elke school aan welke extra achterstand bij het begin van het eerste leerjaar op grond van de samenstelling van de leerlingengroep in het eerste leerjaar verwacht kon worden.
4. Met betrekking tot de omvang van het gecumuleerde groepscompositie-effect op het beginniveau voor wiskunde zijn er tussen scholen significante en grote verschillen. Bij de zeer kansarme scholen gaat het om een te verwachten extra achterstand van gemiddeld ruim een half schooljaar; bij de modale scholen daarentegen om gemiddeld niet meer dan vier weken. De kansarme scholen én de kansrijke scholen (die veel anderstaligen tellen) nemen doorgaans een middenpositie in, met een te verwachten extra achterstand tussen twee en zes maanden (gemiddeld bijna een trimester).
5. De reële extra achterstand waarmee leerlingen bij het begin van L1 (bovenop de effecten van de eigen individuele achtergrond) te kampen hebben, ligt meestal (heel wat) lager dan wat op grond van de samenstelling van de leerlingengroep in het eerste leerjaar te verwachten was. Het gecumuleerde groepscompositie-effect blijkt in veel scholen van meet af aan al meteen “gemilderd”. Ook op dit vlak zijn er tussen scholen grote verschillen. Zo zijn er ook kansarme en zeer kansarme scholen waar de leerlingen ondanks de minder gunstige

samenstelling van de leerlingengroep in het eerste leerjaar met nauwelijks enige (of zelfs helemaal geen) extra achterstand af te rekenen hebben.

6. Op de leerwinst die tussen het begin van het eerste en het einde van het tweede leerjaar voor wiskunde geboekt wordt, blijkt de samenstelling van de betreffende leerlingengroep een omgekeerd effect te hebben - althans wanneer men alle groepscompositie-effecten en beide leerjaren samenneemt. Leerlingen in scholen met veel GOK-leerlingen in hun leeftijdsgroep maken tussen begin L1 en einde L2 méér vooruitgang dan leerlingen met dezelfde individuele instroomkenmerken in scholen met minder GOK-leerlingen in hun groep. Het is niet duidelijk hoe dit omgekeerde groepscompositie-effect precies geïnterpreteerd moet worden. Het zou om een pure afspiegeling kunnen gaan van de inhaalbeweging die bij bepaalde categorieën GOK-leerlingen vastgesteld werd. Het zou ook kunnen betekenen dat leraren een compenserende aanpak hanteren. In elk geval is sprake van een *ongelijkheidsverminderend* effect.
7. Ook de groepscompositie-effecten op de vooruitgang in eerste en tweede leerjaar zijn grotendeels algemeen van aard. Een uitzondering betreft het effect van het percentage Turkse GOK-leerlingen op de leerwinst in L1. Dat is alleen significant voor Turkse en Arabische of Berber GOK-leerlingen (in positieve zin) en voor overige anderstalige GOK-leerlingen (in negatieve zin) en niet voor de andere categorieën.
8. Het positieve gecumuleerde groepscompositie-effect op de leerwinst in L1 en L2 compenseert gedeeltelijk het negatieve gecumuleerde groepscompositie-effect op het beginniveau. Heel wat scholen realiseren in de loop van het eerste en het tweede leerjaar ook nog eens meer toegevoegde waarde dan nodig om te compenseren voor wat rest van het negatieve gecumuleerde GCE op het beginniveau. Die twee elementen, samen met het gegeven dat het GCE op de beginscore meestal al van meet af aan minstens gedeeltelijk gemilderd was, leiden ertoe dat de meeste scholen er tegen het einde van L2 minstens gedeeltelijk in slagen te compenseren voor het negatieve gecumuleerde groepscompositie-effect op het beginniveau voor wiskunde.
9. Ook met betrekking tot de mate waarin in de loop van het eerste en het tweede leerjaar toegevoegde waarde gerealiseerd wordt voor wiskunde, zijn er grote verschillen tussen scholen. Die verschillen doorkruisen de opdeling in zeer kansarme, kansarme, modale en kansrijke scholen. Wanneer we focussen op gemiddelde leerwinst in het eerste en het tweede leerjaar en de toegevoegde waarde die scholen daarbij realiseren, dan blijkt dat er ook in de categorieën kansarme en zeer kansarme ("zwarte") scholen effectieve en heel effectieve scholen te vinden zijn, net zoals als in de categorie modale ("witte") scholen. Omgekeerd blijken er bij alle types scholen ook minder effectieve scholen te zijn.

3.6 Besluit

De bevindingen voor het leerdomein wiskunde kunnen als volgt worden samengevat:

1. Op de meetschaal voor wiskundevaardigheid die in het SiBO-onderzoek wordt gebruikt, maken Vlaamse leerlingen van begin L1 naar einde L2 een gestage groei door.
2. Niveau en evolutie in wiskundevaardigheid worden beïnvloed door allerlei kind- en achtergrondfactoren. Aanvangsniveau én leerwinst zijn groter voor jongens dan voor meisjes, kleiner voor wie schoolse vertraging heeft opgelopen en vanaf het tweede leerjaar ook kleiner voor wie het jaar voordien het eerste leerjaar heeft overgedaan. Er is ook een significante samenhang met aanvankelijke Nederlandse taalvaardigheid. Die samenhang betreft vooral het niveau in wiskundevaardigheid en veel minder de leerwinst.
3. De grootste verschillen in vaardigheidsniveau en leerwinst voor wiskunde treffen we echter aan tussen leerlingen met een verschillende sociale en etnisch-culturele achtergrond. Die verschillen zijn vooral bij het begin van het eerste leerjaar groot. Naargelang van de specifieke categorie varieert de achterstand voor wiskunde t.o.v. modale Vlaamse leerlingen bij het begin van het eerste leerjaar van bijna een trimester (Nederlandstalige GOK-leerlingen) tot bijna een volledig schooljaar (Turkse GOK-leerlingen). Anderstalige GOK-leerlingen maken in de loop van het eerste leerjaar wel een inhaalbeweging. Voor Turkse en voor Arabisch of Berber sprekende GOK-leerlingen bijvoorbeeld bedraagt de achterstand op modale Vlaamse leerlingen op het einde van het eerste leerjaar nog “maar” een goed half schooljaar. Voor Nederlandstalige GOK-leerlingen noteren we op het einde van het eerste leerjaar een status quo. In de loop van het tweede leerjaar boeken de verschillende GOK-categorieën gemiddeld ongeveer eenzelfde leerwinst. Op het einde van het tweede leerjaar zijn hun onderlinge posities bijgevolg nagenoeg dezelfde als op het einde van het eerste leerjaar.
4. Opvallend is dat Nederlandstalige GOK-leerlingen in de loop van de eerste twee leerjaren geen enkele inhaalbeweging maken. Integendeel, hun achterstand dikt nog iets aan. Samen met de vaststelling dat geen enkele van de andere GOK-categorieën die Nederlandstalige GOK-leerlingen blijkt in te halen, leidt dat tot de conclusie dat de achterstand in wiskunde die voor rekening komt van sociaal-economische en/of sociaal-culturele factoren heel hardnekkig is. Achterstand daarentegen die met etnisch-culturele factoren samenhangt, blijkt - althans in het eerste leerjaar - wel voor een belangrijk deel recupereerbaar te zijn.
5. De aanvankelijke achterstand voor wiskunde bij de onderscheiden GOK-categorieën van leerlingen blijkt in belangrijke mate samen te hangen met een geringere initiële Nederlandse taalvaardigheid. Die “belangrijke mate” varieert naargelang van de GOK-categorie van 26% tot 44%. Vooral de initiële achterstand van Nederlandstalige GOK-leerlingen en van Turkse

GOK-leerlingen hangt samen met een gebrekkige Nederlandse taalvaardigheid. Ook die factor blijkt eerder hardnekkig te zijn, getuige de samenhangen met wiskundeprestaties tot op (minstens) het einde van het tweede leerjaar. Toch kan uit de gevonden samenhangen niet zonder meer afgeleid worden dat aanvankelijke Nederlandse taalvaardigheid op zich de oorzaak van een persistente achterstand in wiskunde is. Een derde, meer algemene achterliggende factor kan aan de basis liggen.

6. Tussen scholen blijken er grote verschillen in gemiddelde beginniveau en gemiddelde leerwinst te bestaan. Een groot deel van die verschillen kan door de effecten van instroomkenmerken verklaard worden. De resterende variantie is echter nog altijd significant. Ook na correctie voor instroomkenmerken blijken tussen aanzienlijke aantallen basisscholen nog significante verschillen te bestaan, zowel inzake gemiddelde beginniveau als inzake toegevoegde waarde met betrekking tot de leerwinst die in het eerste en het tweede leerjaar gerealiseerd wordt.
7. Tussen begin en einde van het eerste leerjaar neemt de variantie tussen scholen (na correctie voor de effecten van instroomkenmerken) toe. Daarna blijft ze ongeveer gelijk. Dat ondanks de verschillen in toegevoegde waarde met betrekking tot de leerwinst in het tweede leerjaar de verschillen tussen scholen op het einde van het tweede leerjaar niet groter zijn dan op het einde van het eerste leerjaar, heeft te maken met de negatieve samenhang tussen de toegevoegde waarde met betrekking tot de leerwinst in het tweede leerjaar en de toegevoegde waarde met betrekking tot de leerwinst in het eerste leerjaar. Scholen met een hogere toegevoegde waarde in het eerste leerjaar realiseren eerder een lagere toegevoegde waarde in het tweede leerjaar en omgekeerd.
8. De verschillen tussen scholen zijn niet alleen significant, ze zijn ook behoorlijk groot. Na correctie voor instroomkenmerken komt het verschil in gemiddelde leerwinst die tussen het begin van het eerste en het einde van het tweede leerjaar gerealiseerd wordt, tussen een school op deciel 1 (de laagste 10%) en een school op deciel 9 (de hoogste 10%) overeen met wat in de Vlaamse referentiegroep gemiddeld in een half schooljaar aan leerwinst geboekt wordt. Tegen het einde van het tweede leerjaar zijn (na correctie voor instroomkenmerken) de 10% best presterende scholen al meer dan een half schooljaar uitgelopen op de 10% zwakst presterende scholen.
9. De verschillen die er - na correctie voor instroomkenmerken - tussen scholen inzake gemiddeld beginniveau zijn, zijn ook niet echt klein te noemen. Het verschil tussen de laagste en de hoogste tien procent scholen (deciel 1 vs. deciel 9) komt overeen met bijna 40% van de leerwinst die gemiddeld in het eerste leerjaar voor wiskunde geboekt wordt.

10. Hoe die verschillen in gecorrigeerd beginniveau te verklaren zijn, is op basis van de voorliggende analyses moeilijk te bepalen. Gezien het om de beginmeting gaat, kunnen klas- en schoolfactoren binnen het lager onderwijs moeilijk als verklaring aangewezen worden. Het kan eventueel om effecten van het kleuteronderwijs gaan, maar evengoed om effecten van niet-gedetectedeerde leerlingkenmerken waarop scholen onderling verschillen. De samenstelling van de betreffende leerlingcohorten (in termen van het aandeel leerlingen uit elke GOK-categorie) blijkt een belangrijk deel van die aanvankelijke verschillen te “verklaren”. Een echte verklaring is dat echter niet, gezien ook groepscompositie-effecten t.a.v. het beginniveau beschouwd moeten worden als effecten van hetzij het kleuteronderwijs, hetzij niet-gedetectedeerde gemeenschappelijke kenmerken van subgroepen van leerlingen.
11. Over het geheel van de scholen bekeken is er een duidelijke samenhang tussen de samenstelling van de leerlingcohorten in termen van aandeel leerlingen uit elke GOK-categorie en het gecorrigeerde gemiddelde beginniveau. Die algehele samenhang belet evenwel niet dat scholen met eenzelfde samenstelling naar GOK-categorieën toch sterk kunnen verschillen in gemiddeld gecorrigeerd beginniveau. Vooral bij de “kansarme scholen” (met relatief veel Vlaamse GOK-leerlingen) en de “zeer kansarme” scholen (met veel anderstalige GOK-leerlingen) zijn die verschillen soms zeer groot. Op de vraag welke “beschermende” resp. “bezwarende factoren” hieraan ten grondslag liggen, moeten we voorlopig ook nog het antwoord schuldig blijven. Men kan denken in termen van: effecten van het kleuteronderwijs, verschillen in participatie aan het kleuteronderwijs, buurtgebonden karakteristieken, of andere, niet-gedetectedeerde leerlingkenmerken die binnen (een leerlingcohorten in) een school in grote mate gemeenschappelijk zijn.
12. Een algemene vaststelling is dat de effectiviteit van een school voor wiskunde (in termen van de toegevoegde waarde voor wiskunde die in L1 en L2 samen gerealiseerd wordt) niet samenhangt met de samenstelling van de leerlingcohorten die in het SiBO-onderzoek gevolgd wordt. Zowel onder de kansarme scholen als onder de zeer kansarme scholen treffen we voor wiskunde evengoed effectieve en zeer effectieve scholen aan als onder de modale scholen.
13. Voor wat de leerwinst voor wiskunde in de beginjaren van het lager onderwijs betreft, zijn er bijgevolg geen argumenten die pleiten voor een dringende etnische desegregatie (lees: opdoeken van concentratiescholen). De voorliggende resultaten laten niet toe te concluderen dat een gelijkmatige spreiding van allochtone leerlingen over alle basisscholen binnen een regio automatisch zou leiden tot een verbetering van de leerwinst van de leerlingen. Er dient eerder gezocht te worden naar de factoren die de verschillen in effectiviteit tussen scholen kunnen helpen verklaren. Dat geldt zowel voor de kansarme en de zeer kansarme (“zwarte”) scholen als voor de modale (“witte”) scholen. Die factoren moeten misschien niet

alleen in de school zelf gezocht worden. Wat op het eerste zicht “schoolgerelateerde” verschillen lijken te zijn, zou ook te maken kunnen hebben met verschillen in participatie aan het kleuteronderwijs of met buurt- of wijkgerelateerde kenmerken.

14. Sprekend over het effect van de samenstelling van de leerlingcohorten, moet uiteraard niet alleen gekeken worden naar het effect op de leerwinst. Men dient evenzeer te kijken naar de samenhang met het niveau dat op een gegeven ogenblik bereikt wordt, na correctie voor de effecten van de individuele achtergrond van de leerlingen. In dat verband stellen we vast dat van begin eerste leerjaar naar einde tweede leerjaar de samenhang tussen samenstelling van de betreffende leerlingcohorten en het gemiddelde gecorrigeerde leerniveau drastisch afneemt. Tussen scholen met eenzelfde samenstelling van de leerlingenbevolking zijn de verschillen in het bereikte niveau in die periode alleen maar toegenomen. Dat hangt uiteraard vooral samen met de verschillen in toegevoegde waarde tussen scholen met gelijkaardige leerlingpopulaties, maar het vormt toch een bijkomend argument voor de conclusie die onder punt 13 geformuleerd werd.
15. Meestal realiseren scholen voor de verschillende categorieën van leerlingen eenzelfde of zeer gelijkaardige toegevoegde waarde. Dat kan opgevat worden als een aanwijzing dat (althans voor het leergebied wiskunde) goed onderwijs de belangrijkste basis voor een zorgbreed onderwijs is.

4. Leerwinst en toegevoegde waarde voor technisch lezen

Bij de leerlingen in het SiBO-onderzoek werd op het einde van het eerste leerjaar en daarna ook op het einde van het tweede en enkele navolgende leerjaren de Drie-Minuten-Toets voor Vlaanderen (DMT) afgenomen. De DMT werd oorspronkelijk door het Nederlandse Centraal Instituut voor Toetsontwikkeling (CiTO) voor gebruik in Nederland ontwikkeld en later in opdracht van de Vlaamse overheid voor Vlaanderen aangepast (Rymenans, 2000). We bespreken eerst kort de inhoud van de toets en de scoringswijze. Daarna bespreken we de bevindingen met betrekking tot leerwinst en toegevoegde waarde voor technisch lezen.

4.1 De Drie-Minuten-Toets voor Vlaanderen (DMT)

De DMT bestaat uit drie verschillende leeskaarten die elk gedurende één minuut individueel hardop gelezen moeten worden. De eerste leeskaart bevat woorden van het type ‘KM’, ‘MK’ en ‘MKM’.³ Op de tweede leeskaart staan woorden van het type ‘MMKM’, ‘MKMM’, ‘MMKMM’, ‘MMMKM’ en ‘MKMMM(M)’. De derde leeskaart bevat woorden met twee, drie of vier lettergrepen. Van elke leeskaart werden drie parallelle versies (A, B en C) gemaakt die door elkaar kunnen gebruikt worden. Ze bevatten dezelfde woorden, alleen staan ze telkens in een andere volgorde (Moelands & Rymenans, 2003a). In het SiBO-onderzoek werd de A-versie gebruikt.

De “Drie-Minuten-Toets” werd individueel afgenomen, door een personeelslid van de betrokken scholen: de klassenleerkracht, een andere leerkracht, de zorgcoördinator of de directie. De afname vond plaats op een door de school zelf bepaald moment in de week van 24 tot 28 mei 2004. De verdere verwerking van de registratieformulieren gebeurde door het SiBO-team.

Voor elke leeskaart wordt de score bepaald als het aantal correct gelezen woorden. De totaalscore, gevormd door de som van de drie deelscores, geldt als indicator voor het bereikte niveau voor technisch lezen. Gezien op de verschillende meetmomenten dezelfde toets wordt gebruikt, kunnen de scores voor de verschillende meetmomenten direct op dezelfde schaal geplaatst worden, zonder dat daar enige IRT-calibratie voor nodig is. Meer details betreffende de psychometrische kenmerken van de DMT vindt men bij Rymenans (2000) en Moelands en Rymenans (2003a). Gegevens over de verdelingskenmerken en interne samenhang van de toets bij de SiBO-steekproef vindt men in LOA-rapport 30 (Verachtert e.a., 2005) en LOA-rapport 35 (Hendrikx e.a., 2005).

Met het oog op de verdere analyses werd erop toegezien dat voor het maximaal mogelijke aantal leerlingen een totaalscore voor technisch lezen werd verkregen. Voor het beperkte aantal leerlingen voor wie slechts één of twee deelscores beschikbaar waren, werd daarom door middel van vooraf-

³ M = medeklinker, K = klinker.

gaandelijke regressie-analyses een waarde voor de ontbrekende deelscores geschat. Vervolgens werd van de oorspronkelijk beschikbare en de geëxtrapoleerde deelscore(s) de som gemaakt.

Vanwege de erg scheve verdeling van de scores voor technisch lezen in het eerste leerjaar, werden in eerdere analyses (Van de gaer e.a., 2006) de ruwe scores voor technisch lezen door middel van het nemen van de vierkantswortel getransformeerd zodat een minder scheve verdeling werd verkregen. Omdat uit vooranalyses bleek dat een dergelijke transformatie voor de gegevens uit L2 geen technische voordelen opleverde (de verdeling werd er niet minder scheef op), werd in de voorliggende analyses van dit soort transformatie geen gebruik gemaakt. Dat betekent dat de scores voor technisch lezen - zoals die in dit rapport gepresenteerd worden - direct interpreteerbaar zijn als het totale aantal tijdens drie keer één minuut correct gelezen woordjes.

4.2 Algemene evolutie van einde L1 naar einde L2

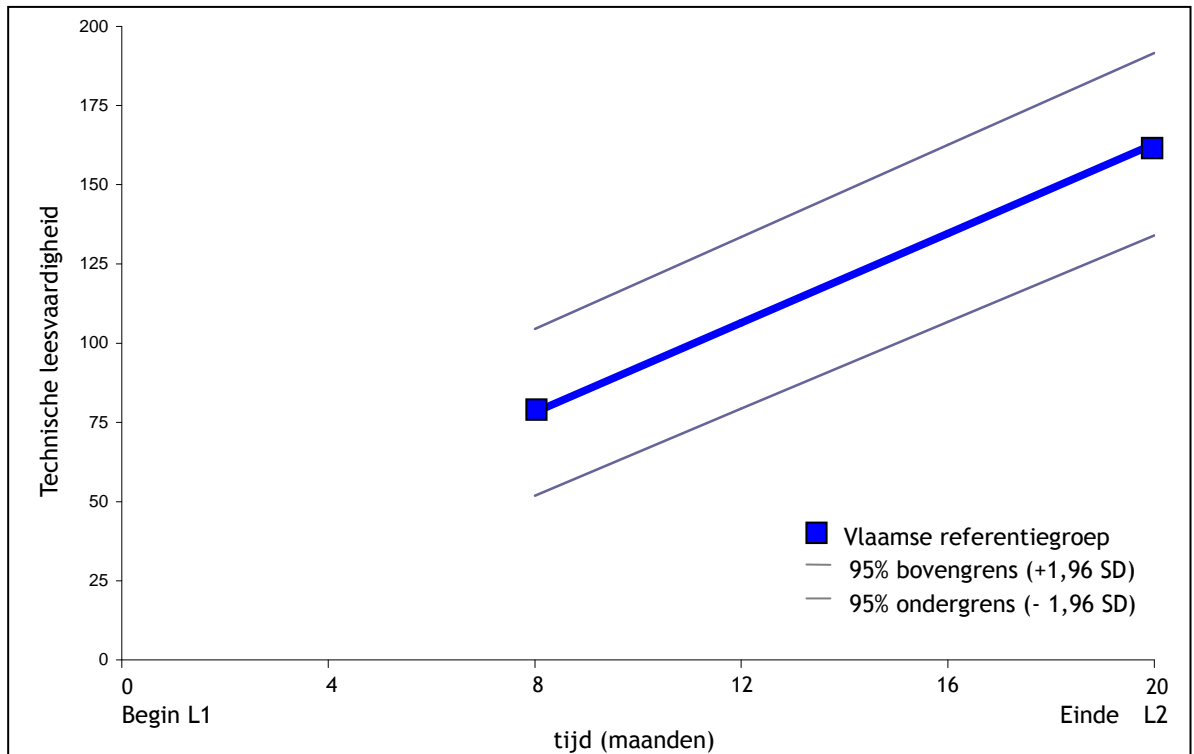
Bij de analyse van de algemene evolutie voor technisch lezen (= zonder onderscheid naar subcategorieën) werd gecorrigeerd voor de effecten van sekse, geboortjaar en geboortemaand van de leerlingen, het al dan niet blijven zitten in L1 of L2, GOK-categorie en score voor aanvankelijke Nederlandse taalvaardigheid (zie inleiding). Figuur 4.1 toont de gemiddelde groeicurve van de *Vlaamse referentiegroep* (zie p. 12). Dat is de groeicurve waarmee de gecorrigeerde gemiddelde groeicurves van afzonderlijke scholen vergeleken kunnen worden.

Om de gelijkheid met de grafieken met de groeicurves voor wiskunde te behouden én om ook grafisch duidelijk te maken dat de eerste score voor technisch lezen geen echte “beginscore” is (maar het resultaat van onderwijs en leren in het eerste leerjaar), loopt de horizontale as van begin L1 tot einde L2.

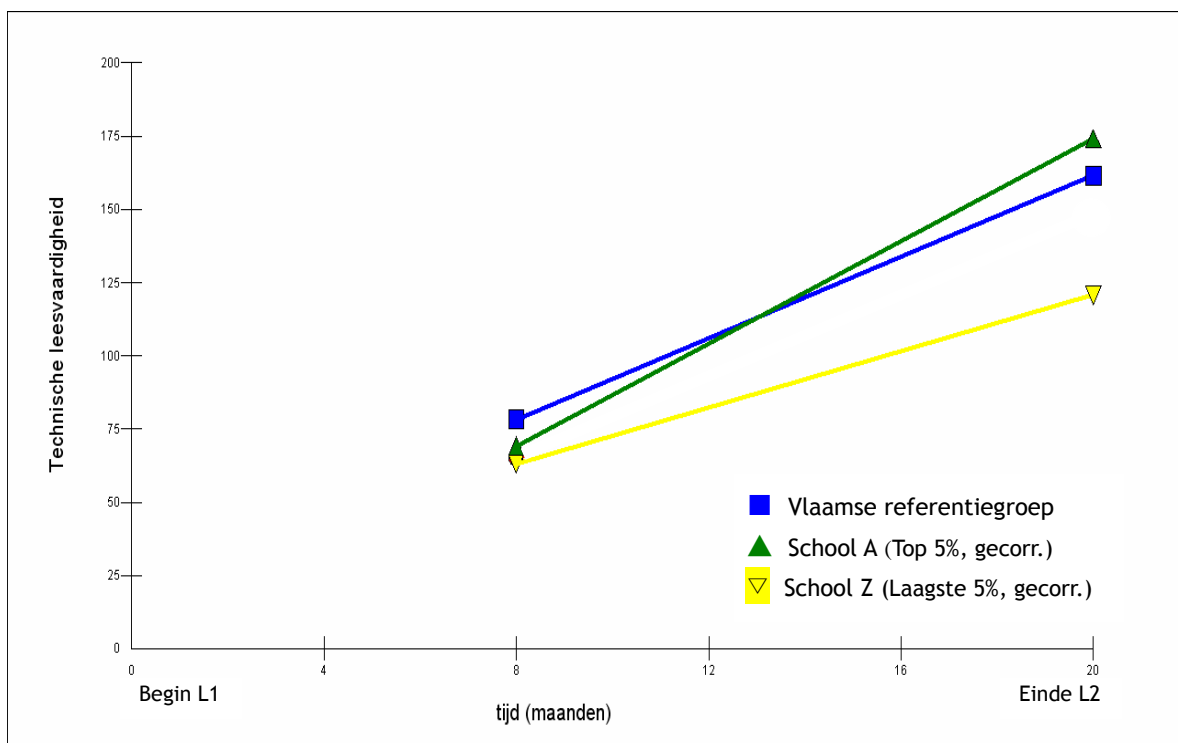
De twee dunnere lijnen boven de gemiddelde groeicurve geven de grenzen aan waarbinnen zich (uitgaande van een normale verdeling) 90% van de schoolgemiddelden bevinden. Die boven- en ondergrens zijn afgeleid van de varianties op schoolniveau op elk van de toetsmomenten en hun onderlinge covarianties. Tussen einde eerste leerjaar en einde tweede leerjaar blijft de variantie tussen scholen nagenoeg constant; de lichte toename is niet significant ($\Delta\text{var} = 35,198$; $\chi^2_{(1)} = 1,346$; ns). In figuur 4.2 geven de gecorrigeerde gemiddelde groeicurves van een school die qua toegevoegde waarde (TW) tot de top 5 behoort (school A) resp. een school die qua TW tot de laagste vijf behoort (school Z) op meer concrete wijze een idee van de variatie binnen het Vlaamse basisonderwijs.

In figuur 4.3 laten we de beginpunten van de verschillende curves samenvallen. Dat laat gemakkelijker toe de hellingsgraad van de verschillende curves te vergelijken.

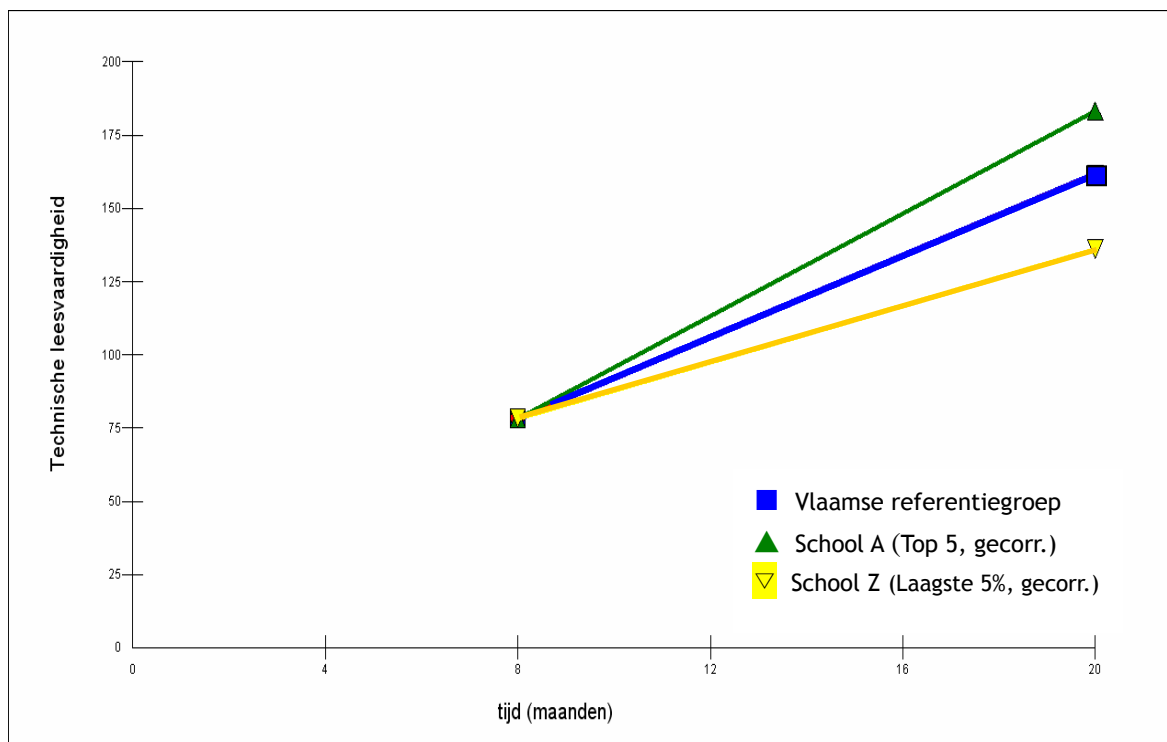
Wat figuur 4.2 op grafische wijze toont, wordt in tabel 4.1 met getallen weergegeven. Een sterretje wijst op een significant verschil met het gemiddelde voor de Vlaamse referentiegroep.



Figuur 4.1 Gemiddelde evolutie voor technisch lezen van einde eerste leerjaar naar einde tweede leerjaar in de Vlaamse referentiegroep (+/- 1,96 x standaarddeviatie op schoolniveau)



Figuur 4.2 Gemiddelde evolutie voor technisch lezen van einde eerste leerjaar naar einde tweede leerjaar in de Vlaamse referentiegroep, in een school met een hoge en in een school met een lage toegevoegde waarde



Figuur 4.3 Gemiddelde evolutie voor technisch lezen van einde eerste leerjaar naar einde tweede leerjaar in het Gentse stedelijke basisonderwijs, Vlaanderen, een school met een hoge TW en een school met een lage TW, na correctie voor verschillen in beginniveau

Tabel 4.1 Gemiddeld niveau einde L1, gemiddelde leerwinst en toegevoegde waarde in L2 voor technisch lezen

	Eerste leerjaar		Tweede leerjaar	
	Gecorrig. gemidd. niveau einde L1	Toegevoegde waarde	Gecorrig. gemidd. leerwinst	Toegevoegde waarde
Vlaamse referentiegroep	78,21	0,00	84,63	0,00
School A	69,22	-8,99	104,93	20,30
School Z	62,99	-15,22	57,76	-26,86

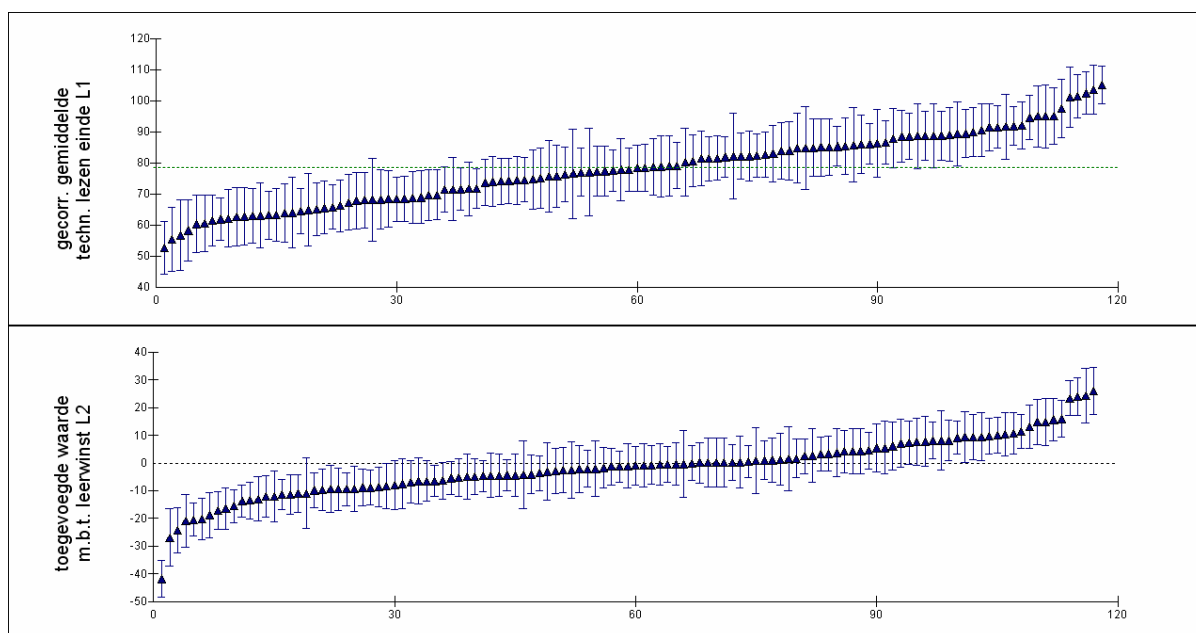
Tabel 4.1 is op dezelfde wijze te interpreteren als tabel 3.2. Er is echter één bijzonderheid: bij het gecorrigeerde gemiddelde niveau voor einde L1 presenteren we ook een “toegevoegde waarde”. De getallen in die kolom geven aan hoeveel hoger (+) of lager (-) het gecorrigeerde gemiddelde einde L1 is dan het gemiddelde in de Vlaamse referentiegroep. Of ook nog: hoeveel hoger of lager school A en school Z gemiddeld scoren dan op grond van de individuele instroomkenmerken van hun leerlingen verwacht kon worden.

Gezien we mogen aannemen dat normaalvorderende leerlingen in het Vlaamse basisonderwijs het eerste leerjaar normaliter starten zonder enige noemenswaardige technische leesvaardigheid, geeft

het niveau einde L1 eigenlijk ook aan hoeveel vooruitgang (“leerwinst”) in de loop van het eerste leerjaar geboekt werd. Wat in de kolom ernaast onder “toegevoegde waarde” staat, geeft daarom aan hoeveel groter of kleiner de bijdrage van school A resp. school Z in die vooruitgang is in vergelijking met de gemiddelde Vlaamse school. Men kan die toegevoegde waarde dus op dezelfde manier interpreteren als de toegevoegde waarde met betrekking tot de leerwinst in het tweede leerjaar.

Tegen het einde van L2 ligt het technische leesniveau in school Z gemiddeld 42 woordjes lager (-15,22 -26,86 = -42,08) dan op grond van de instroomkenmerken van de leerlingen verwacht kon worden. Dat komt overeen met de helft van de leerwinst die modale Vlaamse leerlingen gemiddeld in het tweede leerjaar realiseren. We kunnen dus zeggen dat school Z tegen het einde van het tweede leerjaar een half schooljaar achterop ligt t.a.v. andere scholen met een gelijkaardige instroom gemiddeld bereiken. Dat gecumuleerde verschil (over L1 en L2 samen) is statistisch significant. (Zoals men in tabel 4.1 kan zien, is de TW voor L1 en L2 afzonderlijk niet significant.)

Een iets algemener beeld van de verscheidenheid in prestaties voor technisch lezen krijgen we als we kijken naar het verschil tussen scholen die qua gemiddelde niveau op het eerste deciel staan (de tien procent zwakste scholen) en scholen die op het negende deciel staan (de tien procent sterkste scholen). Op het einde van L1 komt dat verschil overeen met de leerwinst die voor technisch lezen gemiddeld in één trimester geboekt wordt. Tegen het einde van L2 is dat verschil nauwelijks gewijzigd. Voor technisch lezen zijn de verschillen tussen scholen dus minder groot dan voor wiskunde.



Figuur 4.4 Verschillen tussen scholen (referentiesteekproef) voor gemiddeld gecorrigeerd technisch leesniveau einde L1 en toegevoegde waarde m.b.t. leerwinst in L2

Figuur 4.4 toont de scholen uit de referentiesteekproef geordend naar het gecorrigeerde niveau voor technisch lezen einde L1 (boven) resp. de toegevoegde waarde met betrekking tot de leerwinst in het tweede leerjaar (onder). De verticale staafjes geven het betrouwbaarheidsinterval weer. Dat is zo gekozen ($CI = 1,39 \times SE$) dat voor elke twee scholen waarvan de staafjes elkaar niet overlappen, geldt dat ze significant van elkaar verschillen. Voor beide vergelijkingspunten (gecorrigeerd niveau einde L1 en TW L2) zijn er behoorlijk veel scholen die significant verschillen van andere scholen.

De samenhang tussen het niveau einde L1 en de toegevoegde waarde in L2

Scholen met een hoger gecorrigeerd gemiddelde einde L1 halen eerder een lagere TW met betrekking tot de leerwinst in L2 en omgekeerd ($r_{(117)} = -0,463$; $p < 0,001$). Dat belet niet dat de correlatie tussen het gecorrigeerde gemiddelde prestatieniveau einde L1 en het gecorrigeerde gemiddelde prestatieniveau einde L2 toch nog positief is ($r_{(117)} = 0,576$; $p < 0,001$).

4.3 Evolutie voor de verschillende GOK-categorieën

Zoals in hoofdstuk 3 werd uitgelegd, werd bij de analyse van de verschillen tussen de GOK-categorieën niet gecorrigeerd voor de effecten van aanvankelijke Nederlandse taalvaardigheid, wel voor de effecten van de overige individuele instroomkenmerken. Meer nog dan voor wiskunde blijkt ook voor technisch lezen de achterstand van de verschillende GOK-categorieën t.a.v. de modale Vlaamse leerlingen immers in belangrijke mate samen te hangen met verschillen in initiële Nederlandse taalvaardigheid (zie tabel 4.2). Correctie voor de effecten van aanvankelijke Nederlandse taalvaardigheid zou daarom een sterk vertekend beeld geven van de werkelijke achterstand van de onderscheiden categorieën GOK-leerlingen.

Tabel 4.2 Deel van achterstand voor technisch lezen (einde L1) dat samenhangt met een minder goede aanvankelijke Nederlandse taalvaardigheid, per GOK-categorie

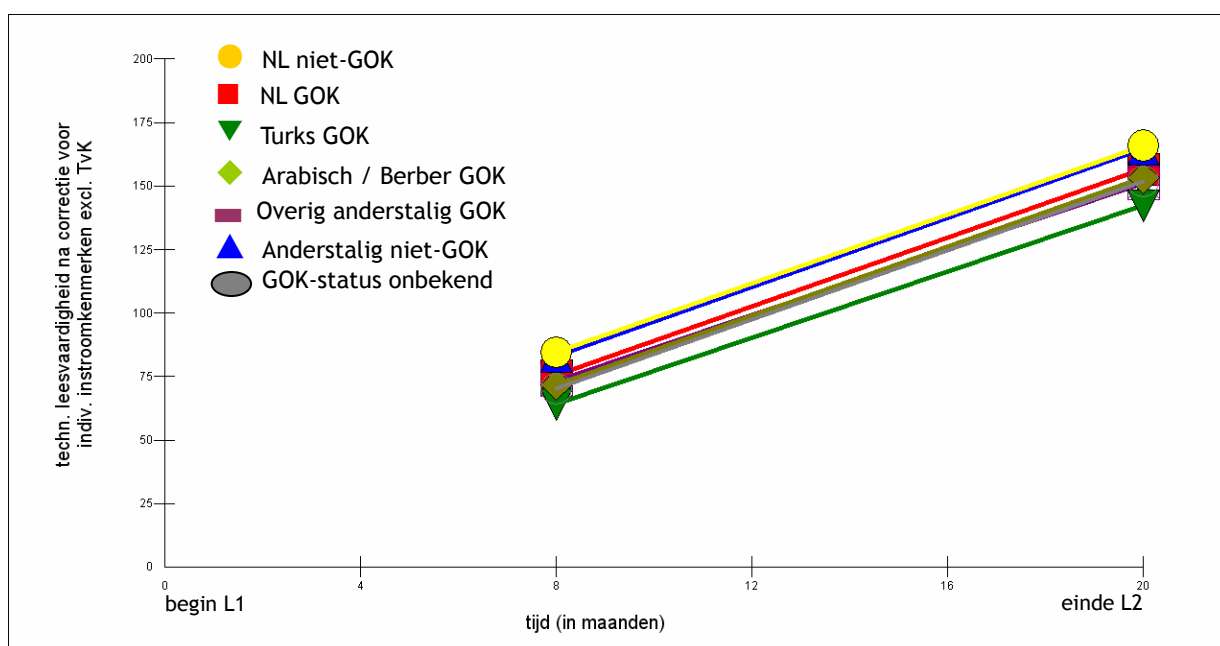
<i>GOK-categorie</i>	<i>Achterstand verklaard door TvK-score</i>
Nederlandstalig GOK	65,4 %
Turks GOK	77,4 %
Arabisch/Berber GOK	73,8 %
Overige anderstalig GOK	90,7 %
Anderstalig niet-GOK	---
GOK-status onbekend	56,6 %

In tabel 4.2 wordt voor de categorie “anderstalige niet-GOK-leerlingen” geen percentage gegeven. Dat komt omdat correctie voor de effecten van de TvK-score de beperkte achterstand van die categorie (zie tabel 4.3) omzet in een voorsprong.

Figuur 4.5 laat zien dat na die correctie de verschillen tussen de GOK-categorieën voor technisch lezen relatief klein zijn wanneer ze afgezet worden tegenover de gemiddelde leerwinst in L2. Grosso

modo kunnen drie posities onderscheiden worden. De koppositie wordt ingenomen door de modale Vlaamse leerlingen en de - qua technische leesvaardigheid nauwelijks van hen te onderscheiden - categorie van de anderstalige niet-GOK-leerlingen. Onderaan vindt men de Turkse GOK-leerlingen, die t.o.v. de modale Vlaamse leerlingen op het einde van L1 gemiddeld een achterstand hebben die overeenkomt met een kwart van de leerwinst die modale Vlaamse leerlingen gemiddeld in L2 voor technisch lezen maken. Daartussenin, met zijn allen ongeveer halverwege de twee andere posities, de overige categorieën: Nederlandstalige GOK-leerlingen, Arabisch of Berber sprekende GOK-leerlingen, overige anderstalige GOK-leerlingen en leerlingen van wie de GOK-status onbekend is. De achterstand van de Turkse GOK-leerlingen t.o.v. de categorie modale Vlaamse leerlingen is verhoudingsgewijs veel minder groot dan voor wiskunde, maar hun lage positie voor technisch lezen bevestigt de tendens die al voor het leergebied wiskunde geobserveerd werd.

Aan de onderlinge posities verandert weinig tussen einde L1 en einde L2. Turkse en overige anderstalige GOK-leerlingen dikken hun achterstand t.o.v. de modale Vlaamse leerlingen nog iets aan; de andere categorieën evolueren zo goed als volledig parallel met de modale Vlaamse leerlingen.



Figuur 4.5 Gemiddelde evolutie voor technisch lezen van einde L1 naar einde L2 voor verschillende GOK-categorieën, na correctie individuele instroomkenmerken behalve aanvankelijke Nederlandse taalvaardigheid

Tabel 4.3 toont de gemiddelde achterstand van de verschillende GOK-categorieën, uitgedrukt als een percentage van de gemiddelde leerwinst voor technisch lezen die de Vlaamse referentiegroep in L2 maakt. Een vergelijking met tabel 3.4 toont dat de achterstanden van GOK-leerlingen voor technisch lezen verhoudingsgewijs veel kleiner zijn dan voor wiskunde. Naargelang van de categorie leerlingen bedraagt de achterstand voor technisch lezen op het einde van L2 slechts een kwart tot de helft van

de achterstand van diezelfde categorie leerlingen voor wiskunde. Bijzonder opvallend is de situatie van de anderstalige niet-GOK-leerlingen die het voor technisch lezen gemiddeld bijna even goed doen als de modale Vlaamse leerlingen.

Tabel 4.3 Achterstand voor technisch lezen van de verschillende GOK-categorieën t.o.v. modale Vlaamse leerlingen¹

GOK-categorie	Toetsmoment	
	Einde L1	Einde L2
Nederlandstalig GOK	-11,0%	-11,1%
Turks GOK	-24,8%	-28,5%
Arabisch/Berber GOK	-15,3%	-15,2%
Overige anderstalig GOK	-13,5%	-17,7%
Anderstalig niet-GOK	-1,7%	-1,5%
GOK-status onbekend	-17,5%	-17,2%

¹Achterstand uitgedrukt als percentage van de gemiddelde leerwinst van de Vlaamse referentiegroep in L2. Berekend op scores gecorrigeerd voor effecten van overige instroomkenmerken behalve initiële Nederlandse taalvaardigheid

Met betrekking tot de GOK-categorieën werden voor technisch lezen geen differentiële school-effecten vastgesteld. Dat betekent dat het patroon van verschillen tussen de GOK-categorieën dat in tabel 4.3 te zien is, op enkele beperkte uitzonderingen na, in alle basisscholen waargenomen wordt.

4.4 Evolutie van de zittenblijvers

Leerlingen die het eerste leerjaar overdoen, eindigen hun “bis-jaar” met een DMT-score die gemiddeld significant hoger ligt dan bij leerlingen met evenveel, maar vóór het eerste leerjaar opgelopen schoolse vertraging (voor zover alle andere individuele instroomkenmerken behalve aanvankelijke Nederlandse taalvaardigheid gelijk gehouden worden). Voor zittenblijvers L1 met slechts één jaar vertraging, ligt het niveau voor technisch lezen op het einde van hun “bis-jaar” (na voornoemde correctie) gemiddeld zelfs even hoog als dat van normaalvorderende leerlingen (zie figuur 2.4 in hoofdstuk 2). Tegen het einde van L2 liggen zittenblijvers L1 al weer flink achterop: één trimester voor zittenblijvers L1 met één jaar schoolse vertraging, voor de anderen nog iets meer.

4.5 Maakt de samenstelling van de leerlingengroep een verschil uit?

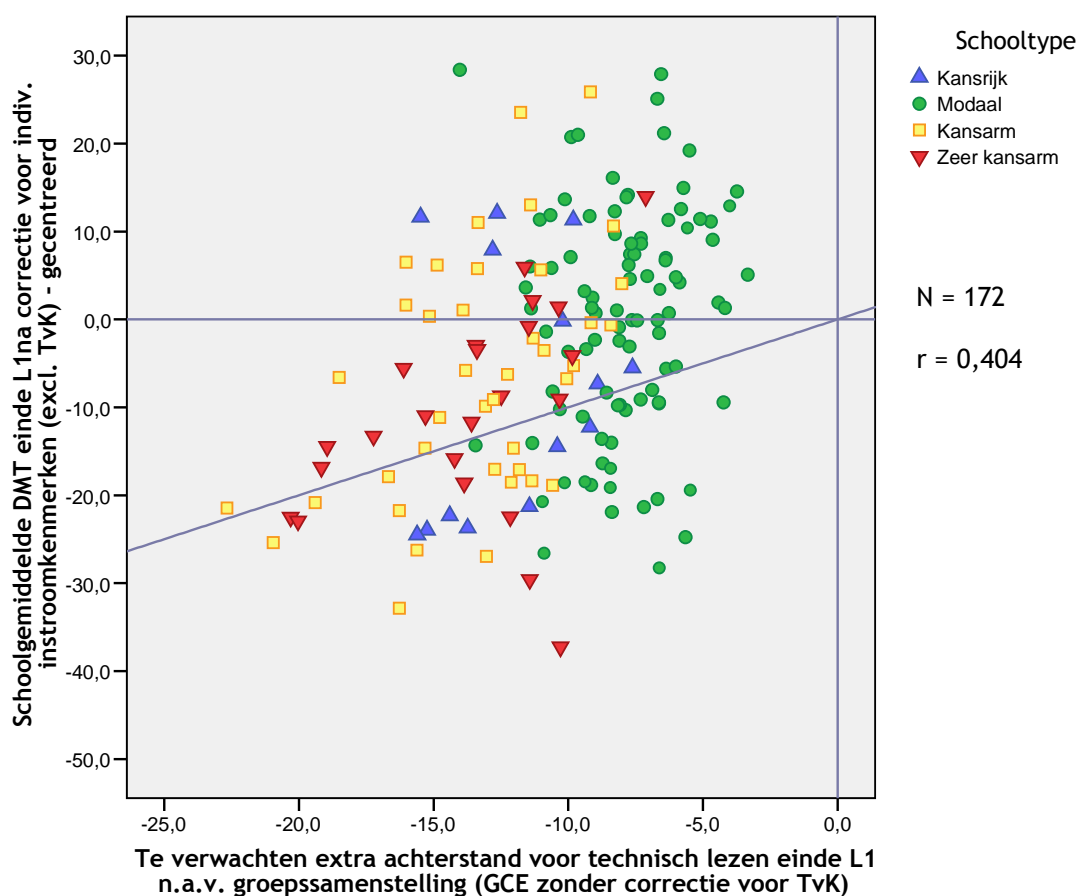
Groepscompositie-effecten werden op dezelfde wijze onderzocht als voor wiskunde (zie hoofdstuk 3). T.a.v. de DMT-score einde L1 werden twee groepscompositie-effecten (GCE) gevonden:

- een negatief GCE met betrekking tot het aandeel Nederlandstalige GOK-leerlingen (hoe meer NL-GOK-leerlingen, hoe lager de gemiddelde DMT-score einde L1),
- een negatief GCE met betrekking tot het aandeel “overige anderstalige” GOK-leerlingen (hoe meer OAT-GOK-leerlingen, hoe lager de gemiddelde DMT-score einde L1).

Het negatieve GCE van het aandeel Nederlandstalige GOK-leerlingen wordt in L2 geneutraliseerd door een evengroot omgekeerd GCE op de leerwinst voor technisch lezen in het tweede leerjaar. Van enige interactie met individuele instroomkenmerken blijkt geen sprake.

Figuur 4.6 geeft een beeld van de omvang van het gecumuleerde groepscompositie-effect op de DMT-score einde L1 en van de samenhang met het gecorrigeerde schoolgemiddelde voor technisch lezen op het einde van L1 (gecentreerd rond het Vlaamse gemiddelde). De scores op de verticale as zijn te interpreteren als een “toegevoegde waarde” en geven aan hoeveel hoger of lager het schoolgemiddelde uitvalt dan op grond van de individuele achtergrondkenmerken van de leerlingen verwacht kon worden. Figuur 4.6 is analoog aan de figuren 3.12 en 3.16 (zie hoofdstuk 3) en ook de interpretatie kan op analoge wijze gebeuren.

Zoals in hoofdstuk 3 al uiteengezet werd, dient het gecumuleerde groepscompositie-effect opgevat te worden als de extra achterstand die in principe verwacht kan worden op basis van de groeps-samenstelling. We spreken van een “extra achterstand” omdat ze bovenop de achterstand komt die GOK-leerlingen al op grond van hun individuele achtergrond kenmerken hebben. We spreken van “in principe te verwachten” omdat in scholen met dezelfde groepssamenstelling de feitelijke extra achterstand soms toch nog blijkt te verschillen.



Figuur 4.6 Gecumuleerd groepscompositie-effect op DMT-score einde L1: verschillen tussen scholen en schooltypes en samenhang met gecorrigeerd schoolgemiddelde einde L1

Die in principe te verwachten extra achterstand in DMT-score op het einde van L1 gaat van ongeveer -3 tot -22. Afgezet tegenover de gemiddelde leerwinst in de Vlaamse referentiegroep gedurende L2 (= 81,2 wanneer niet gecorrigeerd wordt voor de effecten van aanvankelijke Nederlandse taalvaardigheid), is dat verhoudingsgewijs veel kleiner dan het gecumuleerde GCE voor wiskunde bij begin L1 (zie hoofdstuk 3). Het groepscompositie-effect t.a.v. technisch lezen op het einde van L1 bedraagt maximaal een kwart van de gemiddelde leerwinst in technisch lezen; gemiddeld gaat het om een achtste van de gemiddelde leerwinst in L2. De vergelijking met de GCE voor wiskunde begin L1 loopt wel wat mank, omdat er een verschil in tijdstip is: begin L1 versus einde L1.

In figuur 4.6 is een duidelijke opdeling naar schooltype zichtbaar, met de modale scholen aan de rechterkant van het spectrum (zwakker GCE) en de andere scholen aan de linkerkant (sterker GCE). Het is echter duidelijk dat er qua gecumuleerd groepscompositie-effect op de DMT-score einde L1 nauwelijks een onderscheid valt te maken tussen kansarme en zeer kansarme scholen. In dat opzicht is er een duidelijk verschil met figuur 3.12 (zie hoofdstuk 3). Het door elkaar verspreid liggen van kansarme en zeer kansarme scholen in figuur 4.6 komt omdat in het gecumuleerde GCE op de DMT-score het effect van het aandeel Nederlandstalige GOK-leerlingen even zwaar doorweegt als het effect van het aandeel overige anderstalige GOK-leerlingen. Er is voor technisch lezen ook geen sprake van groepscompositie-effecten op grond van het aandeel Turkse GOK-leerlingen of van het aandeel Arabisch of Berber sprekende GOK-leerlingen.

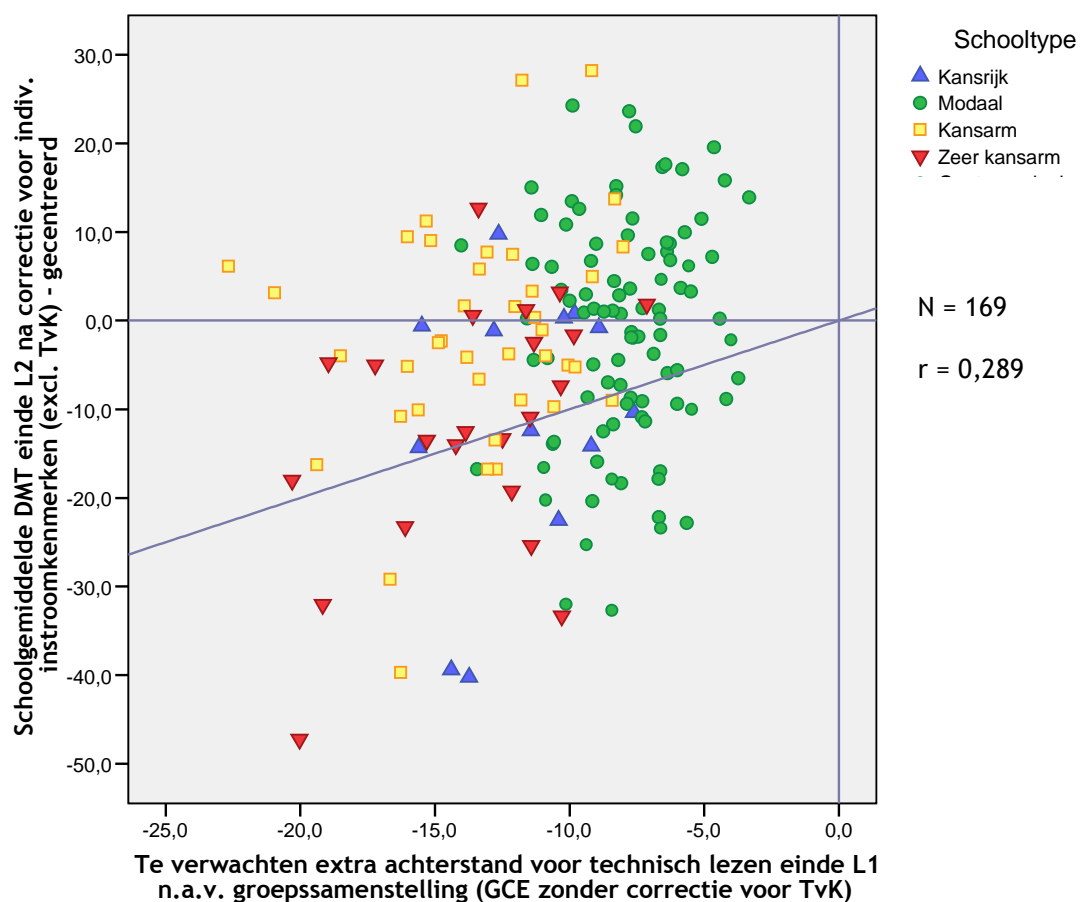
Wat de samenhang met het gecorrigeerde schoolgemiddelde voor DMT einde L1 betreft, valt in figuur 4.6 op dat voor alle schooltypes al sprake is van een vrij ruime spreiding boven en onder de diagonaal en boven en onder de horizontale middellijn. Op dat vlak vertoont figuur 4.6 meer overeenkomsten met figuur 3.16 dan met figuur 3.12. Ook dat heeft natuurlijk te maken met het feit dat de DMT-score einde L1 geen “beginniveau” weergeeft, maar het resultaat van een jaar onderwijs. Een negental (van de 23) zeer kansarme scholen haalt op het einde van het eerste leerjaar een gecorrigeerde gemiddelde DMT-score rond of boven het Vlaamse gemiddelde, evenals 18 van de 40 kansarme scholen. Bij de modale scholen ligt twee derde boven het Vlaamse gemiddelde en bij de kansrijke zes van de veertien. De correlatiecoëfficiënt ($r = 0,404$) geeft aan dat er tussen het gecumuleerde groepscompositie-effect en de gecorrigeerde schoolgemiddelden voor DMT einde eerste leerjaar een matige samenhang bestaat.

De interpretatie van de positie van de scholen is analoog aan wat voor figuur 3.12 werd vermeld: in scholen op de diagonaal hebben de leerlingen op het einde van L1 effectief af te rekenen met de extra achterstand die op grond van de groepssamenstelling te verwachten viel. Scholen tussen de diagonaal en de horizontale nul-lijn, zijn er gedeeltelijk in geslaagd te compenseren voor het groepscompositie-effect. Scholen op of boven de horizontale middellijn zijn er helemaal in geslaagd het GCE te neutraliseren of stijgen er zelfs bovenuit. In scholen onder de diagonaal hebben de leerlingen - ten opzichte van wat op grond van hun individuele instroomkenmerken verwacht kon

worden - tegen het einde van het eerste leerjaar gemiddeld een extra achterstand opgelopen die groter is dan op grond van het gecumuleerde groepscompositie-effect einde L1 verwacht kon worden.

Figuur 4.7 toont dat tegen het einde van L2 de samenhang tussen groepsamenstelling en gemiddeld niveau voor technisch lezen (na correctie voor de effecten van individuele instroomkenmerken behalve aanvankelijke Nederlandse taalvaardigheid) nog verder afgezwakt is. De correlatiecoëfficiënt bedraagt nog slechts 0,289. In vergelijking met figuur 4.6 liggen de scholen t.a.v. de verticale as ook verder uiteen, wat erop wijst dat de verschillen tussen de scholen toegenomen zijn.

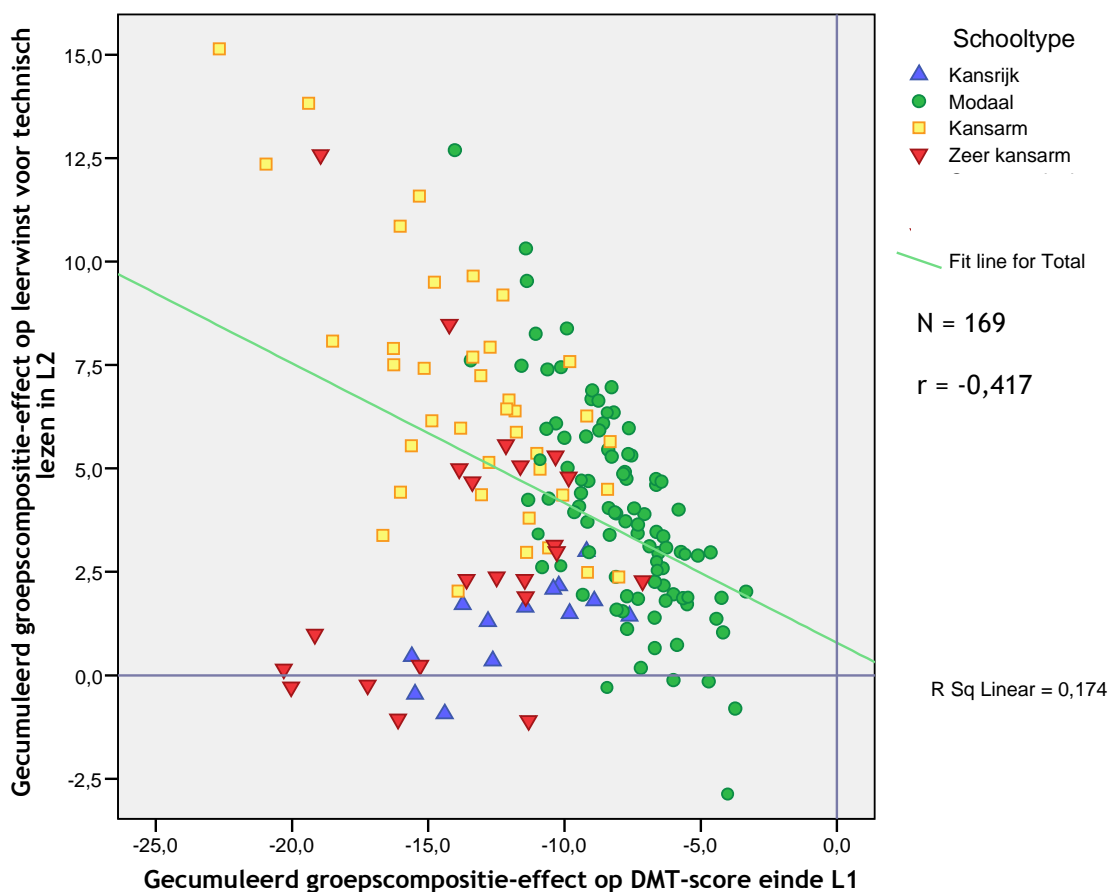
Bij de vergelijking van figuur 4.6 en figuur 4.7 dient men voor ogen te houden dat het gecorrigeerde schoolgemiddelde (verticale as) gecentreerd is rond het Vlaamse gemiddelde van het betreffende meetmoment (einde L1 resp. einde L2). Dat betekent dat de scores op de verticale as niet in absolute zin met elkaar vergeleken kunnen worden. Wel kan men verschuivingen in de positie t.o.v. de horizontale middellijn (het gemiddelde van de Vlaamse referentiegroep) zien.



Figuur 4.7 Mate waarin scholen erin slagen tegen einde L2 het groepscompositie-effect t.a.v. technisch lezen van bij einde L1 te neutraliseren, opgesplitst naar schooltype

Om te zien welke factoren ten grondslag liggen aan het afzwakken van de samenhang die er eind L1 bestond tussen groepsamenstelling en gecorrigeerd gemiddelde voor technisch lezen, is het voor- eerst nuttig te kijken naar de groepscompositie-effecten t.a.v. de leerwinst in L2. We herinneren eraan dat slechts één significant effect gevonden werd: een “omgekeerd” groepscompositie-effect met betrekking tot het aandeel Nederlandstalige GOK-leerlingen in de leerlingengroep. Behalve dat de richting anders is, is dat effect even groot als het effect van het aandeel Nederlandstalige GOK-leerlingen op de DMT-score einde L1. De verschillen in omvang van het gecumuleerde groepscompositie-effect op de leerwinst in L2 (dat dus uitsluitend uit dat ene GCE bestaat) kan men zien in figuur 4.8. Daarin kan men ook zien dat er een negatieve samenhang is met het gecumuleerde groepscompositie-effect op de DMT-score einde L1 (correlatiecoëfficiënt -0,417).

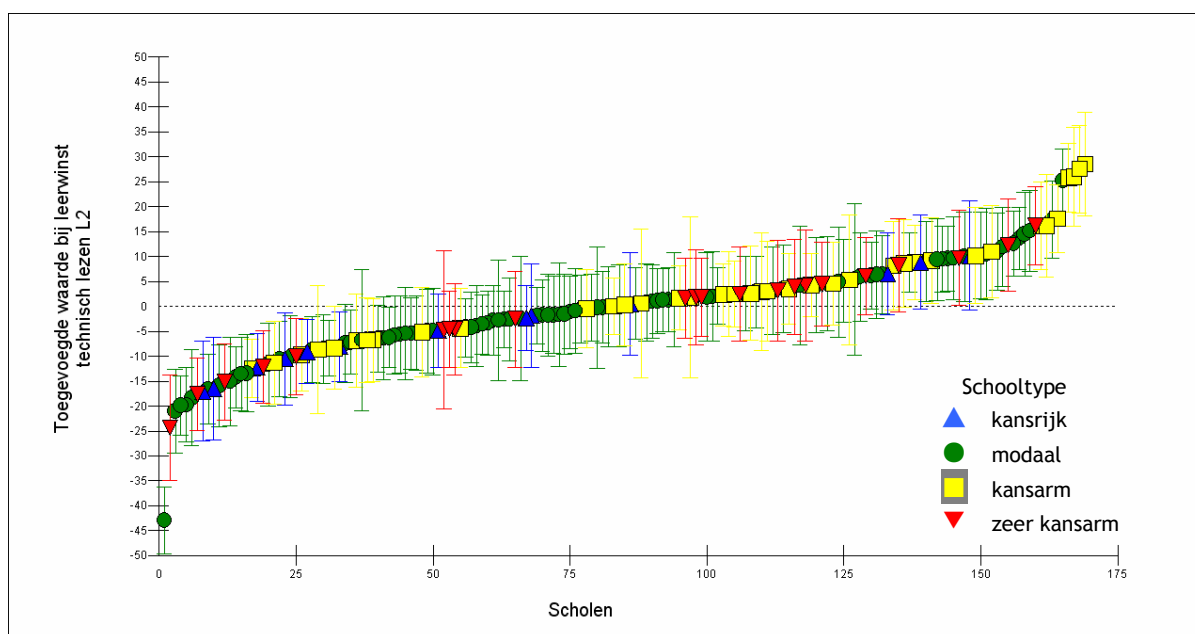
Niet verrassend is dat vooral de kansarme scholen profiteren van het positieve groepscompositie-effect op de leerwinst voor technisch lezen in L2. Die scholen tellen verhoudingsgewijs het grootste aandeel Nederlandstalige GOK-leerlingen.



Figuur 4.8 Gecumuleerd groepscompositie-effect op de leerwinst voor technisch lezen in L2: verschillen tussen scholen en schooltypes en samenhang met gecumuleerd groepscompositie-effect op de DMT-score einde L1

Het *omgekeerde* groepscompositie-effect op de leerwinst voor technisch lezen in L2 verklaart al voor een deel het afzakken van de oorspronkelijke samenhang tussen samenstelling van de leerlingen-cohorte en gemiddelde DMT-score. Een andere verklarende factor is de toegevoegde waarde die scholen realiseren met betrekking tot de leerwinst in het tweede leerjaar. Figuur 4.9 laat zien dat de verschillende schooltypes over het hele spectrum van laag naar hoog verspreid liggen, met zelfs een verhoudingsgewijs licht overwicht van de kansarme en zeer kansarme scholen bij de hogere toegevoegde waarden. Ook voor technisch lezen blijken bij alle schooltypes effectieve en heel effectieve scholen te vinden zijn.

De verticale staafjes in figuur 4.9 geven het zgn. betrouwbaarheidsinterval rond de TW aan ($CI = 1,39 \times SE$). Twee scholen waarvan de staafjes elkaar niet overlappen, verschillen significant van elkaar. Behoorlijk wat scholen zijn in dat geval. In de rups kan een staart- en kopgroep onderscheiden worden die elk ongeveer 20% van de uitgebreide steekproef omvatten. In dat opzicht biedt figuur 4.9 een vrij gelijkaardig beeld als figuur 4.4, waarin alleen de referentiesteekproef opgenomen is. We herinneren dat de TW die in figuur 4.9 afgebeeld wordt, i.t.t. de TW in figuur 4.4, geschat is zónder correctie voor de effecten van initiële Nederlandse taalvaardigheid (maar wel met correctie voor de overige instroomkenmerken).



Figuur 4.9 Verschillen tussen scholen en schooltypes in toegevoegde waarde met betrekking tot de leerwinst voor technisch lezen in het tweede leerjaar

Vier voorbeelden

Bij wijze van illustratie presenteren we in tabel 4.4 de belangrijkste waarden voor vier scholen: twee zeer kansarme en twee modale scholen. (De nummers in de kolom “school” verwijzen *niet* naar

dezelfde scholen als in tabel 3.7). De gegevens in tabel 4.4 laten toe te beoordelen in hoeverre een school erin geslaagd is te compenseren voor de extra achterstand voor technisch lezen die op basis van de groepssamenstelling op het einde van L1 te verwachten viel.

Tabel 4.4 Groepscompositie-effecten en feitelijke extra achterstand of voorsprong voor technisch lezen in vier exemplarische scholen (zonder correctie voor effecten van TvK)

School	Schooltype	GCE einde L1 (verwachte extra achterstand)	Feitelijke extra achterstand / voorsprong einde L1	TW leerwinst L2	Feitelijke extra achterstand / voorsprong einde L2	N einde L1	N einde L2
1	Zeer kansarm	-13,86	-18,63 *	6,09	-12,54	47	36
2	Zeer kansarm	-13,59	-11,70	12,28	0,57	24	20
3	Modaal	-8,44	-19,12 *	-13,54 *	-32,66 *	42	42
4	Modaal	-6,60	3,43	1,22	4,65	74	72
Vlaanderen		-9,60	0,00	0,00	0,00		

* significant verschillend van het gemiddelde van de Vlaamse referentiegroep

Voor elke school kan de graad van kansarmoede afgelezen worden in de derde kolom, die het gecumuleerde groepscompositie-effect einde L1 weergeeft, d.i. de extra achterstand die op grond van de samenstelling van de SiBO-cohorte in de betreffende school te verwachten was. Die extra achterstand is duidelijk groter in beide zeer kansarme scholen. Merk op dat het gemiddelde voor Vlaanderen hier niet nul is. Het GCE einde L1 is maar nul in scholen zonder enige GOK-leerling. Dat is voor Vlaanderen gemiddeld zeker niet het geval. De “kansarmoedegraad” die in kolom 3 van tabel 4.4 aangegeven wordt, is verschillend van de “kansarmoedegraad” die in tabel 3.7 voor wiskunde aangegeven wordt. De groepscompositie-effecten waarvan die “kansarmoedegraad” is afgeleid, zijn immers anders per leergebied.

In de twee zeer kansarme scholen is de feitelijke gemiddelde extra achterstand einde L1 (vierde kolom) al sterk verschillend, ook al is in beide scholen de voor technisch lezen te verwachten extra achterstand nagenoeg even groot. In school 1 is die feitelijke extra achterstand groter, in school 2 is ze kleiner dan wat op grond van de samenstelling van de leerlingengroep te verwachten viel. Ook voor de twee modale scholen is er een (groot) verschil, zowel tussen de twee scholen voor wat betreft de feitelijke extra achterstand einde L1, als in elk van beide scholen tussen de waarde in kolom 4 en de waarde in kolom 3. Dat toont heel duidelijk dat op grond van de groepssamenstelling over individuele scholen geen uitspraak gedaan kan worden over waar de scholen tegen het einde van het eerste leerjaar voor technisch lezen zullen staan (na correctie voor de *individuele* instroomkenmerken).

Tabel 4.4 illustreert hoe de verschillen in gemiddelde prestaties voor technisch lezen tussen scholen met een gelijkaardige samenstelling van de leerlingengroep nog verder aangedikt worden in de loop

van het tweede leerjaar. De leerlingen in zeer kansarme school 1 vertonen op het einde van L2 gemiddeld een extra achterstand die nagenoeg overeenkomt met wat een jaar eerder op grond van de samenstelling van de leerlingengroep te verwachten viel. Ze scoren gemiddeld 12,5 punten lager dan op grond van hun *individuele* instroomkenmerken verwacht mocht worden. Zeer kansarme school 2 echter heeft in de loop van het tweede leerjaar de extra achterstand volledig gecompenseerd. Op het einde van L2 staan de leerlingen in school 2 gemiddeld waar ze op grond van hun individuele achtergrond verwacht mochten worden te staan. Het verschil dat tegen het einde van het tweede leerjaar tussen scholen 3 en 4, twee “witte” scholen met een beperkt aantal GOKleerlingen, tot stand is gekomen, is nog veel groter. Op het einde van L2 liggen de leerlingen in school 3 voor technisch lezen gemiddeld bijna een half jaar achter op de leerlingen in school 4.

Opnieuw moet voor één van de vier scholen voorzichtigheidshalve voorlopig nog enig voorbehoud uitgesproken worden n.a.v. de grote uitstroom van leerlingen bij de overgang van het eerste naar het tweede leerjaar. Anders dan in tabel 3.7 gaat het hier nu echter niet om een zeer kansarme school die een relatief grote TW realiseert, maar om een zeer kansarme school (school 1) die in de loop van L2 een TW realiseert die kleiner is dan die van zeer kansarme school 2 die maar een beperkte uitstroom heeft. Zoals in hoofdstuk 3 uitgelegd (p. 77), mogen we in principe aannemen dat verschillen in uitstroom de schatting van de TW van een school niet beïnvloeden, aangenomen dat uitstromers nadat ze de school verlaten hebben, gemiddeld niet anders evolueren dan doorstromers met dezelfde instroomkenmerken en een gelijke vroegere leerprestaties in dezelfde school of in gelijkaardige scholen.

4.6 Besluit

De bevindingen voor het leergebied technisch lezen kunnen als volgt samengevat worden:

1. Niveau en leerwinst in technisch lezen worden door diverse kind- en achtergrondfactoren beïnvloed. Het niveau einde eerste leerjaar én de leerwinst in het tweede leerjaar liggen lager voor leerlingen met schoolse vertraging; wie het eerste leerjaar heeft moeten overdoen, kijkt tegen het einde van het tweede leerjaar toch al weer tegen een significante achterstand aan. Er is - vooral qua niveau, minder inzake leerwinst - een samenhang met het aanvankelijke Nederlandse taalvaardigheidsniveau van de leerlingen. Tussen jongens en meisjes zijn er geen significante verschillen, niet op het niveau einde eerste leerjaar en evenmin inzake de leerwinst in het tweede leerjaar.
2. De grootste verschillen in vaardigheidsniveau en leerwinst voor technisch lezen treffen we echter aan tussen leerlingen met een verschillende sociale en etnisch-culturele achtergrond. Die verschillen veranderen weinig in de loop van het tweede leerjaar. De grootste kloof met de modale Vlaamse leerlingen vinden we voor de Turkse GOK-leerlingen. Daarna komen de

Arabisch of Berber sprekende GOK-leerlingen, de overige anderstalige GOK-leerlingen en de Nederlandstalige GOK-leerlingen. Tot zover is het beeld voor de kloof inzake technisch lezen niet zo verschillend van wat voor wiskunde vastgesteld werd. Een constante is de zeer precaire achterstandspositie van de Turkse GOK-leerlingen. Een verschil is er wel in de omvang van de kloof met modale Vlaamse leerlingen. Die is voor alle categorieën van leerlingen veel kleiner voor technisch lezen dan voor wiskunde. Opmerkelijk is dat anderstalige niet-GOK-leerlingen het voor technisch lezen gemiddeld bijna even goed doen als modale Vlaamse leerlingen.

3. De achterstand van GOK-leerlingen voor technisch lezen wordt in zeer belangrijke mate (d.i. voor 57% tot 100%) door verschillen in aanvankelijke Nederlandse taalvaardigheid verklaard. Gezien de onderlinge posities van de verschillende GOK-categorieën nauwelijks veranderen, wijzigt ook de samenhang met aanvankelijke Nederlandse taalvaardigheid nauwelijks. Of dit betekent dat die aanvankelijke Nederlandse taalvaardigheid als een persistente factor in de achterstand voor technisch lezen beschouwd moet worden, is daarmee geen uitgemaakte zaak. Er zou ook sprake kunnen zijn van een meer algemene, achterliggende factor die zowel de aanvankelijke Nederlandse taalvaardigheid als latere prestaties voor technisch lezen beïnvloedt.
4. Tussen scholen blijken er voor technisch lezen grote verschillen in gemiddelde niveau einde eerste leerjaar en gemiddelde leerwinst in het tweede leerjaar te bestaan. Een groot deel van die verschillen kan door de effecten van instroomkenmerken verklaard worden. De resterende variantie is echter nog altijd significant. Ook na correctie voor individuele instroomkenmerken blijken basisscholen onderling nog in aanzienlijke mate significant te verschillen. Dat geldt zowel voor het gemiddelde niveau dat op het einde van het eerste leerjaar bereikt wordt als voor de leerwinst die in de loop van het tweede leerjaar gemiddeld geboekt wordt.
5. Tussen einde eerste en einde tweede leerjaar blijft de variantie tussen scholen (na correctie voor de effecten van instroomkenmerken) ongeveer gelijk. Dat ondanks de verschillen in toegevoegde waarde met betrekking tot de leerwinst in het tweede leerjaar de verschillen tussen scholen op het einde van het tweede leerjaar niet groter zijn dan op het einde van het eerste leerjaar, heeft te maken met de negatieve samenhang tussen de toegevoegde waarde met betrekking tot de leerwinst in het tweede leerjaar en de toegevoegde waarde met betrekking tot niveau voor technisch lezen op het einde van het eerste leerjaar. Scholen met een hogere toegevoegde waarde einde eerste leerjaar realiseren vaker een lagere toegevoegde waarde in het tweede leerjaar en omgekeerd.

6. De verschillen tussen scholen zijn voor technisch lezen niet alleen significant, ze zijn ook behoorlijk groot (zij het niet zo groot als voor wiskunde). Na correctie voor instroomkenmerken komt het verschil in gemiddeld niveau op het einde van het eerste leerjaar tussen een school op deciel 1 (de laagste 10%) en een school op deciel 9 (de hoogste 10%) overeen met wat in de Vlaamse referentiegroep voor technisch lezen gemiddeld in één trimester aan leerwinst geboekt wordt. Dat verschil is tegen het einde van het tweede leerjaar niet gewijzigd.
7. Die verschillen in gecorrigeerd niveau einde eerste leerjaar kunnen wellicht grotendeels aan verschillen in onderwijseffectiviteit toegeschreven worden. De prestaties voor technisch lezen op het einde van het eerste leerjaar vormen immers geen “beginniveau”, maar zijn (onder meer) het resultaat van een schooljaar lang onderwijs in technisch lezen.
8. Hoewel statistisch significante groepscompositie-effecten gevonden worden, blijkt de samenstelling van de SiBO-leerlingencohorte (in termen van het aandeel leerlingen uit elke GOK-categorie) maar een beperkte verklaring te bieden voor de verschillen die tussen scholen - ook na correctie voor de effecten van individuele instroomkenmerken - inzake niveau voor technisch lezen nog blijken voor te komen. Binnen alle types van scholen, zowel “kansarme scholen” (met relatief veel Vlaamse GOK-leerlingen) en “zeer kansarme” scholen (met veel anderstalige GOK-leerlingen), als bij “modale scholen” (met vooral veel modale Vlaamse leerlingen) en “kansrijke scholen” (met relatief veel anderstalige niet-GOK-leerlingen) zijn er tussen scholen grote, niet door individuele instroomkenmerken verklaarde verschillen in gemiddelde score voor technisch lezen. Dat geldt zowel voor het einde van het eerste als voor het einde van het tweede leerjaar. In zoverre groepscompositie een effect heeft op het niveau en de vooruitgang die leerlingen boeken, blijken sommige scholen daar dan beter voor te kunnen compenseren dan andere. Op de vraag welke “beschermende” resp. “bezwarende factoren” hieraan ten grondslag liggen, kan op grond van de uitgevoerde analyses niet bepaald worden.
9. Een algemene vaststelling is dat de effectiviteit van een school voor technisch lezen (in termen van de toegevoegde waarde met betrekking tot de leerwinst in het tweede leerjaar) niet samenhangt met de samenstelling van de leerlingencohorte die in het SiBO-onderzoek gevolgd wordt. Zowel onder de kansarme scholen als onder de zeer kansarme scholen treffen we voor wiskunde evengoed effectieve en zeer effectieve scholen aan als onder de modale scholen. Ook voor wat de leerwinst voor technisch lezen in de beginjaren van het lager onderwijs betreft, zijn er bijgevolg geen argumenten die pleiten voor een dringende etnische desegregatie (lees: opdoeken van concentratiescholen). De voorliggende resultaten laten niet toe te concluderen dat een gelijkmatige spreiding van allochtone leerlingen over alle basisscholen binnen een regio automatisch zou leiden tot een verbetering van de

leerwinst van de leerlingen. Er dient eerder gezocht te worden naar de factoren die de verschillen in effectiviteit tussen scholen kunnen helpen verklaren. Dat geldt zowel voor de kansarme en de zeer kansarme (“zwarte”) scholen als voor de modale (“witte”) scholen.

10. Tussen het einde van het eerste leerjaar en het einde van het tweede leerjaar neemt de samenhang tussen samenstelling van de gevolgde leerlingengroep en het gecorrigeerde schoolgemiddelde voor technisch lezen af. Dat heeft enerzijds te maken met de voormelde vaststelling dat er ook kansarme en zeer kansarme scholen zijn die met betrekking tot de leerwinst voor technisch lezen in het tweede leerjaar hoge toegevoegde waarden laten optekenen en anderzijds met het vóórkomen van *omgekeerde* groepscompositie-effecten. We spreken van *omgekeerd* omdat een hoger aandeel van leerlingen met een ongunstig kenmerk (i.c. een hoger aandeel Nederlandstalige GOK-leerlingen) gepaard gaat met een gemiddeld hogere leerwinst i.p.v. met een gemiddeld lagere leerwinst.
11. Met betrekking tot de GOK-categorieën is voor technisch lezen geen sprake van differentiële schooleffecten. De toegevoegde waarde die scholen realiseren, realiseren ze in gelijke mate voor alle GOK-categorieën. Dat kan opgevat worden als een aanwijzing dat (althans voor het leergebied technisch lezen) goed onderwijs de belangrijkste basis voor een zorgbreed onderwijs is. Deze aanwijzing dient uiteraard verder getoetst te worden aan de hand van analyse van de samenhang met procesvariabelen.

5. Leerwinst en toegevoegde waarde voor spelling

Bij de leerlingen in het SiBO-onderzoek werd op het einde van het eerste leerjaar en op het einde van het tweede leerjaar een dictee uit de Schaal Vorderingen Spellingvaardigheid (SVS) afgenomen. De SVS werd oorspronkelijk door het Nederlandse Centraal Instituut voor Toetsontwikkeling (CiTO) ontwikkeld en later in opdracht van de Vlaamse overheid voor Vlaanderen aangepast (Rymenans, 2000). We bespreken eerst kort de inhoud van de toetsen en de scoringswijze. Daarna rapporteren we de resultaten van de analyses.

5.1 De Schaal Vorderingen Spellingvaardigheid voor Vlaanderen (SVS-V)

In de SVS-V (Moelands & Rymenans, 2003b) zijn voor elk leerjaar van het lager onderwijs dictees voorzien voor het midden en voor het einde van het schooljaar. Voor elk afnamemoment zijn er twee dictees beschikbaar: een A- en een B-versie. De B-versie is steeds iets moeilijker dan de A-versie.

Voor gebruik in het SiBO-onderzoek werden er aan de SVS-V geen bijkomende aanpassingen gedaan. Op het einde van het eerste leerjaar legden de leerlingen uit de deelnemende scholen één van de beide voorziene dictees af. De dictee-versies werden toegewezen op basis van de gemiddelde school-scores op de Nederlandse taalvaardigheidstoets die bij het begin van het eerste leerjaar was afgenomen. Alle leerlingen van de 20 gemiddeld laagst scorende scholen legden de A-versie af. De leerlingen van de andere scholen kregen de B-versie. Het beperkte aantal scholen zonder toetsgegevens voor het begin van het eerste leerjaar kreeg de B-versie toegestuurd. Oordeelden die scholen dat die B-versie te moeilijk was voor hun leerlingen, dan konden ze alsnog de A-versie opvragen. De A-versie van het dictee einde L1 telt 41 opgaven, de B-versie 39. De afname gebeurde in twee fasen met een kleine pauze tussenin.

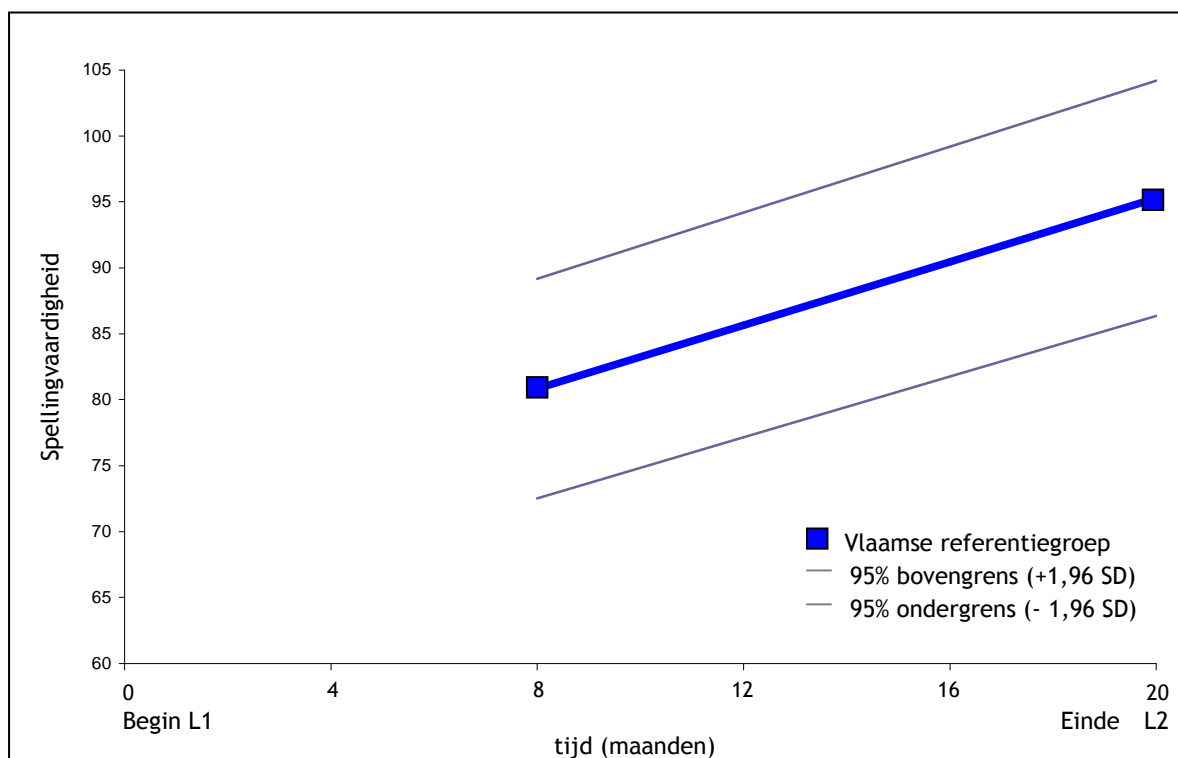
Op het einde van het tweede leerjaar legden alle leerlingen de moeilijker B-versie af. Het dictee bestaat uit één- en tweelettergrepige woorden, die in zinsverband worden aangeboden. Verschillende spellingproblemen komen in het dictee aan bod: onder andere éénlettergrepige woorden met -a, -o, of -u (vb. stro), tweelettergrepige woorden met open lettergreep (vb. dagen), één- of meerlettergrepige woorden met -b (vb. web). De toets einde tweede leerjaar telt 40 opgaven en werd in twee delen afgenomen met een korte pauze. Gegevens over de verdelingskenmerken van de SVS-V bij de SiBO-steekproef vindt men in LOA-rapport 30 (Verachtert e.a., 2005) en LOA-rapport 35 (Hendriks e.a., 2005).

Met behulp van de omzettingstabellen in de handleiding van de SVS-V (Moelands & Rymenans, 2003b) werden de ruwe scores omgezet in een vaardigheidsscore. Die omzettingstabellen zijn gebaseerd op de resultaten van een IRT-calibratie die in het kader van de aanpassing van de SVS voor Vlaanderen uitgevoerd werd (Rymenans, 2000). Door de omzetting naar vaardigheidsscores komen alle scores

(van de verschillende afnamemomenten en van de verschillende versies) op één meetschaal. De vaardigheidsscores van leerlingen die de A-versie aflegden zijn daardoor direct vergelijkbaar met de vaardigheidsscores van de leerlingen die de B-versie aflegden. En de vaardigheidsscores voor eind tweede leerjaar zijn direct vergelijkbaar met de vaardigheidsscores voor eind eerste leerjaar. De vooruitgang die de leerlingen voor spelling maken, kan bijgevolg gemakkelijk bepaald worden door het verschil te nemen tussen de vaardigheidsscore voor het tweede leerjaar en de vaardigheidsscore voor het eerste leerjaar.

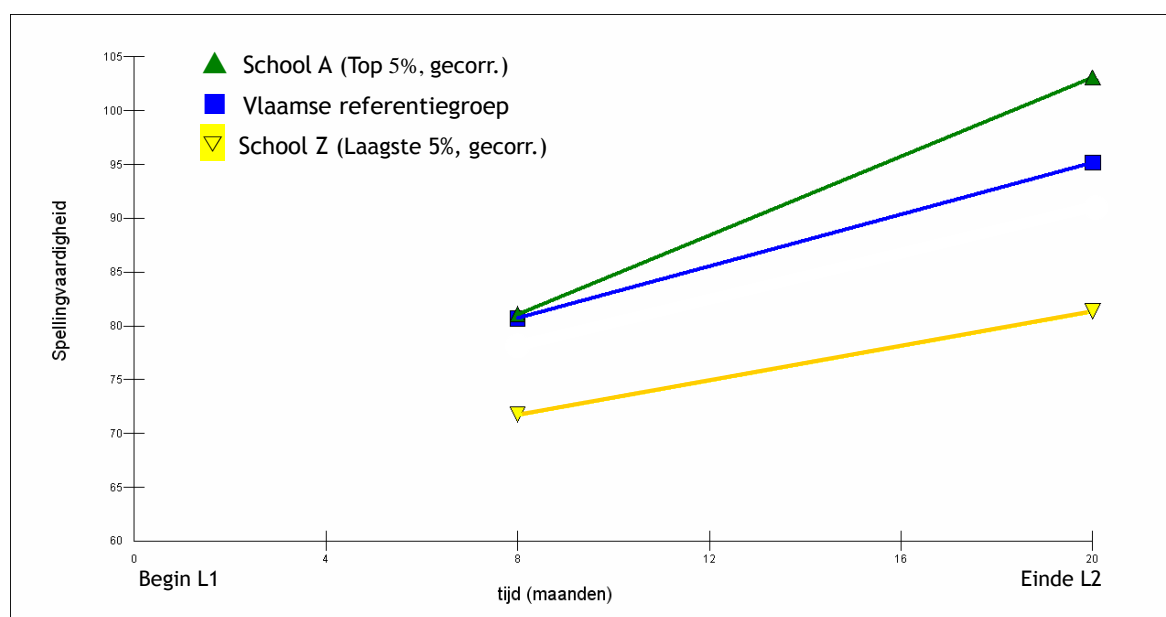
5.2 Algemene evolutie van einde L1 naar einde L2

Bij de analyse van de algemene evolutie voor spelling werd gecorrigeerd voor de effecten van sekse, geboortjaar en geboortemaand van de leerlingen, het al dan niet blijven zitten in L1 of L2, GOK-categorie en score voor aanvankelijke Nederlandse taalvaardigheid (zie inleiding). Figuur 5.1 toont de gemiddelde groeicurve van de *Vlaamse referentiegroep* (zie p. 12). Dat is de groeicurve waarmee de gecorrigeerde gemiddelde groeicurves van afzonderlijke scholen vergeleken kunnen worden. Om de gelijkens met de grafieken met de groeicurves voor wiskunde te behouden én om ook grafisch duidelijk te maken dat de eerste score voor technisch lezen geen echte “beginscore” is (maar het resultaat van onderwijs en leren in het eerste leerjaar), loopt de horizontale as van begin L1 tot einde L2.



Figuur 5.1 Gemiddelde evolutie voor spelling van einde eerste leerjaar naar einde tweede leerjaar in de Vlaamse referentiegroep (+/- 1,96 x standaarddeviatie op schoolniveau)

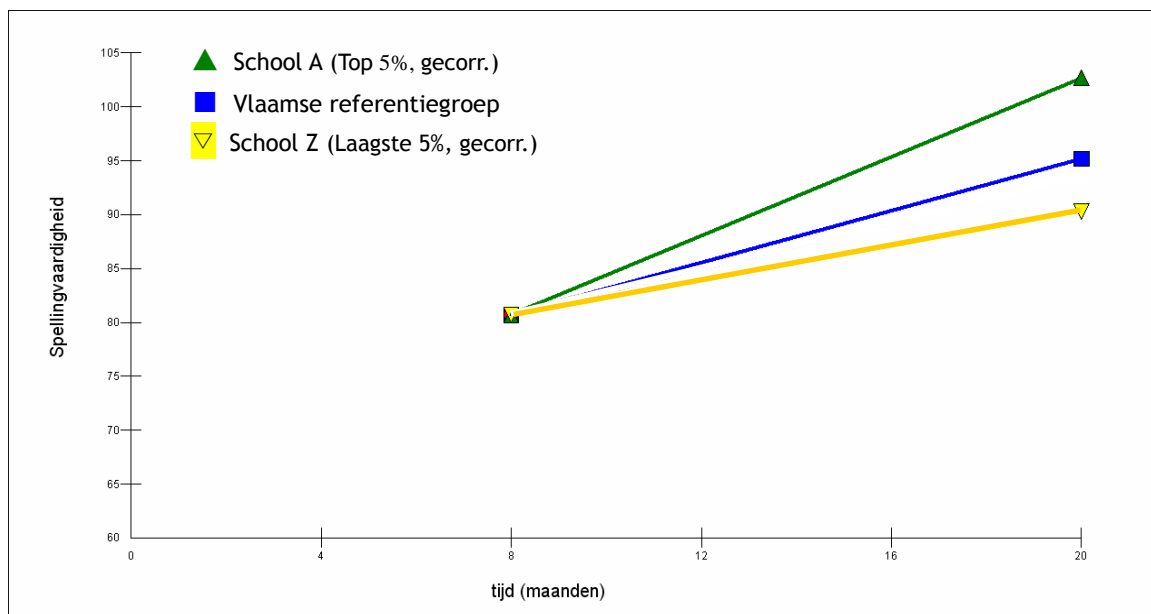
De twee dunnere lijnen boven de gemiddelde groeicurve geven de grenzen aan waarbinnen zich (uitgaande van een normale verdeling) 90% van de schoolgemiddelden bevinden. Die boven- en ondergrens zijn afgeleid van de varianties op schoolniveau op elk van de toetsmomenten en hun onderlinge covarianties. Tussen einde eerste leerjaar en einde tweede leerjaar blijft de variantie tussen scholen nagenoeg constant ($\Delta\text{var} = 2,678$; $\chi^2_{(1)} = 0,492$; ns). In figuur 5.2 geven de gecorrigeerde gemiddelde groeicurves van een school die qua toegevoegde waarde (TW) tot de top 5 behoort (school A) resp. een school die qua TW tot de laagste vijf behoort (school Z) op meer concrete wijze een idee van de variatie binnen het Vlaamse basisonderwijs. De scholen A en Z in figuur 5.2 zijn niet dezelfde scholen als in figuren 3.3 of 4.2.



Figuur 5.2 Gemiddelde evolutie voor spellingvaardigheid van einde eerste leerjaar naar einde tweede leerjaar in de Vlaamse referentiegroep, in een school met een hoge en in een school met een lage toegevoegde waarde

In figuur 5.3 laten we de “beginpunten” (eigenlijk het niveau bij het einde van L1) van de curves samenvallen. Dat maakt het makkelijker om de verschillen in hellingsgraad van de groeicurven te zien.

Tabel 5.1 geeft de cijfers die bij figuur 5.2 horen. Een sterretje wijst op significant verschil met het gemiddelde van de Vlaamse referentiegroep. Net als tabel 4.1 vermeldt tabel 5.1 een “toegevoegde waarde” bij het gemiddelde niveau einde L1. De getallen in die kolom geven aan hoeveel hoger (+) of hoeveel lager (-) het gecorrigeerde gemiddelde einde L1 is dan het gemiddelde van de Vlaamse referentiegroep. Of ook nog: hoeveel hoger of lager scholen A en Z gemiddeld scoren dan op grond van de instroomkenmerken van hun leerlingen verwacht kon worden.



Figuur 5.3 Gemiddelde evolutie in spellingvaardigheid van einde eerste leerjaar naar einde tweede leerjaar in Vlaamse referentiegroep, in een zeer hoog en in een zeer laag scorende school, na correctie voor verschillen in beginniveau

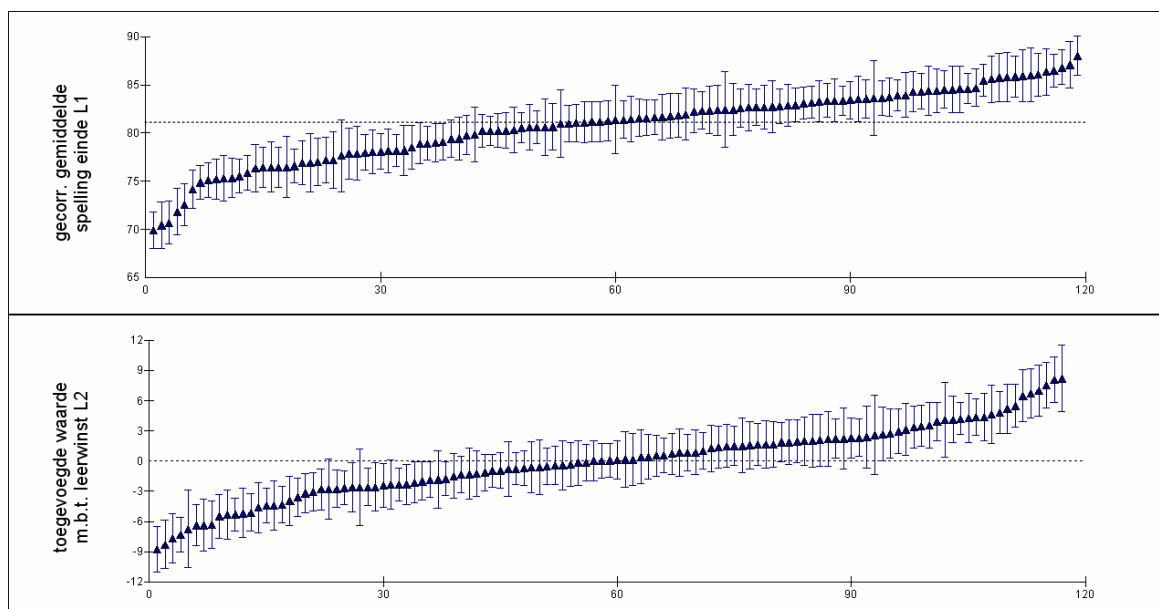
Net als voor technisch lezen mogen we voor spelling aannemen dat normaalvorderende leerlingen in het Vlaamse basisonderwijs het eerste leerjaar beginnen zonder noemenswaardige spellingvaardigheid. Het niveau dat op het einde van het eerste leerjaar bereikt wordt, kan bijgevolg in principe als “leerwinst” beschouwd worden. Maar anders dan de DMT-score werd de SVS-score verkregen door IRT-calibratie. In de meetschaal die daaruit resulteert, hebben beginpunt en meeteenheid geen absolute betekenis. Waar het punt op de SVS-meetschaal ligt dat duidt op “totaal geen Nederlandse woordjes correct kunnen spellen” valt niet te zeggen. (Het kan net zo goed op score 0 als op score 20 of zelfs score -10 liggen.) Daarom kan de spellingvaardigheidsscore die het niveau einde L1 aangeeft, niet geïnterpreteerd worden als een maat voor de vooruitgang of de leerwinst die in de loop van het eerste leerjaar gemaakt werd.

Tabel 5.1 Gemiddeld niveau einde L1, gemiddelde leerwinst en toegevoegde waarde in L2 voor spelling

	Eerste leerjaar		Tweede leerjaar	
	Gecorrig. gemidd. niveau einde L1	Toegevoegde waarde	Gecorrig. gemidd. leerwinst	Toegevoegde waarde
Vlaamse referentiegroep	80,84	0,00	14,46	0,00
School A	81,11	0,27	21,98 *	7,54 *
School Z	71,75 *	-9,09 *	9,60	-4,85

In tabel 5.1 kan men verschillende significante waarden zien. Onder meer de gemiddelde achterstand van school Z (tegenover de Vlaamse referentiegroep) op het einde van L1 is significant. Maar hoe groot is die gemiddelde achterstand? De enige praktische vergelijkingsbasis die tabel 5.1 ons biedt, is de leerwinst die in de loop van het *tweede* leerjaar gemiddeld in de Vlaamse referentiegroep gemaakt wordt (14,46 punten). De 9,09 punten gemiddelde achterstand in school Z vormt 63 % van die gemiddelde leerwinst voor spelling in L2. We kunnen dus zeggen dat de leerlingen uit school Z - na correctie voor instroomkenmerken - op het einde van het eerste leerjaar gemiddeld een achterstand van bijna twee trimesters hebben. Tegen het einde van het tweede leerjaar is die niet door de individuele instroomkenmerken te verklaren achterstand opgelopen tot gemiddeld 13,94 punten ($-9,09 - 4,85 = -13,94$; dit is 96 % van de gemiddelde leerwinst in L2) of bijna een vol schooljaar. School A heeft tegen het eind van het tweede leerjaar een voorsprong van ruim een half schooljaar opgebouwd (tegenover de Vlaamse referentiegroep). De voorsprong van school A tegenover school Z bedraagt op het einde van het tweede leerjaar precies anderhalf schooljaar. Dit allemaal na correctie voor instroomkenmerken.

Kijken we naar het verschil tussen de 10% scholen met het hoogste gecorrigeerde gemiddelde (het 9^e deciel om precies te zijn) en de 10% scholen met het laagste gecorrigeerde gemiddelde (het 1^e deciel), dan zien we op het einde van L1 een verschil van 10,1 punten en op het einde van L2 een verschil van 9,7 punten. Die verschillen komen overeen met met de leerwinst die voor spelling gemiddeld in iets meer dan twee trimesters wordt gerealiseerd.



Figuur 5.4 Verschillen tussen scholen (referentiesteekproef) voor gecorrigeerd gemiddelde voor spelling einde L1 en voor toegevoegde waarde m.b.t. de leerwinst voor spelling in L2

Figuur 5.4 toont de scholen uit de referentiesteekproef geordend naar het gecorrigeerde niveau voor spelling einde L1 (boven) resp. de toegevoegde waarde met betrekking tot de leerwinst voor spelling

in het tweede leerjaar (onder). De verticale staafjes geven het betrouwbaarheidsinterval weer. Dat is zo gekozen ($CI = 1,39 \times SE$) dat voor elke twee scholen waarvan de staafjes elkaar niet overlappen, geldt dat ze significant van elkaar verschillen. Voor beide vergelijkingspunten (gecorrigeerd niveau einde L1 en TW L2) zijn er behoorlijk veel scholen die significant verschillen van andere scholen. Er is een grote, van elkaar significant verschillende kop- en staartgroep te onderscheiden die elk (ruim) een kwart van de scholen telt.

De samenhang tussen het niveau einde L1 en de toegevoegde waarde in L2

Scholen met een hoger gecorrigeerd gemiddelde einde L1 halen eerder een lagere TW met betrekking tot de leerwinst in L2 en omgekeerd ($r_{(117)} = -0,476$; $p < 0,001$). Dat belet niet dat de correlatie tussen het gecorrigeerde gemiddelde einde L1 en het gecorrigeerde gemiddelde einde L2 toch nog positief is ($r_{(117)} = 0,547$; $p < 0,001$).

5.3 Evolutie voor de verschillende GOK-categorieën

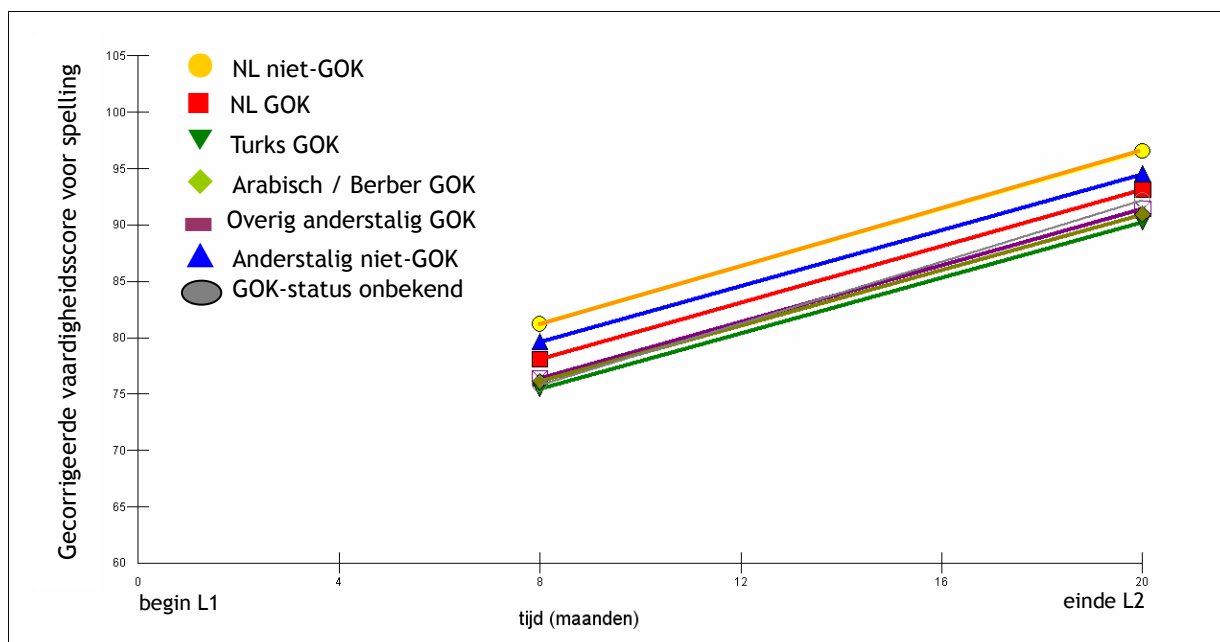
Zoals in hoofdstuk 3 werd uitgelegd, werd bij de analyse van de verschillen tussen de GOK-categorieën niet gecorrigeerd voor de effecten van aanvankelijke Nederlandse taalvaardigheid, wel voor de effecten van de overige individuele instroomkenmerken. De achterstand die de onderscheiden categorieën van leerlingen ook voor spelling gemiddeld t.a.v. modale Vlaamse leerlingen hebben, hangt voor een belangrijk deel samen met hun aanvankelijke Nederlandse taalvaardigheid. Die samenhang verschilt sterk van categorie tot categorie (zie tabel 5.2), maar is steeds aanzienlijk te noemen. Correctie voor de effecten van aanvankelijke Nederlandse taalvaardigheid zou daarom een sterk vertekend beeld geven van de werkelijke achterstand van de onderscheiden categorieën GOK-leerlingen.

Tabel 5.2 Deel van achterstand voor spelling (einde L1) dat samenhangt met een minder goede aanvankelijke Nederlandse taalvaardigheid, per GOK-categorie

<i>GOK-categorie</i>	<i>Achterstand verklaard door TvK-score</i>
Nederlandstalig GOK	44,6 %
Turks GOK	85,3 %
Arabisch/Berber GOK	51,7 %
Overige anderstalig GOK	62,6 %
Anderstalig niet-GOK	92,0 %
GOK-status onbekend	46,7 %

Figuur 5.5 laat zien dat na die correctie de verschillende GOK-categorieën in het Vlaamse basis-onderwijs een duidelijke (zij het niet zo heel grote) achterstand hebben t.o.v. de modale Vlaamse leerlingen. Die achterstand neemt in de loop van het tweede leerjaar in heel minieme mate toe. Er

kunnen grosso modo vier posities onderscheiden worden. Bovenaan bevinden zich de modale Vlaamse leerlingen, onderaan de anderstalige GOK-leerlingen. De verschillen tussen de onderscheiden taalgroepen (Turks, Arabisch of Berber, overige anderstaligen) zijn beperkt. Halverwege die twee posities vinden we de Nederlandstalige GOK-leerlingen. Halverwege de Nederlandstalige GOK-leerlingen en de modale Vlaamse leerlingen, treffen we de anderstalige niet-GOK-leerlingen. De onderlinge posities van de verschillende GOK-categorieën veranderen niet in de loop van L2.



Figuur 5.5 Gemiddelde evolutie voor spelling van einde L1 naar einde L2 voor verschillende GOK-categorieën, na correctie individuele instroomkenmerken behalve aanvankelijke Nederlandse taalvaardigheid

Tabel 5.3 geeft voor beide meetmomenten een overzicht van de achterstand van de verschillende GOK-categorieën t.o.v. de referentiecategorie, uitgedrukt als een percentage van de leerwinst die modale Vlaamse leerlingen gemiddeld in de loop van L2 voor spelling boeken.

Tabel 5.3 Achterstand voor spelling van de verschillende GOK-categorieën t.o.v. modale Vlaamse leerlingen¹

GOK-categorie	Toetsmoment	
	Einde L1	Einde L2
Nederlandstalig GOK	-20,3%	-22,3%
Turks GOK	-37,6%	-40,9%
Arabisch/Berber GOK	-33,3%	-36,4%
Overige anderstalig GOK	-31,5%	-33,2%
Anderstalig niet-GOK	-10,1%	-13,2%
GOK-status onbekend	-35,2%	-28,3%

¹Achterstand uitgedrukt als percentage van de gemiddelde leerwinst van de Vlaamse referentiegroep in L2

Met betrekking tot de GOK-categorieën werden voor spelling geen differentiële schooleffecten vastgesteld. Dat betekent dat het patroon van verschillen tussen de GOK-categorieën dat in tabel 4.3 te zien is, op enkele beperkte afwijkingen na, in alle basisscholen waargenomen wordt.

5.4 Evolutie van de zittenblijvers

Leerlingen die het eerste leerjaar overdoen, eindigen hun “bis-jaar” met een spellingvaardigheidscore die gemiddeld significant hoger ligt dan bij andere leerlingen met evenveel (maar vóór het eerste leerjaar opgelopen) schoolse vertraging, tenminste voor zover alle andere individuele instroomkenmerken behalve aanvankelijke Nederlandse taalvaardigheid gelijk gehouden worden. Bij vergelijking met normaalvorderende leerlingen moet echter ook de achterstand in rekening gebracht worden die samenhangt met schoolse vertraging (zie figuur 2.6 in hoofdstuk 2). Voor zittenblijvers L1 met alles samen slechts één jaar vertraging, ligt het niveau voor spelling op het einde van hun “bis-jaar” (na voornoemde correctie) uiteindelijk gemiddeld 0,7 punten lager dan dat van normaalvorderende leerlingen. Dat komt neer op een achterstand van goed anderhalve week. Tegen het einde van L2 is die beperkte achterstand al weer flink opgelopen. Ex-zittenblijvers L1 met één jaar schoolse vertraging (= enkel het jaar van hun zittenblijven) liggen tegen het einde van het tweede leerjaar gemiddeld 13 weken achter op hun normaalvorderende klasgenoten. Ex-zittenblijvers L1 met meer jaren vertraging kennen tegen einde L2 voor spelling gemiddeld ongeveer dezelfde mate van achterstand. Dat is het algemene patroon in Vlaanderen.

Met betrekking tot het geboortjaar werd voor de spellingscore einde L1 een differentieel school-effect gevonden. Dat werkt uiteraard ook door in de relatieve positie van zittenblijvers (= de positie van zittenblijvers t.o.v. normaalvorderende leerlingen in hun eigen school). Sommige scholen realiseren t.a.v. de leerlingen met vertraging (waaronder ook zittenblijvers L1) een TW die hoger ligt dan de TW die ze t.a.v. hun normaalvorderende leerlingen realiseren, bij andere scholen ligt de TW voor de zittenblijvers lager dan de TW voor de normaalvorderende leerlingen. Dat verschil in TW varieert van -4,69 (lagere TW voor zittenblijvers) tot 5,08 (hogere TW voor de zittenblijvers). Maar dat zijn de twee extreme gevallen uit de groep van 34 scholen met minstens vijf zittenblijvers in het eerste leerjaar (schooljaar 2003-2004) met een SVS-score op het einde van L1. Tabel 5.5 toont de cijfers voor twee minder extreme scholen. In school 1 ligt de TW t.a.v. zittenblijvers L1 met één jaar vertraging 2,53 lager dan voor normaalvorderende leerlingen, in school 2 ligt de TW t.a.v. zittenblijvers met één jaar vertraging 3,97 punten hoger.

Het is niet toevallig de school met de hoogste TW t.a.v. de normaalvorderende leerlingen die een hogere TW voor de zittenblijvers (en andere leerlingen met vertraging) realiseert. In de referentiesteekproef blijkt er een positieve, matig sterke maar toch significante samenhang tussen hoogte van de TW (t.a.v. normaalvorderende leerlingen) en hoogte van de TW t.a.v. zittenblijvers en andere leerlingen met vertraging ($r_{(119)} = 0,378$; $p < 0,001$).

Tabel 5.4 Gemiddeld gecorrigeerd niveau einde L1, gemiddelde gecorrigeerde leerwinst L2 en toegevoegde waarde voor de normaalvorderende leerlingen en voor de zittenblijvers L1 in Vlaanderen en in twee exemplarische scholen

leerling-categorie		Einde eerste leerjaar			Tweede leerjaar		
		Aantal einde L1	Gemiddeld gecorrigeerd niveau	Toegevoegde waarde	Aantal einde L2	Gemiddelde gecorrigeerde leerwinst	Toegevoegde waarde
Vlaanderen	N	3584	81,19	0,00	3404	15,39	0,00
	Z	275	80,46	0,00	243	10,74	0,00
School							
1	N	20	70,31	-10,89 *	21	22,07	6,67
	Z	5	67,04	-13,42 *	2	17,41	6,67
2	N	35	82,74	1,55	27	8,95	-6,44
	Z	7	85,98	5,52	6	2,87	-7,86 *

N = normaalvorderende leerlingen, Z = Zittenblijvers L1 * sign. verschillend van Vlaams gemiddelde voor die categorie

Met betrekking tot de leerwinst in L2 verschilt de TW voor ex-zittenblijvers L1 met één jaar vertraging normaliter niet van de TW voor normaalvorderende leerlingen. Er werd voor de samenhang tussen schoolse vertraging en de leerwinst voor spelling in L2 namelijk geen significant differentieel schooleffect vastgesteld. In het statistisch model werd de TW voor beide categorieën daarom gelijkgezet. Dat we voor school 2 toch een andere TW zien voor de zittenblijvers L1 dan voor de normaalvorderende leerlingen heeft te maken met het gegeven dat die school tot één van de oversamplings behoort. Voor die scholen werd een licht afwijkende samenhang gevonden tussen zittenblijven in L1 en de leerwinst voor spelling in L2.

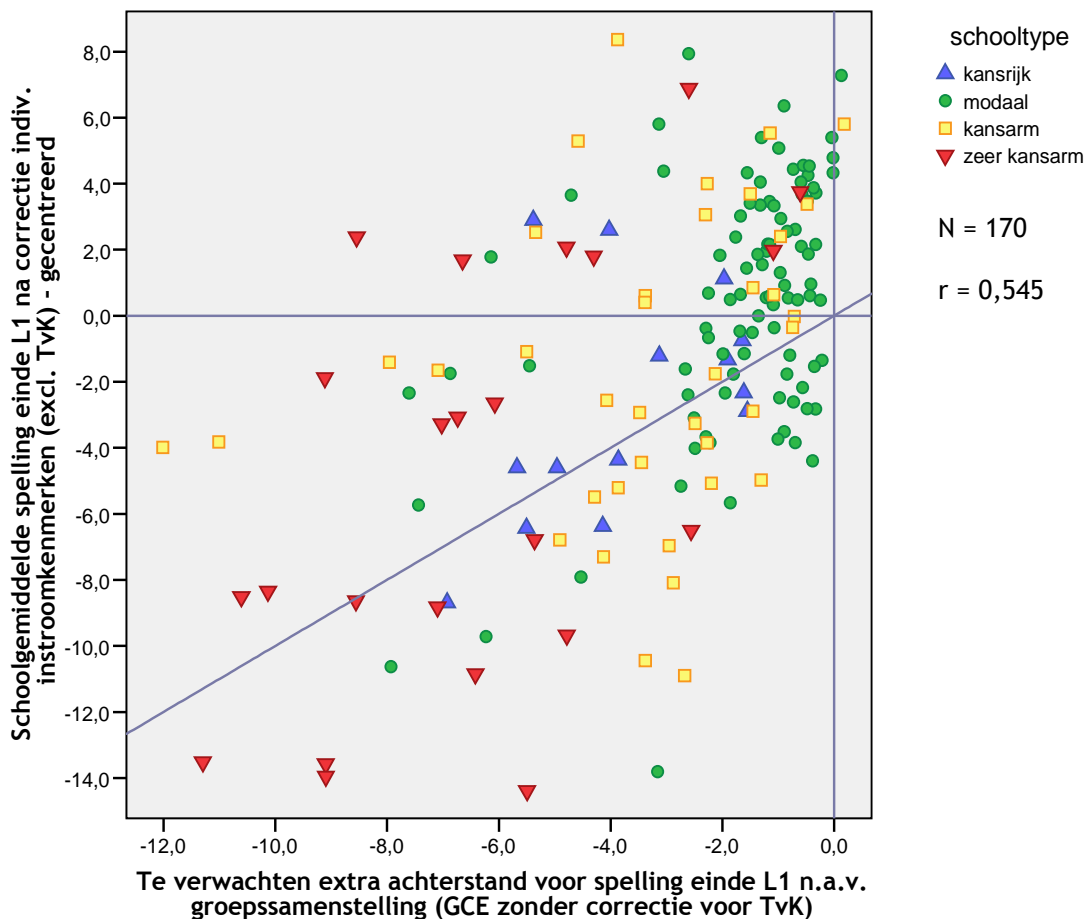
5.5 Maakt de samenstelling van de leerlingengroep een verschil uit?

Groepscompositie-effecten werden op dezelfde wijze onderzocht als voor wiskunde (zie hoofdstuk 3) en voor technisch lezen (zie hoofdstuk 4). T.a.v. de SVS-score einde L1 werden voor Vlaanderen twee groepscompositie-effecten gevonden:

- een negatief GCE met betrekking tot het aandeel “overige anderstalige GOK-leerlingen” (hoe meer van die leerlingen, hoe lager de gemiddelde SVS-score einde L1),
- een negatief GCE met betrekking tot het aandeel leerlingen voor wie de GOK-status onbekend is (hoe meer zulke leerlingen, hoe lager de gemiddelde SVS-score einde L1).

Het eerste GCE is niet gelijk voor iedereen. Het blijkt in veel mindere (uiteindelijk niet meer significante) mate te gelden voor modale Vlaamse leerlingen. Voor hen maakt het m.a.w. niet uit of er in hun leerjaar in hun school veel of weinig overige anderstalige GOK-leerlingen zitten. Van enig GCE t.a.v. de leerwinst voor spelling in L2 was geen sprake.

Figuur 5.6 geeft een beeld van de omvang van het gecumuleerde groepscompositie-effect op de SVS-score op het einde van het eerste leerjaar en van de verschillen op dat vlak tussen de scholen en de hoger onderscheiden schooltypes (kansrijk, modaal, kansarm en zeer kansarm - zie p. 63). Figuur 5.6 toont ook de samenhang tussen dat gecumuleerde GCE einde L1 en de schoolgemiddelden voor spelling einde L1, na correctie voor individuele instroomkenmerken (behalve aanvankelijke Nederlandse taalvaardigheid). Die gemiddelde gecorrigeerde SVS-scores op de verticale as werden gecentreerd rond het Vlaamse gemiddelde en zijn dus te interpreteren als “toegevoegde waarden”. Ze geven per school aan hoeveel hoger of lager het schoolgemiddelde voor SVS uitvalt dan op grond van de instroomkenmerken van de leerlingen verwacht kon worden. Figuur 5.6 is analoog aan figuren 3.12, 3.16 en 4.6 en kan op gelijkaardige wijze geïnterpreteerd worden.



Figuur 5.6 Gecumuleerd groepscompositie-effect op spellingvaardigheid einde L1: verschillen tussen scholen en schooltypes en samenhang met gecorrigeerd schoolgemiddelde einde L1

Zoals ook in hoofdstukken 3 en 4 uiteengezet werd, dient het gecumuleerde GCE opgevat te worden als de extra achterstand die in principe op grond van de groepsamenstelling verwacht kan worden. “Extra” betekent hier: bovenop de achterstand die sommige leerlingen op grond van hun individuele achtergrond al hebben.

Die extra achterstand voor spelling varieert op het einde van het eerste leerjaar tussen 0 en 12 punten, met een gemiddelde van net geen 3 punten. Afgezet tegenover de leerwinst die voor spelling in de loop van L2 gemiddeld in de Vlaamse referentiegroep gemaakt wordt (15,4 wanneer niet gecorrigeerd wordt voor de effecten van aanvankelijke Nederlandse taalvaardigheid), is het gecumuleerd groepscompositie-effect voor spelling einde L1 daarmee qua grootte vergelijkbaar met het gecumuleerde groepscompositie-effect voor wiskunde begin L1 (zie hoofdstuk 3). Vanwege het verschil in tijdstip (einde versus begin L1) loopt de vergelijking tussen het GCE voor wiskunde en het GCE voor spelling uiteraard wat mank.

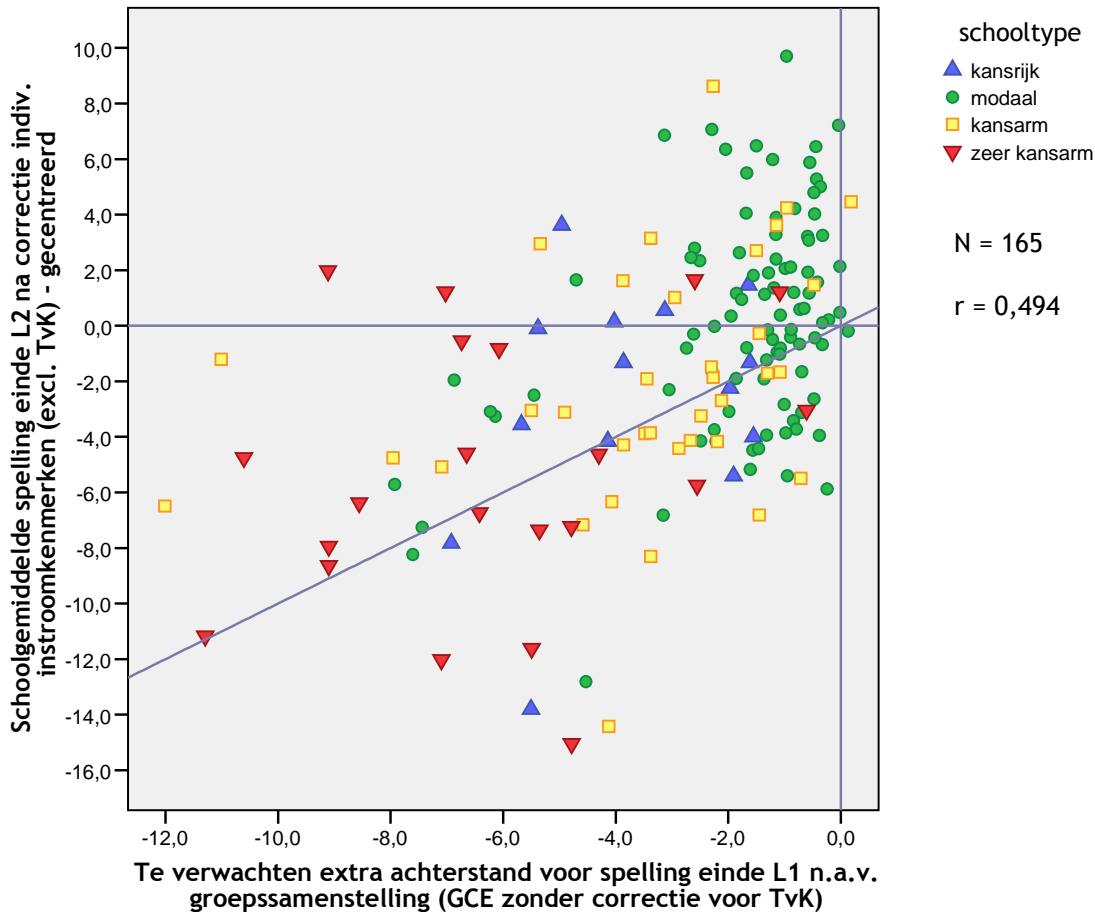
Voor het gros van de scholen is in figuur 5.6 duidelijk een opdeling naar schooltype zichtbaar, met de meeste modale scholen aan de rechterkant (geen tot een zeer klein GCE), de meeste kansarme (én de kansrijke) scholen in het midden (matig GCE) en de meeste zeer kansarme scholen links (sterk GCE) van het spectrum. Sommige scholen doorbreken dat patroon. Dat heeft te maken met een specifiek groepscompositie-effect voor het aandeel anderstalige niet-GOK-leerlingen bij de oversampling.

Wat de samenhang met het gecorrigeerde schoolgemiddelde voor spelling einde L1 betreft, valt in figuur 5.6 op dat voor alle schooltypes al sprake is van een vrij ruime spreiding boven en onder de diagonaal en boven en onder de horizontale middellijn. Net als voor technisch lezen dienen we er hier op te wijzen dat de scores voor spelling geen “beginniveau” weergeven, maar het resultaat vormen van een jaar onderwijs. In die zin is het niet zo gek dat figuur 5.6 (net als figuur 4.6 voor technisch lezen) meer overeenkomsten vertoont met figuur 3.16 (wiskunde einde L2) dan met figuur 3.12 (wiskunde begin L1).

Zeven (van de 23) zeer kansarme scholen en veertien (van de 38) kansarme scholen liggen boven de horizontale middellijn. Zij zijn er tegen einde L1 al volledig in geslaagd te compenseren voor de extra achterstand die samenhangt met het gecumuleerde GCE t.a.v. spelling einde L1 en behalen met hun leerlingen voor spelling zelfs een gemiddelde dat uitsteekt boven wat op grond van de individuele instroomkenmerken verwacht kon worden. Nog eens vijf zeer kansarme en negen kansarme scholen liggen tussen de diagonaal en de middellijn. Zij zijn er tegen einde L1 gedeeltelijk in geslaagd te compenseren voor de te verwachten negatieve effecten van de groepsamenstelling. Voor de kansrijke scholen zien we een zeer gelijkaardige spreiding t.o.v. de verticale as. Van de modale scholen bevinden er zich verhoudingsgewijs méér boven de diagonaal en boven de horizontale middellijn.

Tegen het einde van het tweede leerjaar is de samenhang tussen het gecorrigeerde schoolgemiddelde voor spelling en de groepsamenstelling lichtjes verminderd. De correlatiecoëfficiënt daalt tussen einde L1 en einde L2 van 0,545 naar 0,494. De evolutie in positie ten opzichte van het Vlaamse gemiddelde verschilt (vooral binnen de categorieën kansarme en zeer kansarme scholen) sterk van school tot school. In sommige scholen is de extra achterstand in vergelijking met een jaar

eerder verminderd, maar omgekeerd zijn er ook scholen waarin de extra achterstand in die tijds-
spanne toegenomen is of waarin de relatieve voorsprong afgekalfd is of zelfs omgezet is in een extra
achterstand (zie figuur 5.7).



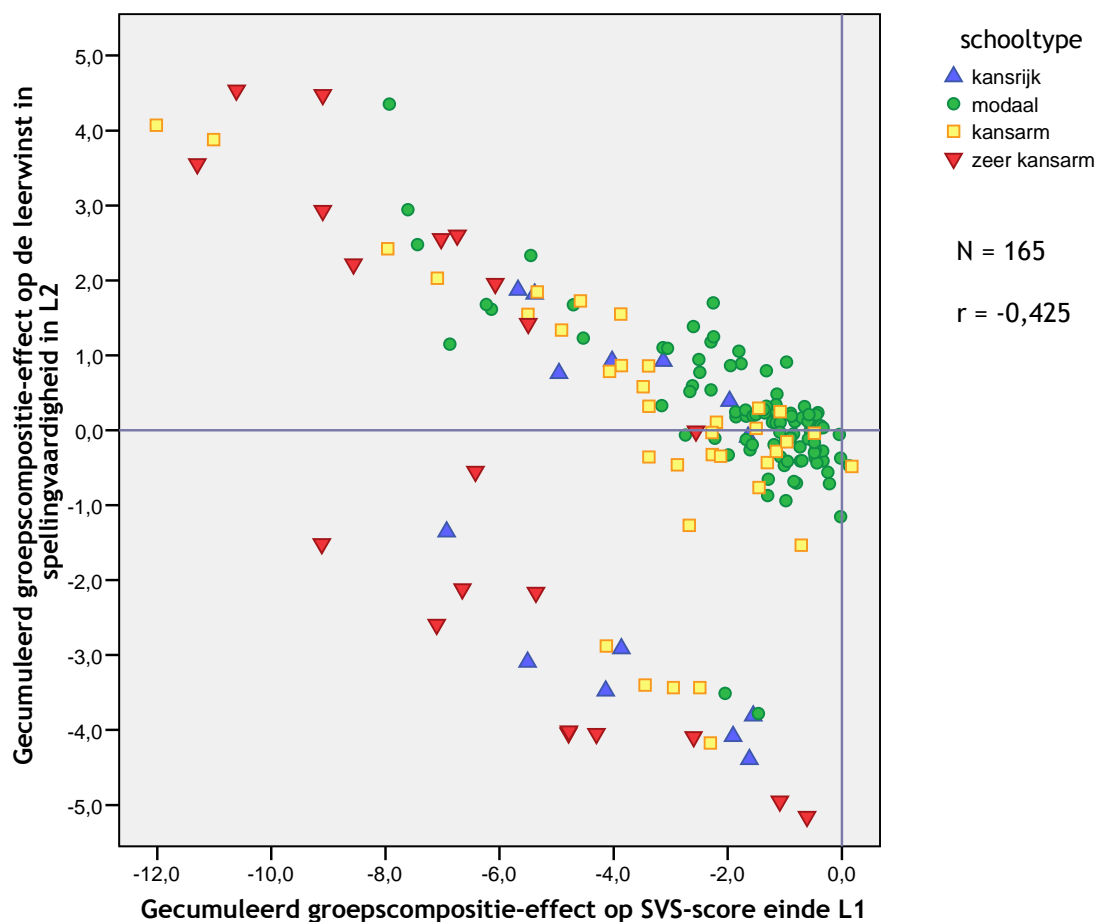
Figuur 5.7 Mate waarin scholen erin slagen tegen einde L2 het groepscompositie-effect t.a.v. spellingvaardigheid van bij einde L1 te neutraliseren, opgesplitst naar schooltype

Om te zien welke factoren ten grondslag liggen aan de gewijzigde posities van scholen kijken we eerst naar het gecumuleerde groepscompositie-effect t.a.v. de leerwinst in L2. We herinneren eraan dat voor Vlaanderen geen van de afzonderlijke GCE t.a.v. de leerwinst in L2 significant was. Specifiek voor de Gentse oversampling werd wel een positief (“omgekeerd”) GCE voor het aandeel anders-talige niet-GOK-leerlingen gevonden.

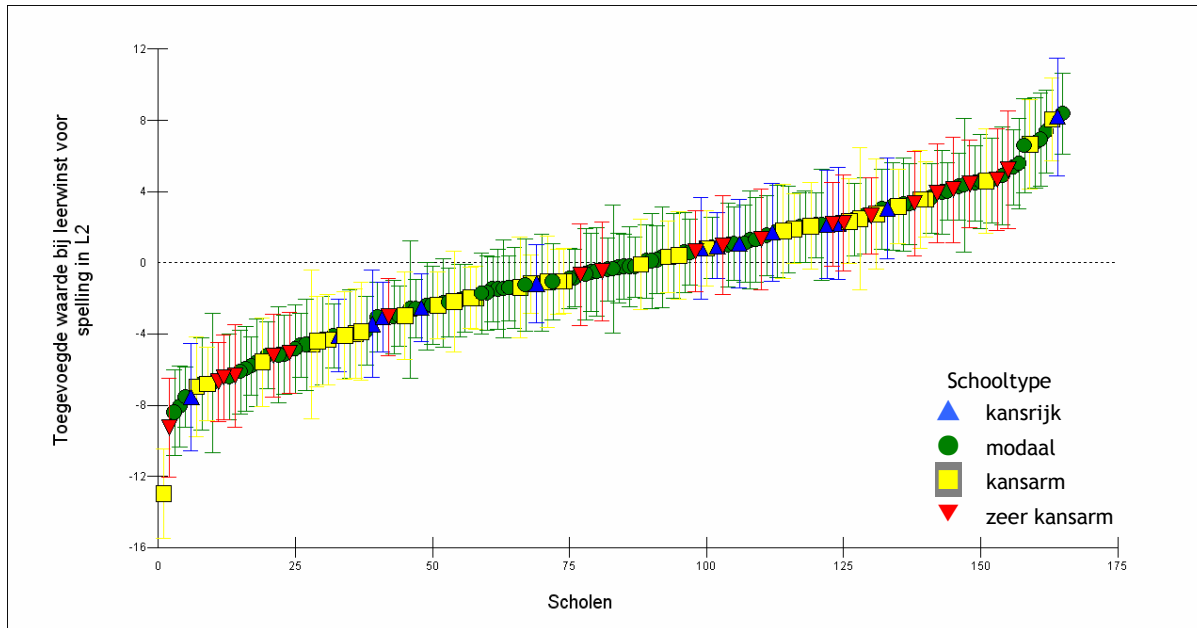
Figuur 5.8 toont voor het leergebied spelling de samenhang tussen het gecumuleerde groepscompositie-effect t.a.v. de leerwinst in L2 en het gecumuleerde groepscompositie-effect einde L1. Die samenhang is negatief (correlatiecoëfficiënt $r = -0,425$). Er is echter een duidelijke tweedeling te zien tussen een grotere reeks scholen bovenaan en een kleinere reeks onderaan. De kleinere reeks onderaan omvat de scholen uit de GOK-oversampling. Bij die groep gaat een gecumuleerde GCE

einde L1 van gemiddeld -4,1 punten gepaard met een gecumuleerd GCE t.a.v. de leerwinst in L2 van gemiddeld -3,3. Voor die scholen stellen we tweemaal een negatief gecumuleerd GCE vast. Bij de groep links bovenaan (waarvan de meeste scholen tot het Gentse stedelijk onderwijs behoren) gaat een gemiddeld sterker negatief gecumuleerd GCE einde L1 (gemiddeld -6,7 punten) gepaard met een *positief* gecumuleerd GCE op de leerwinst in L2 van gemiddeld 2,3 punten. De meeste scholen uit de referentiesteekproef situeren zich aan de rechterkant en rond de horizontale middellijn: zowel voor einde L1 als t.a.v. de leerwinst in L2 is het gecumuleerd GCE beperkt (gemiddeld -1,8 voor einde L1 en 0,19 t.a.v. de leerwinst in L2). Figuur 5.10 laat dus zien dat achter de vaststelling dat er voor de steekproef als geheel t.a.v. de leerwinst in L2 geen significante GCE's zijn, een complexere realiteit lijkt schuil te gaan met positieve en negatieve gecumuleerde GCE's. Een inhoudelijke verklaring voor dit fenomeen ligt echter niet voor de hand, zodat de vraag rijst in hoeverre het beeld dat figuur 5.8 laat zien, op statistische artefacten berust.

Wanneer voor de bovenste en de onderste groep in figuur 5.8 de correlatiecoëfficiënt afzonderlijk berekend wordt, blijkt de samenhang een heel stuk sterker: $r = -0,821$ voor de onderste groep ($N = 25$) en $r = -0,897$ voor de bovenste ($N = 140$).



Figuur 5.8 Gecumuleerd GCE op leerwinst in spellingvaardigheid in L2: verschillen tussen scholen en schooltypes en samenhang met gecumuleerd GCE op de spellingvaardigheid einde L1



Figuur 5.9 Verschillen tussen scholen en schooltypes in toegevoegde waarde bij de leerwinst voor spelling in het tweede leerjaar

Een factor van belang in de relatieve positie die scholen tegen het einde van het tweede leerjaar innemen, is ook de toegevoegde waarde die zij realiseren met betrekking tot de leerwinst voor spelling in L2. Figuur 5.9 toont dat die toegevoegde waarde niet samenhangt met het schooltype. Kansarme en zeer kansarme scholen liggen net als modale en kansrijke scholen over het hele spectrum van lage naar hoge toegevoegde waarde verspreid. In elk van de schooltypes treffen we m.a.w. effectieve en minder effectieve scholen aan. In figuur 5.9 geven de verticale staafjes het betrouwbaarheidsinterval voor de TW weer ($CI = 1,39 \times SE$). Twee scholen met staafjes die elkaar niet overlappen verschillen significant van elkaar in de TW die ze met betrekking tot de leerwinst voor spelling in L2 realiseren.

Tabel 5.5 Groepscompositie-effecten en feitelijke extra achterstand of voorsprong voor spelling in 4 exemplarische basisscholen (zonder correctie voor initiële Nederlandse taalvaardigheid)

School	Schooltype	GCE einde L1 (verwachte extra achterstand)	Feitelijke extra achterstand / voorsprong einde L1	TW leerwinst L2	Feitelijke extra achterstand / voorsprong einde L2	N einde L1	N einde L2
1	Zeer kansarm	-9,11	-1,88	3,85	1,97	21	24
2	Zeer kansarm	-9,10	-13,20*	4,67*	-8,53*	24	23
3	Modaal	-0,02	4,33*	-2,19*	2,14*	20	24
4	Modaal	-0,25	0,47	-6,35	-5,88*	24	24
Vlaanderen		-1,80	0,00	0,00	0,00		

* Significant verschillend van het gemiddelde van de Vlaamse referentiegroep

Vier voorbeelden

Bij wijze van illustratie presenteren we in tabel 5.5 de belangrijkste waarden voor vier scholen: twee zeer kansarme en twee modale scholen. (De nummers in de kolom “school” verwijzen *niet* naar dezelfde scholen als in tabel 3.7 of tabel 4.4). De gegevens in tabel 5.5 laten toe te beoordelen in hoeverre een school erin geslaagd is te compenseren voor de extra achterstand voor technisch lezen die op basis van de groepssamenstelling op het einde van L1 te verwachten viel.

Voor elke school kan de graad van kansarmoede afgelezen worden in de derde kolom, die het gecumuleerde groepscompositie-effect einde L1 weergeeft, d.i. de extra achterstand die op grond van de samenstelling van de SiBO-cohorte in de betreffende school te verwachten was. De “kansarmoedegraad” die in kolom 3 van tabel 5.5 aangegeven wordt, is verschillend van de “kansarmoedegraad” in tabellen 3.7 (wiskunde) en 4.4 (technisch lezen). De groepscompositie-effecten waarvan die “kansarmoedegraad” is afgeleid, zijn immers anders per leergebied.

Voor de twee modale scholen is het gecumuleerde groepscompositie-effect op het einde van L1 zo goed als nul; voor de twee kansarme scholen komt ze overeen met een te verwachten extra achterstand van ruim een half schooljaar. Merk nogmaals op dat het gemiddelde voor Vlaanderen hier niet nul is. Het GCE einde L1 is maar nul in scholen zonder enige GOK-leerling, wat voor de gemiddelde Vlaamse school uiteraard niet het geval is.

Hoewel de twee zeer kansarme scholen een zeer vergelijkbare groepssamenstelling kennen (zie kolom 3), liggen hun gecorrigeerde gemiddelden op het einde van L1 heel sterk uiteen. Waar school 1 een gemiddelde SVS-score haalt die bijna overeenkomt met wat op grond van de *individuele* instroomkenmerken van de leerlingen verwacht kon worden, heeft school 2 een bijzonder grote extra achterstand opgelopen ($\approx 85\%$ van de leerwinst die gemiddeld in L2 voor spelling geboekt wordt). Die niet door individuele achtergrondkenmerken te verklaren achterstand in school 2 is op het einde van L1 anderhalve keer zo groot als op basis van de groepssamenstelling verwacht kon worden. In de loop van het tweede leerjaar wordt dat gedeeltelijk gecompenseerd door een leerwinst die gemiddeld hoger ligt dan wat in Vlaanderen met leerlingen met gelijkaardige instroomkenmerken gemiddeld gerealiseerd wordt (TW = 4,67). Ook school 1 realiseert een positieve TW in L2. Het resultaat is dat tegen einde van L2 tussen beide even kansarme scholen een groot kloof in gemiddeld prestatieniveau voor spelling bestaat, een kloof die overeenkomt met de vooruitgang die gemiddeld in twee derde van een schooljaar geboekt wordt. Terwijl de leerlingen in school 2 op het einde van L2 nog steeds kampen met een extra achterstand die gemiddeld ongeveer overeenkomt met wat een jaar eerder op grond van de groepssamenstelling verwacht had kunnen worden, valt in school 1 van enig groepscompositie-effect in feite niets meer te bespeuren - een situatie die erg gelijkend is aan de situatie in school 3. Tussen de twee modale scholen zijn gelijkaardige, zij het iets minder grote verschillen te noteren.

5.6 Besluit

De bevindingen voor het leergebied spelling kunnen als volgt samengevat worden:

1. Niveau en leerwinst in spellingvaardigheid worden door diverse kind- en achtergrondfactoren beïnvloed. Het niveau einde eerste leerjaar én de leerwinst in het tweede leerjaar zijn iets hoger voor meisjes dan voor jongens; leerlingen met schoolse vertraging halen een lager niveau; wie het eerste leerjaar heeft moeten overdoen, kijkt tegen het einde van het tweede leerjaar toch al weer tegen een significante achterstand aan. Er is een significante en - ondanks een zekere inhaalbeweging bij de zwaksten toch vrij persistente - samenhang met het aanvankelijke Nederlandse taalvaardigheidsniveau van de leerlingen.
2. Ook voor spelling treffen we de grootste verschillen in vaardigheidsniveau en leerwinst aan tussen leerlingen met een verschillende sociale en etnisch-culturele achtergrond. De achterstand die GOK-leerlingen t.o.v. modale Vlaamse leerlingen voor spelling hebben, wijzigt nauwelijks in de loop van het tweede leerjaar. De grootste kloof met de modale Vlaamse leerlingen vinden we voor de anderstalige GOK-leerlingen. Voor spelling zijn er weinig verschillen tussen de onderscheiden taalgroepen. Hun achterstand is anderhalve keer zo groot als die van de Nederlandstalige GOK-leerlingen. Voor spelling situeert de omvang van de kloof met modale Vlaamse leerlingen zich halfweg de grotere kloof die de GOK-categorieën voor wiskunde hebben en de kleinere die er voor technisch lezen is.
3. De achterstand van GOK-leerlingen voor spelling wordt in zeer belangrijke mate (d.i. voor 45% tot 90%) door verschillen in aanvankelijke Nederlandse taalvaardigheid verklaard. Gezien de onderlinge posities van de verschillende GOK-categorieën nauwelijks veranderen, wijzigt ook de samenhang met aanvankelijke Nederlandse taalvaardigheid nauwelijks. Net als voor wiskunde en technisch lezen rijst hier de vraag in hoeverre (en tot hoe lang in de schoolloopbaan) een zwakkere aanvankelijke Nederlandse taalvaardigheid als (persistente) oorzaak van achterstand voor spelling aangewezen kan worden, dan wel een meer algemene achterliggende factor.
4. Tussen scholen blijken er ook voor spelling grote verschillen in gemiddelde niveau einde eerste leerjaar en gemiddelde leerwinst in het tweede leerjaar te bestaan. Een groot deel van die verschillen kan door de effecten van instroomkenmerken verklaard worden. De resterende variantie is echter nog altijd significant. Ook na correctie voor individuele instroomkenmerken blijken basisscholen onderling nog in aanzienlijke mate significant te verschillen. Dat geldt zowel voor het gemiddelde niveau dat op het einde van het eerste leerjaar bereikt wordt als voor de leerwinst die in de loop van het tweede leerjaar gemiddeld geboekt wordt.

5. Tussen einde eerste en einde tweede leerjaar blijft de variantie tussen scholen (na correctie voor de effecten van instroomkenmerken) ongeveer gelijk. Dat ondanks de verschillen in toegevoegde waarde met betrekking tot de leerwinst in het tweede leerjaar de verschillen tussen scholen op het einde van het tweede leerjaar niet groter zijn dan op het einde van het eerste leerjaar, heeft te maken met de negatieve samenhang tussen de toegevoegde waarde met betrekking tot de leerwinst in het tweede leerjaar en de toegevoegde waarde met betrekking tot niveau voor spelling op het einde van het eerste leerjaar. Scholen met een hogere toegevoegde waarde einde eerste leerjaar realiseren vaker een lagere toegevoegde waarde in het tweede leerjaar en omgekeerd.
6. De verschillen tussen scholen zijn voor spelling niet alleen significant, ze zijn ook behoorlijk groot. Afgezet tegenover de leerwinst die gemiddeld geboekt wordt, loopt het verschil - na correctie voor instroomkenmerken - tussen scholen op het eerste resp. het negende deciel tegen het einde van het eerste leerjaar al op tot ruim twee derde van een schooljaar. Dat is groter dan voor technisch lezen en voor wiskunde het geval is. Dat verschil is tegen het einde van het tweede leerjaar niet gewijzigd.
7. Die verschillen in gecorrigeerd niveau einde eerste leerjaar kunnen wellicht grotendeels aan verschillen in onderwijseffectiviteit toegeschreven worden. Net als voor technisch lezen vormen de prestaties voor spelling op het einde van het eerste leerjaar immers geen “beginniveau”, maar zijn ze (onder meer) het resultaat van een schooljaar lang onderwijs.
8. Hoewel statistisch significante groepscompositie-effecten gevonden worden, blijkt de samenstelling (in termen van het aandeel leerlingen uit elke GOK-categorie) van de leerlingengroep die aan het SiBO-onderzoek deelneemt, maar een beperkte verklaring te bieden voor de verschillen tussen scholen in gemiddelde spellingvaardigheid op het einde van het eerste leerjaar, die niet door individuele instroomkenmerken verklaard kunnen worden. Bij alle types van scholen, zowel “kansarme scholen” en “zeer kansarme” scholen, als bij “modale scholen” en “kansrijke scholen” zijn er grote, niet door individuele instroomkenmerken verklaarde verschillen in gemiddelde scores voor spelling. Dat geldt zowel voor het einde van het eerste als voor het einde van het tweede leerjaar. Op elk meetmoment blijken scholen met een gelijke mate van kansarmoede toch erg te kunnen verschillen in de weg die ze met hun leerlingen inzake spellingvaardigheid hebben afgelegd. In zoverre groepscompositie een effect heeft op het niveau en de vooruitgang die leerlingen boeken, blijken sommige scholen daar dus toch beter voor te kunnen compenseren dan andere. Op de vraag welke “beschermende” resp. “bezwarende factoren” hieraan ten grondslag liggen, valt op grond van de hier gerapporteerde analyses niet te bepalen.

9. Een algemene vaststelling is dat de effectiviteit van een school voor spelling (in termen van de toegevoegde waarde met betrekking tot de leerwinst in het tweede leerjaar) niet samenhangt met de samenstelling van de leerlingcohort die in het SiBO-onderzoek gevolgd wordt. Zowel onder de kansarme scholen als onder de zeer kansarme scholen treffen we voor wiskunde evengoed effectieve en zeer effectieve scholen aan als onder de modale scholen.
10. Net als voor beide andere leergebieden, vinden we ook voor het domein spelling geen argumenten die zouden pleiten voor een dringende etnische desegregatie (lees: opdoeken van concentratiescholen). De voorliggende resultaten laten niet toe te concluderen dat een gelijkmatige spreiding van allochtone leerlingen over alle basisscholen binnen een regio automatisch zou leiden tot een verbetering van de leerwinst voor spelling. Er dient eerder gezocht te worden naar de factoren die de verschillen in effectiviteit tussen scholen, “zwarte” zowel als “witte”, kunnen helpen verklaren.
11. Tussen het einde van het eerste leerjaar en het einde van het tweede leerjaar neemt de samenhang tussen samenstelling van de leerlingcohort die aan het SiBO-onderzoek deelneemt en het gemiddelde gecorrigeerde niveau voor spelling lichtjes af.
12. Met betrekking tot de GOK-categorieën is voor spelling geen sprake van differentiële schooleffecten. De toegevoegde waarde die scholen realiseren, realiseren ze in gelijke mate voor alle GOK-categorieën. Differentiële schooleffecten zijn er wel t.a.v. leerlingen met schoolse vertraging.

6. Algemeen besluit en vergelijking over de leergebieden

In dit rapport werden de resultaten gepresenteerd van analyses met betrekking tot het vaardigheidsniveau en de leerwinst in eerste en/of tweede leerjaar voor drie leergebieden: wiskunde, technisch lezen en spelling. Die resultaten zijn gebaseerd op de gegevens verzameld in het kader van het longitudinale SiBO-onderzoek. Behalve de data van de referentiesteekproef (120 scholen), werden daarbij ook de gegevens uit twee oversamplings betrokken (52 scholen).

In de bespreking gingen we in op:

- de effecten van kind- en gezinsfactoren,
- de verschillen tussen scholen in beginniveau, gemiddelde leerwinst en toegevoegde waarde,
- groepscompositie-effecten.

In wat volgt, brengen we eerst een synthese van de bevindingen voor elk van de leergebieden. Daarna gaan we nog even in op de samenhang tussen de leergebieden.

Kleine en grote verschillen tussen leerlingen naar thuistaal en sociale achtergrond

Verschillen in leerprestaties tussen leerlingen met een verschillende sociale en/of etnisch-culturele achtergrond blijken doorgaans van een grotere orde te zijn dan verschillen tussen jongens en meisjes of verschillen tussen leerlingen die vroeger of later op het jaar geboren zijn.

Om de complexe samenhangen tussen sociale achtergrond, thuistaal en leerprestaties wat eenvoudiger in kaart te kunnen brengen, hebben wij de leerlingen opgedeeld in een aantal categorieën, ons daarbij baserend op de criteria die in het Vlaamse GOK-beleid gebruikt worden. We hebben wel wat meer differentiatie aangebracht op basis van de thuistaal.

De achterstand waar sommige categorieën van leerlingen (in vergelijking met modale Vlaamse leerlingen) tegen aankijken, verschilt sterk naargelang van het leergebied. Voor wiskunde varieert de gemiddelde achterstand bij het begin van het eerste leerjaar van één trimester (Nederlandstalige GOK-leerlingen) tot bijna een volledig schooljaar (Turkse GOK-leerlingen). Tegen het einde van het eerste leerjaar hebben de anderstalige (GOK- en niet-GOK-) leerlingen hun aanvankelijke achterstand al voor een substantieel deel goedgemaakt, maar ze schommelt voor de anderstalige GOK-leerlingen gemiddeld nog steeds rond een half schooljaar met nog maar kleine verschillen tussen de onderscheiden taalgroepen. Voor spelling staan anderstalige GOK-leerlingen op het einde van het eerste leerjaar gemiddeld een trimester achter, Nederlandstalige GOK-leerlingen ruim zeven weken. De kleinste achterstand vinden we voor technisch lezen: van vier weken (Nederlandstalige GOK-leerlingen) tot negen weken (Turkse GOK-leerlingen) bij het einde van het eerste leerjaar.

Gemeenschappelijk voor alle leergebieden is de vaststelling dat in de loop van het tweede leerjaar alle categorieën van leerlingen een leerwinst boeken die gemiddeld weinig verschilt van de leerwinst die modale Vlaamse leerlingen gemiddeld realiseren. Dat brengt met zich dat de achterstandposities die op het einde van het eerste leerjaar ingenomen werden, nog maar weinig veranderen in de loop van het tweede leerjaar.

Verschillen tussen scholen

Tussen Vlaamse basisscholen blijken er grote verschillen te zijn.

Die verschillen betreffen in de eerste plaats het publiek dat de school bedient. Op basis van de samenstelling van de leerlingcohort die aan het SiBO-onderzoek deelneemt hebben we de scholen ingedeeld in vier types: “modale scholen” (met gemiddeld 70% modale Vlaamse leerlingen), “kansarme scholen” (met relatief veel Nederlandstalige GOK-leerlingen), “zeer kansarme scholen” (nauwelijks modale Vlaamse leerlingen, veel anderstalige GOK-leerlingen) en “kansrijke” scholen (met gemiddeld 50% anderstalige, meestal Franstalige, niet-GOK-leerlingen).

Na correctie voor de effecten van individuele instroomkenmerken blijken er tussen scholen nog significante en behoorlijk grote effecten te bestaan. Ordenen we de scholen naar het gecorrigeerde gemiddelde op het einde van het eerste leerjaar, dan vinden we tussen het hoogste en het laagste deciel verschillen die - naargelang van het leerdomein - overeenkomen met één trimester leerwinst (technisch lezen), een half schooljaar leerwinst (wiskunde) tot twee trimesters leerwinst (spelling). Op het einde van het tweede leerjaar vinden we gelijkaardige verschillen.

Die resterende, niet door individuele instroomkenmerken verklaarde verschillen lijken gedeeltelijk door groepscompositie-effecten verklaard te kunnen worden. Dat zou betekenen dat bovenop de effecten van de eigen achtergrond van elk kind, de samenstelling van de leerlingengroep nog een bijkomend effect heeft. Bij zulke effecten denkt men spontaan aan de invloed die leerlingen op elkaar hebben en die - naargelang van het geval - het leren bevorderen of afremmen. Of aan - hetzij productieve, hetzij contraproductieve - aanpassingen van de leraar aan kenmerken (b.v. het gemiddelde niveau) van de groep. Maar de vaststelling dat groepscompositie-effecten ook lijken voor te komen bij beginscores voor wiskunde noopt minstens tot aanvullende verklaringen. Die kunnen gezocht worden in effecten van het kleuteronderwijs of de aanwezigheid van gemeenschappelijke, niet in de analyses in rekening gebrachte leerlingenkenmerken.

In onze analyses hebben wij groepscompositie onderzocht in termen van het aandeel leerlingen uit elk van de onderscheiden “GOK-categorieën”. Een merkwaardig vaststelling was dat sommige groepscompositie-effecten gecounterd werden door een navolgend, “omgekeerd” effect, waarbij een groter aandeel van de betreffende GOK-categorie blijkbaar leidt tot méér leerwinst in plaats van minder. Of dit echt de weerspiegeling van een onderwijsrealiteit vormt (b.v. van een aanpak die

succesvol compenseert voor ongunstige instroomkenmerken), dan wel eerder gezien moet worden als een statistisch artefact, blijft voorlopig onduidelijk.

Tabellen 6.1 en 6.2 geven bij wijze van samenvatting voor elk van de drie onderzochte leergebieden een overzicht van de variantie die verklaard wordt door het geheel van individuele achtergrondkenmerken enerzijds en door de gecombineerde groepscompositie-effecten anderzijds. Tabel 6.1 biedt dat overzicht met betrekking tot de variantie op schoolniveau, tabel 6.2. met betrekking tot de totale variantie. Naar dat laatste moet men kijken wanneer men wil zien in welke mate individuele achtergrondkarakteristieken resp. groepscompositie-effecten verschillen tussen leerlingen over de scholen heen helpen verklaren.

Tabel 6.1 Percentage schoolvariantie verklaard door individuele kenmerken resp. groepscompositie-effecten

	Wiskunde ¹		Technisch lezen ²		Spelling ²	
	Individuele kenmerken	GCE	Individuele kenmerken	GCE	Individuele kenmerken	GCE
Begin L1	83,5 %	8,8 %				
Leerwinst L1	25,6 %	7,0 % ³	32,3 %	6,0 %	33,3 %	14,2 %
Leerwinst L2	6,8 %	3,2 % ³	0,0 %	10,6 % ³	5,5 %	4,6 % ³

¹analyses met correctie voor effecten aanvankelijke Nederlandse taalvaardigheid

²analyses zonder correctie voor effecten aanvankelijke Nederlandse taalvaardigheid

³GCE in omgekeerde richting: hoe meer GOK-leerlingen, hoe hoger het schoolgemiddelde

Tabel 6.1 Percentage totale variantie verklaard door individuele kenmerken resp. groepscompositie-effecten

	Wiskunde ¹		Technisch lezen ²		Spelling ²	
	Individuele kenmerken	GCE	Individuele kenmerken	GCE	Individuele kenmerken	GCE
Begin L1	55,8 %	2,3 %				
Leerwinst L1	10,5 %	1,7 % ³	6,8 %	0,8 %	15,4 %	14,2 %
Leerwinst L2	4,7 %	0,9 % ³	4,2 %	1,2 % ³	1,5 %	4,6 % ³

¹analyses met correctie voor effecten aanvankelijke Nederlandse taalvaardigheid

²analyses zonder correctie voor effecten aanvankelijke Nederlandse taalvaardigheid

³GCE in omgekeerde richting: hoe meer GOK-leerlingen, hoe hoger het schoolgemiddelde

Een belangrijke vaststelling is dat voor elk van de leergebieden en op elk van de meetmomenten tussen scholen met een zeer gelijkaardige samenstelling van de leerlingengroep grote verschillen blijken te bestaan in het gecorrigeerde gemiddelde niveau dat bereikt wordt. Voor wiskunde blijken in dat verband vooral de verschillen binnen de categorie kansarme scholen en binnen de categorie zeer kansarme scholen groot te zijn. Op basis van de samenstelling van de leerlingengroep blijkt het m.a.w. moeilijk te voorspellen hoe veel hoger of lager het schoolgemiddelde zal liggen dan op grond

van de individuele achtergrondkenmerken verwacht mocht worden. Als de statistisch aangetroffen groepscompositieeffecten al de weerspiegeling van reële fenomenen vormen, dan moet uit het voorgaande besloten worden dat de ene school beter dan de andere weet te compenseren voor de negatieve effecten van een ongunstige samenstelling van de leerlingengroep. Dat is op zich een belangrijke vaststelling, maar wel een vaststelling die meteen de praktische bruikbaarheid van de gevonden groepscompositie-effecten in vraag stelt. Wat kan men aanvragen met de bevinding dat de samenstelling van de leerlingengroep er over het algemeen gesproken wel toe lijkt te doen, als onmiddellijk daarna gesteld moet worden dat de reële uitwerking ervan toch sterk van school tot school verschilt?

Nemen we de correlatie met het groepscompositie-effect op het vroegste meetmoment als criterium, dan blijkt van meetmoment naar meetmoment de samenhang tussen samenstelling van de leerlingengroep en het (voor individuele instroomkenmerken gecorrigeerde) gemiddelde prestatieniveau bovendien steeds maar zwakker te worden. Dat is vooral bij wiskunde en technisch lezen het geval.

Kijkend naar de toegevoegde waarde die scholen realiseren met betrekking tot de leerwinst in het eerste en/of het tweede leerjaar, dan stellen we tenslotte voor elk van de leergebieden vast dat er helemaal geen verband bestaat tussen die toegevoegde waarde en de samenstelling van de leerlingengroep. In elk van de vier types van scholen, van zeer kansarme tot kansrijke, treffen we ma.w. zowel effectieve als minder effectieve scholen aan.

De bevindingen inzake groepscompositie die uit de analyses op de SiBO-data voor het eerste en het tweede leerjaar naar voren komen, lijken te sporen met recentelijk in de wetenschappelijke literatuur gerapporteerde bevindingen (zie b.v. Driessen, 2007; Harker & Tymms, 2004).

De voorgaande vier punten samen beijkend, vinden we in de resultaten van de analyses geen argumenten die zouden pleiten voor bijvoorbeeld een doorgedreven, dwingende spreiding van allochtone leerlingen over een groter aantal scholen. De resultaten suggereren dat de extra leerwinst die dat zou genereren eerder beperkt zou zijn en wellicht niet zou opwegen tegen de hoge kosten van zo'n maatregel. De vaststelling dat er naast duidelijk minder effectieve, blijkbaar ook effectieve kansarme en zeer kansarme scholen bestaan, suggereert eerder dat het de moeite loont de één van de ander te laten leren hoe betere onderwijskwaliteit gerealiseerd kan worden.

De samenhang tussen de leergebieden

Tot zover de samenvatting van de belangrijkste bevindingen over de drie leergebieden heen. Een vraag die daarbij nog niet aan de orde kwam, is of er een samenhang is tussen de toegevoegde waarde die scholen voor het ene leergebied realiseren en de toegevoegde waarde die ze voor de andere leergebieden realiseren.

Tussen de leergebieden spelling en technisch lezen is die samenhang inzake toegevoegde waarde behoorlijk groot, althans met betrekking tot het *niveau op het einde van het eerste leerjaar* ($r = 0,64$ bij 170 scholen, significant). Inzake de toegevoegde waarde met betrekking tot de leerwinst in het tweede leerjaar is de samenhang tussen beide leergebieden al een stuk minder ($r = 0,32$; 166 scholen, significant).

Heel wat zwakker is de samenhang tussen het leergebied wiskunde en elk van beide andere leergebieden. In verband met de vergelijkbaarheid beperken we ons hier tot de samenhangen tussen de toegevoegde waarden met betrekking tot de leerwinst in het tweede leerjaar. Tussen wiskunde en technisch lezen is die samenhang inzake toegevoegde waarde - alhoewel significant - ronduit zwak: $r = 0,19$ (169 scholen). Tussen wiskunde en spelling is die iets sterker ($r = 0,33$; 167 scholen, significant).

Kijken we naar de positie die scholen op het einde van L2 innemen met betrekking tot de mate waarin de gemiddelde score lager of hoger ligt dan op basis van de individuele instroomkenmerken verwacht kon worden, dan zien we matige tot behoorlijke sterke, in elk geval telkens significante samenhangen tussen de leergebieden: $r = 0,58$ (spelling - technisch lezen, 164 scholen), $r = 0,35$ (wiskunde - technisch lezen, 169 scholen) en $r = 0,66$ (wiskunde - spelling, 165 scholen). Deze cijfers wijzen er enerzijds op dat scholen die effectief blijken te zijn voor een leergebied dat ook eerder zijn voor de andere, maar laten tegelijk ook zien dat een school niet noodzakelijk voor alle leergebieden even effectief is. De effectiviteit van scholen wordt dus best per leergebied afzonderlijk bekeken.

Referenties

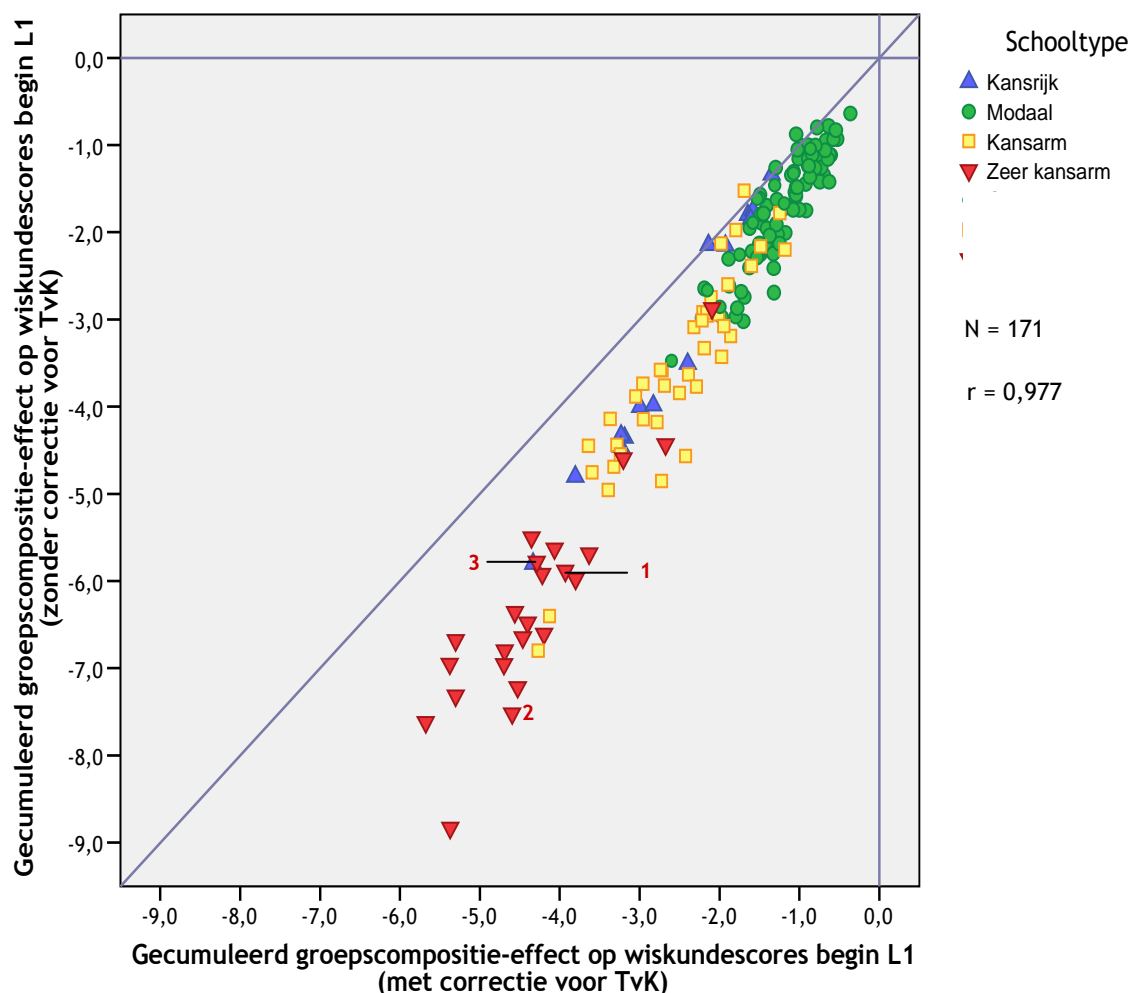
- Driessen, G. (2007). *'Peer group' effecten op onderwijsprestaties. Een internationaal review van effecten, verklaringen en theoretische en methodologische aspecten*. Nijmegen: ITS - Radboud Universiteit (<http://www.its-nijmegen.nl/publicaties/>).
- Harker, R. & Tymms, P. (2004). The Effects of Student Composition on School Outcomes. *School Effectiveness and School Improvement*, 15 (2), 177-199.
- Hendriks, K., Verachtert, P., Ghesquière, P., Maes, F. & Van Damme, J. (2005). *Longitudinaal onderzoek in het basisonderwijs. Toetsen tweede leerjaar (2004-2005)*. Leuven: Steunpunt 'Loopbanen doorheen Onderwijs naar Arbeidsmarkt', Cel 'Schoolloopbanen in het basisonderwijs' (SiBO) (= LOA-rapport nr. 35).
- Moelands, F., & Rymenans, R. (2003a). *Drie-Minuten-Toets voor Vlaanderen. Handleiding*. Arnhem: Citogroep.
- Moelands, F., & Rymenans, R. (2003b). *Schaal Vorderingen in Spellingvaardigheid voor Vlaanderen (SVS-V)*. Handleiding. Arnhem: Citogroep.
- Rymenans, R. (2000). *Leerlingvolgsysteem Taal. OBPWO-project 97.06. Eindrapport*. Antwerpen: Universiteit Antwerpen.
- Van de gaer, E., Verhaeghe, J. P., Reynders, T. & Van Damme, J. (2006). *Longitudinaal onderzoek in het basisonderwijs. GOK-leerlingen in het eerste leerjaar: achterstand en evolutie voor het leergebied technisch lezen*. Leuven: Steunpunt 'Loopbanen doorheen Onderwijs naar Arbeidsmarkt', Cel 'Schoolloopbanen in het basisonderwijs' (SiBO) (= LOA-rapport nr 43).
- Verachtert, P., Ghesquière, P., Hendriks, K., Maes, F. & Van Damme, J. (2005). *Longitudinaal onderzoek in het basisonderwijs. Toetsen eerste leerjaar (2003-2004)*. Leuven: Steunpunt 'Loopbanen doorheen Onderwijs naar Arbeidsmarkt', Cel 'Schoolloopbanen in het basisonderwijs' (SiBO) (= LOA-rapport nr. 30).
- Verachtert, P., Verhaeghe, J.P., Ghesquière, P., Onghena, P., Van Petegem, P. & Van Damme, J. (2006). *Longitudinaal onderzoek in het basisonderwijs. Beginsituatie en leerresultaten in het eerste leerjaar*. Leuven: Steunpunt 'Loopbanen doorheen Onderwijs naar Arbeidsmarkt', Cel 'Schoolloopbanen in het basisonderwijs' (SiBO) (= LOA-rapport nr. 38).
- Verhaeghe, J.P., Van de gaer, E., & Reynders, T. & Van Damme, J. (2006). *Longitudinaal onderzoek in het basisonderwijs. GOK-leerlingen in het eerste leerjaar: achterstand en evolutie voor het leergebied wiskunde. Een benadering vanuit de schooleffectiviteit*. Leuven: Steunpunt 'Loop-

banen doorheen Onderwijs naar Arbeidsmarkt', Cel 'Schoolloopbanen in het basisonderwijs' (SiBO)
(= LOA-rapport nr. 44).

Bijlage 1. GCE voor wiskunde met en zonder correctie voor aanvankelijk Nederlands taalvaardigheidsniveau

Gecumuleerd groepscompositie-effect op de beginscore voor wiskunde

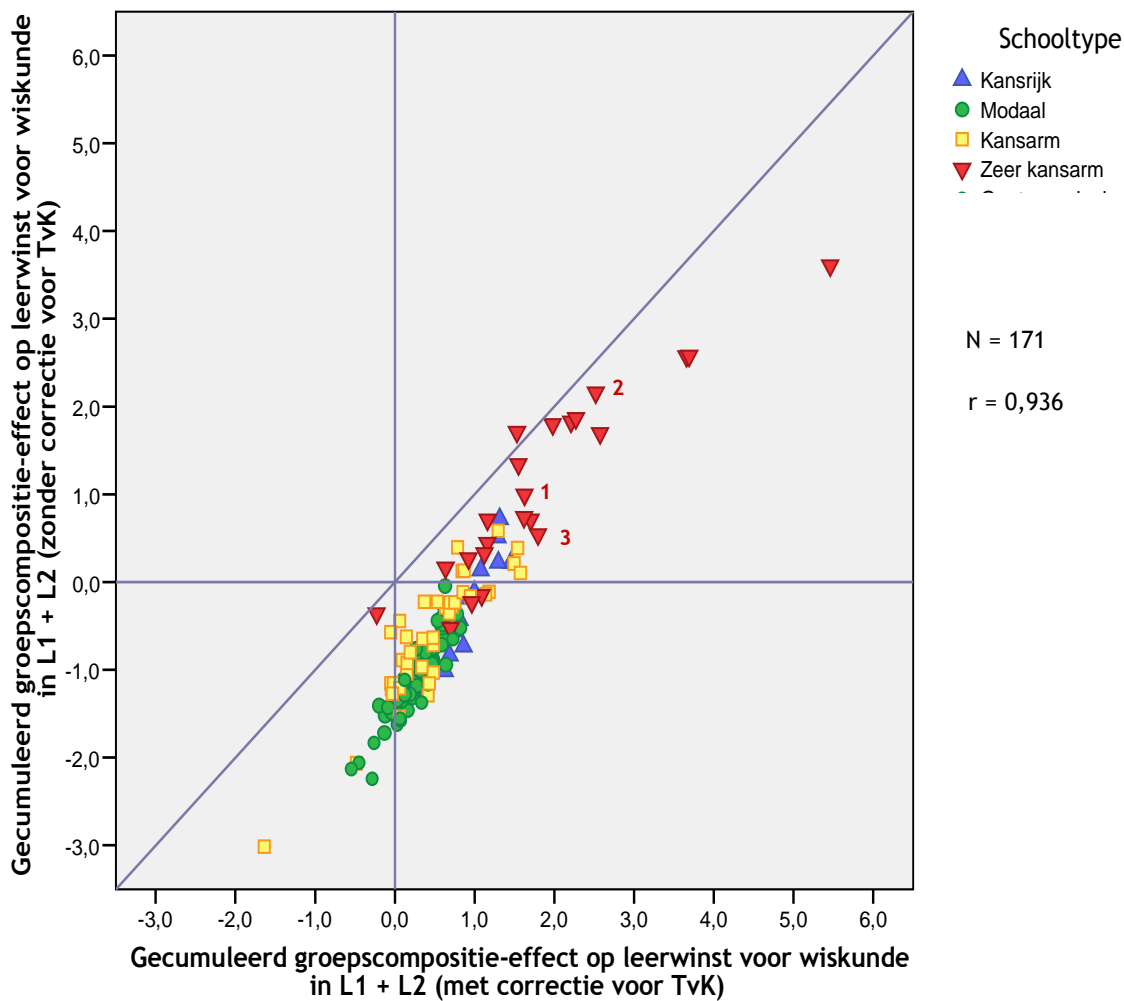
In figuur 3.16c wijst een positie verder rechts onder de diagonaal op een grotere afname van het gecumuleerde groepscompositie-effect wanneer ook gecorrigeerd wordt voor de samenhang met aanvankelijk Nederlands taalvaardigheidsniveau (TvK). Vooral bij de zeer kansarme scholen wordt het gecumuleerde groepscompositie-effect kleiner wanneer gecorrigeerd wordt voor de samenhang met TvK. Dat houdt verband met het grotere aandeel anderstalige leerlingen in die scholen, die doorgaans een veel lagere TvK-score halen.



Figuur 3.16c Gecumuleerd groepscompositie-effect op de wiskundescore begin L1 met en zonder correctie voor effecten van aanvankelijke Nederlandse taalvaardigheid

Gecumuleerd groepscompositie-effect op de leerwinst voor wiskunde in L1+L2

Het beeld in figuur 3.20c is grosso modo het omgekeerde van het beeld in figuur 3.16c. Nu zijn het de zeer kansarme scholen die bovenaan en dicht bij de diagonaal liggen. Dat heeft te maken met het feit dat het GCE zelf ook “omgekeerd” is met een hogere leerwinst voor wiskunde naarmate er meer (anderstalige) GOK-leerlingen zijn.



Figuur 3.20c Gecumuleerd groepscompositie-effect op de leerwinst voor wiskunde L1 + L2 met en zonder correctie voor effecten van aanvankelijke Nederlandse taalvaardigheid

Bijlage 2. Tabellen multiniveau-analyses: verdeling onverklaarde variantie en aandeel verklaarde variantie

De tabellen in deze bijlage geven voor elk leergebied de verdeling van de onverklaarde variantie en het aandeel verklaarde variantie in de belangrijkste modellen aan. De resultaten met betrekking tot de impact van de verschillende predictoren en hun onderlinge interacties zijn verwerkt in de tabellen en figuren in de hoofdtekst van het rapport.

Verdeling van de variantie voor wiskunde

Model 0 = model met enkel constante + tijdcodering voor leerwinst in L1 resp. leerwinst in L2
 Model 1 = model 0 + alle leerlingkenmerken inclusief aanvankelijke Nederlandse taalvaardigheid
 Model 2 = model 0 + alle leerlingkenmerken behalve aanvankelijke Nederlandse taalvaardigheid
 Model 3 = model 2 + groepscompositie-effecten

Tabel 3.8 Verdeling van de onverklaarde variantie voor het leergebied wiskunde

	Model 0		Model 1		Model 2		Model 3	
	Var	%	Var	%	Var	%	Var	%
Beginniveau								
School	17,934	26,2%	2,962	9,8%	6,195	12,7%	2,690	6,0%
Leerling	50,557	73,8%	27,280	90,2%	42,438	87,3%	42,507	94,0%
Totaal	68,491	100,0%	30,242	100,0%	48,633	100,0%	45,197	100,0%
Leerwinst in L1								
School	9,936	23,4%	7,392	19,4%	7,336	19,2%	6,493	17,4%
Leerling	32,520	76,6%	30,620	80,6%	30,785	80,8%	30,773	82,6%
Totaal	42,456	100,0%	38,012	100,0%	38,121	100,0%	37,266	100,0%
Leerwinst in L2								
School	9,791	26,1%	9,128	25,5%	9,195	25,8%	8,784	24,9%
Leerling	27,776	73,9%	26,670	74,5%	26,486	74,2%	26,487	75,1%
Totaal	37,567	100,0%	35,798	100,0%	35,681	100,0%	35,271	100,0%

Var = onverklaarde variantie

Toelichting bij tabel 3.8:

De kolommen “var” tonen de onverklaarde variantie op respectievelijk school- en leerling-niveau. De kolommen “%” tonen hoe die onverklaarde variantie is verdeeld over school- en leerling-niveau.

Zonder enige verklarende variabele in het model (model 0), blijkt zowel voor het beginniveau voor wiskunde als voor de leerwinst in het eerste en voor de leerwinst in het tweede leerjaar ongeveer een kwart van de totale verschillen te maken te hebben met verschillen tussen scholen.

Tabel 3.9 Verklaarde variantie en toename in verklaarde variantie voor wiskunde

	Model 0	Model 1		Model 2		Model 3	
	PEV	PEV	Δ PEV vs model 0	PEV	Δ PEV vs model 1	PEV	Δ PEV vs model 2
Beginniveau							
School		83,5%	83,5%	65,5%	-18,0%	85,0%	19,5%
Leerling		46,0%	46,0%	16,1%	-30,0%	15,9%	-0,1%
Totaal		55,8%	55,8%	29,0%	-26,9%	34,0%	5,0%
Leerwinst in L1							
School		25,6%	25,6%	26,2%	0,6%	34,7%	8,5%
Leerling		5,8%	5,8%	5,3%	-0,5%	5,4%	0,0%
Totaal		10,5%	10,5%	10,2%	-0,3%	12,2%	2,0%
Leerwinst in L2							
School		6,8%	6,8%	6,1%	-0,7%	10,3%	4,2%
Leerling		4,0%	4,0%	4,6%	0,7%	4,6%	0,0%
Totaal		4,7%	4,7%	5,0%	0,3%	6,1%	1,1%

PEV = percentage explained variance, percentage verklaarde variantie

Toelichting bij tabel 3.9

Alle leerlingenkenmerken samen (model 1) verklaren 83,5% van de verschillen tussen scholen inzake beginniveau voor wiskunde. In het eerste leerjaar komt een kwart van de verschillen tussen scholen inzake gemiddelde leerwinst eveneens voor rekening van instroomverschillen; voor de leerwinst in het tweede leerjaar is dat nog maar krap 7%. Vergelijking met figuur 3.2 op p. 44 suggereert ons hoe we die percentages moeten interpreteren. Tussen begin en einde L1 nemen de verschillen tussen scholen nog toe. Dat komt omdat de gemiddelde leerwinst van school tot school sterk verschillend is. En het zijn die laatste verschillen (in gemiddelde leerwinst) die voor een kwart voor rekening van verschillen in instroomkenmerken komen. Figuur 3.2 toont dat in de loop van L2 de verschillen tussen scholen niet meer toenemen. Er zijn misschien nog wel verschillen in gemiddelde leerwinst, maar die verschillen heffen elkaar op zodat tegen einde L2 de verschillen tussen scholen in gemiddeld niveau niet toegenomen zijn. Die (elkaar dus opheffende) verschillen tussen scholen in gemiddelde leerwinst hangen voor 6,8% samen met instroomverschillen.

Voor de verschillen tussen leerlingen binnen scholen zien we een gelijkaardig fenomeen. Verschillen in achtergrond spelen vooral een rol in het beginniveau, veel minder - en in afnemende mate - in de groei.

Met betrekking tot het beginniveau voor wiskunde kan 18% van de variantie op schoolniveau verklaard worden door instroomverschillen betreffende aanvankelijke Nederlandse taalvaardigheid bij de leerlingen (= verschil in verklaarde variantie model 2 vs. model 1). Van de verschillen tussen leerlingen in het beginniveau voor wiskunde komt 30% voor rekening van verschillen in aanvankelijke Nederlandse taalvaardigheid. De impact van aanvankelijke Nederlandse taalvaardigheid op de variantie met betrekking tot de leerwinst in wiskunde is quasi nihil.

Vergelijking van model 3 met model 2 toont dat de gecumuleerde groepscompositie-effecten voor het beginniveau in wiskunde 19,5% van de variantie tussen scholen verklaart, 8,5% van de verschillen tussen scholen betreffende de leerwinst voor wiskunde in L1 en 4,2% van de verschillen tussen scholen met betrekking tot de leerwinst voor wiskunde in L2.

Om te zien in hoeverre het gecumuleerde groepscompositie-effect verschillen tussen leerlingen over alle scholen heen helpt verklaren, moet men naar de verandering in *totale* verklaarde variantie

kijken (model 3 vs model 2). Voor het beginniveau gaat het om 5%, voor de leerwinst in L1 om 2% en voor de leerwinst in L2 om 1,1%. Die percentages kan men vergelijken met de percentages totale variantie verklaard door individuele leerlingkenmerken. Exclusief aanvankelijke Nederlandse taalvaardigheid gaat het om 29% (beginniveau), 10,2% (leerwinst L1) en 5% (leerwinst L2). De invloed van de individuele achtergrond is dus veel groter dan de invloed van de gecumuleerde groepscompositie-effecten.

Table 3.10. Proportie variantie verklaard door individuele leerlingenkenmerken versus gecumuleerde groepscompositie-effecten

	Analyse inclusief TvK		TvK	Analyse exclusief TvK	
	Individuele kenmerken	GCE		Individuele kenmerken	GCE
Beginniveau wiskunde					
School	83,5%	8,8%	18,0%	65,4%	19,6%
Leerling	46,0%	0,0%	30,1%	16,0%	-0,1%
Totaal	55,8%	2,3%	26,9%	29,0%	5,1%
Leerwinst wiskunde in L1					
School	25,6%	7,0%	-1,8%	26,3%	8,4%
Leerling	5,8%	0,1%	0,4%	5,3%	0,0%
Totaal	10,5%	1,7%	-0,1%	10,2%	2,0%
Leerwinst wiskunde in L2					
School	6,8%	3,2%	0,9%	6,2%	4,1%
Leerling	4,0%	0,0%	-1,5%	4,6%	0,0%
Totaal	4,7%	0,9%	-0,9%	5,0%	1,1%

Toelichting tabel 3.10

De % in de kolom "TvK" geven aan hoeveel van de verschillen (tussen scholen resp. tussen leerlingen binnen scholen) verklaard worden door verschillen in initiële taalvaardigheid, zoals kan worden afgeleid uit een vergelijking tussen modellen 2 en 3 (zie tabel 3.9).

Rechts daarvan zien we wat de bijdrage is van de gecumuleerde groepscompositie-effecten in de verklaring van de verschillen tussen scholen / verschillen tussen leerlingen binnen scholen resp. met betrekking tot het beginniveau, de leerwinst in L1 en de leerwinst in L2, in een analyse waarin die GCE geschat zijn op basis van modellen waarin niet gecorrigeerd wordt voor de effecten van initiële Nederlandse taalvaardigheid.

Links zien we hoe groot die bijdragen van de GCE zijn wanneer die GCE bepaald worden op basis van modellen waarin wel gecorrigeerd wordt voor TvK. Men kan zien dat als wel gecorrigeerd wordt voor TvK, het gecumuleerde groepscompositie-effect maar een veel kleinere rol meer speelt. (Het verschil tussen beide analyses is vooral voor het beginniveau opvallend.) Dat illustreert de stelling van o.m. Harker en Tymms (2004) en Driessen (2007) dat groepscompositie-effecten wellicht in belangrijke mate artefacten zijn t.g.v. inadequate modelspecificatie, meer bepaald een onder-specificatie van het model op leerlingniveau, d.w.z. het niet in rekening brengen van relevante leerlingkenmerken.

Merk op dat de bijdrage van de GCE in de verklaring van de verschillen tussen leerlingen binnen scholen 0% is. Dat komt omdat GCE een verklarende variabele op schoolniveau is. Die kan uiteraard geen verschillen *binnen* scholen verklaren.

Verdeling van de variantie voor technisch lezen

Model 0 = model met enkel constante + tijdcodering voor leerwinst in L1 resp. leerwinst in L2
 Model 1 = model 0 + alle leerlingkenmerken inclusief aanvankelijke Nederlandse taalvaardigheid
 Model 2 = model 0 + alle leerlingkenmerken behalve aanvankelijke Nederlandse taalvaardigheid
 Model 3 = model 2 + groepscompositie-effecten

Tabel 4.5 Verdeling van de variantie in technische leesvaardigheid

	Model 0		Model 1		Model 2		Model 3	
	Var	%	Var	%	Var	%	Var	%
Niveau einde L1								
School	307,040	14,1%	179,952	10,7%	207,798	10,2%	189,468	9,4%
Leerling	1873,827	85,9%	1503,174	89,3%	1824,538	89,8%	1824,349	90,6%
Totaal	2180,867	100,0%	1683,126	100,0%	2032,336	100,0%	2013,817	100,0%
Leerwinst in L2								
School	139,576	11,5%	139,833	12,0%	140,020	12,1%	125,174	10,9%
Leerling	1072,250	88,5%	1021,516	88,0%	1020,392	87,9%	1020,278	89,1%
Totaal	1211,826	100,0%	1161,349	100,0%	1160,412	100,0%	1145,452	100,0%

Toelichting bij tabel 4.5

In een model zonder verklarende variabelen (model 0) komt met betrekking tot het niveau voor technisch lezen einde L1 14% van alle verschillen voor rekening van verschillen tussen scholen. Voor de leerwinst in L2 gaat het om 11,5% van alle verschillen.

Tabel 4.6 Verklarde variantie en toename in verklaarde variantie voor technisch lezen

	Model 0	Model 1		Model 2		Model 3	
	PEV	PEV	Δ PEV vs model 0	PEV	Δ PEV vs model 1	PEV	Δ PEV vs model 2
Niveau einde L1							
School		41,4%	41,4%	32,3%	-9,1%	38,3%	6,0%
Leerling		19,8%	19,8%	2,6%	-17,2%	2,6%	0,0%
Totaal		22,8%	22,8%	6,8%	-16,0%	7,7%	0,8%
Leerwinst in L2							
School		-0,2%	-0,2%	-0,3%	-0,1%	10,3%	10,6%
Leerling		4,7%	4,7%	4,8%	0,1%	4,8%	0,0%
Totaal		4,2%	4,2%	4,2%	0,1%	5,5%	1,2%

PEV = percentage explained variance, percentage verklaarde variantie

Toelichting bij tabel 4.6

Instreamverschillen met betrekking tot aanvankelijke Nederlandse taalvaardigheid staan in voor 9% van de verschillen tussen scholen in technische leesvaardigheid einde L1, maar voor slechts 0,1% van de verschillen tussen scholen inzake de leerwinst in L2 (zie vergelijking model 2 met model 1).

GCE verklaren 6% van de verschillen tussen scholen in technische leesvaardigheid einde L1 en 10,6% van de verschillen tussen scholen inzake de leerwinst in L2 (model 3). Van de verschillen tussen leerlingen over de scholen heen verklaren de GCE 0,8% (niveau einde L1) resp. 1,2% (leerwinst L2). De individuele leerlingkenmerken (excl. init. taalvaardigheid - model 2) verklaren 6,8% resp. 4,2%.

Verdeling van de variantie voor spelling

Model 0 = model met enkel constante + tijdcodering voor leerwinst in L1 resp. leerwinst in L2

Model 1 = model 0 + alle leerlingkenmerken inclusief aanvankelijke Nederlandse taalvaardigheid

Model 2 = model 0 + alle leerlingkenmerken behalve aanvankelijke Nederlandse taalvaardigheid

Model 3 = model 2 + groepscompositie-effecten (GCE)

Tabel 5.6 Verdeling van de variantie in spellingvaardigheid

	Model 0		Model 1		Model 2		Model 3	
	Var	%	Var	%	Var	%	Var	%
Niveau einde L1								
School	35,870	26,0%	18,667	19,8%	23,988	20,5%	18,912	16,9%
Leerling	102,356	74,1%	75,730	80,2%	92,972	79,5%	92,901	83,1%
Totaal	138,226	100,0%	94,397	100,0%	116,960	100,0%	111,813	100,0%
Leerwinst in L2								
School	18,873	19,6%	17,406	18,7%	17,836	18,8%	16,969	18,1%
Leerling	77,363	80,4%	75,883	81,3%	76,914	81,2%	76,932	81,9%
Totaal	96,236	100,0%	93,289	100,0%	94,750	100,0%	93,901	100,0%

Toelichting bij tabel 5.6

In een model zonder verklarende variabelen (model 0) komt met betrekking tot het niveau voor spelling einde L1 ruim een kwart van alle verschillen voor rekening van verschillen tussen scholen. Voor de leerwinst in L2 gaat het om net geen 20% van alle verschillen.

Tabel 5.7 Verklaarde variantie en toename in verklaarde variantie voor spelling

	Model 0	Model 1		Model 2		Model 3	
	PEV	PEV	Δ PEV vs model 0	PEV	Δ PEV vs model 1	PEV	Δ PEV vs model 2
Niveau einde L1							
School		48,0%	48,0%	33,1%	-14,8%	47,3%	14,2%
Leerling		26,0%	26,0%	9,2%	-16,9%	9,2%	0,1%
Totaal		31,7%	31,7%	15,4%	-16,3%	19,1%	3,7%
Leerwinst in L2							
School		7,8%	7,8%	5,5%	-2,3%	10,1%	4,6%
Leerling		1,9%	1,9%	0,6%	-1,3%	0,6%	-0,2%
Totaal		3,1%	3,1%	1,5%	-1,5%	2,4%	0,9%

PEV = percentage explained variance, percentage verklaarde variantie

Toelichting bij tabel 5.7

Instroomverschillen met betrekking tot aanvankelijke Nederlandse taalvaardigheid staan in voor bijna 15% van de verschillen tussen scholen in spellingvaardigheid einde L1, maar voor slechts 2,3% van de verschillen tussen scholen inzake de leerwinst in L2 (zie vergelijking model 2 met model 1).

GCE verklaren 14,2% van de verschillen tussen scholen in spellingvaardigheid einde L1 en 4,6% van de verschillen tussen scholen inzake de leerwinst in L2 (model 3). Van de verschillen tussen leerlingen over de scholen heen verklaren de GCE 3,7% (niveau einde L1) resp. 0,9% (leerwinst L2). De individuele leerlingkenmerken (excl. initiële taalvaardigheid - model 2) verklaren 15,4% resp. 1,5%.