

**Wiskunde in het lager onderwijs in TIMSS 2003 en TIMSS 2011
in Vlaanderen, zonder internationaal perspectief:
Een voorlopige vergelijking**



**Katholieke Universiteit Leuven
Onderwijseffectiviteit en -evaluatie**

**Barbara Belfi
Thomas Arkens
Jan Van Damme
Sarah Gielen**

April 2012

Inhoudsopgave

Woord vooraf	2
1. Inleiding	3
1.1 Wat is TIMSS?	3
1.2 Een vergelijking tussen 2003 en 2011	3
1.3 Wie nam er deel aan TIMSS?	3
2. Verschillen tussen leerlingen in wiskundeprestaties: Een vergelijking tussen 2003 en 2011	5
2.1 Vergelijking tussen de gemiddelde wiskundeprestaties in 2003 en 2011	5
2.2 Vergelijking tussen de gemiddelde wiskundeprestaties van jongens en meisjes in 2003 en 2011	5
2.3 Vergelijking tussen de gemiddelde wiskundeprestaties van leerlingen die al dan niet thuis Nederlands spreken in 2003 en 2011	7
2.4 Vergelijking tussen de gemiddelde wiskundeprestaties van leerlingen met verschillende hoeveelheden boeken thuis in 2003 en 2011	8
2.5 Vergelijking tussen de gemiddelde wiskundeprestaties van leerlingen met een verschillende geboortemaand in 2003 en 2011	10
2.6 Vergelijking tussen de gemiddelde wiskundeprestaties van leerlingen met een verschillend geboortjaar in 2003 en 2011	12
3. Verschillen tussen leerlingen in wiskunde zelfconcept: Een vergelijking tussen 2003 en 2011	15
3.1 Vergelijking tussen het gemiddelde wiskunde zelfconcept in 2003 en 2011	15
3.2 Vergelijking tussen het gemiddelde wiskunde zelfconcept van jongens en meisjes in 2003 en 2011	15
3.3 Vergelijking tussen het gemiddelde wiskunde zelfconcept van leerlingen die al dan niet thuis Nederlands spreken in 2003 en 2011	16
3.4 Vergelijking tussen het gemiddelde wiskunde zelfconcept van leerlingen met verschillende hoeveelheden boeken thuis in 2003 en 2011	17
3.5 Vergelijking tussen het gemiddelde wiskunde zelfconcept van leerlingen met een verschillende geboortemaand in 2003 en 2011	18
3.6 Vergelijking tussen het gemiddelde wiskunde zelfconcept van leerlingen met een verschillend geboortjaar in 2003 en 2011	19
4. Verschillen tussen scholen in wiskundeprestaties: Een vergelijking tussen 2003 en 2011	21
5. Verschillen tussen scholen in wiskunde zelfconcept: Een vergelijking tussen 2003 en 2011	27
6. Het effect van schoolsamenstelling naar thuistaal op wiskundeprestaties in 2003 en 2011	31
7. Het effect van schoolsamenstelling naar SES op wiskundeprestaties in 2003 en 2011	32
8. Conclusies en discussie	33
9. Referenties	36
Appendix	37

Woord vooraf

Omdat het bij internationaal onderzoek lang duurt vooraleer resultaten en gegevensbestanden beschikbaar zijn, hebben we wat betreft de Vlaamse deelname aan TIMSS 2011 (grade 4) reeds vooraf enige Vlaamse resultaten berekend. Daartoe werden de wiskundescores van alle Vlaamse leerlingen (die niet allemaal dezelfde toetsvragen beantwoord hadden) op één schaal gebracht. Dit gebeurde in eerste instantie voor 2011 afzonderlijk, en nadien voor 2003 en 2011 samen.

Dit maakte het mogelijk om alvast enige voorlopige resultaten betreffende de vergelijking tussen 2011 en 2003 voor te leggen en de meewerkende scholen enige feedback te geven over de resultaten van hun leerlingen, in vergelijking met een gemiddelde Vlaamse school met een gelijkaardig leerlingenpubliek.

De afwerking van dit eerste rapportje betreffende de Vlaamse deelname aan TIMSS 2011 (grade 4) is een goede aanleiding om sommige instanties en personen te danken. In de eerste plaats danken we de Vlaamse overheid, enerzijds voor de gedeeltelijke financiering van de Vlaamse deelname en anderzijds voor de financiering van analyse-opdrachten, zowel in het kader van het Steunpunt SSL als in het kader van een extra-opdracht. In de tweede plaats danken we de meewerkende scholen, ook enkele BuLO-scholen, in het bijzonder hun directies en leerkrachten en leerlingen van het vierde leerjaar. Last but not least danken we Thomas Arkens voor het uitvoeren van de kalibraties en Barbara Belfi voor de eigenlijke analyses en het rapporteren daarover.

Jan Van Damme en Sarah Gielen

1. Inleiding

In deze bijdrage belichten we de belangrijkste vaststellingen inzake de trends in wiskundeprestaties en wiskunde zelfconcept tussen 2003 en 2011 in Vlaanderen op basis van de TIMSS-resultaten. Hierbij staan drie onderzoeksvragen centraal:

- (1) Hoe verschillen de Vlaamse leerlingen die hebben deelgenomen aan TIMSS 2003 van de Vlaamse leerlingen die hebben deelgenomen aan TIMSS 2011 in wiskundeprestaties en wiskunde zelfconcept?
- (2) Hoe verschillen de Vlaamse scholen die hebben deelgenomen aan TIMSS 2003 van de Vlaamse scholen die hebben deelgenomen aan TIMSS 2011 in wiskundeprestaties en wiskunde zelfconcept?
- (3) Verschilt het effect van de leerlingensamenstelling van een school in termen van gemiddelde SES op wiskundeprestaties in 2003 en 2011?

1.1 Wat is TIMSS?

De 'Trends in International Mathematics and Science Study' (TIMSS) is een studie die de leerlingprestaties voor wiskunde en wereldoriëntatie internationaal in kaart brengt. De 'International Association for the Evaluation of Educational Achievement' (IEA) coördineert TIMSS. In de TIMSS-studie worden leerlingen getoetst na vier en acht jaar formele scholing. In Vlaanderen komt dit overeen met het vierde leerjaar lager onderwijs en het tweede leerjaar secundair onderwijs. Het Vlaamse lager onderwijs nam in 2011 voor de tweede keer deel aan TIMSS. De eerste Vlaamse deelname met het lager onderwijs was in 2003. De analyses die hier gepresenteerd worden hebben enkel betrekking op het domein wiskunde (en het academisch zelfconcept van leerlingen met betrekking tot wiskunde).

1.2 Een vergelijking tussen 2003 en 2011

De vergelijking van de scores van de leerlingen in het vierde leerjaar in 2011 met de scores van de leerlingen in 2003, kan ons een en ander leren over de trends of evoluties in het Vlaamse onderwijssysteem. De hier gepresenteerde analyses willen daartoe een aanzet doen.

Met behulp van de Item Respons Theorie, een statistische techniek, werden de Vlaamse wiskundescores van TIMSS 2003 en TIMSS 2011 op een gezamenlijke schaal gebracht, zodat de prestaties over de jaren heen vergeleken konden worden. In elke nieuwe ronde van TIMSS worden er vragen uit de voorgaande edities terug opgenomen. Door de verschillende jaren via deze ankeritems aan elkaar te koppelen, kan de vaardigheid van leerlingen uit de verschillende edities op een gemeenschappelijke schaal geplaatst worden.

De kalibratie werd in Vlaanderen met het softwareprogramma Bilog gerealiseerd en gebeurde vooraleer de gegevens van de andere landen beschikbaar waren. Er werden geen wegingen toegepast. Bij de kalibratie werd in Vlaanderen enkel naar het statistische plaatje gekeken. Om tot een optimale oplossing te komen werden 9 slecht fittende items stapsgewijs uit de kalibratie verwijderd. Dit is slechts één van vele mogelijke kalibratie-methoden. De in Vlaanderen berekende score blijkt een correlatie van 0.93 te vertonen met de naderhand internationaal berekende Rasch-score.

1.3 Wie nam er deel aan TIMSS?

In 2003 namen 149 scholen van het gewoon lager onderwijs met maximum 2 klassen van het vierde leerjaar deel aan de TIMSS studie. De totale steekproef bestond daarmee uit 4712 leerlingen.

In 2011 namen 141 scholen van het gewoon lager onderwijs met al hun klassen van het vierde leerjaar deel aan de TIMSS studie. De totale steekproef bestond daarmee uit 5000 leerlingen.¹ In 2011 was de getrokken steekproef gestratificeerd naar provincie, net (publiek versus vrij) en GOK-percentage (4 categorieën). Leerlingen die niet in staat waren aan het onderzoek deel te nemen omwille van een fysieke handicap, een mentale achterstand of een gebrek aan beheersing van de Nederlandse taal werden uitgesloten van het onderzoek.

¹ Naast scholen voor het gewoon lager onderwijs namen er in 2011 ook 5 scholen voor het buitengewoon onderwijs type 3 en type 8 deel aan de TIMSS-studie. Met de inclusie van deze types BLO-scholen werd gepoogd te voldoen aan het criterium om minder dan 5% van de leerlingen uit te sluiten, hetgeen in 2006 (bij de Vlaamse participatie aan PIRLS) niet gelukt was. Omdat de analyses in de huidige bijdrage tot doel hadden een 'eerlijke' vergelijking te maken tussen de prestaties in 2003 en 2011, werden de BLO-scholen uit de 2011-steekproef verwijderd. Om dezelfde reden werd voor de huidige analyses de data van een 'outlier-school' (m.a.w. een school die zeer afwijkende resultaten vertoont) verwijderd uit de 2011-steekproef. Voor meer informatie over de prestaties van de BLO-scholen en de outlier-school, zie Appendix.

2. Verschillen tussen leerlingen in wiskundeprestaties: Een vergelijking tussen 2003 en 2011

De TIMSS wiskundetoets heeft betrekking op de volgende deelgebieden: getallen, patronen en relaties, metingen, meetkunde en data (Meelissen & Drent, 2008). Bij de samenstelling van de toets wordt uitgegaan van een soort ‘internationaal curriculum’ waarin is aangegeven wat leerlingen in de betreffende fase van het onderwijs veelal leren (Mulis et al., 2005). Dit internationale curriculum is in eerste instantie vormgegeven door een internationale commissie van vakspecialisten vervolgens voorgelegd aan alle deelnemende landen. Als een onderwerp in de meeste landen niet onderwezen wordt, wordt dit onderdeel niet of slechts als een klein onderdeel opgenomen in de toets. Als een onderwerp in een beperkt aantal landen niet onderwezen wordt, dan wordt het wel opgenomen in de toets. Dit betekent dat leerlingen toetsopgaven kunnen krijgen die ze (nog) niet onderwezen hebben gekregen.

2.1 Vergelijking tussen de gemiddelde wiskundeprestaties in 2003 en 2011

In Tabel 1 worden de gemiddelde, minimum en maximum scores op het wiskundeonderdeel van de TIMSS 2003 en TIMSS 2011 toetsen weergegeven. De gemiddelde scores blijken niet significant van mekaar te verschillen² (zie Tabel 2). Gemiddeld genomen doen leerlingen het dus even goed in wiskunde in 2003 en 2011.

Tabel 1. Gemiddelde scores voor wiskunde in 2003 en 2011.

	Gemiddelde	SD	Minimum	Maximum
TIMSS 2003	49.84	9.70	9.68	91.43
TIMSS 2011	50.15	10.17	11.55	88.45

Tabel 2. Ongepaarde t-toets resultaat ter vergelijking van de gemiddelde wiskundescores in 2003 en 2011.

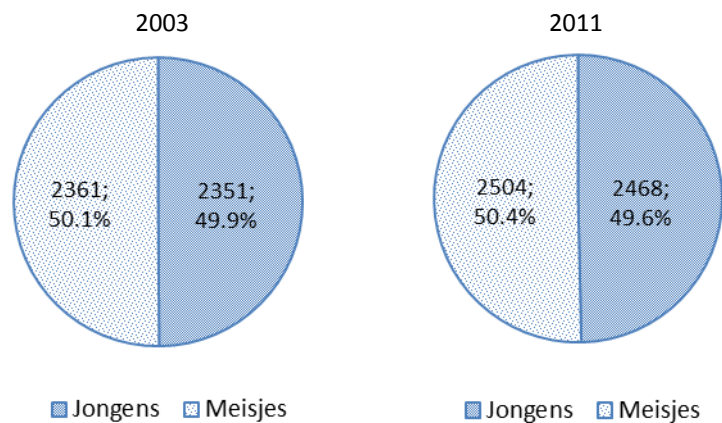
TIMSS 2003-TIMSS 2011	
t-waarde	-1.55
df	9710
P-waarde	.12

2.2 Vergelijking tussen de gemiddelde wiskundeprestaties van jongens en meisjes in 2003 en 2011

In Figuur 1 wordt een overzicht gegeven van de aantallen/percentages meisjes en jongens die deelnamen in TIMSS 2003 en TIMSS 2011. Het is duidelijk te zien dat in allebei de afnamejaren het aandeel meisjes/jongens gelijk verdeeld is over de steekproef.

In Tabel 3 worden de gemiddelde wiskundeprestaties weergegeven van jongens en meisjes in 2003 en 2011. Opvallend is dat in 2011 jongens een significant hogere score behalen dan meisjes, terwijl in 2003 de scores van jongens en meisjes niet significant van elkaar verschillen (zie Tabel 4). Dit komt doordat gemiddeld genomen de meisjes nagenoeg hetzelfde zijn blijven presteren in 2011 terwijl de jongens significant beter presteren in 2011 dan in 2003 (zie Tabel 5).

² Uitspraken over statistische significantie gebeuren altijd met een zekere mate van onzekerheid. In de analyses doorheen dit rapport gebruiken wij het zogenaamde “5 %-niveau”. Als we een verschil “statistisch significant” noemen, bedoelen we dat er ten hoogste 5% kans is dat het verschil wél op toeval berust. Als we de meting 100 keer zouden uitvoeren, zouden we minstens 95 keer tot de conclusie komen dat er zo’n verschil is en hoogstens 5 keer tot de conclusie komen dat er geen verschil is of een verschil in de omgekeerde richting.



Figuur 1. Verdeling meisjes/jongens in 2003 en in 2011.

Tabel 3. Gemiddelde scores voor wiskunde van jongens en meisjes in 2003 en in 2011.

	Gem. meisjes (SD)	Gem. jongens (SD)	Verschil (jongens-meisjes)
TIMSS 2003	49.62	50.06	0.44
TIMSS 2011	49.58	50.80	1.22

Tabel 4. Ongepaarde t-toets resultaten ter vergelijking van het verschil tussen de gemiddelde wiskundescores van jongens en meisjes in 2003 en 2011.

	Jongens-Meisjes 2003	Jongens-Meisjes 2011
t-waarde	1.56	4.20
df	4710	4970
P-waarde	.12	<.01

Tabel 5. Ongepaarde t-toets resultaten ter vergelijking van de gemiddelde wiskundescores van jongens (en van meisjes) in 2003 en 2011.

	Jongens 2003-2011	Meisjes 2003-2011
t-waarde	-2.49	0.14
df	4817	4863
P-waarde	.01	.89

Echter, bovenstaande resultaten geven slechts gedeeltelijk inzicht in de werkelijke prestatieverschillen tussen jongens en meisjes, aangezien jongens doorgaans veel vaker blijven zitten dan meisjes (en ook sterker vertegenwoordigd zijn in het BLO). Om een meer faire vergelijking te maken tussen de wiskundeprestaties van jongens en meisjes kan men overwegen alleen te kijken naar de normaalvorderende leerlingen. In Tabel 6 staan de gemiddelde wiskundeprestaties van de normaalgevorderde jongens en meisjes in 2003 en 2011. Te zien is dat in beide afnamejaren jongens iets hoger scoren dan meisjes. Dit is niet echt verrassend aangezien de normaalgevorderden bij de jongens een meer selectieve groep vormen dan bij de meisjes. Het verschil tussen jongens en meisjes is significant in 2011 en marginaal significant in 2003 (zie Tabel 7).

Tabel 6. Gemiddelde scores voor wiskunde van normaalvorderende jongens en meisjes in 2003 en 2011.

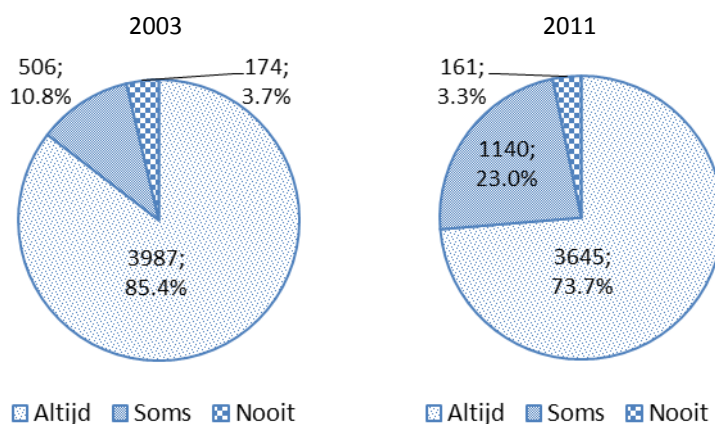
	Gem. meisjes (SD)	Gem. jongens (SD)	Verschil (jongens-meisjes)
TIMSS 2003	50.80 (9.24)	51.37 (9.57)	0.57
TIMSS 2011	50.77 (9.37)	51.99 (10.25)	1.22

Tabel 7. Ongepaarde t-toets resultaten ter vergelijking van het verschil tussen de gemiddelde wiskundescores van normaalvorderende jongens (en van meisjes) in 2003 en 2011.

	Jongens-Meisjes 2003	Jongens-Meisjes 2011
t-waarde	1.90	3.98
df	3923	4104
P-waarde	.058	< .01

2.3 Vergelijking tussen de gemiddelde wiskundeprestaties van leerlingen die al dan niet thuis Nederlands spreken in 2003 en 2011

In Figuur 2 wordt een overzicht gegeven van de aantallen/percentages leerlingen die thuis altijd (of bijna altijd), soms en nooit thuis Nederlands spreken in 2003 en 2011. Het is duidelijk te zien dat in allebei de afnamejaren het grootste aandeel van de leerlingen thuis (bijna) altijd Nederlands spreekt, een aanzienlijk kleiner aandeel slechts soms thuis Nederlands spreekt en een nog kleiner aandeel thuis nooit Nederlands spreekt. De exacte verdeling tussen de drie groepen verschilt echter per afnamejaar. Hoewel het aandeel leerlingen dat thuis nooit Nederlands spreekt nagenoeg identiek is in beide afnamejaren, is er anno 2011 sprake van een aanzienlijk kleiner aandeel leerlingen dat thuis (bijna) altijd Nederlands spreekt, en een aanzienlijk groter aandeel leerlingen dat slechts soms Nederlands spreekt, dan in 2003.



Figuur 2. Verdeling van leerlingen die thuis (bijna) altijd, soms en nooit Nederlands spreken in 2003-2011.

In Tabel 8 worden de gemiddelde wiskundescores weergegeven van leerlingen die thuis altijd (of bijna altijd), soms en nooit thuis Nederlands spreken in 2003 en 2011. In beide afnamejaren doen de leerlingen die thuis (bijna) altijd Nederlands spreken, het significant beter dan leerlingen die dat soms of helemaal niet doen. Opmerkelijk is evenwel dat in 2011, de scores van de leerlingen die slechts soms of helemaal nooit thuis Nederlands spreken niet significant van elkaar verschillen, terwijl in 2003 de leerlingen die nooit thuis Nederlands spreken het juist significant beter doen dan leerlingen die dat soms doen (zie Tabel 9).

Wanneer wordt gekeken of de scores per categorie verschillend zijn per afnamejaar, blijkt dat de leerlingen die (bijna) altijd thuis Nederlands spreken in 2011 significant beter presteren in wiskunde dan in 2003. Ook de leerlingen die soms Nederlands thuis spreken doen het in 2011 significant beter dan in 2003. De leerlingen die thuis nooit Nederlands spreken verschillen echter niet significant van elkaar in beide afnamejaren (zie Tabel 10).

Tabel 8. Gemiddelde scores voor wiskunde van leerlingen die thuis altijd (of bijna altijd), soms en nooit thuis Nederlands spreken in 2003 en 2011.

	Gem. (bijna) altijd NL (SD)	Gem. soms NL (SD)	Gem. nooit NL (SD)
TIMSS 2003	50.62 (9.66)	44.84 (8.45)	46.90 (9.67)
TIMSS 2011	51.45 (10.13)	46.89 (9.03)	46.78 (10.09)

Tabel 9. Ongepaarde t-toets resultaten ter vergelijking van de gemiddelde wiskundescores van leerlingen die altijd, soms of nooit thuis Nederlands spreken in 2003 en 2011: vergelijking binnen afnamejaren.

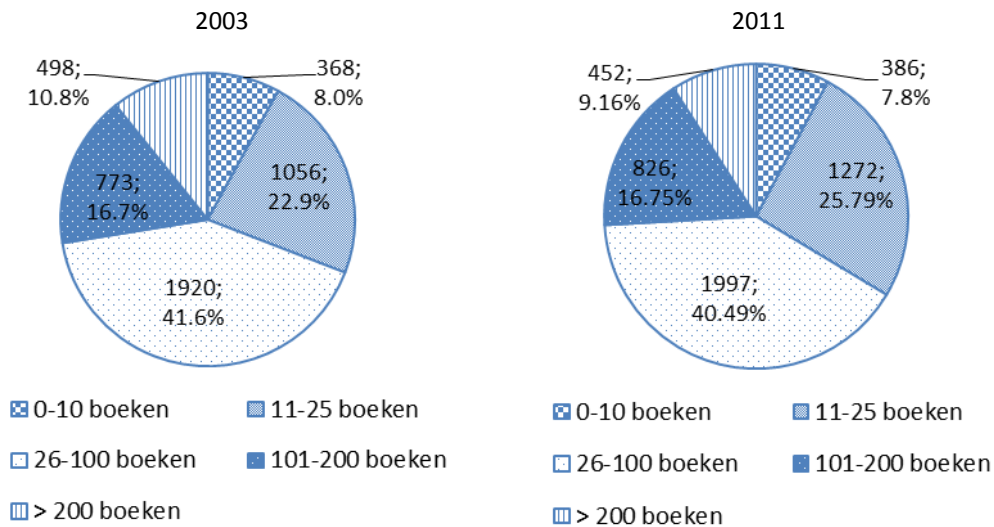
	Altijd-Soms 2003	Altijd-Soms 2011	Soms-Nooit 2003	Soms-Nooit 2011
t-waarde	12.85	13.43	-2.67	0.14
df	4491	4783	678	1299
P-waarde	< .01	< .01	< .01	.89

Tabel 10. Ongepaarde t-toets resultaten ter vergelijking van de gemiddelde wiskundescores van leerlingen die altijd, soms of nooit thuis Nederlands spreken in 2003 en 2011: vergelijking tussen afnamejaren.

	Altijd 2003-2011	Soms 2003-2011	Nooit 2003-2011
t-waarde	-3.66	-4.33	0.11
df	7630	1644	333
P-waarde	< .01	< .01	.91

2.4 Vergelijking tussen de gemiddelde wiskundeprestaties van leerlingen met verschillende hoeveelheden boeken thuis in 2003 en 2011

Omdat uit eerdere TIMSS-studies gebleken is dat het achterhalen van het opleidingsniveau van de ouders bij leerlingen van het vierde leerjaar veel ontbrekende waarden oplevert, is al sinds 1995 in de TIMSS-leerlingvragenlijst de proxy-indicator voor SES 'het aantal boeken thuis' opgenomen (Meelissen & Drent, 2008). In Figuur 3 wordt een overzicht gegeven van de aantallen/percentages leerlingen die thuis geen of erg weinig boeken hebben (0-10 boeken), genoeg boeken hebben om een boekenplank mee te vullen (11-25 boeken), genoeg boeken hebben om een boekenkast mee te vullen (26-100 boeken), genoeg boeken hebben om twee boekenkasten mee te vullen (101-200 boeken) en genoeg boeken hebben om drie of meer boekenkasten mee te vullen (meer dan 200 boeken). Uit Figuur 3 blijkt allereerst dat de verdeling tussen de 5 categorieën in beide afnamejaren nagenoeg dezelfde is. Het grootse aandeel van de leerlingen beschikt thuis over 26-100 boeken (1 boekenkast), een iets kleiner aandeel van de leerlingen beschikt thuis over 11-25 boeken (1 boekenplank), een nog iets kleiner aandeel van de leerlingen beschikt thuis over 101-200 boeken (2 boekenkasten), hierop volgt een nog kleiner aandeel van leerlingen dat thuis meer dan 200 boeken heeft (>2 boekenkasten). Tenslotte beschikt het kleinste aandeel leerlingen thuis over 0-10 boeken.



Figuur 3. Aantal en percentage leerlingen die thuis geen of zeer weinig boeken hebben (0-10 boeken), genoeg boeken hebben om een boekenplank mee te vullen (11-25 boeken), genoeg boeken hebben om een boekenkast mee te vullen (26-100 boeken), genoeg boeken hebben om twee boekenkasten mee te vullen (101-200 boeken) en genoeg boeken hebben om drie of meer boekenkasten mee te vullen (>200 boeken) in 2003-2011.

In Tabel 11 worden de gemiddelde wiskundescores weergegeven van leerlingen die thuis (0-10 boeken, genoeg boeken hebben om een boekenplank mee te vullen (11-25 boeken), genoeg boeken hebben om een boekenkast mee te vullen (26-100 boeken), genoeg boeken hebben om twee boekenkasten mee te vullen (101-200 boeken) en genoeg boeken hebben om drie of meer boekenkasten mee te vullen (>200 boeken) in 2003 en in 2011.

In beide afnamejaren doen leerlingen die thuis over 11-25 boeken beschikken het significant beter in wiskunde dan leerlingen die thuis over 0-10 boeken beschikken. Ook doen in beide afnamejaren leerlingen die thuis over 26-100 boeken beschikken het significant beter in wiskunde dan leerlingen die thuis over 11-25 boeken beschikken. In 2003 doen leerlingen die thuis 101-200 boeken hebben het significant beter dan leerlingen die thuis 26-100 boeken hebben, in 2011 is er echter geen sprake van een significant verschil tussen deze twee categorieën. Tenslotte is er in beide afnamejaren geen sprake van een significant verschil in de wiskundescores van leerlingen die thuis 101-200 boeken hebben en leerlingen die thuis meer dan 200 boeken hebben (zie Tabel 12). Wanneer wordt gekeken of de scores per categorie verschillend zijn per afnamejaar, blijkt dat geen enkele categorie significant slechter of beter scoort in 2003 of in 2011 (zie Tabel 13).

Tabel 11. Gemiddelde scores voor wiskunde van leerlingen die thuis over verschillende hoeveelheden boeken beschikken in 2003 en 2011.

	0-10 (SD)	11-25 (SD)	26-100 (SD)	101-200 (SD)	>200 (SD)
TIMSS 2003	44.76 (9.45)	47.91 (9.14)	50.48 (9.51)	52.09 (9.42)	51.62 (9.89)
TIMSS 2011	45.33 (9.70)	48.38 (9.30)	51.32 (10.04)	51.75 (9.94)	52.69 (10.69)

Tabel 12. Ongepaarde t-toets resultaten ter vergelijking van de gemiddelde wiskundescores van leerlingen met verschillende hoeveelheden boeken 2003 en 2011: vergelijking binnen afnamejaren.

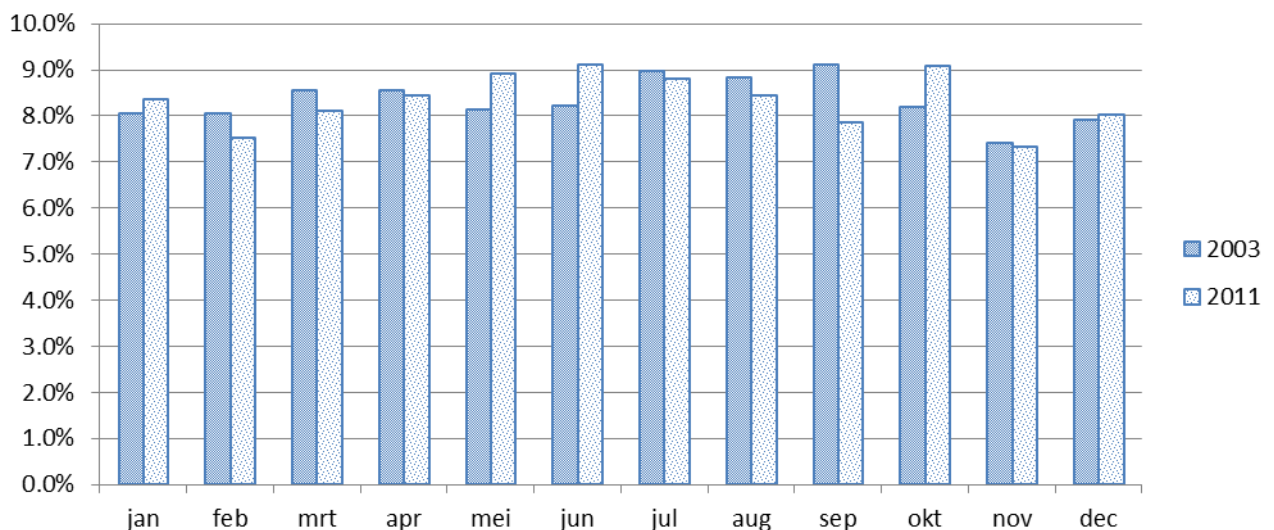
	0-10/11-25 2003	11-25/26-100 2003	26-100/101-200 2003	101-200/>200 2003
t-waarde	-5.66	-7.07	-2.92	0.87
df	1422	2110	1827	1569
P-waarde	< .01	< .01	< .01	.38
	0-10/11-25 2011	11-25/26-100 2011	26-100/101-200 2011	101-200/>200 2011
t-waarde	-5.58	-8.38	-1.04	-1.57
df	1655	3249	2803	1276
P-waarde	< .01	< .01	.30	.12

Tabel 13. Ongepaarde t-toets resultaten ter vergelijking van de gemiddelde wiskundescores van leerlingen met verschillende hoeveelheden boeken 2003 en 2011: vergelijking tussen afnamejaren.

	0-10 2003-2011	11-25 2003-2011	26-100 2003-2011	101-200 2003-2011	>200 2003-2011
t-waarde	-0.82	-1.22	-1.82	0.70	-1.62
df	751	2326	3897	1597	948
P-waarde	.41	.22	.07	.48	.11

2.5 Vergelijking tussen de gemiddelde wiskundeprestaties van leerlingen met een verschillende geboortemaand in 2003 en 2011

In Figuur 4 wordt een overzicht gegeven van de percentages leerlingen die in de afzonderlijke maanden zijn geboren in 2003 en 2011. Het is duidelijk te zien dat in beide afnamejaren de leerlingen redelijk evenredig verdeeld zijn over de twaalf maanden. Ook tussen de afnamejaren is er sprake van slechts kleine verschillen in percentages per maand.



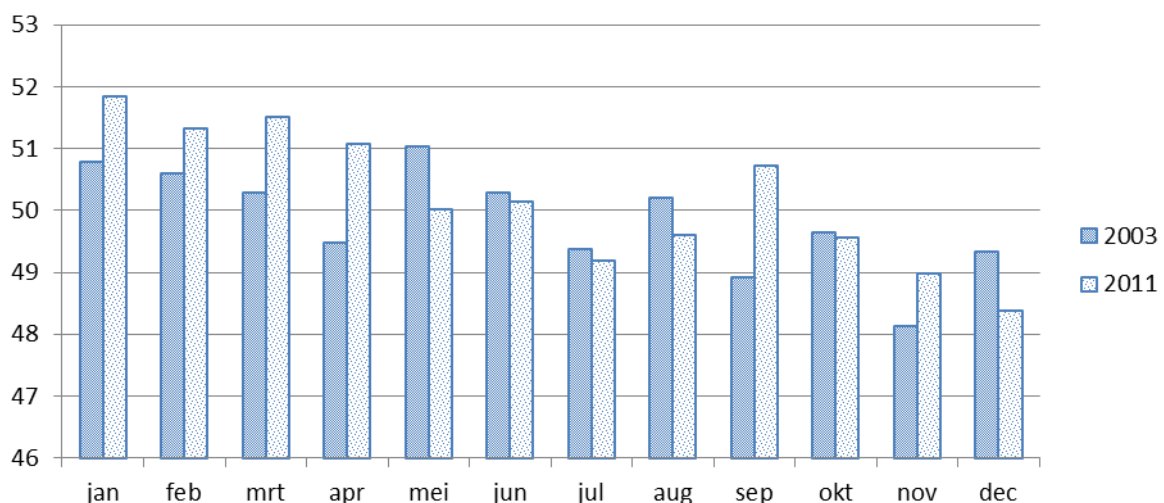
Figuur 4. Verdeling van leerlingen over de 12 geboortemaanden in 2003 en in 2011.

In Figuur 5 en Tabel 14 worden de gemiddelde wiskundeprestaties weergegeven van leerlingen die in de afzonderlijke maanden geboren zijn in 2003 en 2011. We hebben naar de prestaties gekeken van alle leerlingen, niet slechts naar die van de normaalvorderende leerlingen. Vanwege het relatieve leeftijdsverschil zullen december-leerlingen eerder vertraagd zijn en januari-leerlingen eerder versneld. Een analyse gebaseerd op enkel de normaalvorderende leerlingen kan zodoende andere resultaten opleveren.

In 2003 zijn het de leerlingen die in mei zijn geboren die gemiddeld het hoogst presteren op de wiskundetoets en de leerlingen die in november zijn geboren die gemiddeld het laagst presteren. Het verschil in prestatie tussen beide groepen leerlingen bedraagt 3.47 punten en is statistisch significant. In 2011 is een ander patroon zichtbaar. In dit afnamejaar presteren de leerlingen die in januari geboren zijn gemiddeld het best, en de leerlingen die in december geboren zijn gemiddeld het laagst. Het verschil tussen beide groepen leerlingen bedraagt 2.92 punten en is statistisch significant.

De vaststelling dat de januari-leerlingen relatief het best presteren en de december-leerlingen relatief het laagst, is een welbekend fenomeen binnen het onderwijsonderzoek en wordt ook wel het ‘geboortemaand-effect’ genoemd. Het geboortemaand-effect houdt in dat de relatieve leeftijdspositie die leerlingen innemen binnen hun klasgroep, hun leerprestaties beïnvloedt. Meer specifiek voorspelt het geboortemaand-effect dat de jongste leerlingen binnen een klasgroep (in dit geval de december-leerlingen) het beduidend slechter zullen doen dan de oudste leerlingen (in dit geval de januari-leerlingen). De data laten zien dat dit in 2011 ook duidelijk het geval is. Hoewel in 2003 de januari- en december-leerlingen er niet uitspringen als respectievelijk de ‘beste’ en ‘zwakste’ presteerders van de hele leerlingengroep, doen ook in dit jaar de januari-leerlingen het significant beter op de wiskundetoets dan de december-leerlingen.

Wanneer we de gemiddelde wiskundeprestaties van de leerlingen die in de afzonderlijke maanden zijn geboren vergelijken tussen de afnamejaren, blijkt dat er vrijwel geen verschillen zijn (zie Tabel 15). Enkel in de maand september zijn de verschillen tussen de afnamejaren statistisch significant.



Figuur 5. Gemiddelde scores voor wiskunde van leerlingen die in de afzonderlijke maanden geboren zijn in 2003 en in 2011.

Tabel 14. Gemiddelde scores voor wiskunde van leerlingen die in de afzonderlijke maanden geboren zijn in 2003 en 2011.

	jan (SD)	feb (SD)	mrt (SD)	apr (SD)
TIMSS 2003	50.79 (9.43)	50.60 (9.21)	50.29 (9.80)	49.48 (9.55)
TIMSS 2011	51.84 (11.22)	51.33 (10.68)	51.51 (10.68)	51.07 (10.34)
	mei (SD)	jun (SD)	jul (SD)	aug (SD)
TIMSS 2003	51.04 (10.34)	50.29 (9.87)	49.37 (10.20)	50.20 (10.06)
TIMSS 2011	50.02 (9.35)	50.14 (10.69)	49.19 (9.37)	49.59 (9.67)
	sep (SD)	okt (SD)	nov (SD)	dec (SD)
TIMSS 2003	48.91 (9.73)	49.64 (9.06)	48.12 (9.72)	49.33 (8.98)
TIMSS 2011	50.72 (10.22)	49.55 (9.56)	48.98 (9.87)	48.37 (9.57)

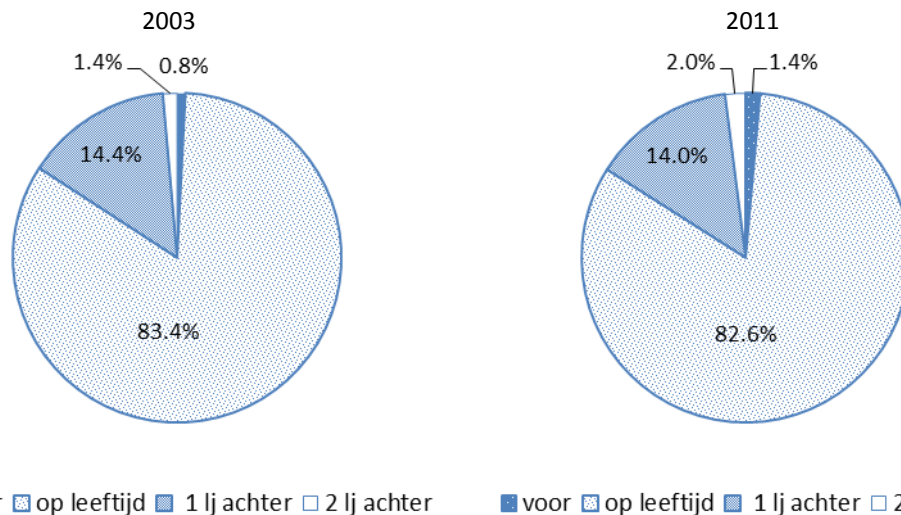
Tabel 15. Ongepaarde t-toets resultaten ter vergelijking van de gemiddelde wiskundeprestaties van leerlingen die in de afzonderlijke maanden geboren zijn: vergelijking tussen 2003 en 2011.

	jan 2003-2011	feb 2003-2011	mrt 2003-2011	apr 2003-2011	mei 2003-2011	jun 2003-2011
t-waarde	-1.42	-1.01	-1.69	-1.80	1.49	0.21
df	794	752	804	821	824	838
P-waarde	.16	.31	.09	.07	.14	.83
	jul 2003-2011	aug 2003-2011	sep 2003-2011	okt 2003-2011	nov 2003-2011	dec 2003-2011
t-waarde	0.27	0.89	-2.60	0.14	-1.17	1.43
df	859	833	817	836	712	769
P-waarde	.79	.37	< .01	.89	.24	.15

2.6 Vergelijking tussen de gemiddelde wiskundeprestaties van leerlingen met een verschillend geboortjaar in 2003 en 2011

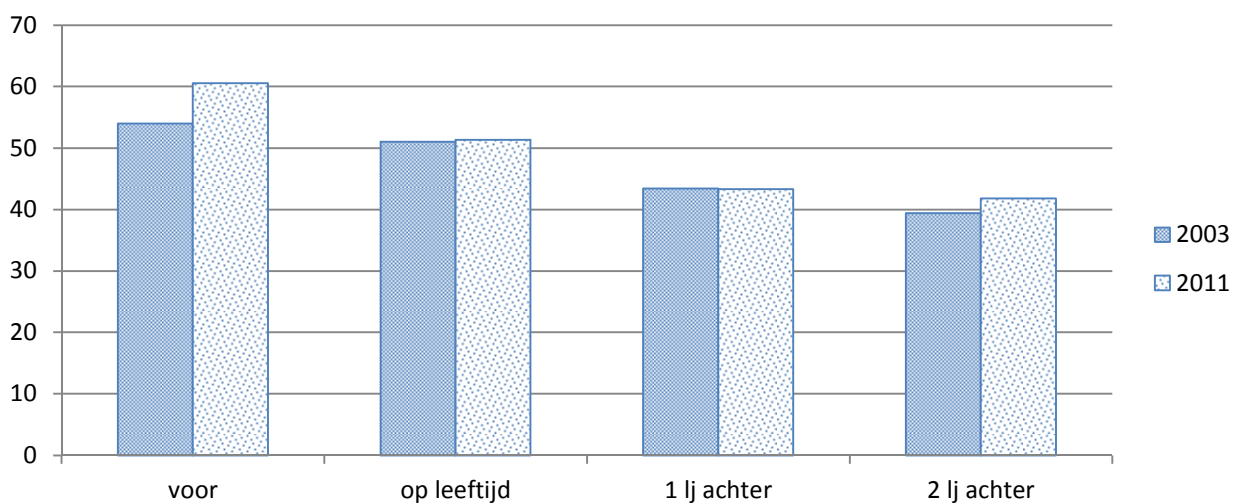
Zowat alle Vlaamse leerlingen zitten in een bepaald leerjaar. Bij dat leerjaar hoort een 'normale' leeftijd. Door de combinatie van het geboortjaar van een leerling met het leerjaar waarin hij/zij in een bepaald schooljaar zit, kunnen we bepalen of een leerling 'op tijd' is of niet. Leerlingen die 'op leeftijd' zitten zijn de normaalvorderende leerlingen. Leerlingen die niet 'op leeftijd' zitten kunnen voorsprong hebben of één of meerdere leerjaren vertraging (Van Landeghem & Van Damme, 2004).

In deze paragraaf gaan we na, in hoeverre leerlingen die vertraagd of versneld zijn, gemiddeld beter of slechter presteren. We kijken daarbij niet naar in welk leerjaar de leerling is blijven zitten. Het kan dus zijn dat de leerling in het jaar van de toetsafname in hun bisjaar zitten of dat de vertraging al eerder is opgelopen. In Figuur 6 wordt allereerst getoond welk aandeel leerlingen er één of meerdere leerjaren vertraagd of versneld is de afnamejaren 2003 en 2011. Te zien is dat het merendeel van de leerlingen in beide afnamejaren gewoon 'op leeftijd' zit (+/-83%). In beide afnamejaren bestaat de op één na grootste groep leerlingen uit leerlingen die één leerjaar vertraagd zijn (+/- 14%). De op één na kleinste groep bestaat uit leerlingen die twee of meer leerjaren vertraagd zijn (+/- 2%). Het aandeel leerlingen dat één of meerdere leerjaren versneld is, is in beide afnamejaren zeer klein (+/- 1%). De meest opvallende wijziging is de toename van het aandeel versnelde leerlingen.



Figuur 6. Verdeling van leerlingen die voor (versneld), op leeftijd, 1 leerjaar of 2 leerjaren achter (vertraagd) zijn, in 2003 en 2011.

In Figuur 7 en Tabel 16 worden de gemiddelde wiskundeprestaties van de verschillende categorieën leerlingen in 2003 en 2011 weergegeven. Allereerst valt op dat in beide afnamejaren leerlingen die één leerjaar vertraagd zijn, gemiddeld +/- 8 punten lager scoren op de wiskundetoets dan leerlingen die 'op leeftijd' zitten. Dit verschil is statistisch significant (zie Tabel 17). Leerlingen die meerdere leerjaren vertraagd zijn presteren op hun beurt in beide afnamejaren lager dan leerlingen die slechts één leerjaar vertraagd zijn. Het verschil tussen beide categorieën is in 2003 statistisch significant maar niet in 2011 (zie Tabel 17). In beide afnamejaren doen leerlingen die één of twee leerjaren voorliggen ('versneld zijn') het beter dan leerlingen die 'op leeftijd' zitten. De voorsprong van versnelde leerlingen is echter veel groter in 2011 dan in 2003. Het verschil tussen 'versnelde' leerlingen en leerlingen 'op leeftijd' is dan ook significant in 2011 en marginaal significant in 2003. Wanneer we de prestaties van de vier categorieën leerlingen tussen de afnamejaren met elkaar vergelijken, valt op dat enkel de prestaties van de versnelde leerlingen significant verschillend zijn tussen 2003 en 2011 (zie Tabel 18).



Figuur 7. Gemiddelde scores voor wiskunde van leerlingen die vertraagd, versneld, of op leeftijd zitten, in 2003 en 2011.

Tabel 16. Gemiddelde scores voor wiskunde van leerlingen die 1 of 2 leerjaren vertraagd, versneld, of op leeftijd zijn in 2003 en 2011.

	Voor (SD)	Op leeftijd (SD)	Achter (1 lj) (SD)	Achter (2+ lj) (SD)
TIMSS 2003	54.00 (8.60)	51.08 (9.41)	43.47 (8.23)	39.42 (9.51)
TIMSS 2011	60.54 (11.18)	51.38 (9.83)	43.34 (8.22)	41.85 (8.39)

Tabel 17. Ongepaarde t-toets resultaten ter vergelijking van de gemiddelde wiskundeprestaties van leerlingen die 1 of 2 leerjaren vertraagd, versneld, of op leeftijd zijn in 2003 en 2011: vergelijking binnen afnamejaren.

	voor-op leeftijd 2003	op leeftijd-achter (1lj) 2003	achter (1 lj)- achter (+2 lj) 2003
t-waarde	1.96	19.79	3.79
df	3963	4603	745
P-waarde	.05	< .01	< .01
	voor-op leeftijd 2011	op leeftijd-achter (1lj) 2011	achter (1 lj)- achter (+2 lj) 2011
t-waarde	7.60	20.41	1.69
df	4172	4799	794
P-waarde	< .01	< .01	.09

Tabel 18. Ongepaarde t-toets resultaten ter vergelijking van de gemiddelde wiskundeprestaties van leerlingen die 1 of 2 leerjaren vertraagd, versneld of op leeftijd zijn in 2003 en 2011: vergelijking tussen afnamejaren.

	Voor 2003-2011	Op leeftijd 2003-2011	Achter (1 lj) 2003-2011	Achter (+2 lj) 2003-2011
t-waarde	-3.18	-1.40	0.07	-1.74
df	106	8029	1373	166
P-waarde	< .01	.16	.95	.08

3. Verschillen tussen leerlingen in wiskunde zelfconcept: Een vergelijking tussen 2003 en 2011

In de TIMSS leerlingvragenlijst van 2003 en 2011 werd middels vier items getoetst hoe zelfzeker de leerlingen waren over hun eigen wiskundige vaardigheden. Deze items waren: (1) Meestal ben ik goed in wiskunde, (2) Wiskunde is voor mij moeilijker dan voor veel van mijn klasgenoten, (3) Ik ben gewoon niet goed in wiskunde, (4) Ik leer vrij snel dingen over wiskunde. Er waren telkens 4 antwoordcategorieën, gaande van 'helemaal niet akkoord' tot 'helemaal akkoord'³. Het gemiddelde van de vier items werd genomen als indicatie van het wiskunde zelfconcept van de leerling (Cronbach's alpha 2003=0.82; Cronbach's alpha 2011=0.82). Hoe hoger de score op deze schaal, hoe positiever een leerling over zijn/haar wiskundige vaardigheden denkt.

3.1 Vergelijking tussen het gemiddelde wiskunde zelfconcept in 2003 en 2011

In Tabel 19 worden de gemiddelde, minimum en maximum scores op de schaal 'wiskunde zelfconcept' van TIMSS 2003 en TIMSS 2011 weergegeven. De gemiddelde scores verschillen significant van elkaar (zie Tabel 20). Gemiddeld genomen dachten de leerlingen in 2011 positiever over hun eigen wiskundevaardigheden dan in 2003.

Tabel 19. Gemiddeld wiskunde zelfconcept in 2003 en 2011.

	Gemiddelde	SD	Minimum	Maximum
TIMSS 2003	2.99	0.67	1	4
TIMSS 2011	3.10	0.75	1	4

Tabel 20. Ongepaarde t-toets resultaten ter vergelijking van het gemiddelde wiskunde zelfconcept in 2003 en 2011.

	TIMSS 2003- TIMSS 2011
t-waarde	-7.49
df	9421
P-waarde	< .01

3.2 Vergelijking tussen het gemiddelde wiskunde zelfconcept van jongens en meisjes in 2003 en 2011

In Tabel 21 wordt het gemiddelde wiskunde zelfconcept weergegeven van jongens en meisjes in 2003 en 2011. In beide afnamejaren hebben jongens een significant hoger wiskunde zelfconcept dan meisjes (zie Tabel 22). Deze bevinding is overeenkomstig met wat doorgaans in onderwijsonderzoek wordt vastgesteld. Vermeldenswaard is dat het verschil een toenemende tendens vertoont. Daarnaast ligt zowel het wiskunde zelfconcept van de jongens als dat van de meisjes significant hoger in 2011 dan in 2003 (zie Tabel 23).

Tabel 21. Gemiddeld wiskunde zelfconcept van jongens en meisjes in 2003 en 2011.

	Gem. meisjes (SD)	Gem. jongens (SD)	Vershil (jongens-meisjes)
TIMSS 2003	2.87 (0.68)	3.12 (0.65)	0.25
TIMSS 2011	2.94 (0.77)	3.26 (0.69)	0.32

³ De negatieve stellingen werden omgezet naar positieve stellingen voor de creatie van de schaal.

Tabel 22. Ongepaarde t-toets resultaten ter vergelijking van het verschil tussen het gemiddelde wiskunde zelfconcept van jongens en meisjes in 2003 en 2011.

	Jongens-Meisjes 2003	Jongens-Meisjes 2011
t-waarde	13.14	14.98
df	4576	4843
P-waarde	< .01	< .01

Tabel 23. Ongepaarde t-toets resultaten ter vergelijking van het gemiddelde wiskunde zelfconcept van jongens en meisjes tussen 2003 en 2011.

	Jongens 2003-2011	Meisjes 2003-2011
t-waarde	-6.77	-3.62
df	4671	4748
P-waarde	< .01	< .01

3.3 Vergelijking van het gemiddelde wiskunde zelfconcept van leerlingen die al dan niet thuis Nederlands spreken tussen 2003 en 2011

In Tabel 24 wordt het gemiddelde wiskunde zelfconcept weergegeven van leerlingen die thuis altijd (of bijna altijd), soms en nooit Nederlands spreken in 2003 en 2011. Als eerste valt op dat alle leerlingen die thuis (bijna) altijd of soms Nederlands spreken een significant hoger wiskunde zelfconcept vertonen in 2011 dan in 2003. Het wiskunde zelfconcept van leerlingen die thuis nooit Nederlands spreken verschilt echter niet tussen 2003 en 2011, mede omwille van de kleine aantallen (zie Tabel 25).

Wanneer de drie categorieën onderling worden vergeleken, blijkt dat in beide afnamejaren leerlingen die thuis nooit Nederlands spreken een nagenoeg even hoog wiskunde zelfconcept hebben als leerlingen die thuis (bijna) altijd Nederlands spreken. Leerlingen die thuis slechts soms Nederlands spreken vertonen in beide afnamejaren een significant lager wiskunde zelfbeeld dan de leerlingen die thuis altijd Nederlands spreken. Het wiskunde zelfconcept van leerlingen die soms en nooit Nederlands spreken verschilt overigens niet significant, eveneens vooral omwille van de kleine aantallen (zie Tabel 26).

Tabel 24. Gemiddeld wiskunde zelfconcept van leerlingen die thuis altijd (of bijna altijd), soms en nooit thuis Nederlands spreken in 2003 en 2011.

	Gem. (bijna) altijd (SD)	Gem. soms (SD)	Gem. nooit (SD)
TIMSS 2003	3.00 (0.66)	2.92 (0.69)	3.02 (0.72)
TIMSS 2011	3.11 (0.76)	3.04 (0.72)	3.10 (0.77)

Tabel 25. Ongepaarde t-toets resultaten ter vergelijking van het gemiddelde wiskunde zelfconcept van leerlingen die altijd, soms of nooit thuis Nederlands spreken tussen 2003 en 2011.

	Altijd 2003-2011	Soms 2003-2011	Nooit 2003-2011
t-waarde	-6.72	-3.00	-1.02
df	7460	1590	323
P-waarde	< .01	< .01	.31

Tabel 26. Ongepaarde t-toets resultaten ter vergelijking van het gemiddelde wiskunde zelfconcept van leerlingen die altijd, soms of nooit thuis Nederlands spreken in 2003 en 2011 (binnen afnamejaren-vergelijking).

	Altijd-Soms 2003	Soms-Nooit 2003	Altijd-Nooit 2003
t-waarde	2.16	-1.28	-0.19
df	4380	651	4065
P-waarde	.03	.20	.85
	Altijd-Soms 2011	Soms-Nooit 2011	Altijd-Nooit 2011
t-waarde	2.54	-0.97	0.16
df	4670	1262	3718
P-waarde	.01	.33	.87

3.4 Vergelijking tussen het gemiddelde wiskunde zelfconcept van leerlingen met verschillende hoeveelheden boeken thuis in 2003 en 2011

In Tabel 27 wordt het gemiddelde wiskunde zelfconcept weergegeven van leerlingen die thuis geen of bijna geen boeken hebben (0-10 boeken), genoeg boeken hebben om een boekenplank mee te vullen (11-25 boeken), genoeg boeken hebben om een boekenkast mee te vullen (26-100 boeken), genoeg boeken hebben om twee boekenkasten mee te vullen (101-200 boeken) en genoeg boeken hebben om drie of meer boekenkasten mee te vullen (>200 boeken) in 2003-2011.

Als eerste valt op dat op alle leerlingcategorieën, behalve de leerlingen die thuis de minste boeken (0-10 boeken) en de meeste boeken (>200 boeken) hebben, over een significant hoger wiskunde zelfconcept beschikken in 2011 dan in 2003 (zie Tabel 28).

Wanneer we kijken naar de categorieën onderling, valt op dat in zowel 2003 als 2011 de leerlingen met de minste boeken thuis (0-10 boeken) gemiddeld het laagste wiskunde zelfconcept hebben en de leerlingen met de meeste boeken thuis (≥ 200 boeken) gemiddeld het hoogste wiskunde zelfconcept.

In allebei de afnamejaren blijken verder de laagste twee groepen nauwelijks van elkaar te verschillen in wiskunde zelfconcept (zie Tabel 29). De middelste groep (26-100 boeken) heeft een significant hoger wiskunde zelfconcept dan de laagste twee groepen in beide afnamejaren. In beide afnamejaren heeft verder de op een na hoogste groep (101-200 boeken) een significant hoger wiskunde zelfconcept dan de middelste groep. Het wiskunde zelfconcept van de hoogste groep (≥ 200 boeken) verschilt niet van dat van de op één na hoogste groep in beide afnamejaren.

Tabel 27. Gemiddeld wiskunde zelfconcept van leerlingen die thuis over verschillende hoeveelheden boeken beschikken in 2003 en 2011.

	Gem. 0-10 (SD)	Gem. 11-25 (SD)	Gem. 26-100 (SD)	Gem. 101-200 (SD)	Gem. >200 (SD)
TIMSS 2003	2.90 (0.67)	2.93 (0.66)	2.99 (0.66)	3.06 (0.69)	3.14 (0.70)
TIMSS 2011	3.00 (0.77)	3.01 (0.75)	3.11 (0.74)	3.18 (0.74)	3.23 (0.76)

Tabel 28. Ongepaarde t-toets resultaten ter vergelijking van het gemiddelde wiskunde zelfconcept van leerlingen met verschillende hoeveelheden boeken tussen 2003 en 2011.

	0-10 2003-2011	11-25 2003-2011	26-100 2003-2011	101-200 2003-2011	>200 2003-2011
t-waarde	-1.47	-2.63	-5.56	-3.31	-1.94
df	727	2268	3810	1559	926
P-waarde	.14	< .01	< .01	< .01	.053

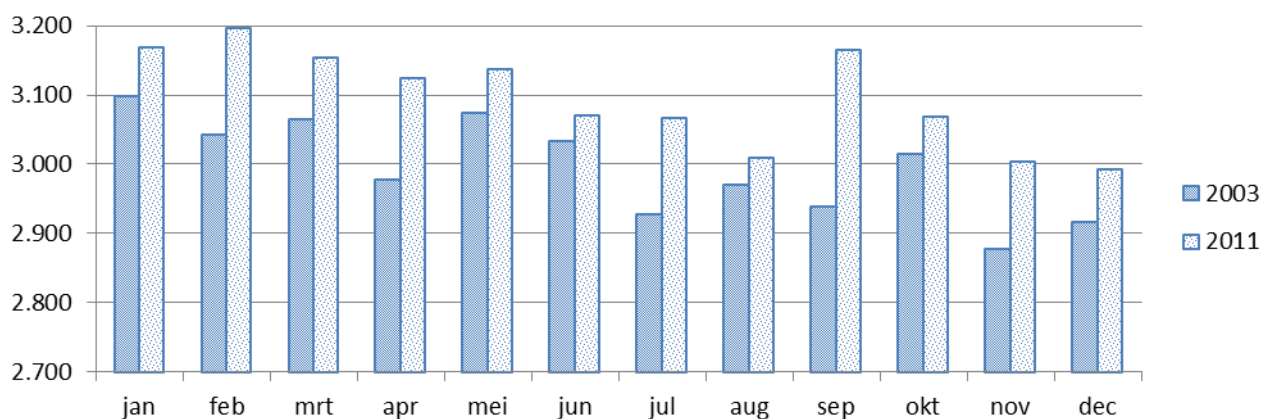
Tabel 29. Ongepaarde t-toets resultaten ter vergelijking van het gemiddelde wiskunde zelfconcept van leerlingen met verschillende hoeveelheden boeken 2003 en 2011: vergelijking binnen de afnamejaren.

	0-10/11-25 2003	11-25/26-100 2003	26-100/101-200 2003	101-200/>200 2003
t-waarde	-0.90	-2.11	-2.40	-1.92
df	1385	2901	2625	1238
P-waarde	.37	.04	.02	.055

	0-10/11-25 2011	11-25/26-100 2011	26-100/101-200 2011	101-200/>200 2011
t-waarde	-0.83	-3.76	-2.04	-1.14
df	1610	3177	2744	1247
P-waarde	.41	< .01	.04	.26

3.5 Vergelijking tussen het gemiddelde wiskunde zelfconcept van leerlingen met een verschillende geboortemaand in 2003 en 2011

In Figuur 8 en Tabel 30 worden het gemiddelde wiskunde zelfconcept weergegeven van leerlingen naargelang hun geboortemaand in 2003 en 2011. We hebben daarbij gekeken naar de leerlingen uit alle geboortejaren en niet enkel naar de normaalvorderende leerlingen. Opvallend is dat ook hier weer duidelijk het geboortemaand-effect zichtbaar is: in allebei de afnamejaren hebben leerlingen die in januari geboren zijn een significant hoger wiskunde zelfconcept dan leerlingen die in december geboren zijn. Wanneer we het gemiddelde wiskunde zelfconcept van de verschillende categorieën leerlingen vergelijken tussen de afnamejaren, blijkt dat de leerlingen die in de maanden februari, april, juli, september en november zijn geboren gemiddeld een hoger wiskunde zelfconcept hebben in 2011 dan in 2003. In de overige maanden zijn de verschillen in wiskunde zelfconcept tussen de afnamejaren niet statistisch significant (zie Tabel 31).



Figuur 8. Gemiddeld wiskunde zelfconcept van leerlingen die in de afzonderlijke maanden geboren zijn in 2003 en in 2011.

Tabel 30. Gemiddeld wiskunde zelfconcept van leerlingen die in de afzonderlijke maanden zijn geboren in 2003 en in 2011.

	jan (SD)	feb (SD)	mrt (SD)	apr (SD)
TIMSS 2003	3.10 (0.62)	3.04 (0.67)	3.06 (0.68)	2.98 (0.71)
TIMSS 2011	3.17 (0.75)	3.20 (0.69)	3.15 (0.74)	3.12 (0.74)
	mei (SD)	jun (SD)	jul (SD)	aug (SD)
TIMSS 2003	3.07 (0.66)	3.03 (0.64)	2.93 (0.67)	2.97 (0.66)
TIMSS 2011	3.14 (0.76)	3.07 (0.73)	3.07 (0.73)	3.02 (0.75)
	sep (SD)	okt (SD)	nov (SD)	dec (SD)
TIMSS 2003	2.94 (0.70)	3.01 (0.68)	2.88 (0.67)	2.88 (0.67)
TIMSS 2011	3.16 (0.74)	3.07 (0.80)	3.00 (0.80)	2.99 (0.75)

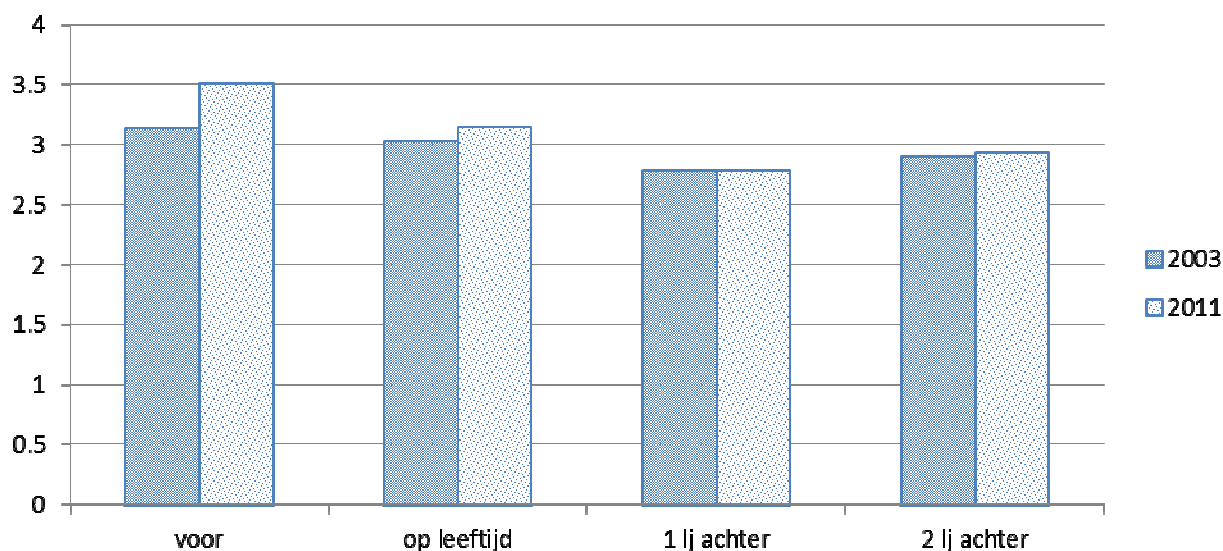
Tabel 31. Ongepaarde t-toets resultaten ter vergelijking van het gemiddelde wiskunde zelfconcept van leerlingen die in de afzonderlijke maanden zijn geboren tussen 2003 en 2011.

	jan 2003-2011	feb 2003-2011	mrt 2003-2011	apr 2003-2011	mei 2003-2011	jun 2003-2011
t-waarde	-1.41	-3.18	-1.77	-2.74	-1.39	-0.82
df	778	727	782	804	805	813
P-waarde	.16	< .01	.08	< .01	.17	.41
	jul 2003-2011	aug 2003-2011	sep 2003-2011	okt 2003-2011	nov 2003-2011	dec 2003-2011
t-waarde	-2.89	-1.00	-4.30	-1.15	-2.14	-1.32
df	839	811	791	819	689	747
P-waarde	< .01	0.32	< .01	.25	.03	.19

3.6 Vergelijking tussen het gemiddelde wiskunde zelfconcept van leerlingen met een verschillend geboortjaar in 2003 en 2011

In Figuur 9 en Tabel 32 wordt het gemiddelde wiskunde zelfconcept weergegeven van leerlingen die één leerjaar vertraagd zijn, twee of meer leerjaren vertraagd zijn, op leeftijd zitten en één of meerdere leerjaren versneld zijn in 2003 en 2011.

Allereerst valt op dat in beide afnamejaren de leerlingen die één leerjaar vertraagd zijn een lager wiskunde zelfconcept vertonen dan leerlingen die 'op leeftijd' zitten. Dit verschil is statistisch significant (zie Tabel 33). Het verschil in wiskunde zelfconcept tussen leerlingen die één en twee of meer leerjaren vertraagd zijn is niet significant verschillend in zowel 2003 als 2011. In beide afnamejaren hebben leerlingen die één of twee leerjaren voorliggen ('versneld zijn') een hoger wiskunde zelfconcept dan leerlingen die op leeftijd zitten. De voorsprong in wiskunde zelfconcept van versnelde leerlingen is echter veel groter in 2011 dan in 2003; het verschil tussen 'versnelde' leerlingen en leerlingen 'op leeftijd' is dan ook alleen significant in 2011. Wanneer we het wiskunde zelfconcept van de vier categorieën leerlingen tussen de afnamejaren met elkaar vergelijken, valt op dat enkel de wiskunde zelfconcepten van de leerlingen die op leeftijd zitten en de versnelde leerlingen significant verschillend zijn tussen 2003 en 2011 (zie Tabel 34).



Figuur 9. Gemiddeld wiskunde zelfconcept van leerlingen die vertraagd, versneld, of op leeftijd zitten, in 2003 en 2011.

Tabel 32. Gemiddeld wiskunde zelfconcept van leerlingen die 1 of 2 leerjaren vertraagd, versneld, of op leeftijd zijn in 2003 en 2011.

	Voor (SD)	Op leeftijd (SD)	Achter (1 lj) (SD)	Achter (2+ lj) (SD)
TIMSS 2003	3.14 (.64)	3.03 (.67)	2.78 (.68)	2.90 (.61)
TIMSS 2011	3.51 (.61)	3.15 (.73)	2.78 (.77)	2.93 (.72)

Tabel 33. Ongepaarde t-toets resultaten ter vergelijking van het gemiddelde wiskunde zelfconcept van leerlingen die 1 of 2 leerjaren vertraagd, versneld of op leeftijd zijn in 2003 en 2011: vergelijking binnen afnamejaren.

	voor-op leeftijd 2003	op leeftijd-achter (1 lj) 2003	achter (1 lj)- achter (+2 lj) 2003
t-waarde	1.03	8.78	-1.35
df	3863	4473	711
P-waarde	.30	< .01	.18
	voor-op leeftijd 2011	op leeftijd-achter (1 lj) 2011	achter (1 lj)- achter (+2 lj) 2011
t-waarde	4.04	12.05	-1.80
df	4075	4678	765
P-waarde	< .01	< .01	.07

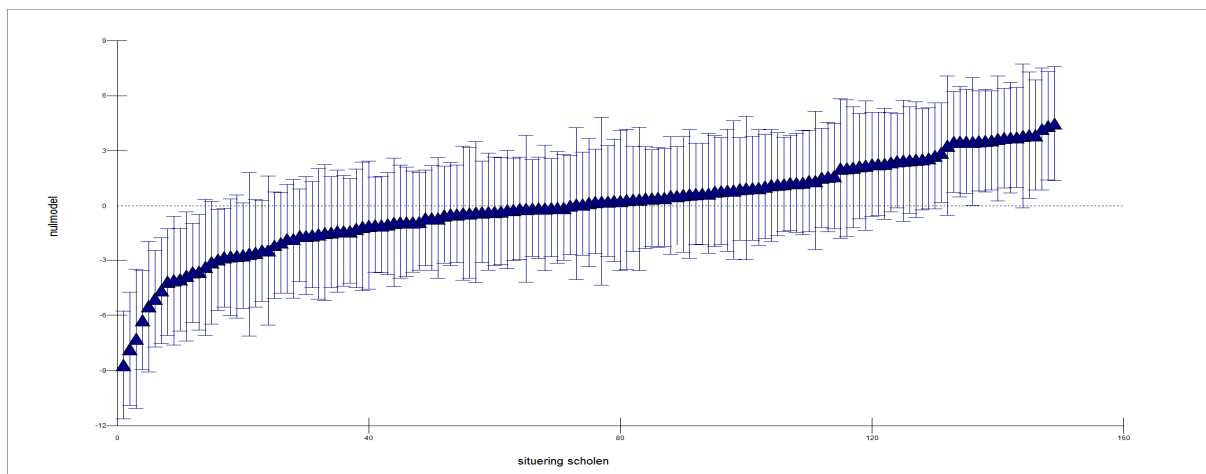
Tabel 18. Ongepaarde t-toets resultaten ter vergelijking van het gemiddelde wiskunde zelfconcept van leerlingen die 1 of 2 leerjaren vertraagd, versneld of op leeftijd zijn in 2003 en 2011: vergelijking tussen afnamejaren.

	Voor 2003-2011	Op leeftijd 2003-2011	Achter (1 lj) 2003-2011	Achter (+2 lj) 2003-2011
t-waarde	-3.00	-7.57	0.00	-0.27
df	106	7832	1319	157
P-waarde	< .01	< .01	1.00	.79

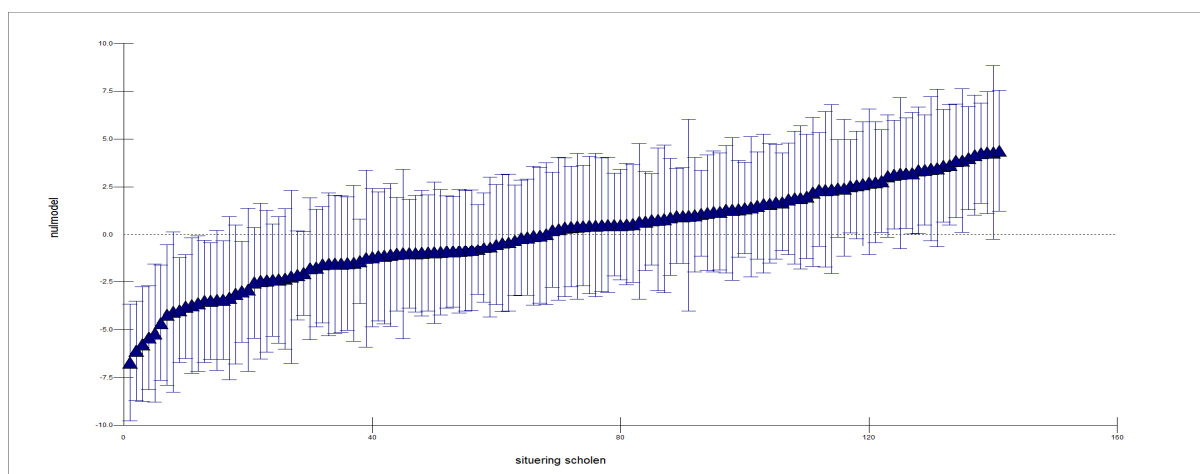
4. Verschillen tussen scholen in wiskundeprestaties: Een vergelijking tussen 2003 en 2011

Sommige scholen doen het voor wiskunde beter dan andere scholen. Onderwijsonderzoek poogt te achterhalen waarom bepaalde scholen het beter doen dan andere om op die manier de kwaliteit van het onderwijs te kunnen verbeteren. Er is echter een aantal kenmerken van leerlingen waarop scholen geen invloed hebben, maar die wel een invloed op prestaties hebben. Dit zijn de achtergrondkenmerken van leerlingen zoals bijvoorbeeld de sociaaleconomische thuissituatie, de taal die thuis gesproken wordt en, zij het in mindere mate, de sekse van de leerlingen. We kunnen verwachten dat een school die bestaat uit leerlingen met een hogere sociaaleconomische achtergrond hoger scoort op een wiskundetoets. Het is dus niet fair om scholen te beoordelen enkel en alleen op basis van hun gemiddelde toetsscore voor wiskunde. Het is eerlijker om scholen ten opzichte van elkaar te situeren nadat gecorrigeerd is voor het leerlingenpubliek van een school. Dan is immers de situering niet zozeer een weerspiegeling van de leerlingenpopulatie van een school maar van andere aspecten zoals de kwaliteit van het onderwijs (Van Damme et al., 2004).

In Figuur 10 en 11 worden de scholen die aan respectievelijk TIMSS 2003 en TIMSS 2011 hebben deelgenomen, gerangschikt naar hun gemiddelde wiskunde prestaties, wanneer er geen correctie wordt toegepast voor de instroomkenmerken.



Figuur 10. Situering van de deelnemende scholen uit TIMSS 2003 op basis van de wiskundescores zonder rekening te houden met de instroomkenmerken van de leerlingen.



Figuur 11. Situering van de deelnemende scholen uit TIMSS 2011 op basis van de wiskundescores zonder rekening te houden met de instroomkenmerken van de leerlingen (exclusief outlier- en BLO-scholen).

Figuur 10 en 11 zijn als volgt te interpreteren. Elk verticaal lijntje stelt een school voor. Het driehoekje in het verticale lijntje geeft de gemiddelde prestatie van die bepaalde school weer. De lengte van het verticaal lijntje geeft het 95%-betrouwbaarheidsinterval aan van het gemiddelde van die school. Dit wil zeggen dat met 95% zekerheid aangenomen mag worden dat het gemiddelde van de school zich in het aangegeven bereik bevindt. De horizontale stippellijn door 0.0 geeft het algemeen gemiddelde van de Vlaamse leerlingen voor de prestaties wiskunde weer. Wanneer een verticaal lijntje de horizontale stippellijn kruist dan wil dat zeggen dat het gemiddelde van die school niet significant verschilt van het algemeen gemiddelde. Wanneer een verticaal lijntje de horizontale stippellijn niet kruist, wil dat zeggen dat het gemiddelde van die school significant lager of hoger is dan het algemeen gemiddelde. Telkens bevinden de scholen met het laagste gemiddelde zich links in de figuur en die met het hoogste gemiddelde rechts.

Uit Figuur 10 wordt zodoende duidelijk dat wanneer er geen correctie voor instroomkenmerken plaats vindt, in 2003, 10.7% van de scholen een bovengemiddelde wiskundescore heeft behaald, en 10.1% van de scholen een ondergemiddelde wiskundescore. In 2011 zijn de resultaten vergelijkbaar. Figuur 11 toont dat in 2011, 11.3% van de scholen een bovengemiddelde score behaalde en 10.6% van de scholen een ondergemiddelde score.

In Tabel 35 en 36 wordt weergegeven in hoeverre de wiskundeprestaties van leerlingen zijn toe te schrijven aan verschillen tussen leerlingen binnen eenzelfde klas, tussen klassen binnen eenzelfde school en tussen scholen. De resultaten laten zien dat in beide afnamejaren de wiskundeprestaties van leerlingen voor ongeveer 90% zijn toe te schrijven aan verschillen tussen leerlingen binnen eenzelfde klas en voor ongeveer 9% aan verschillen tussen scholen. Opvallend is dat er alleen in 2011 sprake is van prestatieverschillen tussen klassen binnen eenzelfde school. In dit jaar zijn de wiskundeprestaties van leerlingen voor ongeveer 1.2% toe te schrijven aan verschillen tussen klassen.

Tabel 35. Leeg model TIMSS 2003

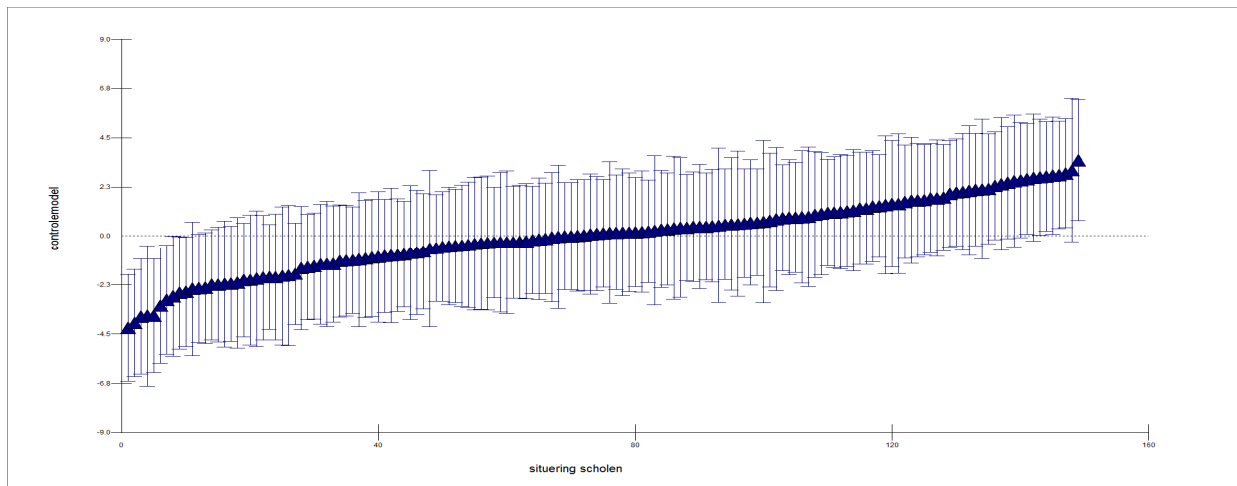
Tabel 36. Leeg model TIMSS 2011

<i>Random</i>	<i>b</i>	<i>S.E.</i>	<i>Random</i>	<i>b</i>	<i>S.E.</i>
Intercept	49.772*	0.286	Intercept	50.125*	0.306
Variantie school	8.669*	1.477	Variantie school	8.937*	1.622
Variantie klas	0.836	0.757	Variantie klas	1.727*	0.848
Variantie leerling	84.941*	1.798	Variantie leerling	93.008*	1.911
% Variantie school	9.3%		% Variantie school	8.6%	
% Variantie klas	<i>n.s.</i>		% Variantie klas	1.7%	
% Variantie leerling	90.7%		% Variantie leerling	89.7%	
Deviantie	34535.216		Deviantie	37100.103	

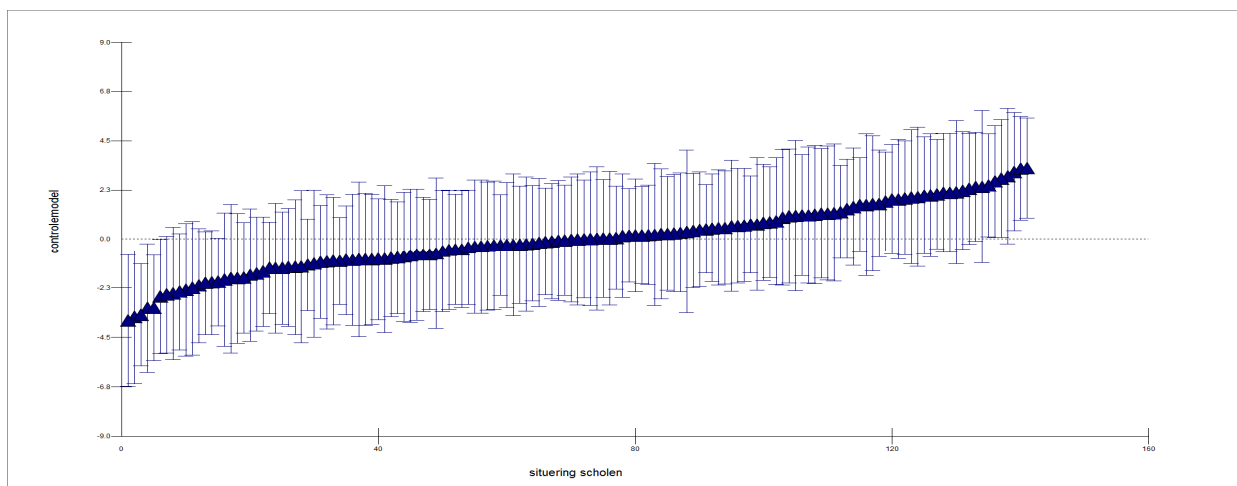
* Deze coëfficiënten zijn statistisch significant bij $\alpha = .05$.

In Figuur 12 en 13 worden de scholen die aan respectievelijk TIMSS 2003 en TIMSS 2011 hebben deelgenomen, gerangschikt naar hun wiskunde prestaties, wanneer er wel correctie wordt toegepast voor de instroomkenmerken. Deze instroomkenmerken zijn: (1) sekse, (2) thuistaal, (3) aantal boeken thuis, (4) geboortemaand en (5) geboortjaar⁴.

⁴ We hebben er voor gekozen om het geboortjaar van de leerling mee op te nemen als een predictor in het model, aangezien er een verband bestaat tussen de wiskundeprestaties van een leerling en zijn/haar geboortjaar. Deze keuze is echter discutabel omdat het al dan niet laten zittenblijven van een leerling deel uitmaakt van het gevoerde schoolbeleid.



Figuur 12. Situering van de deelnemende scholen uit TIMSS 2003 op basis van de wiskundescores rekening houdend met de instroomkenmerken van de leerlingen.



Figuur 13. Situering van de deelnemende scholen uit TIMSS 2011 op basis van de wiskundescores rekening houdend met de instroomkenmerken van de leerlingen (exclusief outlier- en BLO-scholen) .

Figuur 12 laat zien dat wanneer er controle plaatsvindt voor de instroomkenmerken, het aandeel bovengemiddeld presterende scholen in 2003 daalt naar 4.7% en het aandeel ondergemiddeld presterende scholen naar 5.4%. Figuur 13 toont dat na controle voor de instroomkenmerken, in 2011 4.3% van de scholen een bovengemiddelde score behaalt en tevens 4.3% een ondergemiddelde score. In 2011 is dus zowel het aandeel boven- als ondergemiddeld presterende scholen iets kleiner dan in 2003.

In Tabel 37 en 38 worden de coëfficiënten en de standaardfouten van de leerlingkenmerken weergegeven wanneer ze samen in een model zitten. Te zien is dat wanneer alle kenmerken gezamenlijk in een model zitten, alle coëfficiënten, behalve de coëfficiënt voor geboortemaand, statistisch significant zijn. Dit geeft aan, dat wanneer rekening wordt gehouden met de overige leerlingkenmerken, alle leerlingkenmerken (behalve geboortemaand) beduidend samenhangen met de wiskundescores van leerlingen.

Tabel 37. Controlemodel TIMSS 2003

	<i>b</i>	<i>S.E.</i>	<i>Random</i>	<i>b</i>	<i>S.E.</i>
Intercept	53.253*	0.500	Intercept	53.921*	0.513
<i>Fixed</i>			<i>Fixed</i>		
Sekse	0.804*	0.265	Sekse	1.237*	0.263
Thuis taal	-2.100*	0.299	Thuis taal	-1.856*	0.277
Aantal boeken thuis	1.171*	0.128	Aantal boeken thuis	1.201*	0.131
Geboortemaand	-0.055	0.039	Geboortemaand	-0.150*	0.039
Geboortjaar	-5.152*	0.313	Geboortjaar	-5.277*	0.299
<i>Random</i>			<i>Random</i>		
Variatie school	4.650*	1.024	Variatie school	4.300*	1.032
Variatie klas	1.126	0.745	Variatie klas	1.746*	0.781
Variatie leerling	76.658*	1.646	Variatie leerling	82.449*	1.712
Deviantie	33119.579		Deviantie	35709.189	

Noot. Het intercept verwijst hier naar de gemiddelde wiskundescore van een meisje dat 'op leeftijd zit', in januari is geboren, thuis (bijna) altijd Nederlands spreekt, thuis over een gemiddeld aantal boeken beschikt (26 à 100 boeken) en een school bezoekt met een leerlingenpubliek met een gemiddelde SES.

* Deze coëfficiënten zijn statistisch significant bij $\alpha = .05$.

Door telkens een leerlingkenmerk apart aan het nulmodel toe te voegen, kan worden nagegaan in hoeverre dat leerlingkenmerk de variantie op de verschillende niveaus kan verklaren wanneer er geen rekening wordt gehouden met de andere leerlingkenmerken. Daarnaast is het interessant om te zien hoeveel variantie er door de leerlingkenmerken 'aantal boeken thuis' en 'thuis taal' gezamenlijk wordt verklaard, aangezien een lage SES en een andere thuis taal vaak samengaan, en hoeveel variantie er door alle leerlingkenmerken tezamen kan worden verklaard.

Om een geldige vergelijking tussen twee of meerdere modellen te kunnen maken, is het van belang dat de modellen op dezelfde groep leerlingen zijn gebaseerd. Om deze reden hebben we in deze analyses alleen de gegevens gebruikt van leerlingen van wie we informatie hebben over al hun leerlingkenmerken. Zodoende zijn de huidige analyses gebaseerd op de gegevens van 4588 leerlingen in 2003 en 4898 leerlingen in 2011. De resultaten van deze analyses worden weergegeven in Tabel 39 en Tabel 40.

Tabel 39. Percentage verklaarde variantie per leerlingkenmerk in TIMSS 2003

	<i>Leerling %</i>	<i>Klas %</i>	<i>School %</i>	<i>Totaal %</i>	<i>Deviantie</i>	<i>Vershil in deviantie*</i>
Sekse	0.1%	<i>n.s.</i>	0.0%	0.1%	33617.999	4.565
Thuis taal	1.5%	<i>n.s.</i>	14.9%	2.8%	33536.381	86.183
Aantal boeken thuis	3.1%	<i>n.s.</i>	15.9%	3.9%	33474.811	147.753
Thuis taal + Aantal boeken thuis	4.3%	<i>n.s.</i>	27.8%	6.1%	33401.248	221.316
Geboortemaand	0.3%	<i>n.s.</i>	0.0%	0.4%	33606.881	15.683
Geboortjaar	7.0%	<i>n.s.</i>	35.9%	9.1%	33266.360	356.204
Alle leerlingkenmerken	9.7%	<i>n.s.</i>	46.3%	9.0%	33119.580	502.984

* In vergelijking met het nulmodel.

Tabel 40. Percentage verklaarde variantie per leerlingkenmerk in TIMSS 2011

	<i>Leerling %</i>	<i>Klas %</i>	<i>School %</i>	<i>Totaal %</i>	<i>Deviantie</i>	<i>Vershil in deviantie*</i>
Sekse	0.4%	0.0%	0.3%	0.4%	36253.166	19.512
Thuis taal	1.1%	8.1%	25.3%	3.3%	36187.847	84.831
Aantal boeken thuis	2.6%	1.0%	12.3%	3.4%	36133.434	139.244
Thuis taal + Aantal boeken thuis	3.5%	7.6%	33.1%	6.1%	36065.523	207.155
Geboortemaand	0.8%	3.3%	0.1%	0.8%	36232.961	39.726
Geboortjaar	7.1%	2.1%	34.4%	9.3%	35884.074	388.604
Alle leerlingkenmerken	10.0%	0.9%	50.6%	13.3%	35709.189	563.489

* In vergelijking met het nulmodel.

In Tabel 39 en 40 is te zien dat het leerlingkenmerk 'sekse' het kleinste percentage van de totale variantie in wiskundescores verklaart, en het leerlingkenmerk 'geboortjaar' het grootste. Verder is uit de tabellen af te lezen dat alle leerlingkenmerken tezamen ruim 4% meer van de totale variantie in wiskundescores verklaren in 2011 dan in 2003. De vijf leerlingkenmerken hebben dus gezamenlijk een iets grotere impact op de wiskundescores in 2011 dan in 2003. Het feit dat de optelsom van de percentages totaal verklaarde variantie van de afzonderlijke leerlingkenmerken niet te herleiden is tot het percentage totaal verklaarde variantie van alle leerlingkenmerken samen, komt door de onderlinge samenhang tussen bepaalde leerlingkenmerken. Wanneer twee of meerdere variabelen sterk samenhangen heffen deze elkaar in zekere zin deels op.

Wanneer we de percentages verklaarde variantie op leerlingniveau bekijken is te zien dat het geboortjaar van leerlingen het meest in staat is om de verschillen tussen leerlingen binnen eenzelfde klas te verklaren en de sekse van leerlingen het minst. Bij een vergelijking tussen de twee afnamejaren valt op dat de leerlingkenmerken 'sekse', 'geboortemaand' en 'geboortjaar' iets meer van de verschillen tussen leerlingen binnen eenzelfde klas kunnen verklaren in 2011 dan in 2003. Dit wil zeggen dat de verschillen in wiskundescores tussen deze typen leerlingen binnen eenzelfde klas iets groter zijn in 2011 dan in 2003. In 2003 verklaren daarentegen de leerlingkenmerken 'aantal boeken thuis' en 'thuis taal' iets meer van de verschillen tussen leerlingen binnen eenzelfde klas dan in 2011.

Aangezien we eerder bij het nulmodel al hadden gezien dat er in 2003 geen sprake is van variantie op het klasniveau kunnen we het wat betreft de invloed van leerlingkenmerken op klasverschillen alleen maar over 2011 hebben. In dit jaar blijken de verschillen in wiskundescores tussen klassen binnen eenzelfde school het meest verklaard te kunnen worden door verschillen in de thuis taal van leerlingen en het minst door verschillen in de sekse van leerlingen. Met name de eerste vaststelling is frappant. Dit doet vermoeden dat er scholen zijn die hun leerlingen (mede) op basis van hun thuis taal in klassen groeperen.

Wat betreft de percentages verklaarde variantie op schoolniveau is het treffend dat de verschillen in wiskundescores tussen scholen het meest worden verklaard door het geboortjaar van leerlingen en het minst door de geboortemaand van leerlingen. Het hoge percentage verklaarde variantie door het geboortjaar van leerlingen duidt op grote verschillen tussen scholen in percentage vertraagde leerlingen. Dit heeft wellicht te maken met het beleid dat scholen omtrent zittenblijven voeren. Sommige scholen kiezen ervoor om hun zwakke leerlingen een leerjaar te laten dubbelen terwijl andere scholen ervoor kiezen om zwakke leerlingen te laten doorstromen en te voorzien van extra begeleiding. Bij een vergelijking tussen de twee afnamejaren, valt op dat in 2003 iets meer van de verschillen tussen scholen in wiskundescores worden verklaard door verschillen tussen scholen in het gemiddelde aantal boeken dat de leerlingen thuis hebben, dan in 2011.

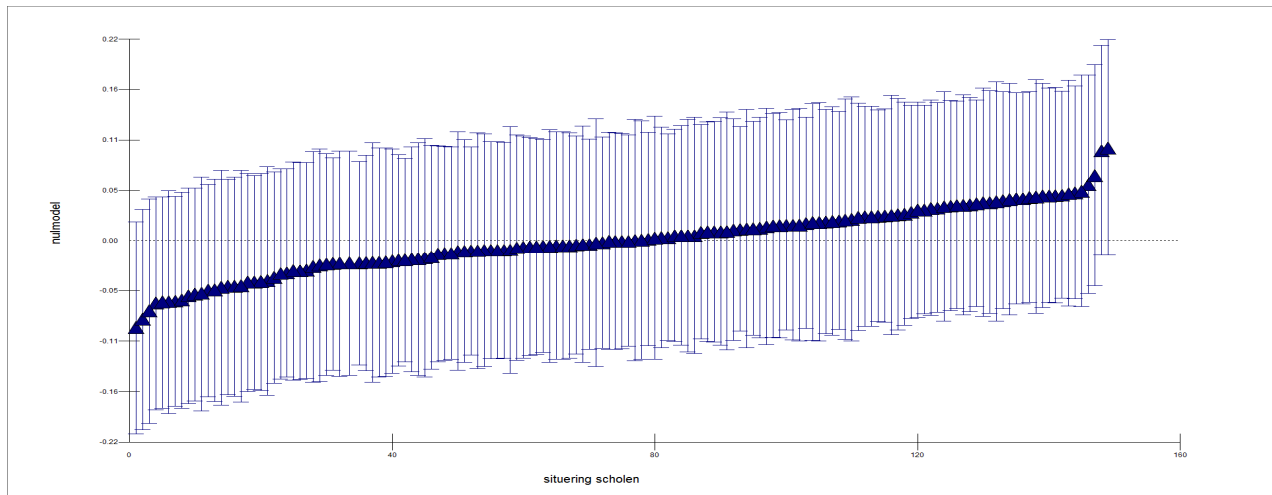
Opmerkelijk is dat in 2011 de verschillen in wiskundescores tussen scholen voor ruim 10% meer worden verklaard door verschillen tussen scholen in de thuistaal van leerlingen dan in 2003. Voorzichtig kan dus worden gesteld dat terwijl de segregatie tussen scholen voor wat betreft SES ('aantal boeken thuis') iets is afgenomen, de segregatie voor wat betreft etniciteit ('thuistaal') juist is toegenomen tussen 2003 en 2011. Gezamenlijk verklaren de leerlingkenmerken 'aantal boeken thuis' en 'thuistaal' ruim 5% meer van de verschillen in wiskundescores tussen scholen in 2011 dan in 2003.

Bij het vergelijken van de in totaal verklaarde variantie naargelang van het afnamemoment, blijkt dat vooral de sekse, de geboortemaand en de thuistaal in 2011 relatief sterker samenhangen met het prestatieniveau dan in 2003, terwijl de SES-indicator (het aantal boeken thuis) aan effect iets aan belang ingeboet heeft.

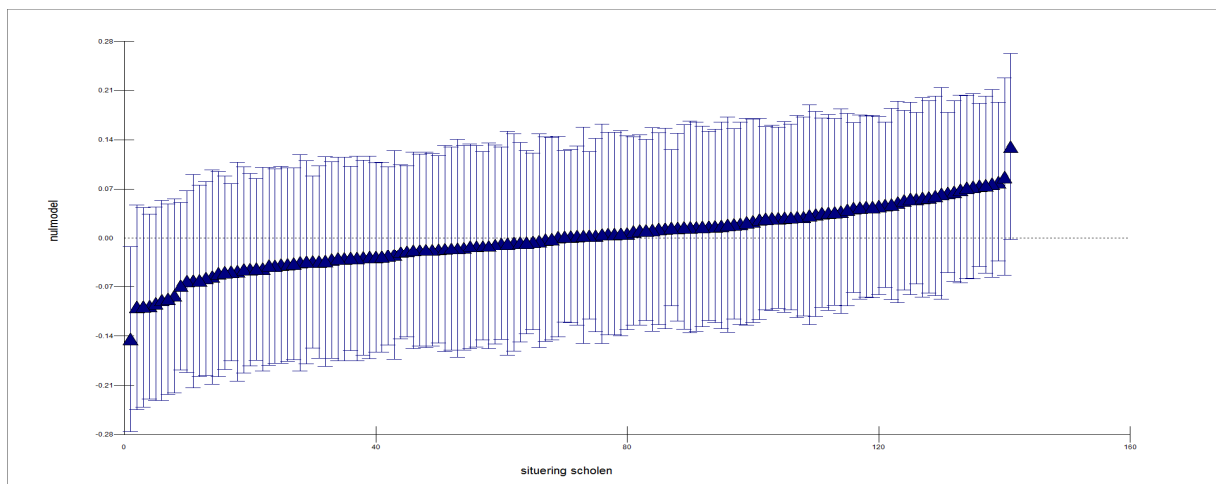
5. Verschillen tussen scholen in wiskunde zelfconcept: Een vergelijking tussen 2003 en 2011

Scholen verschillen niet alleen in gemiddelde wiskunde prestaties, maar ook in gemiddelde wiskunde zelfconcept. Dit heeft opnieuw te maken met de instroomverschillen tussen scholen maar ook met de mate waarin een school aandacht besteedt aan en/of invloed uitoefent op het niet-cognitieve functioneren van leerlingen, waaronder het wiskunde zelfconcept.

In Figuur 14 en 15 worden de scholen die aan respectievelijk TIMSS 2003 en TIMSS 2011 hebben deelgenomen, gerangschikt naar hun gemiddelde wiskunde zelfconcept, wanneer er geen correctie wordt toegepast voor de instroomkenmerken.



Figuur 14. Situering van de deelnemende scholen uit TIMSS 2003 op basis van het wiskunde zelfconcept zonder rekening te houden met de instroomkenmerken van de leerlingen.



Figuur 15. Situering van de deelnemende scholen uit TIMSS 2011 op basis van het wiskunde zelfconcept zonder rekening te houden met de instroomkenmerken van de leerlingen (exclusief outlier- en BLO-scholen).

Uit Figuur 14 wordt duidelijk dat er in 2003 geen enkele school afwijkt van het Vlaamse gemiddelde qua wiskunde zelfconcept, wanneer er geen correctie voor instroom-kenmerken plaats vindt. In 2011 is er een school die er uitspringt met een significant ondergemiddeld wiskunde zelfconcept, in vergelijking met de andere scholen. Opvallend is dat deze school niet significant onderpresteert in wiskunde (wanneer er geen rekening wordt gehouden met instroomverschillen). Dit suggereert dat

een school die effectief is op het gebied van wiskunde, niet noodzakelijk ook effectief hoeft te zijn in het bevorderen van het wiskunde zelfconcept van leerlingen.

In Tabel 41 en 42 wordt weergegeven in hoeverre het wiskunde zelfconcept van leerlingen is toe te schrijven aan verschillen tussen leerlingen binnen eenzelfde klas, tussen klassen binnen eenzelfde school en tussen scholen. De resultaten laten zien dat in beide afnamejaren het wiskunde zelfconcept van leerlingen voor ongeveer 99% is toe te schrijven aan verschillen tussen leerlingen binnen eenzelfde klas, helemaal niet is toe te schrijven verschillen tussen klassen binnen eenzelfde school, en voor slechts ongeveer 1% aan verschillen tussen scholen. Het wiskunde zelfconcept van leerlingen is dus iets waar scholen zeer weinig invloed op hebben, althans in die zin dat hun gemiddelde invloed op dit punt gelijkaardig is.

Tabel 41. Leeg model TIMSS 2003

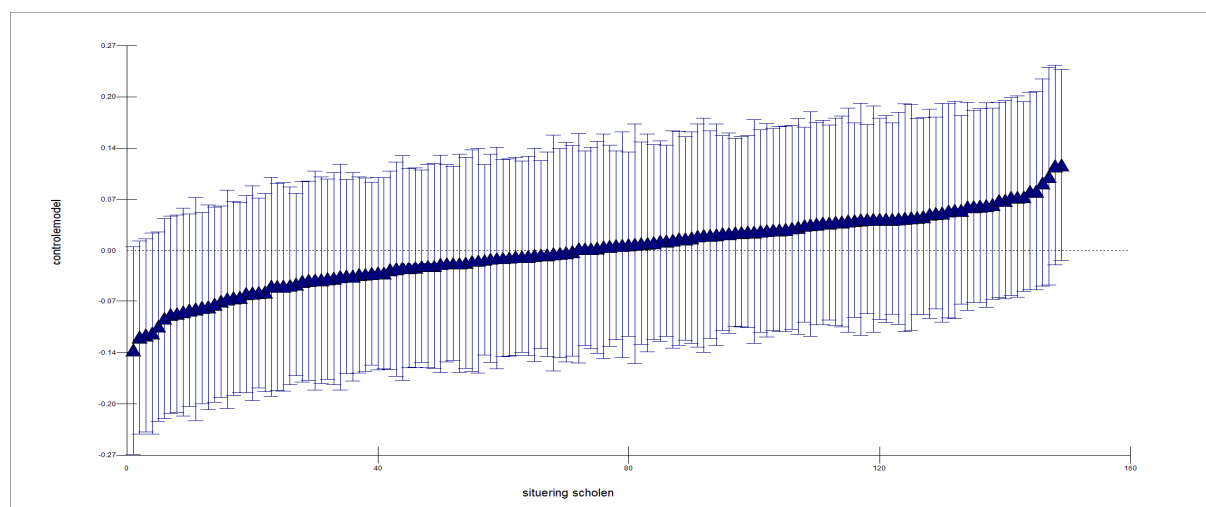
<i>Random</i>	<i>b</i>	<i>S.E.</i>
Intercept	2.995*	0.012
Variantie school	0.005*	0.002
Variantie klas	0.000	0.000
Variantie leerling	0.450*	0.010
% Variantie school	1.1%	
% Variantie klas	<i>n.s.</i>	
% Variantie leerling	98.9%	
Deviantie	9383.391	

Tabel 42. Leeg model TIMSS 2011

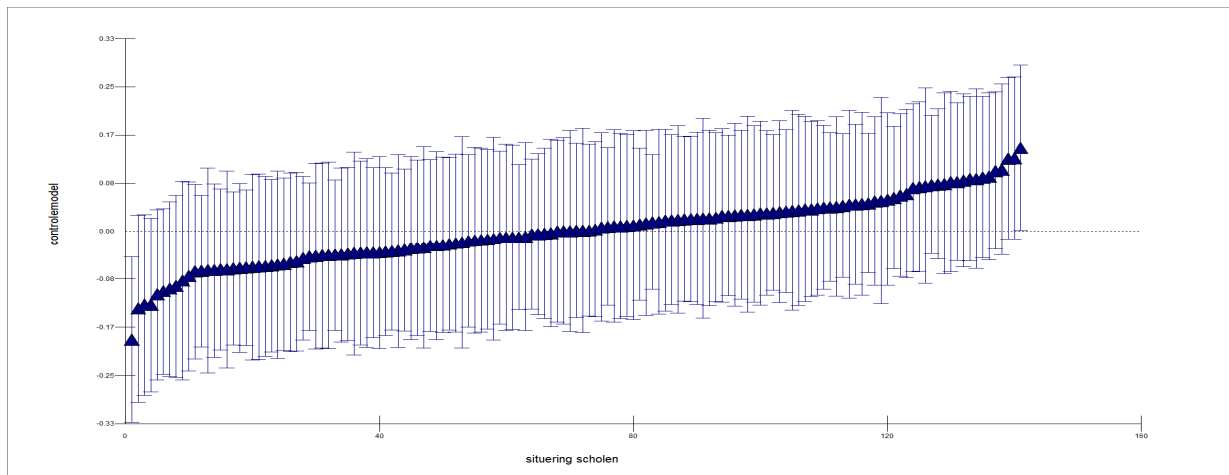
<i>Random</i>	<i>b</i>	<i>S.E.</i>
Intercept	3.093*	0.013
Variantie school	0.007*	0.003
Variantie klas	0.000	0.000
Variantie leerling	0.557*	0.011
% Variantie school	1.2%	
% Variantie klas	<i>n.s.</i>	
% Variantie leerling	98.8%	
Deviantie	10965.866	

* Deze coëfficiënten zijn statistisch significant bij $\alpha = .05$.

In Figuur 16 en 17 worden de scholen die aan respectievelijk TIMSS 2003 en TIMSS 2011 hebben deelgenomen, gerangschikt naar hun gemiddelde wiskunde zelfconcept, wanneer er correctie wordt toegepast voor de instroomkenmerken. Deze instroomkenmerken zijn: (1) sekse, (2) thuistaal, (3) aantal boeken thuis, (4) geboortemaand en (5) geboortjaar.



Figuur 16. Situering van de deelnemende scholen uit TIMSS 2003 op basis van het wiskunde zelfconcept, rekening houdend met de instroomkenmerken van de leerlingen.



Figuur 17. Situering van de deelnemende scholen uit TIMSS 2011 op basis van het wiskunde zelfconcept, rekening houdend met de instroomkenmerken van de leerlingen (exclusief outlier- en BLO-scholen).

Figuur 16 laat zien dat er ook wanneer er controle plaatsvindt voor de instroomkenmerken, er geen school significant afwijkt van het Vlaamse gemiddelde qua wiskunde zelfconcept in 2003.

Figuur 17 toont verder dat in 2011 na controle voor de instroomkenmerken, de school met het significante ondergemiddelde wiskunde zelfconcept, het nog slechter doet dan wanneer er geen controle plaats vond voor de instroomkenmerken. Het lagere wiskunde zelfconcept van de leerlingen binnen deze school lijkt dus vooral te wijten te zijn aan processen die zich binnen deze school afspelen. Verder is er nu een school bijgekomen die een significant bovengemiddeld wiskunde zelfconcept bij zijn leerlingen weet te behalen.

In Tabel 43 en 44 worden de coëfficiënten en de standaardfouten van de leerlingkenmerken weergegeven wanneer ze samen in een model zitten. Te zien is dat wanneer alle kenmerken gezamenlijk in een model zitten, alle coëfficiënten, behalve de thuistaal van leerlingen, statistisch significant zijn. Dit indiceert dat de thuistaal van leerlingen geen verband heeft met het wiskunde zelfconcept van leerlingen wanneer er rekening wordt gehouden met de overige leerlingkenmerken.

Tabel 43. Controlemodel TIMSS 2003

Tabel 44. Controlemodel TIMSS 2011

	<i>b</i>	<i>S.E.</i>	<i>Random</i>	<i>b</i>	<i>S.E.</i>
Intercept	2.847*	0.045	Intercept	3.044*	0.038
<i>Fixed</i>			<i>Fixed</i>		
Sekse	0.262*	0.020	Sekse	0.318*	0.021
Thuistaal	-0.022	0.022	Thuistaal	-0.005	0.021
Aantal boeken thuis	0.049*	0.009	Aantal boeken thuis	0.054*	0.010
Geboortemaand	-0.011*	0.003	Geboortemaand	-0.010*	0.003
Geboortejahr	-0.156*	0.023	Geboortejahr	-0.234*	0.024
<i>Random</i>			<i>Random</i>		
Variantie school	0.007*	0.002	Variantie school	0.009*	0.003
Variantie klas	0.000	0.000	Variantie klas	0.000	0.000
Variantie leerling	0.419*	0.009	Variantie leerling	0.511*	0.011
Deviantie	8870.193		Deviantie	10426.469	

* Deze coëfficiënten zijn statistisch significant bij $\alpha = .05$.

In de volgende analyses trachten we de variantie in wiskunde zelfconcept op de verschillende niveaus te verklaren door telkens een leerlingkenmerk apart aan het nulmodel toe te voegen. Daarnaast gaan we na hoeveel variantie er door de leerlingkenmerken 'aantal boeken thuis' en 'thuisstaal' gezamenlijk wordt verklaard en hoeveel variantie er door alle leerlingkenmerken tezamen kan worden verklaard.

Om een geldige vergelijking tussen de modellen te kunnen maken, hebben we in deze analyses alleen de gegevens gebruikt van leerlingen van wie we informatie hebben over al hun leerlingkenmerken. Zodoende zijn de huidige analyses gebaseerd op de gegevens van 4478 leerlingen in 2003 en 4781 leerlingen in 2011. De resultaten van deze analyses worden weergegeven in Tabel 45 en Tabel 46.

Tabel 45. Percentage verklaarde variantie per leerlingkenmerk in TIMSS 2003

	<i>Leerling %</i>	<i>Klas %</i>	<i>School %</i>	<i>Totaal %</i>	<i>Deviantie</i>	<i>Vershil in deviantie*</i>
Sekse	3.8%	<i>n.s.</i>	0.0%	3.5%	8988.380	166.426
Thuisstaal	0.0%	<i>n.s.</i>	0.0%	0.0%	9152.670	2.136
Aantal boeken thuis	0.9%	<i>n.s.</i>	16.7%	1.1%	9111.363	43.443
Thuisstaal + Aantal boeken thuis	0.9%	<i>n.s.</i>	0.0%	0.9%	9110.473	44.333
Geboortemaand	0.4%	<i>n.s.</i>	0.0%	0.4%	9128.960	25.846
Geboortejaar	1.6%	<i>n.s.</i>	0.0%	1.3%	9090.244	64.562
Alle leerlingkenmerken	6.3%	<i>n.s.</i>	0.0%	6.0%	8870.193	284.613

* In vergelijking met het nulmodel.

Tabel 46. Percentage verklaarde variantie per leerlingkenmerk in TIMSS 2011

	<i>Leerling %</i>	<i>Klas %</i>	<i>School %</i>	<i>Totaal %</i>	<i>Deviantie</i>	<i>Vershil in deviantie*</i>
Sekse	4.9%	<i>n.s.</i>	0.0%	4.3%	10593.453	222.177
Thuisstaal	0.0%	<i>n.s.</i>	0.0%	0.0%	10811.183	4.447
Aantal boeken thuis	0.9%	<i>n.s.</i>	0.0%	0.9%	10769.955	45.675
Thuisstaal + Aantal boeken thuis	0.9%	<i>n.s.</i>	0.0%	0.7%	10767.937	47.693
Geboortemaand	0.4%	<i>n.s.</i>	0.0%	0.4%	10791.661	23.969
Geboortejaar	2.4%	<i>n.s.</i>	14.3%	2.5%	10692.867	122.763
Alle leerlingkenmerken	8.1%	<i>n.s.</i>	0.0%	7.6%	10426.481	389.149

* In vergelijking met het nulmodel.

Tabel 45 en 46 tonen dat in 2011 alle leerlingkenmerken samen 1.6% meer kunnen verklaren van de totale variantie in het wiskunde zelfconcept van leerlingen. In beide afnamejaren verklaart de sekse van de leerlingen het grootste percentage van de totale variantie in wiskunde zelfconcept en de thuisstaal van leerlingen geen.

Wat betreft de variantie op leerlingniveau is hetzelfde patroon zichtbaar: de meeste variantie wordt verklaard door de sekse van de leerlingen en geen variantie wordt verklaard door de thuisstaal. Op schoolniveau wordt er in beide afnamejaren slechts door één leerlingkenmerk (gedeeltelijk) de variantie in het wiskunde zelfconcept verklaard. In 2003 was dat het 'aantal boeken thuis'. Kennelijk hingen in 2003 verschillen tussen scholen in wiskunde zelfconcept sterker samen met onderlinge verschillen in het gemiddelde aantal boeken dat de leerlingen thuis hebben dan in 2011. In 2011 worden verschillen tussen scholen in wiskunde zelfconcept enkel deels verklaard door onderlinge verschillen in het gemiddelde geboortejaar van leerlingen. Dit duidt opnieuw op verschillen tussen scholen in percentage zittenblijvers.

6. Het effect van schoolsamenstelling naar SES op wiskundeprestaties in 2003 en 2011

Leerlingen interageren met hun onderwijsomgeving en dus ook met hun medeleerlingen op school. Onder meer daarom heeft de samenstelling van het leerlingenpubliek van een school invloed op individuele leerlingen. In hoofdstuk 4 hebben we gekeken naar hoe de individuele kenmerken van leerlingen de wiskundeprestaties⁵ beïnvloedden in TIMSS 2003 en TIMSS 2011. In het huidige hoofdstuk onderzoeken we of de gemiddelde SES (i.c. gemiddeld aantal boeken thuis) van de leerlingen van de school die een leerling bezoekt, een bijkomend effect heeft op zijn/haar wiskundeprestaties. Bovendien gaan we na of het eventuele effect van de gemiddelde SES van een school verschillend is in 2003 en 2011. Met name deze laatste vraag is bijzonder interessant, omdat deze enigszins inzicht biedt in de mogelijke effectiviteit van het GOK-beleid dat in september 2002 werd ingevoerd. De TIMSS-data laten toe om te onderzoeken of de impact van de SES-samenstelling van een school al dan niet is verkleind acht jaar nadat het GOK-beleid in werking trad.

Om na te gaan of de impact van de gemiddelde SES van een school in 2011 is veranderd ten aanzien van 2003, toetsten we het volgende multiniveau regressiemodel:

$$\text{Prestaties wiskunde} = \text{intercept} + \text{afnamejaar (2003/2011)} + \text{seks} + \text{geboortejaar} + \text{geboortemaand} + \text{thuismaal} + \text{SES} + \text{gem SES school} + \text{gem SES school} * \text{afnamejaar (2003/2011)}$$

In Tabel 47 staan de resultaten van bovenstaand regressiemodel vermeld. Aangezien de interactie tussen de gemiddelde SES van een school en het afnamejaar niet statistisch significant is, kunnen we vaststellen dat het effect van de gemiddelde SES van school niet verschillend is in 2003 en in 2011. Sterker nog, in beide afnamejaren is er geen effect van de gemiddelde SES van een school zichtbaar.

Tabel 47. Model met schoolsamenstelling naar SES in TIMSS 2003 en 2011

<i>Fixed</i>	<i>b</i>	<i>S.E.</i>
Intercept	53.899*	0.401
<i>Leerlingniveau</i>		
Afnamejaar	-0.656	0.332
Sekse	1.032*	0.186
Geboortejaar	-5.192*	0.217
Geboortemaand	-0.105*	0.028
Thuismaal	-1.960*	0.203
SES (aantal boeken thuis)	1.151*	0.093
<i>Schoolniveau</i>		
Gem SES school	0.820	0.684
<i>Interactieterm</i>		
Gem SES school *Afnamejaar	0.302	0.923
Variantie school	4.294*	0.716
Variantie klas	1.467*	0.543
Variantie leerling	79.705*	1.190
Deviantie	68836.129	

Noot. Het intercept verwijst hier naar de gemiddelde wiskundescore van een meisje dat aan TIMSS 2011 mee deed, 'op leeftijd zit', in januari is geboren, thuis (bijna) altijd Nederlands spreekt, thuis over een gemiddeld aantal boeken beschikt (26 à 100 boeken) en een school bezoekt met een leerlingenpubliek met een gemiddelde SES.

* Deze coëfficiënten zijn statistisch significant bij $\alpha = .05$

⁵ Omdat we in het vorige hoofdstuk vaststelden dat er voor het wiskunde zelfconcept nagenoeg geen variantie op het schoolniveau gesitueerd is, onderzoeken we in dit hoofdstuk enkel het effect van schoolsamenstelling op wiskunde-prestaties.

7. Het effect van schoolsamenstelling naar thuistaal op wiskundeprestaties in 2003 en 2011

Om na te gaan of de impact van het percentage anderstalige leerlingen⁶ van een school in 2011 is veranderd ten aanzien van 2003, toetsten we het volgende multiniveau regressiemodel:

$$\text{Prestaties wiskunde} = \text{intercept} + \text{afnamejaar (2003/2011)} + \text{seks} + \text{geboortjaar} + \text{geboortemaand} + \text{thuistaal} + \text{SES} + \% \text{ anderstaligen op school} + \% \text{ anderstaligen op school} * \text{afnamejaar (2003/2011)}$$

In Tabel 48 staan de resultaten van bovenstaand regressiemodel vermeld. Aangezien de interactie tussen het percentage anderstalige leerlingen van een school en het afnamejaar niet statistisch significant is, kunnen we vaststellen dat het negatieve effect van het percentage anderstalige leerlingen van een school in beide afnamejaren gelijk is⁷.

Tabel 48. Model met schoolsamenstelling naar thuistaal in TIMSS 2003 en 2011

<i>Fixed</i>	<i>b</i>	<i>S.E.</i>
Intercept	53.713*	0.397
<i>Leerlingniveau</i>		
Afnamejaar	-1.127*	0.328
Sekse	1.031*	0.186
Geboortjaar	-5.153*	0.216
Geboortemaand	-0.106*	0.028
Thuistaal	-1.646*	0.211
SES (aantal boeken thuis)	1.178*	0.091
<i>Schoolniveau</i>		
% anderstalige leerlingen	-4.905*	1.073
<i>Interactieterm</i>		
% anderstalige leerlingen* Afnamejaar	-0.838	1.750
Variantie school	3.616*	0.659
Variantie klas	1.446*	0.539
Variantie leerling	79.689*	1.189
Deviantie		68807.002

Noot. Het intercept verwijst hier naar de gemiddelde wiskundescore van een meisje dat aan TIMSS 2011 mee deed, 'op leeftijd zit', in januari is geboren, thuis (bijna) altijd Nederlands spreekt, thuis over een gemiddeld aantal boeken beschikt (26 à 100 boeken) en een school bezoekt met een leerlingenpubliek met een gemiddeld percentage anderstalige leerlingen (=20.4%).

* Deze coëfficiënten zijn statistisch significant bij $\alpha = .05$.

⁶ Onder 'anderstalige leerlingen' worden leerlingen verstaan die thuis soms of altijd een andere taal spreken dan Nederlands.

⁷ In een ander multiniveau regressiemodel dat hier niet wordt getoond, werd onderzocht of het percentage anderstalige leerlingen van een school dat thuis altijd een andere taal spreekt, soms wel een differentieel school compositie-effect opleverde in 2011 ten aanzien van in 2003. Dit bleek niet het geval.

8. Conclusies en discussie

In deze bijdrage werden de TIMSS 2003 en TIMSS 2011 resultaten inzake wiskundeprestaties en wiskunde zelfconcept met elkaar vergeleken. We herinneren eraan dat het om voorlopige resultaten gaat, gezien de internationaal berekende scores nog niet beschikbaar zijn. In deze paragraaf worden de belangrijkste overeenkomsten en verschillen tussen de vaststellingen betreffende de TIMSS-data in 2003 en 2011 opgesomd. Daarnaast staan we stil bij enkele beperkingen van het huidige onderzoek en doen we suggesties voor toekomstige analyses.

Voornaamste overeenkomsten tussen TIMSS 2003 en TIMSS 2011

- In beide afnamejaren presteren Vlaamse leerlingen gemiddeld nagenoeg even goed op de TIMSS wiskundetoets.
- In beide afnamejaren presteren leerlingen die thuis (bijna) altijd Nederlands spreken het hoogst op de TIMSS wiskundetoets in vergelijking met leerlingen die dat slechts soms of nooit doen.
- In beide afnamejaren hebben de leerlingen met de minste boeken thuis (0-10 boeken) gemiddeld de laagste wiskundescore.
- In beide afnamejaren presteren leerlingen die één of meerdere leerjaren vertraging hebben significant slechter op de TIMSS wiskundetoets dan leerlingen die op leeftijd zitten en leerlingen die één of meerdere leerjaren versneld zijn significant beter.
- In beide afnamejaren hebben jongens een significant hoger wiskunde zelfconcept dan meisjes.
- In beide afnamejaren hebben leerlingen die thuis nooit Nederlands spreken een nagenoeg even hoog wiskunde zelfconcept als leerlingen die thuis (bijna) altijd Nederlands spreken.
- In beide afnamejaren is een duidelijk geboortemaand-effect zichtbaar voor zowel de wiskundeprestaties als het wiskunde zelfconcept. Dit houdt in dat januari-leerlingen beduidend hogere wiskundeprestaties en een beduidend hoger wiskunde zelfconcept vertonen dan december-leerlingen.
- In beide afnamejaren hebben leerlingen die één of meerdere leerjaren vertraging hebben een significant lager wiskunde zelfconcept dan de leerlingen die op leeftijd zitten. De leerlingen die één of meerderde leerjaren versneld zijn, vertonen op hun beurt weer een significant hoger wiskunde zelfbeeld dan de leerlingen die op leeftijd zitten.
- In beide afnamejaren bevindt zich ongeveer 90% van de variantie in wiskundeprestaties op leerlingniveau en ongeveer 9% op schoolniveau.
- In beide afnamejaren verklaren leerlingkenmerken vooral verschillen in wiskundeprestaties tussen scholen en in mindere mate verschillen tussen leerlingen binnen een school.
- In beide afnamejaren wordt het grootste gedeelte van de totale variantie in wiskundeprestaties verklaard door leerlingkenmerk 'geboortemaand' en het kleinste gedeelte door het leerlingkenmerk 'sekses'.
- In beide afnamejaren worden de verschillen in wiskundescores tussen scholen het meest worden verklaard door het geboortemaand van leerlingen. Dit duidt op grote verschillen tussen scholen in percentage zittenblijvers.
- In beide afnamejaren bevindt zich ongeveer 99% van de variantie in wiskunde zelfconcept op leerlingniveau, 0% op klasniveau en ongeveer 1% op schoolniveau.
- In beide afnamejaren verklaart de sekses van de leerlingen het grootste percentage van de totale variantie in het wiskunde zelfconcept van leerlingen en de thuistaal van leerlingen geen.

- Het effect van de gemiddelde SES van een school is hetzelfde in 2003 en in 2011; en is onbestaande.
- Het percentage leerlingen met een andere thuistaal van een school heeft een negatief effect op de wiskundeprestaties van leerlingen. Dit negatieve effect is even groot in 2003 en 2011.

Voornaamste verschillen tussen TIMSS 2003 en TIMSS 2011

- Jongens presteerden in 2011 significant beter op de TIMSS wiskundetoets dan in 2003.
- In 2011 presteerden jongens significant beter op de TIMSS wiskundetoets dan meisjes. In 2003 was er geen sprake van een significant sekseverschil in wiskundescore.
- Er is een dalende trend zichtbaar in het spreken van Nederlands thuis. In 2003 sprak nog 85% van de leerlingen altijd of (bijna) altijd Nederlands thuis, in 2011 was dat nog maar 74%.
- Leerlingen die thuis (bijna) altijd Nederlands spreken doen het net als leerlingen die slechts soms thuis Nederlands spreken beter op de TIMSS wiskundetoets in 2011 dan in 2003.
- In 2011 hebben leerlingen gemiddeld een hoger wiskunde zelfconcept dan in 2003.
- Leerlingen met een versnelde schoolloopbaan presteren in 2011 significant hoger op de wiskundetoets dan in 2003.
- Het percentage leerlingen met een versnelde schoolloopbaan is bijna verdubbeld in 2011 ten aanzien van 2003.
- Enkel in 2011 is er sprake van prestatieverschillen tussen klassen binnen eenzelfde school. In dit jaar zijn de wiskundeprestaties van leerlingen voor ongeveer 1.2% toe te schrijven aan verschillen tussen klassen.
- In 2011 verklaren alle leerlingkenmerken tezamen ruim 4% meer van de totale variantie in wiskundescores dan in 2003.
- In 2011 blijken de verschillen in wiskundescores tussen klassen binnen eenzelfde school het meest verklaard te kunnen worden door onderlinge verschillen in de thuistaal van leerlingen. Dit doet vermoeden dat er scholen zijn die hun leerlingen (mede) op basis van hun thuistaal in klassen groeperen
- Voorzichtig kan worden gesteld dat terwijl de segregatie tussen scholen voor wat betreft SES ('aantal boeken thuis') iets is afgenomen, de segregatie voor wat betreft etniciteit ('thuistaal') juist is toegenomen tussen 2003 en 2011.
- In 2003 worden verschillen tussen scholen in het wiskunde zelfconcept van leerlingen enkel deels verklaard door onderlinge verschillen in het gemiddelde aantal boeken dat de leerlingen thuis hebben.
- In 2011 worden verschillen tussen scholen in het wiskunde zelfconcept van leerlingen enkel deels verklaard door onderlinge verschillen in het gemiddelde geboortjaar van de leerlingen.
- In 2011 verklaren alle leerlingkenmerken tezamen 1.6% meer van de totale variantie in wiskunde zelfconcept dan in 2003

Beperkingen en suggesties voor vervolgonderzoek

In de loop van 2012 worden de officiële Vlaamse resultaten voor TIMSS 2011 vrijgegeven. De officiële analyses zullen dan pas plaatsvinden. Op dat moment zal ook worden onderzocht of analyses ter vergelijking van de prestaties van leerlingen op de TIMSS wereldoriëntatietoets in 2003 en 2011 tot vergelijkbare resultaten leiden ten opzichte van de hier gepresenteerde resultaten voor wiskunde. Bovendien kan dan worden nagegaan hoe de gemiddelde Vlaamse prestaties zich verhouden tot die

van de andere deelnemende landen aan de TIMSS-studie. Verder lenen de uitgebreide TIMSS-data zich voor meer uitvoerige analyses waarin de invloed van kenmerken van het onderwijssysteem, de school, de klas en de leerlingen op zowel de prestaties als verschillende non-cognitieve variabelen zoals het schools welbevinden en de leerattitude wordt onderzocht. Ook verschillende variabelen op leerkracht- en schoolniveau kunnen met de TIMSS-data worden bestudeerd.

Een van de beperkingen van de huidige analyses betreft de kwaliteit van de proxy-indicator voor SES, 'aantal boeken thuis'. De kwaliteit van deze variabele als SES-indicator is twijfelachtig te noemen aangezien deze variabele vooral informatie biedt over het culturele kapitaal van een leerling en bijvoorbeeld niet over het economische kapitaal (bv. schatting van het gezinsinkomen), een eveneens belangrijk onderdeel van SES. Bovendien is het maar de vraag in hoeverre leerlingen van het vierde leerjaar in staat zijn om een betrouwbare schatting te maken van het aantal boeken waarover ze thuis beschikken. In veel gezinnen zijn de boeken thuis immers niet allemaal netjes opgeslagen in één kast, maar verspreid over meerdere kasten en plekken in het huis. Dit maakt het voor kinderen wellicht lastig om een totaalschatting te maken. In TIMSS 2011 hebben we middels de toevoeging van twee nationale items die de kinderen vroeg naar de werkende status zowel de vader als de moeder, getracht meer inzicht te krijgen in de SES van leerlingen. Door toevoeging van deze twee items steeg het percentage totaal verklaarde variantie in wiskundescores van 13.3% naar 14.3%. Echter, ook hier is weer de vraag in hoeverre leerlingen uit het vierde leerjaar een juist beeld van de tewerkstelling van hun ouders hebben. Een betere oplossing zou daarom zijn om de TIMSS-data te koppelen aan SES-gegevens van het Vlaamse Ministerie van Onderwijs zoals het opleidingsniveau van de ouders en/of het gemiddelde gezinsinkomen van de leerlingen. Deze koppeling zou leiden tot meer betrouwbare resultaten met betrekking tot de invloed van SES op leerlingprestaties. Bij eventuele deelname aan IEA-onderzoeken lijkt het aangewezen om telkens een korte oudervragenlijst af te nemen (zoals in het verleden gebeurd is bij deelname aan TIMSS (grade 8) in 1999 en 2003). Het voorgaande betekent ook dat bij het geven van schoolfeedback (cfr. het afzonderlijk toegestuurd schoolfeedbackrapport) vooralsnog in een beperkte mate kon gecorrigeerd worden voor de verschillende instroom tussen scholen. Dit impliceert dat vooralsnog vooral scholen met een sterk leerlingenpubliek vernomen hebben dat ze effectief zijn, en omgekeerd. Een sterkere correctie voor de instroomverschillen zou wellicht een ander beeld geven.

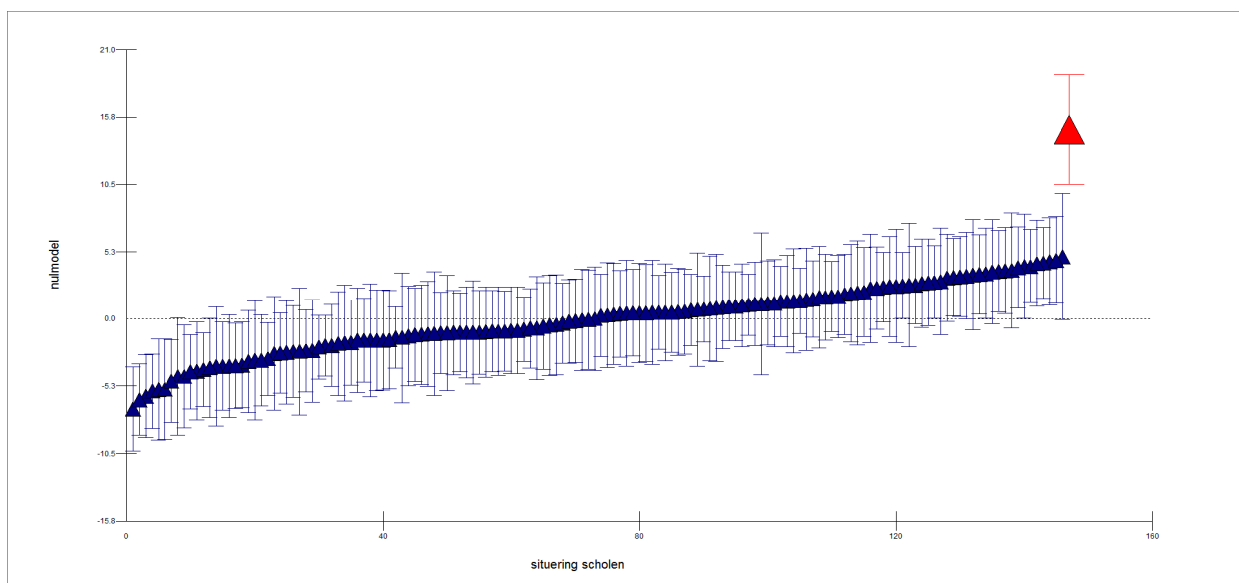
9. Referenties

- Meelissen, M. R. M., & Drent, M. (2008). *TIMSS-2007 Nederland. Trend in leerprestaties in exacte vakken in het basisonderwijs*. Enschede: Universiteit Twente, Vakgroep Onderwijsorganisatie en -management.
- Mullis, I. V. S., Martin, M. O., Ruddock, G. J., O'Sullivan, C., Arora A., & Erberber, E. (2005). *TIMSS 2007 Assessment Frameworks*. Boston: Boston College, TIMSS & PIRLS International Study Center.
- Van Damme, J., Van Den Broeck, A., Brusselmans-Dehairs, C., & Valcke, M. (2004). *Vlaanderen in TIMSS 2003*. Brussel: Ministerie van de Vlaamse Gemeenschap.
- Van Landeghem, G., & Van Damme, J. (2004). *De ongekwalificeerde uitstroom - Een geactualiseerde schatting* (LOA-rapport nr. 18). Leuven: Steunpunt 'Loopbanen doorheen Onderwijs naar Arbeidsmarkt', Cel voor Analyse van Onderwijsloopbanen.

Appendix

Outlier-school

In onderstaande figuur (Figuur 1) zijn de gemiddelde wiskundescores van alle 147 scholen die aan TIMSS 2011 hebben deelgenomen gerangschikt. De scholen zijn geordend van zwakst presterend (helemaal links in de figuur) tot best presterend (helemaal rechts in de figuur). De zogenaamde outlier-school is als een grotere driehoek weergegeven. In de figuur is duidelijk te zien dat de gemiddelde wiskunde-prestatie van deze school heel wat hoger ligt dan de gemiddelde wiskunde-prestaties van de overige scholen. In Tabel 1 is verder te zien dat het verschil in gemiddelde wiskundescore tussen de outlier-school en de overige scholen maar liefst 24.13 punten bedraagt. Dit verschil is statistisch significant (zie Tabel 2). Wanneer we er toch voor hadden gekozen om de outlier-school in de analyses mee te nemen dan had dat er wellicht toe geleid dat gemiddelde wiskundescore van alle leerlingen onterecht hoger kwam te liggen.



Figuur 1. Situering van alle 147 deelnemende scholen uit TIMSS 2011 op basis van de gemiddelde wiskundescore, niet rekening houdend met de instroomkenmerken van de leerlingen. Grote driehoek: de outlier-school.

Tabel 1. Gemiddelde score voor wiskunde van de leerlingen van de outlier-school en de overige scholen.

	N leerlingen	Gemiddelde	SD	Minimum	Maximum
Outlier-school	17	74.84	9.31	57.82	88.45
Overige scholen*	5107	50.06	10.17	11.55	88.45

* Dit zijn alle scholen zonder de outlier-school.

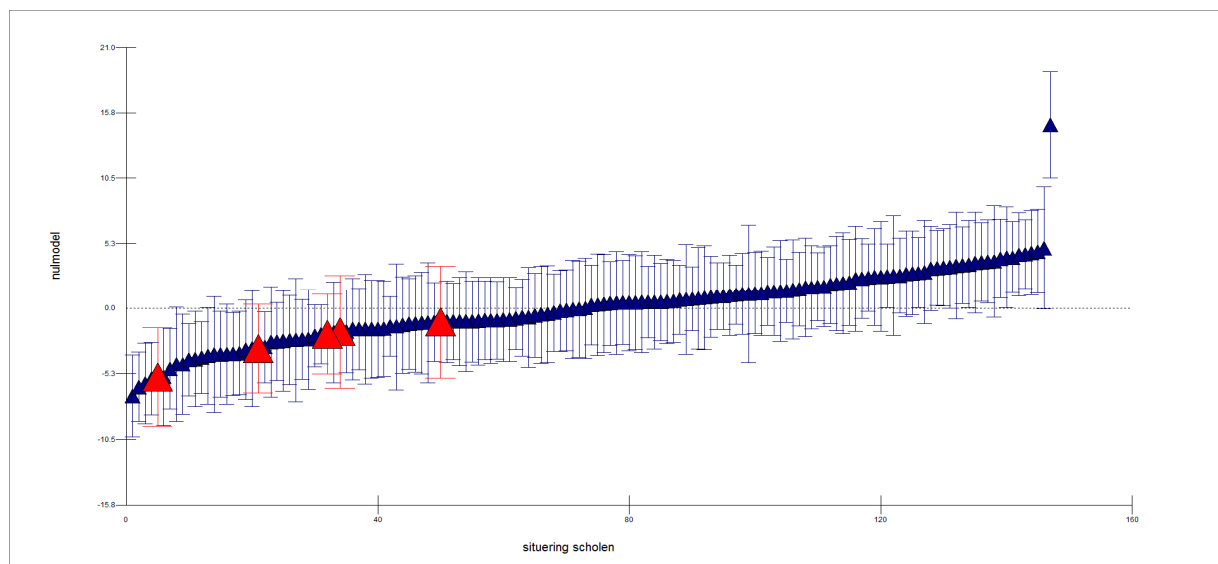
Tabel 2. Ongepaarde t-toets resultaten ter vergelijking van de gemiddelde wiskundescore van de leerlingen van de outlier-school en die van de overige scholen in TIMSS 2011.

Outlier-school/overige scholen	
t-waarde	10.03
df	5122
P-waarde	< .01

BLO-scholen

Vijf BLO-scholen (type 3/type 8) namen deel aan de TIMSS 2011 wiskundetoets met een leerlingenpubliek dat de leerstof van het vierde leerjaar voor wiskunde minstens gedeeltelijk verwerkt had. Dit had als consequentie dat de gemiddelde leeftijd van de deelnemende leerlingen van de BLO-scholen hoger lag dan die van de leerlingen van de gewone lagere scholen. Meer specifiek hadden de leerlingen in het BLO gemiddeld tweeëneenhalf jaar vertraging (zie Tabel 3).

In Figuur 2 zijn opnieuw de 147 scholen uit TIMSS 2011 gerangschikt naar gemiddelde wiskundeprestatie. De vijf BLO-scholen zijn als grotere driehoeken aangeduid. Uit deze figuur blijkt dat slechts één BLO-school een beneden gemiddelde wiskundeprestatie vertoonde in vergelijking met het Vlaamse gemiddelde. Meer informatie over de gemiddelde wiskundescores van de BLO-scholen is te vinden in Tabel 3. Wanneer de gemiddelde wiskundescore van de BLO-scholen wordt vergeleken met gemiddelde wiskundescore van de scholen voor gewoon lager onderwijs, blijkt dat gemiddeld genomen, de BLO-scholen een significant lagere wiskundescore vertoonden dan de gewone scholen (zie Tabel 4).



Figuur 2. Situering alle 147 deelnemende scholen uit TIMSS 2011 op basis van de gemiddelde wiskundescore, niet rekening houdend met de instroomkenmerken van de leerlingen. Grote driehoeken: de vijf BLO-scholen.

Tabel 3. Gemiddelde score voor wiskunde en gemiddelde vertraging van de leerlingen van de BLO-scholen en de overige scholen.

	N leerlingen	Gem. aantal jaren vertraagd	Gemiddelde	SD	Minimum	Maximum
BLO-school a	11	2.27	46.91	6.67	33.74	56.54
BLO-school b	36	2.86	47.48	7.53	30.05	62.88
BLO-school c	13	2.77	48.12	10.99	29.58	66.88
BLO-school d	19	1.74	41.50	10.33	12.52	64.08
BLO-school e	28	2.81	45.56	10.58	23.79	76.72
Alle BLO-scholen	107	2.49	45.94	9.40	12.52	76.72
Overige scholen*	5017	0.18	50.24	10.26	11.55	88.45

* Dit zijn alle scholen zonder de BLO-scholen.

Tabel 4. Ongepaarde t-toets resultaten ter vergelijking van de gemiddelde wiskundescore van de leerlingen van de BLO-scholen en die van de scholen voor gewoon lager onderwijs in TIMSS 2011.

BLO-scholen/overige scholen	
t-waarde	-4.30
df	5122
P-waarde	< .01