

Universiteit Antwerpen
Instituut voor Onderwijs- en Informatiewetenschappen

**Heeft klasgrootte een effect op de leesvaardigheid van leerlingen
in het 4^e leerjaar basisonderwijs?**

Analyse van de PIRLS 2006-data

Sofie De Bondt

Masterproef voorgelegd met het oog op het
behalen van de graad van master in de
Opleidings- en Onderwijswetenschappen

Promotor: prof. dr. S. De Maeyer
Co-promotor: /

Abstract

Klasgrootte is een onderwerp dat in het verleden reeds veel aandacht kreeg. Het is nog steeds een actief thema in Vlaanderen, gezien het beleid inzet op kleinere klassen. Dit schept de verwachting dat klassenverkleining oplevert. Is dit wel zo? Er is weinig Vlaams onderzoek dat dit nagaat. Bovendien zijn er enkele problemen met bestaand onderzoek: ze gebruiken regelmatig een vertekende meting van de variabele klasgrootte, er worden vaak ongeschikte analysemethoden gehanteerd en studies verschillen op basis van de gekozen prestatie maat, de onderzochte populatie en het al dan niet opnemen van controlevariabelen. Deze studie analyseert data van de internationaal vergelijkende PIRLS-studie uit 2006. Ze wil hierbij onderzoeken of er een verband vastgesteld kan worden tussen de klasgrootte en de behaalde leesprestaties van Vlaamse leerlingen uit het vierde leerjaar basisonderwijs. Om rekening te houden met de hiërarchische structuur die aanwezig is in de data, wordt gebruik gemaakt van multilevelanalyses. Er worden modellen opgezet waarbij controlevariabelen opgenomen zijn voor relevante kenmerken op leerling-, klas-, en schoolniveau. Op basis van de resultaten wordt geconcludeerd dat er in geen van de gevallen een significant effect aanwezig is van de klasgrootte. Deze variabele blijkt geen voorspeller te zijn voor de resultaten die de leerlingen behalen op de leesvaardigheidstest.

Dankwoord

Deze masterproef zou niet zijn wat ze nu is, zonder de steun en hulp van enkele personen. Ik wil dan ook een woordje van dank richten aan iedereen die op zijn manier geholpen heeft om dit werk te kunnen realiseren.

Vooreerst gaat mijn dank uit naar mijn promotor, prof. dr. Sven De Maeyer. Zijn kritische blik en waardevolle feedback hebben me steeds aangezet om dat stapje hoger te zetten op de ladder. Mede dankzij zijn begeleiding ben ik erin geslaagd tot boven te klimmen, hoewel het een proces was waarbij ik wel eens terug naar beneden klom, of enkele treden miste.

Ik wil ook prof. dr. Vincent Donche danken voor het tussentijds nalezen van de masterproef. Hij kon vanop een afstand het werk bekijken en zo terechte opmerkingen ter verbetering maken.

Mijn familie maakte het mogelijk dat ik deze studies kon aanvatten en tot een goed einde kon brengen. Hun steun en geduld betekenden erg veel. Bedankt voor de kansen die ik van jullie heb gekregen!

Hoewel ze niet bewust betrokken geweest zijn bij dit werk, wil ik toch de leerkrachten, leerlingen en onderzoekers van de PIRLS-studie bedanken die ervoor zorgden dat gegevens verzameld werden. Zonder deze grootschalige data was het onmogelijk geweest de analyses uit te voeren.

Tot slot wil ik alle medestudenten en vrienden bedanken die hun steentje hebben bijgedragen door het proces mee vorm te geven en te ontdoen van obstakels. Het was fijn om af en toe mijn hart eens te kunnen luchten bij jullie, om bemoedigende woorden te horen, en tips en opmerkingen uit te wisselen.

Inhoud

1	Inleiding en probleemstelling	1
2	Theoretisch kader	8
2.1	Het effect van klasgrootte	8
2.2	Het begrip ‘klasgrootte’	9
2.3	Prestaties	10
2.4	Controlevariabelen	11
2.5	Differentiële effecten	12
2.6	Curvilineariteit	13
3	Onderzoeksvragen	14
4	Methodologie	17
4.1	Databestand PIRLS 2006	17
4.2	Variabelen	18
4.2.1	Afhankelijke variabele: leesprestatie	18
4.2.2	Onafhankelijke variabele: klasgrootte	20
4.2.3	Onafhankelijke controlevariabelen	21
4.3	Analysemethode	22
5	Onderzoeksresultaten	25
5.1	Is er een effect van klasgrootte op individuele leesprestaties van leerlingen?	25
5.2	Is er sprake van een effect als individuele leerlingkenmerken meegenomen worden?	27
5.3	Is er sprake van een effect als gecontroleerd wordt voor klaskenmerken?	27
5.4	Is er een effect van klasgrootte op individuele leesprestaties na controle voor schoolkenmerken?	28
5.5	Verschilt het effect van klasgrootte op leesprestaties tussen klassen en scholen wanneer er gecontroleerd wordt voor de leerling-, klas-, en schoolkenmerken?	29
5.6	Analyses op basis van het 3-level model	30
5.7	Onderzoeken van een curvilineair verband	31
5.8	Overzichtstabel van alle analysemodellen	33
6	Conclusie en discussie	35
7	Referenties	38
8	Bijlagen	41
8.1	Analyse van het nulmodel level 2= school	42
8.2	Analyse van het nulmodel level 2= klas	42
8.3	Analyse van het RImodel	43
8.4	Analyse van het model met leerlingkenmerken toegevoegd: geslacht, leeftijd, SES, cultureel kapitaal, thuistaal	43
8.5	Analyse van het model met klaskenmerken toegevoegd: geslacht en ervaring van de leerkracht, gemiddeld leesniveau, gemiddelde SES, gemiddelde thuistaal	44
8.6	Analyse van het model met schoolkenmerken toegevoegd: schoolgrootte, urbanisatiegraad	45
8.7	Analyse van het model met alle kenmerken samen	46
8.8	Analyse om een curvilineair verband te testen	48

1 Inleiding en probleemstelling

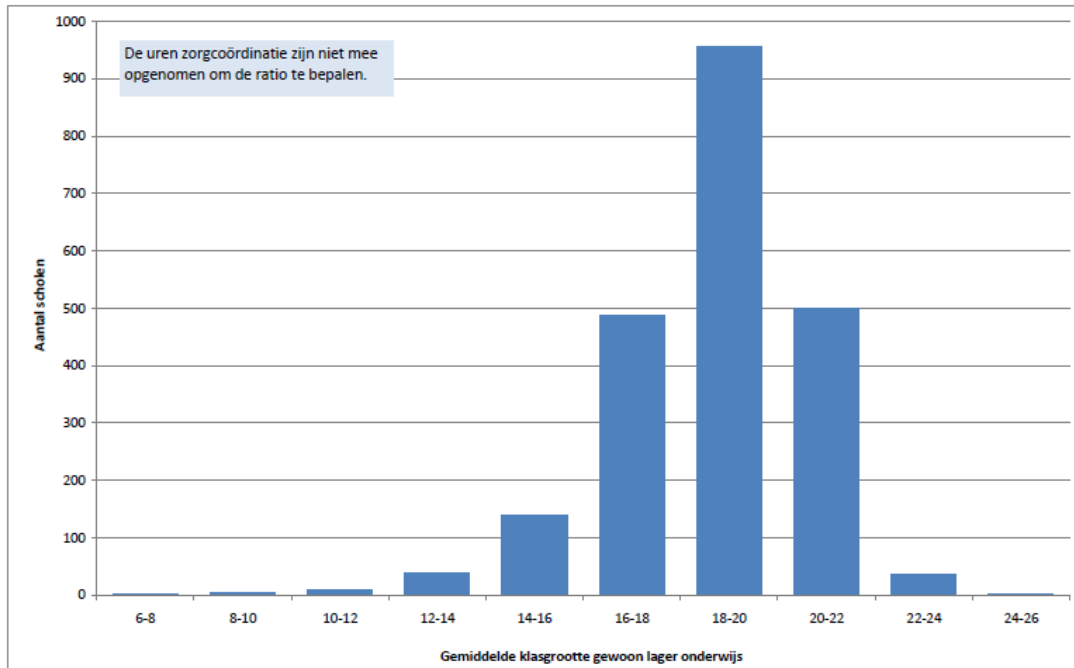
Klasgrootte is een onderwerp dat reeds veel aandacht kreeg in de bestaande onderzoeksliteratuur. Hierbij zijn verschillende effecten nagegaan, waaronder een mogelijke relatie tussen klassengroottes en de prestaties van leerlingen. In 1978 maakten Glass en Smith al een meta-analyse van het tot dan toe geleverde onderzoek. Zij concludeerden dat klasgrootte significante voordelen kon opleveren wanneer de grootte minder dan 20 leerlingen per klas bedraagt. Toch blijkt deze materie nog niet unaniem uitgeklaard en evenmin in de vergetelheid geraakt. Het is nog steeds een actief thema in Vlaanderen, gezien het beleid inzet op kleinere klassen. Vlaams Minister van Onderwijs, Jeugd, Gelijke kansen en Brussel, Pascal Smet, maakte op 20 juli 2011 bekend dat het nieuwe omkaderingssysteem voor het basisonderwijs principieel is goedgekeurd. Dit betekent dat voor kleuter- en lager onderwijs extra middelen worden voorzien. Zo zouden er met een budget van 52,7 miljoen euro tegen september 2012 liefst 1300 extra voltijdse leerkrachten in dienst worden genomen in het kleuteronderwijs. De extra aangeworven leerkrachten moeten ervoor zorgen dat het gemiddeld aantal kinderen per kleuterklas daalt van 17,5 naar 16,2. Voor het lager onderwijs reflecteert deze optrekking van de omkadering in een daling in het gemiddelde leerlingen/leraar-ratio van 17,0 naar 16,8 (Amkreutz, 2001). Hierbij kan de vraag gesteld worden naar de oorzaken die tot zulke beslissing leiden; vanwege een roep van leerkrachten om de werkdruk te verminderen, vanwege een visie die gelooft in de baten van kleinere klassen of vanwege andere motieven. Een centrale doelstelling van onderwijs zou moeten zijn een zo optimaal mogelijk leef- en leerklimaat te creëren voor leerlingen. Het Vlaams onderwijsbeleid lijkt ervan uit te gaan dat kleinere klasgroottes leiden tot betere leerprestaties. Maar is dat wel zo? Klopt de verwachting dat het verkleinen van klassen oplevert? In deze studie wordt nagegaan of er sprake is van zo'n verband.

Op de website van Onderwijs en Vorming Vlaanderen (2012) stelt men gegevens ter beschikking over de gemiddelde klasgroottes in Vlaamse scholen voor kleuter- en lager onderwijs.

“Om transparantie te brengen biedt het Ministerie van Onderwijs en Vorming een aantal relevante gegevens omtrent de gemiddelde klasgrootte van alle scholen van het gewoon basisonderwijs in Vlaanderen. Het gaat om de situatie op 1 februari 2011, zoals gekend bij het Agentschap voor Onderwijsdiensten (AgODi) op 4 mei 2011. Uit deze cijfers blijkt dat 94% van de scholen in hun kleuteraanbod klassen kunnen realiseren met gemiddeld 24 leerlingen of minder. In het lager onderwijs kunnen 96% van de scholen gemiddelde klassen vormen met 22 leerlingen of minder” (Agentschap voor Onderwijsdiensten, 2011).

Zoals te zien is op Figuur 1 (Agentschap voor Onderwijsdiensten, 2011), is er een redelijke spreiding van de gemiddelde klasgroottes van Vlaamse lagere scholen. Er is een minimumgrootte van 6 tot 8 leerlingen per klas, de grootste klassen tellen 24 à 26 leerlingen. Het grootste aandeel scholen heeft

gemiddelde klasgroottes van 18 tot 20 leerlingen. Er kan hier gesteld worden dat er wel degelijk uiteenlopende klasgroottes zijn. Indien er sprake is van een effect van klasgrootte, dan kan dit duidelijk ook echt iets uitmaken. Het is niet zo dat klassen niet sterk van elkaar verschillen. Een effect kan zich dus manifesteren, de vraag is of dit daadwerkelijk zo is.



Figuur 1: Grafiek met de gemiddelde klasgroottes van Vlaamse scholen in het lager onderwijs

Bestaand Vlaams/Nederlands onderzoek

In tegenstelling tot het feit dat er wel de (recente) belangstelling is vanuit het onderwijsbeleid, bestaat er in de Vlaamse onderzoekswereld weinig literatuur die zich toespitst op klasgrootte. Onderzoek met betrekking tot klasgroottes is veelal in de Verenigde Staten en het Verenigd Koninkrijk te situeren. In Vlaanderen is er in dit domein weinig onderzoek verricht. Er is enig aanknopingspunt met Vlaams onderzoek dat het effect bekeek van leeftijd, socio-economische status en/of klasgrootte op de gemiddelde leesprestaties, en dit op landenniveau (Van Damme, Liu, Vanhee, & Pustjens, 2010). De data van PIRLS 2001 en 2006 is door hen onderhevig geweest aan het testen van onderzoeksvragen over zulke effecten. Zowel bij de gegevens uit 2001 en 2006 werden negatieve correlaties gevonden tussen klasgrootte en leesprestaties, maar deze bleken niet significant (Van Damme et al., 2010). De gegevens van België werden niet verwerkt in de vergelijking tussen het effect klasgrootte-prestaties over de twee meetperiodes heen, omdat België in 2001 niet deelnam aan de PIRLS-meting. Daarnaast is er een probleem met de wijze van de verwerking. Die gebeurde aan de hand van meervoudige regressie-analyses. Daardoor stelt het probleem van aggregatie zich hier. Gegevens op leerlingniveau werden samengevoegd om uitspraken te doen op nationaal niveau. Het gevaar van ‘ecological fallacy’ (Robinson, 1950) treedt op deze manier op. Wester, Renckstorf en Scheepers (2006) stelden:

“Wanneer onderzoekers een samenhang op een boven-individueel niveau vaststellen, kunnen ze daarmee nog niets zeggen over deze samenhang op het individuele niveau” (p.237). Het onderzoek van Van Damme, Liu, Vanhee en Pustjens (2010) kan bijgevolg geen betrouwbare conclusies trekken op het lagere individueel niveau. De interesse (onder andere vanuit het beleid) en scope van dit onderzoek liggen wel bij dit niveau. We willen namelijk weten wat het effect is van klasgrootte op individuele leerlingen, niet op gemiddelden per klas, school of zelfs land. Door met dit laagste hiërarchische niveau te werken, kunnen we nagaan of de effecten mee beïnvloed worden door kenmerken van deze leerlingen en of daar trends uit af te leiden zijn (bijvoorbeeld: klasgrootte heeft een significant effect en dit is sterker voor jongens).

Verder is er nog weinig Nederlandstalig of Belgisch onderzoek gevoerd naar de mogelijke effecten van klasgrootte. Bosker (1998) bracht een onderzoek uit naar het klasgroottevraagstuk in basisscholen. Hij voerde een multilevelanalyse uit op basis van data over klasgrootte en prestaties verzameld in 416 scholen in graad 2, 4, 6 en 8 in Nederland. Er werden covariaten als sociaal-etnische status, IQ, welbevinden en geslacht betrokken. Hij maakte een onderscheid tussen taal- en rekenprestaties. Klasgrootte blijkt een negatief verband te hebben met de leesprestaties wanneer klassen meer dan 20 leerlingen bevatten. In de tweede graad telt klasgrootte voor 13% van de variatie in leesprestaties (Bosker, 1998).

In 1999 verscheen daarnaast een Nederlandse studie van Dobbelsteen, Levin en Oosterbeek (1999) waarin ze tot de constatacie komen dat “verkleining van klassen niet gepaard gaat met betere leerprestaties” (Dobbelsteen, Levin, & Oosterbeek, 1999, p.2). Zij zijn niet de enige auteurs die op basis van Vlaamse data tot deze conclusie komen. Er werd onderzoek verricht door Woessmann en West (2006) naar het effect van klasgrootte op basis van data van de Trends in International Mathematics and Science Study (TIMSS), ‘het broertje’ van PIRLS dat wiskunde- en wetenschappenprestaties meet bij leerlingen in het basisonderwijs. Het onderzoek toonde aan dat in negen schoolsystemen, waaronder de Vlaamse en Waalse, grootschalige effecten op zowel wiskunde en wetenschappen eveneens uitgesloten kunnen worden. Dit roept de vraag op of dit ook geldt voor andere cognitieve leerprestaties, zoals bijvoorbeeld leesprestaties, als afhankelijke variabele.

Bestaand internationaal onderzoek

Internationaal is heel wat aan onderzoek beschreven over klasgrootte en mogelijke effecten, maar eenduidige conclusies zijn er niet. Een meta-analyse van Hanushek (1997) herbekeek 227 schattingen over de impact van de leerkracht-leerling ratio op studentenprestaties, en rapporteerde 15% als significant en positief, 13% als significant en negatief en 72% als statistisch insignificant (Mishel & Rothstein, 2002). Op deze inconsistenties werd reeds vroeger gewezen. Al in 1978 stelde Porwell dat de onderzoeksbevindingen over het effect van klasgrootte op studentenprestaties contradictorisch en niet overtuigend zijn. Meer dan dertig jaar later is de literatuur nog aanzienlijk aangedikt en blijven er

vele verschillende onderzoekspistes die leiden tot uiteenlopende conclusies. Deze verschillende resultaten kunnen verklaard worden door velerlei achterliggende oorzaken: andere onderzoeksmethodes, verschillende populaties, al dan niet meegenomen controlevariabelen, enzovoort. In wat volgt wordt de bestaande literatuur met verschil- en struikelpunten in een overzicht beschreven en de belangrijkste oorzaken voor tegenstrijdige bevindingen worden overlopen. Bepaalde methodologieën en conclusies kregen al meer vertrouwen dan andere. Er werd dan ook door verscheidene auteurs kritiek geuit op voorgaand onderzoek. Deze kritieken worden hier beschreven.

Eén van de eerste bekommernissen is het feit dat veel van de bestaande onderzoeken die men nog steeds als sleutelpublicaties ziet binnen de literatuur over klasgrootte, dateren uit de jaren '80. Denken we bijvoorbeeld aan de meta-analyse van Glass en Smith uit 1978 en het grootschalige STAR-experiment dat werd opgestart in 1985. Hierbij reist de vraag of deze resultaten nog wel relevant zijn voor de huidige onderwijscontext, want het onderwijs ziet er dezer dagen aanzienlijk anders uit dan toen. Onderzoek wordt daarnaast vaak gekenmerkt door de tijdsperiode waarin het werd uitgevoerd. Volgens Glass en Smith (1978) kent onderzoek naar klasgrootte en zijn relatie met prestaties vier stadia. De verschillen in onderzoeksaanpak kunnen voortkomen uit deze fasen. Bij elke nieuwe fase verhoogde namelijk de complexiteit van de onderzoeksmethodologie steeds, en de vraag naar het effect van klasgrootte op prestaties werd onderzocht met verschillende motieven (Glass & Smith, 1978).

Beginnende onderzoeken naar klasgrootte en prestaties worden verweten dat ze te weinig kwantitatieve technieken gebruikten en eerder narratief waren. Studies werden gekenmerkt door hun niet-experimentele logica. Daarom werden deze studies geklasseerd onder de pre-experimentele fase van 1895 tot 1920 (Glass & Smith, 1978). Begin jaren 1920, waarbij Glass en Smith (1978) spreken over de primitief experimentele fase, werd de vraag over klasgrootte en prestaties benaderd met betere methoden zoals experimenten. In de jaren '40 was het klasgrootte-onderzoek inactief vanwege Wereldoorlog II, maar het herleefde vanaf de jaren '50, benoemd als de grote-groep technologiefase. Er verschenen in deze tijd verschillende grootschalige empirische studies om het nationaal onderwijsbeleid te informeren. Deze grote studies bevatten data over de relatie tussen klasgrootte en prestaties. Studies die verschenen in de jaren '70 richtten zich op het onderzoeken van de voordelen van individualisatie. Tijdens deze fase werden dan ook verscheidene experimenten opgezet met erg kleine groepen, beperkt tot één of enkele leerlingen (Glass & Smith, 1978).

Een tweede probleem met bestaand onderzoek dat verdient om bediscussieerd te worden, is de gebruikte onderzoeksmethode. Een aanzienlijk deel van het gevoerde onderzoek over klasgrootte en prestaties bestudeert gegevens door middel van regressie-analyses. Veelal wordt daartoe de kleinste-kwadratenmethode, oftewel de ordinary least squares (OLS) methode, gehanteerd. Dit is voor deze onderzoeksopzet echter geen goede techniek. Er wordt namelijk geen rekening gehouden met een belangrijk kenmerk van de gegevens: er bestaat hiërarchie in de data. Omdat leerlingen genest zijn

binnen klassen en scholen, dienen deze niveaus mee in ogenschouw genomen te worden. Dit pleit ervoor om multilevelanalyses als onderzoeksmethode aan te wenden. Resultaten op basis van onderzoek dat deze methode niet hanteert, zijn daarom mogelijk vertekend. In 4.3 wordt het gebruik van multilevelanalyses in deze studie nader besproken.

Een andere reden die de oorzaak kan zijn voor onderzoeksresultaten die elkaar tegenspreken, is het gegeven of onderzoek gebeurde in een experimentele dan wel niet-experimentele context. Onderzoek kan een effect nagaan op basis van bestaande verschillende klasgroottes. Anderzijds kan men een experiment opzetten. Project STAR, dat staat voor Student/Teacher Achievement Ratio, is zo'n grootschalig onderzoek dat gevoerd werd in Tennessee. Het is een random-assignment experiment met als onderliggend idee dat men valide evidentie kan krijgen over de impact van een bepaalde wel gedefinieerde behandeling, door subjecten willekeurig toe te wijzen aan testgroepen die een behandeling ondergaan en aan controlegroepen. Met deze vorm van experiment wil men effecten van andere factoren kunnen uitzuiveren (Mishel & Rothstein, 2002). Mishel en Rothstein (2002) stellen dat het zorgvuldig ontwerp van het STAR-experiment maakt dat deze resultaten overtuigender zijn dan de rest van de literatuur over klasgrootte. Effecten die afgeleid zijn van experimenten kunnen dan wel zuiver toe te schrijven zijn aan de variabele die men constant houdt (de klasgrootte), ze zeggen echter niets over de huidige onderwijssituatie. Men onderzoekt niet hoe klasgrootte effectief werkzaam is in de realiteit.

Ten derde schuilt er een probleem in de operationalisatie van de klasgroottevariabele. Door ontbrekende data over de werkelijke klasgrootte per student, gebruikte vroeger onderzoek vaak de 'pupil-teacher ratio' voor een hele school als de indicator van klasgrootte (Akerhielm, 1995). Coleman et al. (1966), Coleman, Hoffer en Killoore (1982), Chubb en Moe (1990), Card en Krueger (1992) en Van Damme et al. (2010), baseerden zich voor hun studies allemaal op geaggregeerde ratio's. De leerling-leraar ratio's worden gedefinieerd als het aantal studenten in de school gedeeld door het aantal full-time leerkrachten voor een hele school, en dit kan erg afwijkend zijn van de werkelijke grootte van de klas (Akerhielm, 1995). Onderzoek dat deze leerling-leraar ratio's gebruikte, gaf dus mogelijk geen juist beeld van de reële klasgrootte, en er zijn bijgevolg vraagtekens te plaatsen bij de conclusies die getrokken zijn op basis van deze gegevens.

Bovendien wordt opgemerkt dat voor een meer eerlijke en betrouwbare analyse, men eigenlijk het gemiddelde zou moeten nemen van alle groottes van de klassen waar de leerling in gezeten heeft tot het moment waarop de test voor het meten van de prestaties, wordt afgenomen (Hoxby, 2000). Wanneer een leerling toevallig één jaar in een grote klas zit (het jaar dat het effect wordt onderzocht), terwijl hij daarvoor steeds in kleine klassen zat, dan zal de leerling misschien hoog scoren op de prestatietoets dankzij de vele begeleiding van de leerkrachten die hij in de kleinere klassen kreeg. Een mogelijk effect van deze kleine klassen waar hij voordien in zat, kan zich dan niet manifesteren. Men

zou verkeerdelijk kunnen besluiten dat leerlingen in een grote klas, even goed presteren. De prestaties van leerlingen zijn met andere woorden een accumulatie van alle voorgaande leerervaringen in alle andere jaren, met dus effecten van gevarieerde klasgroottes (Bosker, 1998).

Er stelt zich nog een ander probleem in verband met de klasgrootte. Eerder onderzoek gebruikte vaak data waarin de leerlingenverdeling over verschillende klasgroottes mogelijk geen willekeurig proces was. Als een school een bewust beleid heeft om moeilijke of minder vaardige studenten aan kleinere klassen toe te wijzen, dan kan elk positief effect van een kleine klasgrootte op prestaties van leerlingen vermomd zijn, omdat deze studenten doorgaans lager scoren op tests. Evenzo, als men selectief en bewust gevorderde studenten aan de kleinere klassen toewijst, dan kan een positief verband tussen kleine klasgrootte en de prestaties van leerlingen worden overschat. Ouders die meer betrokken zijn, kunnen hun kinderen in kleine klassen plaatsen of bij parallelklassen kan men leerlingen op basis van hun verwachte prestaties indelen. Het sorteereffect bestaat dus zowel tussen scholen als binnen scholen, door ouders, leerkrachten of administratoren (Woessmann & West, 2006). Men spreekt in dit verband ook wel over het probleem van selectiviteit of endogeniteit: klassenverdeling gebeurt niet altijd per toeval. Verschillende auteurs wezen erop dat dit sorteereffect voor vertekende resultaten kan zorgen (Akerhielm, 1995; Dobbelsesteen et al., 1999; Hoxby, 2000; Woessmann & West, 2006).

Naast de onafhankelijke variabele die op verscheidene manieren invulling krijgt, kan de gekozen maat voor prestaties ook sterk variëren. Leerlingenprestaties kunnen gemeten worden voor verschillende vakgebieden en zelfs binnen vakken kan het effect mogelijk verschillen. De prestaties van leerlingen kunnen gemeten worden in een toets, of gevat worden in een graadletter die gebaseerd is op examencijfers maar daarnaast ook nog op andere elementen die minder met de prestaties te maken hebben (zoals participatie en attitude, de sympathie die de leerkracht heeft voor de leerling, en dergelijke). Er wordt daarom vaak verkozen om met toetsen te werken. De outputmaat werd door verscheidene auteurs aangehaald als belangrijke oorzaak voor tegenstrijdige resultaten (Arias & Walker, 2004; Hancock, 1996).

Bij onderzoek naar het effect van schoolkenmerken (waarvan klasgrootte er één van de velen is) op leerlingenprestaties, wordt vaak gebruik gemaakt van controlevariabelen. In de literatuur wordt meermaals vermeld dat het al dan niet aanwezig zijn van een effect van klasgrootte afhankelijk is van welke extra variabelen men inbrengt (Arias & Walker, 2004; Krueger, 1999). Dit aspect wordt uitgewerkt onder 2.4, waar vanuit de literatuur evidentie is aangebracht voor de variabelen die in deze studie betrokken zijn.

Tot slot is het al dan niet vinden van een effect van klasgrootte nog afhankelijk van de populatie.

Ten eerste heeft de leeftijd van de onderzochte groep hier zijn aandeel bij. Volgens Glass en Smith (1978) is er een gevonden verband tussen klasgrootte en prestatie dat sterker blijkt te zijn in de secundaire graden dan in de elementaire. Mosteller (1995) concludeerde op basis van de data van Project STAR dat er overtuigende evidentie is dat kleinere klassen helpen, ten minste in vroege klassen. Krueger (1999) stelt dan weer dat er een toename is in testscore van de leerlingen in kleine klassen en dat deze uitbreidt met ongeveer een percentiepunt per jaar in de daaropvolgende jaren. Deze bevinding dat de prestatiekloof tussen leerlingen in kleine en normale klassen blijft vergroten door een zwakke aangroei per jaar, wijst op een cumulatief voordeel. Nye, Hedges en Konstantopoulos (1999) gaven aan dat de effecten van klasgrootte minstens vijf jaar aanhouden. De effecten van de kleine klassen waarin leerlingen in de kleuterklas zitten, op de prestaties wanneer ze in Graad 8 zitten, blijken niet uitgedoofd te zijn ten opzichte van die van Graad 3.

Ten tweede wordt ook aangenomen dat onderzoekresultaten doorgaans niet te veralgemenen zijn naar zomaar eender welke populatie. Het effect van klasgrootte op 'achievement' is in de VS onderwerp van veel onderzoek, maar daarom kan het nog niet gegeneraliseerd worden naar andere landen of schoolsystemen (Woessmann & West, 2006). Zo zou men dus kunnen stellen dat buitenlands onderzoek niet noodzakelijk opgaat in de Vlaamse onderwijscontext. Onderzoek uit andere landen, die immers een verschillend onderwijsbeleid en andere methoden hanteren, kan met andere woorden niet zonder meer als geldend beschouwd worden voor de Vlaamse onderwijssituatie. In Vlaanderen zelf is er geen duidelijk systematisch onderzoek aanwezig.

Al deze voorgaande lacunes in het verhaal van de klasgroottes geven stof voor nieuw onderzoek.

2 Theoretisch kader

Om de probleemstelling verder te kaderen, worden hier enkele aspecten en begrippen uitgewerkt. Hieruit vloeien in hoofdstuk 3 de onderzoeksvragen voort. Eerst wordt er ingegaan op het effect dat klasgrootte al dan niet heeft op prestaties en mogelijke inhoudelijke verklaringen die daarvoor te vinden zijn. Daarna wordt stilgestaan bij het begrip ‘klasgrootte’, de voornaamste onafhankelijke variabele. De gemeten leerlingenprestaties dienen als endogene variabele die verklaard moet worden. Het is steeds de vraag welke prestaties men wil meten, en hoe men dat gaat doen. Onderzoekers zijn vrij in de keuze van de outputmaat waarop een effect van klasgrootte gemeten wordt. Het is echter wel belangrijk te erkennen dat deze keuze mede bepalend is voor de resultaten die men verkrijgt. In 2.3 wordt daarom uitgeweid over de afhankelijke variabele die in deze studie gebruikt zal worden. Eerder stelden we reeds dat de keuze van controlevariabelen een belangrijk aspect is in het opzetten van onderzoek naar het effect van klasgrootte. In 2.4 gaan we dieper in op dit punt. Daarnaast verdient ook het aspect van differentiële effectiviteit de aandacht. Tot slot wordt nog een woordje uitleg gegeven over curvilineariteit.

2.1 Het effect van klasgrootte

In de probleemstelling werd duidelijk dat verschillende oorzaken leiden tot het gegeven dat er geen unanieme conclusies luiden over het positieve, negatieve of insignificante effect van klasgrootte. Er werden in de literatuur verschillende inhoudelijke verklaringen aangehaald die uitleggen waarom klasgrootte een (positief dan wel negatief) effect heeft of niet. Deze verklaringen worden hier uiteengezet, samen met een korte opsomming van auteurs die zich aansluiten bij de bevinding dat er al dan niet een effect is van klasgrootte.

“Size matters”. Deze uitspraak komt uit het kamp van verscheidene onderzoekers (Bausell, Moody, & Walze, 1972; Glass & Smith, 1978; Krueger, 1999; Mishel & Rothstein, 2002; Molnar, Smith, Zahorik, & Palmer, 1999; Mosteller, 1995; Nye, Hedges, & Konstantopoulos, 2000; Tope, Groom, & Beeson, 1924). Misschien lijkt dit ook wel het meest logisch. Leraren kunnen de individuele leerlingen meer begeleiden wanneer ze hun tijd over minder leerlingen moeten verspreiden. In het boek van Glass, Cahen, Smith en Filby (1982) wordt onderzoeksevidentie samengevat om te demonstreren dat kleinere klassen meer bevorderlijk zijn voor verbeterde leerlingenprestaties dan grotere klassen; dat kleinere klassen meer kansen bieden om leerprogramma’s aan te passen aan individuele noden; dat leerlingen in kleinere klassen meer interesse hebben in leren; en tenslotte dat de moraal van leerkrachten beter is in kleinere klassen (Glass et al., 1982). Eerder toonden Smith en Glass (1980) al aan dat zo’n vastgesteld effect gemedieerd kan zijn door instructieverschillen tussen kleine en grote klassen. Deze stelling is gebaseerd op het model van schoolleren van Carroll (1963, 1989) (Bosker, 1998). In dit model zijn er vier intermediaire variabelen die mee een rol spelen in het verband tussen de bekwaamheid om te leren en prestaties. Eén van die vier variabelen is de kwaliteit van de

instructies. Men stelt hier dus dat leerkrachten mogelijk anders doceren aan kleine dan aan grotere klassen en dat dit het verschil in prestaties onrechtstreeks kan beïnvloeden. Bovendien zijn er minder verstoringen tijdens de les als de klassen kleiner zijn, en Arias en Walker (2004) halen aan dat leerlingen zich veiliger voelen in kleine klassen om vragen te stellen en te beantwoorden. Deze aspecten zouden positievere prestaties in de hand kunnen werken.

Uit het tegenkamp klinkt dat klassenverkleining een positieve (in plaats van inverse) samenhang vertoont met de prestaties: kleinere klassen leiden niet tot hogere prestaties, integendeel. Vanuit de literatuur over de sociaal-cognitieve leerpsychologie kan hiervoor een mogelijke verklaring geopperd worden. In kleinere klassen kan de leerkracht enerzijds wel meer tijd besteden per individuele leerling, maar anderzijds neemt het aantal leerlingen in de klas ook af. Er van uitgaande dat leerlingen ook van elkaar leren, en niet enkel van de leerkracht, kan dit ervoor zorgen dat kleinere klassen geen positieve effecten opleveren (Dobbelsteen et al., 1999). De auteurs komen immers tot de vaststelling dat de leerprestaties van leerlingen significant positief beïnvloed worden door het aantal medeleerlingen met een vergelijkbaar IQ. In kleinere klassen zijn er echter minder leerlingen met een gelijkaardig IQ waar men van kan leren (Dobbelsteen et al., 1999).

Tot slot is er de scheidingslijn tussen beide kampen, die weerspiegelt dat klasgrootte geen significant effect heeft. Ook hier zijn verschillende aanhangers voor deze bewering (Hancock, 1996; Rice, 1902; Siegfried & Kennedy, 1995). De bevinding dat klasgrootte niet beïnvloedend is voor de leerlingenprestaties kan te verklaren zijn doordat leraren hun docerstijl niet aanpassen (Siegfried & Kennedy, 1995). Als ze op dezelfde manier, bijvoorbeeld plenair, lesgeven aan kleine klassen dan kan er geen effect van de kleine klassen optreden. Hoxby (2000) komt ook tot deze conclusie en suggereert daarom dat leerkrachten gestimuleerd of getraind moeten worden om gepaste doceertechnieken voor kleinere klassen te gebruiken.

2.2 Het begrip 'klasgrootte'

Klasgrootte kan op verschillende manieren gedefinieerd worden, maar verschilt hierin weinig naar inhoudelijke betekenis. Het begrip klasgrootte werd door Glass en Smith (1978) gedefinieerd als de verhouding van leerlingen tot leraren, ook wel instructieve groepsgrootte genoemd. Een andere variabele die regelmatig gehanteerd wordt in onderzoek naar het effect van klasgrootte, is de pupil-teacher ratio of de student-teacher ratio. Hoewel men dit vaak als equivalent van het begrip klasgrootte aanschouwt, is er toch een belangrijke nuance die gemaakt moet worden. Waar het bij de klasgrootte gaat om de verhouding van het aantal leerlingen tot de leraar voor één klas, berekent men de pupil/student-teacher ratio veelal door het totale aantal leerlingen in een school of universiteit te delen door het aantal leerkrachten. Hierdoor krijg je gemiddelde klasgroottes per school. Het is echter verkeerdelijk te veronderstellen dat alle klassen in een school even groot zijn. In het secundair

onderwijs maakt men bijvoorbeeld geen gebruik van vaste klassen. Leerlingen worden gemengd en zitten voor bepaalde vakken samen met een kleine groep van de klas (bijvoorbeeld voor specialisatievakken), terwijl ze voor andere vakgebieden in een grote klas les krijgen met eventueel nog leerlingen uit andere studierichtingen. Met dit gegeven houdt Bosker (1998) rekening in zijn definitie van klasgrootte. Hij stelt dat klasgrootte operationeel gedefinieerd wordt als de grootte van de klas waarin de studenten meer dan de helft van hun tijd op school besteden.

We kunnen vanuit de literatuur dus concluderen dat valide en betrouwbare uitspraken over de impact van klasgrootte, enkel gebaseerd kunnen zijn op onderzoek dat de werkelijke klasgroottes per leerling gebruikt en niet één of andere vorm van geaggregeerde variabelen.

2.3 Prestaties

Wanneer we uitspraken willen doen over het effect dat klasgrootte heeft, dan moeten we dat afzetten tegen een maat. In dit soort onderzoek zouden verschillende afhankelijke variabelen kunnen worden gekozen. We zouden bijvoorbeeld kunnen onderzoeken welk effect klasgrootte heeft op het al dan niet slagen op examens, op het instructiegedrag van leerkrachten, enzoverder. In deze studie gaan we het effect na van klasgrootte op leerlingenprestaties. De vraag stelt zich dan hoe er gemeten moet worden wat de leerlingen presteren. Gaan we hun leesvaardigheid meten, de wiskundeprestaties, een niet-cognitieve outputmaat of het geheel? Naast die keuze moet ook bepaald worden op welke manier dat gedaan wordt. We kunnen hiervoor onder meer beroep doen op gestandaardiseerde of curriculumspecifieke testen. De gevonden effectgrootte kan afhankelijk zijn van deze keuzes. Zo vinden Scheerens en Bosker (1997) in schooleffectiviteitsonderzoek dat de netto-effecten steeds groter zijn voor wiskunde dan voor taal. Zoals aangehaald door Opdenakker en Van Damme (2000), Thomas (2001) en Verhoeven et al. (1992), worden er bovendien vaak grotere schoolverschillen bij cognitieve maten dan bij niet-cognitieve afhankelijke variabelen vastgesteld (De Maeyer & Rymenans, 2004). Het is bijgevolg onmogelijk om generalisaties te maken. Onderzoek op dezelfde gegevens met een andere outputmaat (bijvoorbeeld. wiskunde- in plaats van leesprestaties) zou mogelijk kunnen leiden tot andere conclusies. Waarom wordt er in dit onderzoek voor gekozen om de leesvaardigheidprestaties te meten als afhankelijke variabele?

Leesvaardigheid speelt een onmiskenbare rol in het schoolsucces van leerlingen, en blijkt zelfs mede bepalend te zijn voor het verdere succes in de maatschappij. Eén van de uitspraken van minister van onderwijs als eerste reactie op de PISA 2006-resultaten voor Vlaanderen luidde: “leesvaardigheid is van cruciaal belang voor succes op school omdat vaardigheid in de schooltaal ook bepalend is voor goede leerresultaten in alle andere vakken en inhouden” (Kabinet Vlaams minister van Onderwijs en Vorming, 2007). Uit een Nederlandse meta-analyse van Mol en Bus (2001) blijkt dat lezers hoger scoren dan niet-lezers op intelligentietesten en maten van academisch succes, zoals het toelatingsexamen voor de universiteit en het eindexamencijfer van de middelbare school. Kortom, de

meta-analyse laat zien dat lezen (minstens) tot in de jongvolwassenheid een grote rol speelt. Een leesroutine draagt eveneens bij aan het vergroten van intellectuele vaardigheden en succes in de samenleving (Mol & Bus, 2011).

Leerlingen in het eerste leerjaar lager onderwijs focussen zich op het technische lezen en het hardop lezen. Ze moeten vaardig worden in deze aspecten, vooraleer ze kunnen begrijpen wat ze lezen en er ook plezier aan beleven (Colpin et al., 2002). Het is dus moeilijk om bij deze jonge leeftijdsgroep het begrijpend lezen al te meten. Leerlingen hebben voor het vierde leerjaar ‘leren lezen’ en vanaf nu lezen ze ‘om te leren’ (Chall, 1983). Dit maakt dat lezen een extra dimensie krijgt, lezen is niet langer louter een doel op zich, maar vormt ook het middel tot kennisvergarig.

Uit dit alles mag blijken dat het belang van lezen de aanzet vormt om dit als afhankelijke variabele te gebruiken in deze studie.

2.4 Controlevariabelen

Het gegeven dat de klassen waar leerlingen in zitten verschillen van grootte, is niet het enige wat leerlingen van elkaar onderscheidt. Er zijn andere klas- en schoolafhankelijke elementen die sterk kunnen variëren. Leerlingen hebben bovendien allemaal een eigen persoonlijkheid en achtergrond die mede bepalend kunnen zijn voor hun prestaties. Er kan daarom gepleit worden om voor deze variabelen te controleren.

De meta-analyse van Scheerens en Bosker (1997) van 168 studies over de gehele wereld wees uit dat scholen gemiddeld genomen 19% van de variantie tussen leerlingen kunnen verklaren wanneer er geen rekening gehouden wordt met de instroomverschillen van leerlingen en 8% wanneer daar wel voor wordt gecorrigeerd (De Maeyer & Rymenans, 2004). Dit pleit ervoor om leerlingkenmerken als extra variabelen mee op te nemen in de analyses. Wil men uitspraken doen over de leerwinst die scholen verwezenlijken, dan moet men ze op een eerlijke manier met elkaar vergelijken door rekening te houden met de instroomverschillen. Dit eerlijkheidsprincipe wordt meestal aangeduid met de term ‘added value’ of toegevoegde waarde. Bij schooleffectiviteitsonderzoek wordt de samenhang tussen schoolkenmerken en leerlingprestaties bestudeerd. Wanneer hierbij rekening gehouden wordt met de leerlingeninstream, spreekt men van ‘nettoprestaties’. Zoniet, gaat het om ‘bruto-effecten’. Beide effecten zijn interessant om na te gaan, omdat de netto-effecten het eerlijkheidsprincipe toepassen maar we anderzijds ook een beeld van de huidige realiteit willen verkrijgen (De Maeyer & Rymenans, 2004).

De geselecteerde leerlingkenmerken die mogelijk een deel van de verschillen tussen scholen kunnen verklaren, worden als volgende variabelen meegenomen: geslacht, leeftijd, socio-economische status

(SES), cultureel kapitaal en thuistaal van de leerling. Evidentie om te controleren voor deze kenmerken vinden we eveneens in talrijke andere onderzoeken (e.g. Keil & Partell, 1997; Bosker, 1998; Dobbelsteen, Levin, & Oosterbeek, 1999; Van Damme et al., 2010). Het belang van cultureel kapitaal wordt onder meer onderschreven in een studie van Purves en Elley (1994). Deze wees uit dat directe toegang tot verscheidene soorten geschreven teksten een positieve invloed heeft op de leesvaardigheid. Wanneer bij leerlingen thuis veel leesmateriaal aanwezig is, wat er vaak op wijst dat ouders belang hechten aan geletterdheid, zullen ze sneller geprikkeld worden om te lezen. Een studie van Linnakyla, Malin en Taube (2004) bekeek persoonlijke, socio-economische en culturele determinanten van lage leesvaardigheidsprestaties. Dit werd onderzocht bij Finse en Zweedse studenten op basis van de PISA-data verzameld in 2000. De onderzoekers kwamen zo tot het resultaat dat studenten met een lagere sociaal-economische achtergrond, gemeten aan de hand van de beroepsstatus van de ouders, geconfronteerd worden met een hoger risico op een lage prestatie. De culturele factoren van de thuisomgeving bleken eveneens significant te zijn in het voorspellen van lage prestaties (Linnakyla et al., 2004).

Naast leerlingkenmerken is het niet ondenkbaar dat ook invloeden op klasniveau hun aandeel hebben in de verschillen tussen de resultaten van leerlingen. Als voornaamste voorspeller nemen we uiteraard de klasgrootte mee. Daarnaast controleren we ook op mogelijke effecten die door de leerkracht veroorzaakt kunnen worden, met name het geslacht en de ervaring van de leerkracht. Onderzoek van Christiaens (2009) over de invloed van de leerkracht op leesscores, toont immers aan dat vrouwelijke leerkrachten een significant negatieve invloed hebben op de leesresultaten van leerlingen. Samen met de gemiddelde leesscores per klas, de gemiddelde socio-economische status, en het aandeel anderstaligen in de klas, vormen dit de variabelen op klasniveau.

De schoolkenmerken die in dit onderzoek ter controle meegenomen worden, zijn enerzijds de schoolgrootte en anderzijds de ligging van de school, de zogenaamde urbanisatiegraad.

2.5 Differentiële effecten

Er werden zonet verschillende kenmerken besproken die een effect kunnen hebben op leerlingprestaties. In onderzoek wordt er vaak vanuit gegaan dat gevonden effecten van een bepaald kenmerk (zoals klasgrootte) in gelijke mate gelden voor de hele populatie. Men kan zich echter de vraag stellen of men er van uit kan gaan dat een vastgesteld verband tussen een schoolkenmerk en de leerlingprestaties wel effectief gelijkaardig is voor alle leerlingen. Dit betreft de kern van differentiële effectiviteit (De Maeyer & Rymenans, 2004). Het wordt algemeen aanvaard dat kleine klassen meer voordeel bieden voor kansarme studenten en leerlingen uit lagere graden, zoals aangehaald door Krueger en Whitmore (2001), Finn en Achilles (1999) en Molnar (1999) (Van Damme et al., 2010). Finn en Achilles (1990) voerden grootschalig experimenteel onderzoek waarbij kleuters en

leerkrachten willekeurig werden toegewezen aan grote en kleine klassen in elke participerende school. De leerlingen bleven in deze klassen gedurende twee jaar. Aan het einde van elke graad werden de lees- en wiskundevaardigheden gemeten door middel van gestandaardiseerde en curriculum-gebaseerde testen. Dit onderzoek resulteerde in twee conclusies: een significant voordeel komt ten goede aan studenten in verkleinde klassen voor beide vakgebieden, en er is evidentie dat sociaal achtergestelde studenten in het bijzonder voordeel ervaren van het kleinere klassenklimaat, vooral wanneer toetsen gebaseerd op het curriculum gebruikt worden als leercriterium (Finn & Achilles, 1990). Toch meet oud onderzoek vaak louter de rechtstreekse effecten, zonder oog te hebben voor differentieeffecten. De vraag naar differentiële effectiviteit wordt in dit onderzoek opgenomen en zal hieronder aan bod komen bij de onderzoeksvragen.

2.6 Curvilineariteit

Onderzoek van Bosker (1998) toonde aan dat klasgrootte negatief samenhangt met de leesprestaties wanneer klassen meer dan 20 leerlingen bevatten. Bovendien is er een extra (negatieve) coëfficiënt wanneer klassen groter worden dan 30 leerlingen. Deze bevinding stelt dat er geen lineair verband is van klasgrootte met prestaties, maar wel een niet-lineair, curvilineair effect. Het is bijvoorbeeld denkbaar dat er pas een uitgesproken effect is wanneer de klasgrootte erg veel verschilt. Het kan zijn dat vanaf een bepaalde stijging in klasgroottes, dit niet langer een significant (negatief?) effect heeft op de prestaties van leerlingen. Of omgekeerd, dat vanaf een bepaalde klasgrootte, het afnemen met een klein aantal leerlingen, er een groter effect is van deze daling. Concreet: het effect van tien leerlingen extra zou in een klas met 15 leerlingen een veel grotere negatieve impact kunnen hebben dan het toevoegen van tien studenten aan een grote klas waar 200 leerlingen samen les krijgen (Keil & Partell, 1997).

Fox (1997) stelt dat klassieke statistische modellen sterke assumpties maken over de structuur van data, terwijl deze assumpties vaak niet vasthouden aan de praktijk. Als alternatief kan men data transformeren opdat zij nauwer in overeenstemming is met de beperkende aannames van klassieke statistische modellen. Een bruikbare groep van transformaties is de 'familie' van machten en wortels (Fox, 1997). Om niet-lineariteit in een relatie tussen x en y te testen, kan een term x^2 aan het model toegevoegd worden. Een significante parameter in de x^2 -term duidt op curvilineariteit in de relatie tussen x en y (Crawley, 2007). Om na te gaan of er in deze studie sprake is van zo'n curvilineair verband, zal gebruik gemaakt worden van de gekwadrateerde variabele van klasgrootte.

3 Onderzoeksvragen

Tabel 1 brengt de informatie uit de probleemstelling en het literatuurstuk schematisch bijeen. Er werden meerdere mogelijke verklaringen aangehaald voor al dan niet gevonden effecten van klasgrootte op leerlingenprestaties. Er zijn een aantal variabelen die hier de oorzaak voor kunnen zijn. Deze zijn opgesomd en worden nadien gekoppeld aan hoe ze in deze studie uitwerking krijgen. Hierbij werd geprobeerd aan zoveel mogelijk kritiek tegemoet te komen. Bepaalde keuzes zijn bepaald door de dataset waarop voortgebouwd wordt. In hoofdstuk 4 worden deze keuzes bij de methodologie beargumenteerd.

Tabel 1

Mogelijke Verklaringen en Oorzaken voor al dan niet gevonden Effecten van Klasgrootte op de Leerlingenprestaties

Effect	Wel	Geen
Verklaringen	- Leraren hebben meer tijd voor de individuele leerlingen - Er zijn minder verstoringen tijdens de les	- Leraren passen doceerstijl niet aan - Leerlingen kunnen niet leren van medestudenten (met een gelijkaardig IQ)
Oorzaken	- Er wordt niet met de eigenlijke klasgrootte gewerkt - Er wordt geen rekening gehouden met endogeniteit - Er wordt geen rekening gehouden met voorgaande klasgroottes - Effect is afhankelijk van de onderzoeksmethode - Effect is afhankelijk van de populatie (jong-oud) - Effect is afhankelijk van de maat van prestaties - Effect is afhankelijk van de controlevariabelen - Effect is afhankelijk van de plaats	

Het mag duidelijk zijn dat bestaand onderzoek geproblematiseerd werd door een aantal aspecten: er treden verschillen op in operationalisatie van de variabele ‘klasgrootte’, door het al dan niet rekening houden met het gegeven van endogeniteit en het probleem van aggregatie, door verscheidene onderzoeksmethoden, populaties, prestatiematen, inbreng van controlevariabelen, plaatsafhankelijkheid en het al dan niet onderzoeken van differentiële effectiviteit en (non-)lineariteit. In dit onderzoek krijgen deze aspecten als volgt invulling: vooreerst wordt er gewerkt met de eigenlijke klasgrootte. Het effect hiervan op prestaties wordt berekend door middel van multilevelanalyses om de aanwezige hiërarchie in de data te respecteren. De gegevens werden verzameld in een realistische, niet-experimentele, setting. De populatie omvat leerlingen uit het vierde leerjaar basisonderwijs in Vlaamse scholen. Ter controle worden variabelen op zowel leerling-, klas- als schoolniveau meegenomen.

De probleemstelling en literatuur kaderden de focus van dit onderzoek. Er is reeds gewezen op het belang van lezen, de inbreng van controlevariabelen, en het fenomeen van curvilineariteit en differentiële effecten. Voorgaande beschrijving gaf aan hoe in deze studie onderzoek gedaan zal worden op een betrouwbare, valide manier door zoveel mogelijk te anticiperen op de gebreken van voorgaand onderzoek. Concreet probeert deze studie een antwoord te geven op volgende onderzoeksvragen:

1. **Is er een effect van klasgrootte op individuele leesprestaties van leerlingen?**
2. **Is er sprake van een effect als er individuele leerlingkenmerken meegenomen worden?**
3. **Is er sprake van een effect als gecontroleerd wordt voor klaskenmerken?**
4. **Is er een effect van klasgrootte na controle voor schoolkenmerken?**
5. **Verschilt het effect van klasgrootte op leesprestaties tussen klassen en scholen wanneer er gecontroleerd wordt voor de leerling-, klas-, en schoolkenmerken?**
6. **Bestaat er in plaats van een lineair, een curvilineair verband tussen klasgrootte en leerlingenprestaties?**

Onderzoeksvraag 1 meet het bruto-effect van klasgrootte op leesprestaties. Bij het controleren voor relevante kenmerken, in onderzoeksvragen 2 tot 5, wordt zicht verworven op de netto-effecten. Tot slot gaat onderzoeksvraag 6 de mogelijkheid tot het bestaan van een curvilineair verband na. Bij deze onderzoeksvragen horen enkele verwachtingen over de tendensen die kunnen bestaan. Deze hypothesen zijn gestoeld op vaststellingen uit eerder onderzoek.

Naar de bevindingen van verschillende auteurs (Glass & Smith, 1978; Krueger, 1999; Mosteller, 1995; Nye et al., 2000) wordt een hypothese opgezet over het verband tussen klasgrootte en prestaties:

Hypothese 1: leerlingen in kleine klassen behalen hogere prestaties dan leerlingen in een klas met meer leerlingen.

Indien klasgrootte een significant effect heeft, kunnen we de vraag stellen of dit eenzelfde effect is voor alle leerlingen. Of heeft klasgrootte een differentieel effect voor een bepaald type leerlingen? Dit is de vraag naar differentiële effectiviteit. Is er voor alle leerlingen eenzelfde effect en is het effect in elke school ook even sterk? Heeft klasgrootte bijvoorbeeld eenzelfde effect voor leerlingen met een hoge of lage SES? Het is met andere woorden zinvol om interactie-effecten tussen de onafhankelijke variabelen na te gaan. Gezien het belang van socio-economisch en cultureel kapitaal, zoals eerder aangehaald door Purves en Elley (1994) en Linnakyla, Malin en Taube (2004), en de vaststelling van Finn en Achilles (1990) en Molnar et al. (1999) dat sociaal achtergestelde studenten extra voordeel ervaren van kleinere klassen, wordt volgende hypothese opgesteld:

Hypothese 2: voor leerlingen met een lagere socio-economische status en cultureel kapitaal, is het effect van de klasgrootte op de prestaties groter.

Deze onderzoeksvragen en hypothesen zullen onderzocht worden bij leerlingen in het vierde leerjaar basisonderwijs die deelnamen aan de PIRLS 2006-studie. In het volgende hoofdstuk wordt uitgebreid ingegaan op de methodologie van het onderzoek.

4 Methodologie

In wat volgt wordt de aanpak uitgelegd om na te gaan of er evidentie wordt gevonden voor bovenstaande hypothesen. Allereerst wordt de aandacht gevestigd op de gegevens die gebruikt worden om de onderzoeksvragen te onderzoeken. Dit resulteert in een concrete vertaling naar de selectie en opmaak van de variabelen waarmee gewerkt wordt. Vervolgens wordt ingegaan op de onderzoeksmethode die gehanteerd wordt om een antwoord te kunnen vinden op de vooropgestelde onderzoeksvragen, en de stappen die gezet worden in de uitvoering van de analyses.

4.1 Databestand PIRLS 2006

Voor de analyses maken we gebruik van Vlaamse data die beschikbaar is via het PIRLS-onderzoek uit 2006 dat werd uitgevoerd door de IEA, International Association for the Evaluation of Educational Achievement. ‘Progress in International Reading Literacy Study’ -kortweg PIRLS- verzamelt internationaal gegevens over de leesvaardigheid van leerlingen in het vierde leerjaar lager onderwijs om trends te kunnen meten en informatie te verzamelen over het beleid en de gebruiken rond leren lezen en leesinstructie (International Association for the Evaluation of Educational Achievement, 2012).

De keuze om te werken met deze datagegevens is zowel vanuit inhoudelijke als praktische redenen te staven. Zowel voor het feit dat we lezen gebruiken als outputmaat, als voor de keuze tot leerlingen uit het vierde leerjaar als onderzoeksgroep, is vanuit theoretisch standpunt evidentie gebleken. Vanuit die theoretische achtergrond is het gebruik van de PIRLS-data met andere woorden een logische voortvloeiing. Gezien lezen als zeer belangrijk vakgebied op school wordt aanschouwd, wordt de data gebruikt die deze vaardigheden meet (er bestaat bijvoorbeeld ook gelijkaardige data over wiskunde- en wetenschappenprestaties via het internationaal vergelijkend TIMSS-onderzoek). De jongst beschikbare steekproefgroep om dit te meten, is die van de leerlingen uit het vierde leerjaar. Gezien de beperkte tijdsspanne voor het uitvoeren van dit onderzoek, was het onmogelijk om zelf een nieuwe dataverzameling op te zetten. De theoretische en praktische overwegingen geven een goede verantwoording voor de keuze tot het gebruik van de PIRLS-data.

Aan de tweede editie van het PIRLS-onderzoek, uitgevoerd in 2006, namen 40 landen deel. De 4479 leerlingen uit 137 scholen vormden de Vlaamse steekproef. In Vlaanderen werden relatief veel leerlingen tot deelname uitgesloten, met name 6,1% op schoolniveau. Zeer kleine schooltjes en scholen in het Buitengewoon Lager Onderwijs (BLO-scholen), werden namelijk geweerd uit de steekproef (Faculteit Psychologie en Pedagogische Wetenschappen KU Leuven, n.d.). De PIRLS-studies volgen elkaar op in een vijfjarige cyclus, wat maakt dat in 2011 nieuwe data verzameld werd.

Vlaanderen, in tegenstelling tot Wallonië, neemt niet deel aan deze derde editie van de studie. Daarom vallen we terug op de Vlaamse data van 2006.

4.2 Variabelen

In het theoretisch kader is aangegeven hoe klasgrootte op een goede manier gemeten moet worden. Het gebruik van leesscores als prestatie maat is eveneens verantwoord. Daarnaast is bestudeerd welke variabelen relevant zijn om op te nemen als controlevariabelen. Op basis hiervan komen we tot de volgende selectie en operationalisatie van de afhankelijke en onafhankelijke (controle-)variabelen.

4.2.1 Afhankelijke variabele: leesprestatie

PIRLS 2006 wil informatie verzamelen over de leesvaardigheid van leerlingen en factoren die hier mogelijk een invloed op uitoefenen. Daarom werden vragenlijsten afgenomen van de leerlingen zelf, de leerkrachten, de ouders en de directies. Aanvullend is er per land ook een vragenlijst afgenomen over het leescurriculum.

Met dit internationaal vergelijkend onderzoek wil men bij de leerlingen nagaan in hoeverre de doelen van lezen bereikt worden en de processen van het begrijpen plaatsvinden. De twee leesdoelen zijn het ‘lezen voor de literaire ervaring’ en het ‘lezen om informatie te bekomen en gebruiken’. Voor het begrijpen, wordt naar vier processen gekeken: ‘expliciete informatie verkrijgen’ en ‘eenvoudige gevolgtrekkingen maken’, die samen het onderdeel tekstbegrip vormen, en ‘ideeën en informatie interpreteren en integreren’ en het ‘onderzoeken en evalueren van inhoud, taal en tekstuele elementen’ die overkoepelend benoemd worden als ‘tekstinterpretatie’.

Om dit alles te bevragen werden tien tekstfragmenten met bijhorende vragen opgesteld waarvan vijf literaire en vijf informatieve. Er waren dertien verschillende tekstboekjes met telkens één informatieve en één literaire tekst waarbij de leerlingen achteraf vragen moesten beantwoorden. Om de jonge leerlingen niet te overbelasten werd een maximum werktijd opgelegd van tachtig minuten, veertig minuten per tekst.

De totale beoordelingsschaal overheen de tien teksten bevat 126 items, waarbij een maximumscore van 167 punten behaald kan worden. Die punten zijn verspreid over de vier processen van het begrijpen. Aangezien de leerlingen slechts twee teksten en vraagreeksen krijgen, beantwoordt elke leerling maar een beperkt aantal items van het totale pakket aan beoordelingsitems. PIRLS gebruikt daarom Item Response Theory (IRT)-modellen om voor alle leerlingen geschatte scores te berekenen die ze zouden behalen indien ze wel alle items zouden invullen. Op deze manier krijgen leerlingen algemene leesscores, op de twee leesdoelen en op de twee onderdelen van de leesprocessen namelijk tekstbegrip en tekstinterpretatie. Om onderzoekers de kans te geven het effect van de geschatte scores

op de analyses te beoordelen, werden door PIRLS telkens vijf schattingen gemaakt voor elke score. Op die manier kan er steeds een andere schatting gebruikt worden om analyses te vergelijken, of de vijf schattingen kunnen ook gecombineerd worden tot één somscore die de algemene leesvaardigheid weergeeft (Foy & Kennedy, 2008).

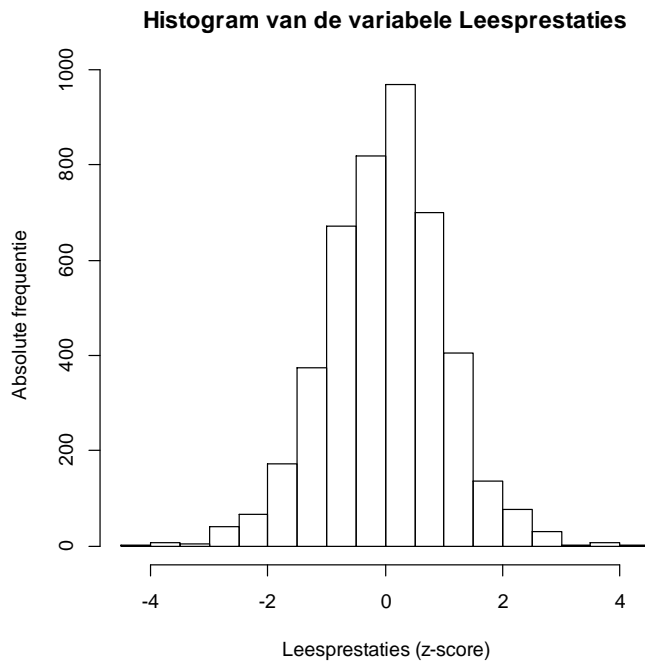
Voor deze studie is geopteerd om als afhankelijke variabele deze totaalscore op leesvaardigheid te hanteren. Zoals eerder gezegd, werd deze ontwikkeld op basis van vijf geschatte scores, ‘plausible values’, die onderworpen werden aan IRT. Een betrouwbaarheidsanalyse gaf een Cronbach’s alpha van 0.96 waardoor besloten kan worden dat deze geschatte scores een zeer betrouwbare schaal vormen. Zo komen we tot één somscore die gestandaardiseerd werd.

Er werd tevens getest of de correlaties tussen de totaalscore en de deelvaardigheden -‘reading for literary purposes’ (LIT), ‘reading for informational purposes’ (INF), ‘retrieving and straightforward inferencing processes’ (RSI) en ‘interpreting, integrating and evaluating processes’ (IIE)- wel hoog genoeg zijn om met deze totale score te werken. Indien de correlaties niet voldoende hoog zijn, zou dit ervoor pleiten om de deelvaardigheden apart te bekijken. Zoals af te lezen is in onderstaande tabel, zijn de correlaties allemaal sterk positief. Voor deze interpretatie werd de norm van effectgroottes volgens Cohen (1998) gebruikt. Deze stelt dat correlaties sterk positief zijn vanaf een waarde 0.50. Deze referentiewaarde wordt hier telkens ruim overschreden, waardoor besloten kan worden dat de deelvaardigheden hoog genoeg correleren om ze tot één totaalscore te maken. Daarom maken we in de analyses gebruik van de totaalvariabele voor de leesprestaties.

Tabel 2
Correlaties tussen de Totale Leesscore en de Deelvaardigheden (n=4479)

	Totale score	LIT	INF	RSI	IIE
Totale score	1.0000000	0.9322535	0.9226416	0.9379194	0.9435803
LIT	0.9322535	1.0000000	0.8865681	0.9128038	0.9146437
INF	0.9226416	0.8865681	1.0000000	0.9006243	0.9111868
RSI	0.9379194	0.9128038	0.9006243	1.0000000	0.9366778
IIE	0.9435803	0.9146437	0.9111868	0.9366778	1.0000000

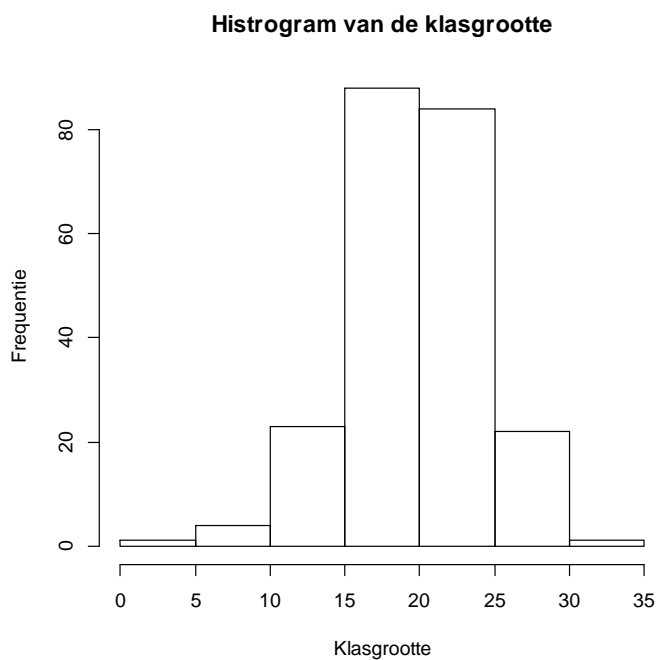
De negatieve kurtosiswaarde van -0.1269862 die berekend werd, duidt erop dat deze gestandaardiseerde variabele lichtjes linksscheef (platykurtic) verdeeld is. Het histogram (Figuur 2) toont dat de variabele echter bij benadering de normaalverdeling volgt. De variabele wordt dus als normaal verdeeld meegenomen in de analyses.



Figuur 2: Histogram die de verdeling van de variabele 'Leesprestaties' toont

4.2.2 Onafhankelijke variabele: klasgrootte

Deze onafhankelijke variabele werd tot een z-score gemaakt zodat we per leerling kunnen bepalen of ze in een grotere dan wel kleinere klas zitten dan gemiddeld. De gemiddelde klas in de steekproef telt 21 leerlingen. Er is in de dataset een minimumgrootte van vier leerlingen per klas, de grootste klas telt 32 leerlingen. In de steekproef is er dus wel degelijk enige spreiding in de variabele klasgrootte. Indien er een erg kleine variantie in de klasgrootte zou zijn, is het veel moeilijker om een effect te vinden.



Figuur 3: Histogram die de spreiding van de klasgroottes weergeeft

4.2.3 Onafhankelijke controlevariabelen

Er zijn zowel controlevariabelen opgenomen op leerling-, klas- als schoolniveau. Volgende tabel geeft deze variabelen weer, met bijhorend niveau en operationalisatie.

Tabel 3
Overzicht van de Controlevariabelen

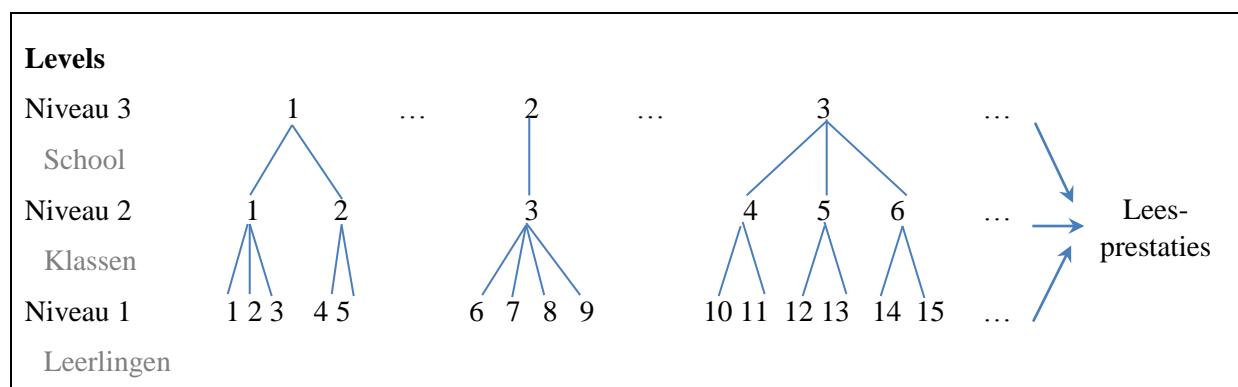
Variabele	Niveau	Operationalisatie
Geslacht	Leerling	Er wordt een dummyvariabele aangemaakt die aanstaat voor meisjes.
Leeftijd	Leerling	Deze variabele wordt gestandaardiseerd tot een z-score.
Socio-economische status	Leerling	Er werd één score berekend voor de socio-economische status op basis van een categorische principale componentenanalyse (catpca). In de vragenlijst aan de ouders werd onder andere gepeild naar het hoogste diploma van de vader en de moeder en de job die de vader en moeder van de leerling uitoefenen. Afhankelijk van de antwoorden op deze vragen krijgt de leerling een hogere of lagere score op SES. Deze individuele score op socio-economische status wordt gestandaardiseerd.
Cultureel kapitaal	Leerling	Een score voor het cultureel kapitaal wordt gebaseerd op het aantal boeken dat ouders aangeeft thuis te hebben. Deze variabele is in de dataset aanwezig met vijf categorieën: 0-10 boeken, 11-25, 26-100, 101-200 of meer dan 200 boeken. Dit wordt vertaald naar vier dummyvariabelen met als referentiecategorie 0 tot 10 boeken.
Thuis taal	Leerling	PIRLS stelde de ouders ook de vraag of hun kind al dan niet Nederlands sprak voordat ze naar school gingen. Hiermee willen we controleren voor de mogelijke achterstand in leesprestaties doordat deze leerlingen bij de start in het lager onderwijs thuis geen Nederlands spraken. Er wordt een dummyvariabele gemaakt die aanstaat voor Nederlands.
Geslacht van de leerkracht	Klas	Er wordt een dummyvariabele aangemaakt die aanstaat voor vrouwen.
Ervaring van de leerkracht	Klas	Een gestandaardiseerde score geeft aan of de leerkracht meer of minder ervaring heeft dan het gemiddeld aantal jaren ervaring van de deelnemende Vlaamse leerkrachten.

Gemiddeld leesniveau	Klas	Een gemiddelde score van de leesprestaties (score op de PIRLS-toets) per klas wordt berekend en gestandaardiseerd.
Gemiddelde socio-economische status	Klas	Een gestandaardiseerde score geeft aan hoe SES per klas is.
Gemiddelde thuistaal	Klas	Het percentage anderstaligen per klas.
Schoolgrootte	School	Deze variabele wordt tot een z-score gemaakt zodat we per leerling kunnen bepalen of ze in een school zitten met meer dan wel minder leerlingen dan gemiddeld.
Urbanisatiegraad	School	Deze categorische variabele beschrijft het gebied waarin de school gelegen is als stedelijk (urban), voorstedelijk (suburban) of landelijk (rural). Er worden twee dummyvariabelen van gemaakt, de referentiecategorie is landelijk.

De dataset van PIRLS 2006 bevat alle nodige variabelen, en dit voor de 4479 Vlaamse leerlingen uit 237 klassen in 137 verschillende scholen. Omdat er bij sommige leerlingen bepaalde ontbrekende waarden zijn, moesten deze verwijderd worden. Dit betekent dat we voor de analyses de gegevens van 3415 leerlingen overhielden.

4.3 Analysemethode

Het gegeven dat kinderen in klassen zitten en deze op hun beurt gegroepeerd zijn over scholen, maakt dat we met informatie te maken hebben op verschillende niveaus. Leerlingenkenmerken kunnen beïnvloed worden door deze hiërarchie. Denken we aan de onderzoeksvragen, dan gaan we hier ook expliciet van uit; we onderzoeken of de klas, meer bepaald de grootte, een invloed heeft op de leesprestaties van de leerlingen. We hebben te maken met leerlingen (niveau 1), die genest zijn in klassen (niveau 2), en die klassen zijn op hun beurt geclusterd binnen scholen (niveau 3), zoals gevisualiseerd in Figuur 4.



Figuur 4: Hiërarchische structuur

Dit gegeven leidt tot de keuze om multilevelanalyses uit te voeren.

We gaan na of de leesprestaties van individuele leerlingen beïnvloed worden door de klasgrootte indien we rekening houden met kenmerken op school-, klas-, en leerlingniveau. Onderstaande formule geeft de vergelijking van het model weer:

$$Y_{ijk} = \beta_0 + (\beta_1 * x_{1ijk} + \mu_{0j} + \varepsilon_{0ij}) + (\beta_2 * x_{2ijk} + \mu_{0j} + \varepsilon_{0ij}) + \dots + (\beta_z * x_{zijk} + \mu_{0j} + \varepsilon_{0ij})$$

Hierbij staat Y_{ijk} voor de score van leerling i in klas j en in school k op de afhankelijke variabele. β_0 staat voor het intercept, het algemeen gemiddelde, over alle leerlingen, klassen en scholen heen. De variabelen x_{ijk} geven de scores op een verklarend kenmerk van leerling i in klas j en school k . De variabelen die opgenomen zijn in de analyses werden in 4.2 reeds verduidelijkt.

Hoewel er zonet drie hiërarchische niveaus werden besproken, is er in de data bij bepaalde leerlingen maar sprake van twee niveaus. Omdat er scholen zijn die slechts één klas hebben in het vierde leerjaar, kunnen we voor die scholen geen verschillen toeschrijven aan het klasniveau, maar enkel aan de schoolkenmerken. Daarom negeren we het schoolniveau in eerste instantie. In een latere fase (onder 5.6) zullen we dit derde niveau wel mee opnemen om eventuele verschillen in kaart te kunnen brengen.

We willen nagaan of de verschillende niveaus er eigenlijk wel toe doen. Daartoe schatten we een **nulmodel** waarin geen verklarende variabelen worden opgenomen, om te kijken of er significante variantie is op klas- en schoolniveau.

Het fixed gedeelte en random gedeelte zal informatie geven over de parameterschattingen. In het nulmodel wordt in het fixed gedeelte enkel een intercept (β_0) geschat. Dit geeft een schatting van het populatiegemiddelde en de standaardfout. Het random gedeelte beschrijft drie residuele scores: de afwijking van de score van school k ten opzichte van het algemene gemiddelde (V_{0k}), de afwijking van de score van klas j ten opzichte van het gemiddelde van zijn school (U_{0jk}) en de afwijking van de score van leerling i ten opzichte van het gemiddelde van zijn klas (E_{0ijk}), telkens met de standaardfout erbij. Op basis van de schattingen van de varianties kunnen we voor elk niveau de intra-class correlatie (ICC of Rho) berekenen. Zo kunnen we in een percentage uitdrukken hoeveel van de variantie in de leesprestaties toe te schrijven is aan de klassen of scholen waarin leerlingen geclusterd zijn.

De eerste onderzoeksvraag gaat het zuiver verband tussen de onafhankelijke en afhankelijke variabele na. De volgende onderzoeksvragen betrekken leerling-, klas- en schoolkenmerken om te toetsen of controleren hiervoor al dan niet een effect heeft. Op basis van literatuur weten we dat leerlingkenmerken hun invloed uitoefenen op leerresultaten. In deze tweede fase van de analyse voegen we daarom naast klasgrootte als voorspeller ook **leerlingkenmerken** toe. Op basis van de significantietoetsen zullen we uitspraken doen over welke variabelen effectief een invloed hebben op de leesprestaties. We veronderstellen hierbij dat de verklarende variabelen een gelijk effect hebben op de

prestaties, ongeacht de klas of school waarin ze zitten. De intra-class correlaties zullen uitwijzen of er nog onverklaarde variantie blijft bestaan op het klas- en schoolniveau. Indien dit het geval is kunnen we de hypothese formuleren dat er nog andere factoren dan de individuele verschillen tussen leerlingen een invloed uitoefenen. We voegen daarom in een derde en vierde fase ook nog de **klas- en schoolkenmerken** toe. Bij de analyses is de hoofdvraag steeds of klasgrootte een significant effect heeft. Het uiteindelijke model zal naast klasgrootte ook alle andere controlevariabelen bevatten en hierbij nagaan hoe het effect van scholen en klassen verschilt. Er kan een lineair effect vastgesteld worden, maar het is eveneens mogelijk dat er een curvilineair verband bestaat, zoals gesuggereerd door Bosker (1998) en Keil en Partell (1997). Deze mogelijkheid wordt eveneens getest. Hiertoe werd de variabele klasgrootte gekwadraterd, op basis van wat we leerden over curvilineariteit in 2.6.

5 Onderzoeksresultaten

In wat volgt worden de resultaten weergegeven en beschreven. De analyses werden uitgevoerd volgens de beschrijving die zonet in 4.3 werd gegeven over de onderzoeksstappen. Er wordt dus gerapporteerd op basis van de verschillende opbouwende modellen, waarbij het effect van klasgrootte op de leesprestaties steeds wordt aangevuld met controlevariabelen op de verschillende niveaus. De relevantste waarden, namelijk de parameterschatting (afgekort als Est.), de p-waarde en de onder- en bovengrens van het betrouwbaarheidsinterval, zijn opgenomen in tabellen. In 5.8 vatten we deze conclusies samen in een overzichtstabel.

Uit het nulmodel dat het klasniveau in rekening brengt, blijkt dat de variantie tussen klassen significant is, dat er met andere woorden genoeg zekerheid is om te stellen dat er een effect is van klassen in het verklaren van de invloed van klasgrootte op de leesprestaties. Dit gezien de waarde nul niet in het betrouwbaarheidsinterval ligt. Er is hierbij geen rekening gehouden met het schoolniveau omdat, zoals eerder aangehaald, er slechts weinig klassen per school zijn waardoor de variantie van klassen binnen scholen hier onbesproken blijft. (In een latere fase, bij 5.6, komen we hier nog op terug.) De ICC-waarde duidt erop dat 4.41% van de variantie in de leesscores kan toegeschreven worden aan de klassen waarin leerlingen zitten.

Tabel 4
Nulmodel op Klasniveau (n=3415)

	Est.	p-waarde	95% B.I.	
			onder	boven
FIXED				
Intercept	0.0583	0.0077*	0.019	0.1004
RANDOM				
Variantie tussen klassen	0.0428			
Variantie binnen klassen	0.9271		0.9441	0.9912
ICC	4.41%			

* p < 0.05: significant

5.1 Is er een effect van klasgrootte op individuele leesprestaties van leerlingen?

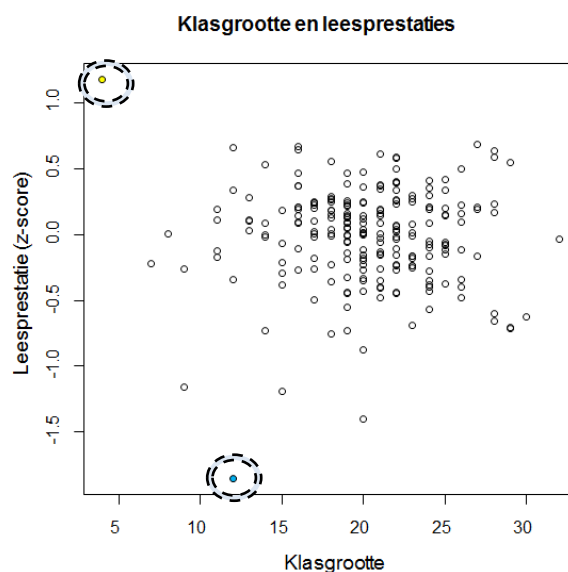
Het model met klasgrootte als enige verklarende variabele (Model 1), zegt ons dat de **klasgrootte geen significante voorspeller** is. De p-waarde van de parameterschatting is immers 0.995, beduidend hoger dan de kritische waarde van 0.05. De ICC bij dit model bedraagt 0.044, anders gezegd kan 4.4% van de variantie toegeschreven worden aan de hogere niveaus, na controle voor de klasgrootte.

Tabel 5
Model met Klasgrootte als Voorspeller (n=3415)

Variabelen	Est.	p-waarde	95% B.I.	
			onder	boven
FIXED				
Intercept	0.0583	0.0079*	0.0190	0.1008
Klasgrootte	0.0001	0.9952	-0.0413	0.0400

* p <0.05: significant

De eerste hypothese die we vooropstelden, namelijk dat klasgrootte een effect heeft op de leerlingenprestaties in de vorm van een omgekeerd verband, wordt dus door de data verworpen. Als we kijken naar het verband tussen de klasgroottes en de bijhorende gemiddelde leesprestaties voor de klassen, dan zien we een verspreide puntenwolk (zie Figuur 5). Er zijn hier enkele klassen te bemerken die ‘uitschieters’ zijn ten opzichte van de algemene tendens, die een vreemde score behalen gezien de klasgrootte. Nemen we bijvoorbeeld het omcirkelde bolletje bovenaan, staande voor een klas met slechts vier leerlingen, die toch meer dan één standaardafwijking hoger scoort op de leestoets dan gemiddeld. We zouden een verklaring kunnen opperen door het feit dat dit misschien een speciale klas is met hoogbegaafde leerlingen. Het omcirkelde bolletje onderaan is een klas met ook relatief weinig leerlingen, waar dan weer erg ondermaats gescoord werd op de PIRLS-toets. Ook hier kunnen we bijvoorbeeld bedenken dat het gaat om een klas waar alle zwakkere leerlingen met opzet in een kleine klas samenzitten (bijvoorbeeld een klas voor Anderstalige Nieuwkomers), zodat ze extra aandacht en begeleiding kunnen krijgen. Deze hypothetische verklaringen wijzen erop dat er mogelijk selectiviteit is bij het toewijzen van leerlingen aan klassen. Dit fenomeen van endogeniteit kan misschien deels verklaren waarom er geen effect werd gevonden van klasgrootte op de leerlingenprestaties. In dit onderzoek kon hier echter niet expliciet rekening mee worden gehouden, gezien de dataset geen gegevens bevat over de manier waarop leerlingen werden toegewezen aan de klassen.



Figuur 5: Puntenwolk van klassen met bijhorende gemiddelde leesprestaties

5.2 Is er sprake van een effect als individuele leerlingkenmerken meegenomen worden?

De leerlingkenmerken die we toevoegden om deze onderzoeksvraag te beantwoorden, zijn geslacht en leeftijd van de leerling, socio-economische status, cultureel kapitaal en thuistaal (Model 2). We zien dat **klasgrootte** na controle voor deze leerlingkenmerken, met een p-waarde van 0.375, **een zwak negatief maar niet-significant effect** heeft op de leesscores. Variabelen die hier wel een significante invloed op uitoefenen zijn het geslacht, de leeftijd, de socio-economische status, het cultureel kapitaal bij kinderen met meer dan 26 boeken thuis en de thuistaal. We zien dat de gemiddelde leesscore met 0.117 punten stijgt voor meisjes, zij behalen dus hogere leesscores dan jongens. Leerlingen die één standaardafwijking hoger scoren op leeftijd, oudere leerlingen dus, blijken 0.087 punten lager te scoren dan de waarde van het intercept. Dit kan mogelijk verklaard worden door het feit dat oudere leerlingen vaak zittenblijvers zijn die meer kans hebben op moeilijkheden met betrekking tot leesvaardigheid. Leerlingen met een socio-economische status die één standaardafwijking hoger is dan gemiddeld, scoren 0.266 punten beter dan gemiddelde kinderen. Leerlingen die thuis in het bezit zijn van meer dan 26 boeken, hebben statistische positievere resultaten. Ten slotte scoren leerlingen die thuis Nederlands spraken voor ze naar school gingen, 0.442 punten hoger dan anderstalige leerlingen.

Tabel 6

Model met Klasgrootte en Leerlingkenmerken als Controlevariabelen (n=3415)

Variabelen	Est.	p-waarde	95% B.I.	
			onder	boven
FIXED				
Intercept	-0.5486	0.0000 *	-0.7242	-0.3905
Klasgrootte	-0.0168	0.3745	-0.0413	0.0177
Geslacht meisje	0.1169	0.0002 *	0.0557	0.1773
Leeftijd	-0.0873	0.0000 *	-0.1220	-0.0537
SES	0.2663	0.0000 *	0.2328	0.3020
11-25 boeken	0.0581	0.2912	-0.0441	0.1684
26-100 boeken	0.0994	0.0468 *	-0.0037	0.1954
101-200 boeken	0.2442	0.0001 *	0.1259	0.3603
Meer dan 200 boeken	0.2562	0.0000 *	0.1396	0.3742
Thuistaal Nederlands	0.4420	0.0000 *	0.2992	0.5967

* p <0.05: significant

5.3 Is er sprake van een effect als gecontroleerd wordt voor klaskenmerken?

De klaskenmerken die we toevoegden om deze onderzoeksvraag te beantwoorden, zijn geslacht en ervaring van de leerkracht, het gemiddelde leesniveau van de klas, de gemiddelde socio-economische status en het percentage Nederlandstaligen in de klas (Model 3). Ook hier duidt de p-waarde van 0.389 erop dat er **geen effect is van klasgrootte**. Het geslacht en de ervaring van de leerkracht blijken ook geen voorspellers te zijn van de leesprestaties, evenals de gemiddelde socio-economische status en het

percentage Nederlandstaligen-Anderstaligen in de klas. De gemiddelde leesscore per klas blijkt wel een significant positief effect te hebben. Klassen die één standaardafwijking hoger scoren dan de gemiddelde leesprestatie, zorgen voor een stijging van 0.953 punten in leesprestaties bij de individuele leerlingen. We zouden dus kunnen concluderen dat leerlingen die omringd worden door veel medeleerlingen met een goede leesvaardigheid, zelf beter scoren. We zouden zo kunnen stellen dat grotere klassen niet nadelig zijn, en dat ze misschien zelfs een positief effect hebben op de leesprestaties voor leerlingen. Met zulke uitspraken dienen we echter zorgzaam om te springen. Als dit al zou opgaan in deze situatie, wil dit nog niet zeggen dat dit ook zo is voor andere vakken of voor alle leerlingen van alle leeftijden.

Tabel 7
Model met Klasgrootte en Klaskenmerken als Controlevariabelen (n=3415)

Variabelen	Est.	p-waarde	95% B.I.	
			onder	boven
FIXED				
Intercept	0.1454	0.4330	-0.2038	0.5357
Klasgrootte	0.0143	0.3885	-0.0181	0.0475
Geslacht leerkracht vrouw	0.0153	0.7000	-0.0629	0.0925
Ervaring leerkracht	-0.0007	0.9652	-0.0336	0.0325
Gemiddelde leesniveau	0.9532	0.0000 *	0.8417	1.0685
Gemiddelde SES	-0.0121	0.7756	-0.0935	0.0723
Percentage Nederlandstaligen	-0.1143	0.5556	-0.4927	0.2812

* p < 0.05: significant

5.4 Is er een effect van klasgrootte op individuele leesprestaties na controle voor schoolkenmerken?

De schoolkenmerken die we in rekening brachten om deze onderzoeksvraag te beantwoorden, zijn de schoolgrootte en de ligging van de school, ook wel urbanisatiegraad genoemd (Model 4). Klasgrootte heeft **geen significant effect**, de p-waarde bedraagt namelijk 0.586. Zo ook heeft de schoolgrootte geen significante invloed. De urbanisatiegraad daarentegen is wel van significante tel. Een gemiddelde leerling uit een school die stedelijk of voorstedelijk gelegen is, scoort respectievelijk 0.077 of 0.084 punten lager dan een vergelijkbare leerling in een landelijke school.

Tabel 8
Model met Klasgrootte en Schoolkenmerken als Controlevariabelen (n=3415)

Variabelen	Est.	p-waarde	95% B.I.	
			onder	boven
FIXED				
Intercept	0.1053	0.0018 *	0.0438	0.1718
Klasgrootte	0.0121	0.5862	-0.0304	0.0531
Schoolgrootte	0.0364	0.0918	-0.0059	0.0767
Ligging stedelijk	-0.0766	0.1929	-0.1856	0.0359
Ligging voorstedelijk	-0.0842	0.0930	-0.1765	0.0091

* p < 0.05: significant

5.5 Verschilt het effect van klasgrootte op leesprestaties tussen klassen en scholen wanneer er gecontroleerd wordt voor de leerling-, klas-, en schoolkenmerken?

Wanneer we alle onafhankelijke kenmerken tezamen in één model brengen (Model 5), zien we dat de p-waarde van klasgrootte wel kleiner geworden is, 0.335, maar nog steeds ver boven de significantiegrens van 0.05 ligt. Er blijkt op basis van deze data duidelijk **geen effect** te zijn **van klasgrootte op de leesprestaties van leerlingen in het vierde leerjaar lager onderwijs**. Bij de inbreng van alle individuele, klas- en schoolkenmerken in het model, zien we dat volgende variabelen een significante invloed hebben op de leesscores: het geslacht en de leeftijd van de leerling, de socio-economische status en het cultureel kapitaal wanneer er meer dan 100 boeken in huis aanwezig zijn, de thuistaal van de leerling, de gemiddelde klasscore op de leestoets, de gemiddelde SES en het percentage Nederlandstaligen in de klas. De ligging van de school blijkt niet langer een significant effect uit te oefenen. Een gemiddeld meisje scoort 0.111 hoger dan een jongen. Leerlingen die één standaardafwijking hoger scoren op leeftijd (die dus ouder zijn), scoren 0.082 punten lager dan het intercept. Leerlingen die een SES-waarde hebben van één standaardafwijking hoger dan gemiddeld, behalen 0.265 punten minder op de leesscore. Hoe meer boeken er thuis aanwezig zijn (als indicator van cultureel kapitaal), hoe hoger de stijging in leesscores ten opzichte van de gemiddelde score. Leerlingen die thuis Nederlands spreken, hebben een voordeel ten aanzien van Anderstaligen, ze scoren namelijk 0.359 punten hoger. Dit geldt ook op klasniveau. Leerlingen in een klas met een gemiddeld hogere socio-economische status, scoren beter. Het gemiddelde leesniveau in de klas blijkt bovendien een sterke invloed te hebben; leerlingen in een klas die gemiddeld één standaardafwijking hoger ligt dan het gemiddelde klasniveau voor lezen, scoren individueel 0.950 punten hoger op de leestoets dan leerlingen in een klas met een gemiddelde leesscore.

Gezien klasgrootte geen significant effect heeft op de leesprestaties, is het niet relevant om interactie-effecten te onderzoeken die nagaan of er sprake is van differentiële effectiviteit. Er zal met andere woorden niet onderzocht kunnen worden of hypothese 2 opgaat voor deze data, omdat hypothese 1 niet bevestigd werd op basis van de resultaten uit 5.1.

Tabel 9

Model met Klasgrootte en Leerling-, Klas- en Schoolkenmerken als Controlevariabelen (n=3415)

Variabelen	Est.	p-waarde	95% B.I.	
			onder	boven
FIXED				
Intercept	-0.0987	0.6107	-0.4701	0.2824
Klasgrootte	0.0154	0.3349	-0.0157	0.0464
Geslacht meisje	0.1111	0.0002 *	0.0538	0.1708
Leeftijd	-0.0816	0.0000 *	-0.1146	-0.0488
SES	0.2653	0.0000 *	0.2292	0.2999
11-25 boeken	0.0346	0.5144	-0.0708	0.1388
26-100 boeken	0.0873	0.0697	-0.0076	0.1834
101-200 boeken	0.2381	0.0000 *	0.1246	0.3483
Meer dan 200 boeken	0.2689	0.0000 *	0.1583	0.3882
Thuis taal Nederlands	0.3592	0.0000 *	0.2123	0.5156
Geslacht leerkracht vrouw	0.0075	0.8412	-0.0649	0.0801
Ervaring leerkracht	0.0009	0.9534	-0.0299	0.0330
Gemiddelde leesniveau	0.9503	0.0000 *	0.8372	1.0567
Gemiddelde SES	-0.3522	0.0000 *	-0.4354	-0.2623
Percentage Nederlandstaligen	-0.3991	0.0475 *	-0.8065	-0.0171
Schoolgrootte	0.0093	0.5385	-0.0215	0.0378
Ligging stedelijk	-0.0116	0.7920	-0.0974	0.0734
Ligging voorstedelijk	-0.0031	0.9300	-0.0724	0.0678

* p <0.05: significant

5.6 Analyses op basis van het 3-level model

Eerder is aangegeven dat er geopteerd werd om het schoolniveau te laten varen, omdat er slechts weinig klassen per school waren. De analyses werden ook uitgevoerd waarbij het schoolniveau wel mee in rekening werd gebracht. Hiertoe werd eerst een nieuw nulmodel geschat.

Het nulmodel op schoolniveau leert ons dat de variantie tussen scholen significant is, dat er met andere woorden een effect is van scholen in het verklaren van de invloed van klasgrootte op de leesprestaties. 4.05% van de variantie in de leesscores kan toegeschreven worden aan de scholen waarin leerlingen zitten.

Tabel 10

Nulmodel op Schoolniveau (n=3415)

Variabelen	Est.	p-waarde	95% B.I.	
			onder	boven
FIXED				
Intercept	0.0513	0.0363*	0.0057	0.0989
RANDOM				
Variantie tussen klassen	0.0393			
Variantie binnen scholen	0.9312		0.9434	0.9900
ICC school	4.05%			

* p <0.05: significant

In onderstaande tabel is het volledige model (Model 5) terug te vinden, zowel zonder als met de inbreng van het schoolniveau. Conclusie luidt dat er geen verschillen optreden tussen de twee manieren van analyseren. De parameterschattingen en p-waarden zijn in beide modellen identiek, enkel de betrouwbaarheidsintervallen verschillen licht. Het gebruiken van een 2-level dan wel 3-level model heeft hier dus geen gevolgen voor de schattingen.

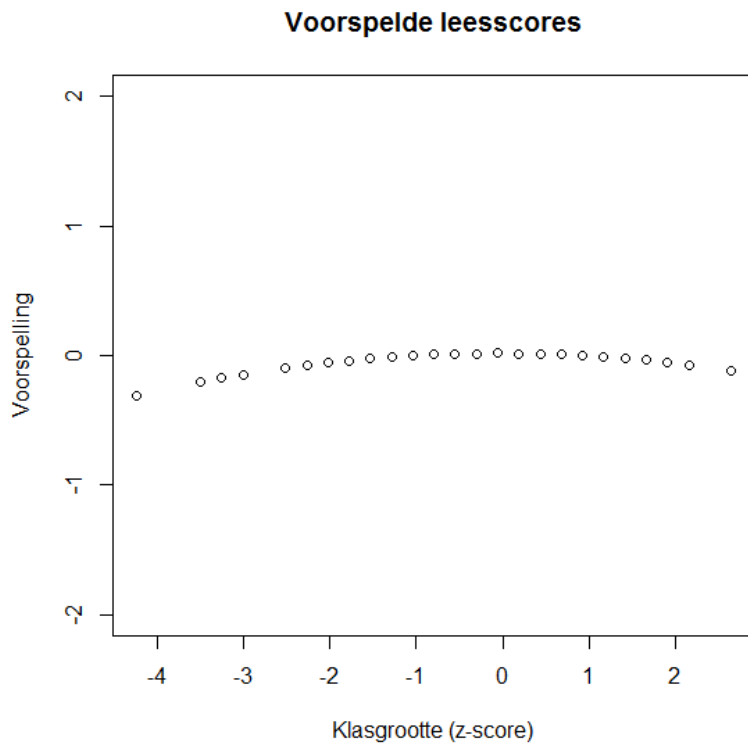
Tabel 11
Model met Klasgrootte en Leerling-, Klas- en Schoolkenmerken als Controlevariabelen (n=3415)

Variabelen	2-level model				3-level model			
	Est.	p-waarde	95% B.I.		Est.	p-waarde	95% B.I.	
			onder	boven			onder	boven
FIXED								
Intercept	-0.0987	0.6107	-0.4701	0.2824	-0.0987	0.6107	-0.4799	0.2908
Klasgrootte	0.0154	0.3349	-0.0157	0.0464	0.0154	0.3349	-0.0165	0.0457
Geslacht meisje	0.1111	0.0002 *	0.0538	0.1708	0.1111	0.0002 *	0.0524	0.1695
Leeftijd	-0.0816	0.0000 *	-0.1146	-0.0488	-0.0816	0.0000 *	-0.1132	-0.0492
SES	0.2653	0.0000 *	0.2292	0.2999	0.2653	0.0000 *	0.2281	0.2992
11-25 boeken	0.0346	0.5144	-0.0708	0.1388	0.0346	0.5144	-0.0698	0.1391
26-100 boeken	0.0873	0.0697	-0.0076	0.1834	0.0873	0.0697	-0.0038	0.1839
101-200 boeken	0.2381	0.0000 *	0.1246	0.3483	0.2381	0.0000 *	0.1292	0.3598
Meer dan 200 boeken	0.2689	0.0000 *	0.1583	0.3882	0.2689	0.0000 *	0.1542	0.3815
Thuis taal Nederlands	0.3592	0.0000 *	0.2123	0.5156	0.3592	0.0000 *	0.2002	0.5035
Geslacht leerkracht V	0.0075	0.8412	-0.0649	0.0801	0.0075	0.8412	-0.0665	0.0796
Ervaring leerkracht	0.0009	0.9534	-0.0299	0.0330	0.0009	0.9534	-0.0291	0.0339
Gemiddeld leesniveau	0.9503	0.0000 *	0.8372	1.0567	0.9503	0.0000 *	0.8412	1.0617
Gemiddelde SES	-0.3522	0.0000 *	-0.4354	-0.2623	-0.3522	0.0000 *	-0.4372	-0.2653
Percentage NL	-0.3991	0.0475 *	-0.8065	-0.0171	-0.3991	0.0475 *	-0.7944	0.0020
Schoolgrootte	0.0093	0.5385	-0.0215	0.0378	0.0093	0.5385	-0.0204	0.0398
Ligging stedelijk	-0.0116	0.7920	-0.0974	0.0734	-0.0116	0.7920	-0.0967	0.0738
Ligging voorstedelijk	-0.0031	0.9300	-0.0724	0.0678	-0.0031	0.9300	-0.0746	0.0655

* p < 0.05: significant

5.7 Onderzoeken van een curvilineair verband

Zoals eerder besproken, kan een term x^2 aan het analysemodel toegevoegd worden om niet-lineariteit in een relatie tussen x en y te testen. Via de variabele die klasgrootte kwadrateerde, bekeken we of er een curvilineair verband bestond (ter beantwoording van onderzoeksvraag 6). Dit bleek er niet te zijn. In Figuur 6 is te zien dat er wel degelijk sprake is van een lineair verband. Op basis van het model met klasgrootte en klasgrootte² als voorspellers, kan besloten worden dat er geen significant effect is op de leesscores.



Figuur 6: Voorspelling van de leesscores op basis van de klasgrootte

In onderstaande tabel is af te lezen dat de variabele ‘klasgrootte²’ niet significant is. Het blijkt dus geen voorspeller te zijn voor de leesprestaties.

Tabel 12

Model met Klasgrootte en Klasgrootte² als Voorspellers (n=4410)

Variabelen	Est.	p-waarde	95% B.I.	
			onder	boven
FIXED				
Intercept	0.0156	0.5780	-0.0349	0.0683
Klasgrootte	-0.0014	0.9533	-0.0449	0.0446
Klasgrootte ²	-0.0187	0.1918	-0.0457	0.0075
RANDOM				
Variantie tussen klassen	0.0803			
Variantie binnen klassen	0.9210			
ICC school	8.02%			

5.8 Overzichtstabel van alle analysemodellen

Volgende tabel geeft een overzicht van de parameterschattingen, p-waardes en betrouwbaarheidsintervallen van alle modellen weer: het nulmodel, model 1 waar klasgrootte als enige verklarende variabele is opgenomen, model 2 waar de individuele leerlingkenmerken zijn toegevoegd, model 3 waar de klaskenmerken zijn toegevoegd, model 4 waar de schoolkenmerken zijn toegevoegd en model 5, dat het integrale model is van alle kenmerken op leerling-, klas- en schoolniveau.

Tabel 13
Analysemodellen (n=3415)

Variabelen	Nulmodel		Model 1				Model 2					
	Est.	p-waarde	95% B.I.		Est.	p-waarde	95% B.I.		Est.	p-waarde	95% B.I.	
			onder	boven			onder	boven			onder	boven
Intercept	0.0583	0.0077	0.0190	0.1004	0.0583	0.0079	0.0190	0.1008	-0.5486	0.0000	-0.7242	-0.3905
Klasgrootte					0.0001	0.9952	-0.0413	0.0400	-0.0168	0.3745	-0.0413	0.0177
Geslacht meisje									0.1169	0.0002	0.0557	0.1773
Leeftijd									-0.0873	0.0000	-0.1220	-0.0537
SES									0.2663	0.0000	0.2328	0.3020
11-25 boeken									0.0581	0.2912	-0.0441	0.1684
26-100 boeken									0.0994	0.0468	-0.0037	0.1954
101-200 boeken									0.2442	0.0001	0.1259	0.3603
Meer dan 200 boeken									0.2562	0.0000	0.1396	0.3742
Thuis taal Nederlands									0.4420	0.0000	0.2992	0.5967

Variabelen	Model 3		Model 4				Model 5					
	Est.	p-waarde	95% B.I.		Est.	p-waarde	95% B.I.		Est.	p-waarde	95% B.I.	
			onder	boven			onder	boven			onder	boven
Intercept	0.1454	0.4330	-0.2038	0.5357	0.1053	0.0018	0.0438	0.1718	-0.0987	0.6107	-0.4701	0.2824
Klasgrootte	0.0143	0.3885	-0.0181	0.0475	0.0121	0.5862	-0.0304	0.0531	0.0154	0.3349	-0.0157	0.0464
Geslacht meisje									0.1111	0.0002	0.0538	0.1708
Leeftijd									-0.0816	0.0000	-0.1146	-0.0488
SES									0.2653	0.0000	0.2292	0.2999
11-25 boeken									0.0346	0.5144	-0.0708	0.1388
26-100 boeken									0.0873	0.0697	-0.0076	0.1834
101-200 boeken									0.2381	0.0000	0.1246	0.3483
Meer dan 200 boeken									0.2689	0.0000	0.1583	0.3882
Thuis taal Nederlands									0.3592	0.0000	0.2123	0.5156
Geslacht leerkracht												
vrouw	0.0153	0.7000	-0.0629	0.0925					0.0075	0.8412	-0.0649	0.0801
Ervaring leerkracht	-0.0007	0.9652	-0.0336	0.0325					0.0009	0.9534	-0.0299	0.0330
Gemiddelde leesniveau	0.9532	0.0000	0.8417	1.0685					0.9503	0.0000	0.8372	1.0567
Gemiddelde SES	-0.0121	0.7756	-0.0935	0.0723					-0.3522	0.0000	-0.4354	-0.2623
Percentage												
Nederlandstaligen	-0.1143	0.5556	-0.4927	0.2812					-0.3991	0.0475	-0.8065	-0.0171
Schoolgrootte					0.0364	0.0918	-0.0059	0.0767	0.0093	0.5385	-0.0215	0.0378
Ligging stedelijk					-0.0766	0.1929	-0.1856	0.0359	-0.0116	0.7920	-0.0974	0.0734
Ligging voorstedelijk					-0.0842	0.0930	-0.1765	0.0091	-0.0031	0.9300	-0.0724	0.0678

6 Conclusie en discussie

Er wordt hier concluderend een antwoord gegeven op de geformuleerde onderzoeksvragen. Dit gebeurt aan de hand van de resultaten uit de analyses in hoofdstuk 5. Tevens worden limitaties van het onderzoek aangestipt, met mogelijkheden tot vervolgonderzoek.

De probleemstelling en het theoretisch kader leerden dat onderzoek naar klasgrootte en prestaties erg verscheiden is, zowel in resultaten maar zeker ook naar de wijze en kwaliteit van de uitvoering. Hier wordt zo dadelijk nog op teruggekomen. Over de kwantiteit van bestaand onderzoek valt op te merken dat er nauwelijks Belgisch onderzoek uitgevoerd is. In deze studie werd de PIRLS 2006-data gebruikt om de relatie tussen klasgrootte en leesprestaties te onderzoeken. Dit gebeurde aan de hand van multilevelanalyses op de gegevens van 3415 Vlaamse leerlingen uit het vierde leerjaar basisonderwijs.

Is er een effect van klasgrootte op individuele leesprestaties van leerlingen?

Op basis van de analyses kan er met veel zekerheid gesteld worden dat uit deze steekproefgegevens geen effect afgeleid kan worden van de klasgrootte op de leesprestaties die Vlaamse leerlingen uit het vierde leerjaar behaalden op de test van PIRLS 2006. Bij de inleiding en probleemstelling werd reeds gezegd dat onderzoek op basis van de TIMSS-data aantoonde dat er geen grootschalige effecten gevonden werden op de wiskunde- en wetenschappenprestaties bij Vlaamse leerlingen (Woessmann & West, 2006). Er wordt nu dezelfde conclusie getrokken voor het effect van klasgrootte op de leesprestaties van leerlingen.

Verschilt het effect van klasgrootte op leesprestaties tussen klassen en scholen wanneer er gecontroleerd wordt voor de leerling-, klas-, en schoolkenmerken?

Zoals bij het zuiver verband, werd er geen significant effect gevonden tussen klasgrootte en leesprestaties wanneer er werd gecontroleerd voor individuele leerlingkenmerken, klas- en schoolkenmerken die op basis van literatuur als relevante controlevariabelen naar voren kwamen. Klasgrootte treedt niet als voorspeller op voor de leesprestaties, noch wanneer de leerling-, klas- en schoolkenmerken afzonderlijk in een model werden bekeken, noch wanneer ze in een integraal model werden geanalyseerd. Van de controlevariabelen blijken volgende variabelen in het integraal model wel een significant effect te vertonen: geslacht en leeftijd van de leerlingen, de socio-economische status, het cultureel kapitaal (wanneer er meer dan 100 boeken in huis aanwezig zijn) en de thuistaal van de leerlingen, de gemiddelde klasscore op de leestoets en de gemiddelde SES en het percentage Nederlandstaligen in de klas.

Indien klasgrootte wel significant bleek, zou het interessant geweest zijn niet enkel te kijken of klassen en scholen een effect hebben, maar ook of dit effect verschilt naargelang de niveaus. Hiertoe zouden Random Slopes-modellen opgesteld kunnen worden. Gezien de insignificantie van de variabele, was dit echter niet zinvol en wenselijk.

Bestaat er in plaats van een lineair, een curvilineair verband tussen klas-grootte en leerlingenprestaties?

Via de variabele die klasgrootte kwadrateerde, werd onderzocht of er zo'n curvilineair verband bestaat. Dit bleek niet het geval te zijn; er is wel degelijk sprake van een lineair verband. Op basis van het model met klasgrootte en klasgrootte² als voorspellers, kan besloten worden dat er geen significant effect is op de leesscores.

Limitaties en mogelijkheden tot vervolgonderzoek

De onderzoeksmethode, populatie, de aard en de gebruikte maat om de prestaties te meten, de controlevariabelen en de operationalisatie van de klasgroottevariabele, vormen allemaal factoren die de onderzoeksresultaten kunnen beïnvloeden. Dit maakt dat uitspraken zeer moeilijk gegeneraliseerd kunnen worden. Doorheen de literatuur werden allerlei kritieken aangehaald op bepaalde onderzoekswijzen. In deze studie werd getracht zoveel mogelijk op deze zaken te anticiperen, wat maakt dat bepaalde, maar niet alle, 'valkuilen' vermeden konden worden.

Eén van de zwaktes die doorheen het lezen van literatuur naar boven kwam, is dat men vaak geen gebruik maakte van de eigenlijke klasgrootte. Als afhankelijke variabele werden in deze studie de werkelijke klasgroottes gebruikt, en niet de gemiddelde klasgrootte door de schoolgrootte te delen door het aantal klassen. Zo is er in de dataset een correcte toewijzing van de klasgrootte bij elke individuele leerling. Dit kan dus niet voor vertekende resultaten hebben gezorgd. Het feit dat werd geanalyseerd op basis van multileveltechnieken, maakt dat er rekening is gehouden met invloeden op het niveau van de klas en de school. Ook dit zorgt voor zuiverdere resultaten omdat rekening is gehouden met de hiërarchische structuur die er bestaat. Ten slotte geven de resultaten ook zicht op de huidige situatie van het lager onderwijs in Vlaanderen. Doordat de gegevens verzameld zijn in de realiteit en niet op basis van een experiment waar bepaalde interventies werden opgezet, vertellen ze iets over de onderwijsrealiteit in plaats van over een experimentele context.

Het was desalniettemin onmogelijk om aan alle kritiek op de onderzoeksmethoden tegemoet te komen, onder andere gezien de aard van de gegevens. De data die door PIRLS verzameld werd, bevat geen gegevens over de groottes van klassen waarin de leerlingen de voorgaande jaren zaten. Er van uitgaande dat deze mee de huidige prestaties kunnen beïnvloeden, is dit een limitatie aan het onderzoek.

Door gebrek aan informatie over de toewijzing van leerlingen aan klassen, kon er geen rekening gehouden worden met de selectiviteit of endogeniteit van klasgroottes. Misschien werden leerlingen bewust gegroepeerd in een kleine dan wel grote klas, maar hier is geen zicht op. Het zou erg zinvol zijn om dit in toekomstig onderzoek na te gaan. Gezien de beperkte tijdsspanne en de nood aan een grootschalige dataverzameling om kwantitatieve analyses op een betrouwbare manier te kunnen uitvoeren, werd geopteerd voor het gebruik van de PIRLS 2006-data. Indien middelen en tijd het toelieten, zou een eigen dataverzameling zinvol zijn zodat de uitwerking op eigen manier invulling kan krijgen. Er zou zo rekening kunnen worden gehouden met het issue van endogeniteit en andere factoren die in het huidig onderzoek een limitatie vormen.

Naast de invulling van klasgrootte als voorspellende variabele, is ook de operationalisatie van de voorspelde variabele mee bepalend geweest voor de onderzoeksresultaten. Het feit dat hier de leesprestaties als afhankelijke variabele werden gebruikt, beperkt het onderzoek tot een cognitieve outputmaat. Men zou in de toekomst ook de effecten op prestaties op sociaal-affectief vlak of voor meerdere vakgebieden samen kunnen onderzoeken.

De leerlingen van de steekproef zaten allemaal in het vierde leerjaar. Mogelijks is er een verschillend effect naargelang de leeftijd van de leerlingen. Vervolgonderzoek zou zich kunnen toespitsen op de vraag of er wel of geen, meer of minder effect is voor jongere dan wel oudere leerlingen. In het secundair onderwijs dient rekening gehouden te worden met het feit dat de klasgrootte per vak kan verschillen, omwille van keuzevakken en dergelijke. Bovendien zou men, indien er een effect vastgesteld wordt, kunnen onderzoeken hoe lang en in welke mate dit klasgrootte-effect doorwerkt.

Een andere beperking die het gebruik van de PIRLS-data met zich meebracht, is het feit dat er geen gegevens in opgenomen zijn over leerlingen in het buitengewoon onderwijs. Het zou niet verwonderlijk zijn dat leerlingen met extra zorgbehoeften positievere resultaten boeken wanneer ze in kleinere klassen zitten. Dit valt te verklaren door het gegeven dat leerkrachten hun tijd en aandacht dan over minder leerlingen hoeven te spreiden. Toekomstig onderzoek zou zich kunnen toespitsen op deze leerlingen en zou eventueel het individuele kenmerk IQ mee kunnen opnemen in de analyses.

7 Referenties

- Agentschap voor Onderwijsdiensten. (2011). Gemiddelde klasgrootte per basisschool. *Ministerie van Onderwijs en Vorming*. Retrieved from <http://www.ond.vlaanderen.be/wegwijs/agodi/cijfermateriaal/klasgrootte/default.htm>
- Akerhielm, K. (1995). Does class size matter? *Economics of Education Review*, 14(3), 229–241. Elsevier.
- Amkreutz, R. (2001). 1.300 extra leraars voor overvolle kleuterklassen. Retrieved September 20, 2011, from <http://www.pascalsmet.be/article/1300-extra-leraars-voor-overvolle-kleuterklassen/>
- Arias, J., & Walker, D. M. (2004). Additional evidence on the relationship between class size and student performance. *The Journal of Economic Education*, 35(4), 311–329. Heldref Publications.
- Bausell, R. B., Moody, W. B., & Walze, F. N. (1972). A factorial study of tutoring versus classroom instruction. *American Educational Research Journal*, 9, 591-598.
- Bosker, R. J. (1998). The class size question in primary schools: policy issues, theory, and empirical findings from the Netherlands. *International Journal of Educational Research*, 29(8), 763–778. Elsevier.
- Card, D., & Krueger, A. B. (1992). School quality and black-white relative earnings: A direct assessment. *The Quarterly Journal of Economics*, 107, 151-200.
- Chall, J. (1983). *Stages of reading development*. New York: McGraw-Hill.
- Christiaens, W. (2009). *De invloed van de leerkracht op de leesvaardigheid in het basisonderwijs (Masterthesis)*. Universiteit Antwerpen.
- Chubb, J. E., & Moe, T. M. (1990). *Politics, markets & America's schools*. Washington, D.C.: The Brookings Institution.
- Coleman, J. S., Campbell, E. Q., Hobson, C. J., McPartland, J., Mood, A. M., Weinfeld, F. D., & York, R. L. (1966). *Equality of educational opportunity*. Washington, D.C.: National Center for Education Statistics.
- Coleman, J. S., Hoffer, T., & Kilgore, S. (1982). *High school achievement: Public, Catholic and private schools compared*. New York: Basic Books.
- Colpin, M., Ramaut, G., Timmermans, S., Van den Branden, K., Vandenbroucke, M., & Van Gorp, K. (2002). *Leesrijk school- en klasklimaat*. Antwerpen - Appeldoorn: Garant.
- Crawley, M. J. (2007). *The R book* (pp. 356-357). West Sussex: John Wiley & Sons Ltd.
- De Maeyer, S., & Rymenans, R. (2004). *Onderzoek naar kenmerken van effectieve scholen* (pp. 10, 17-18, 159-161, 164). Gent: Academia Press.
- Dobbelsteen, S., Levin, J., & Oosterbeek, H. (1999). Waarom klassenverkleining niet helpt. *Economisch Statistische Berichten*, 84(4197), 250–252.

- Faculteit Psychologie en Pedagogische Wetenschappen KU Leuven. (n.d.). PIRLS 2006 Vlaanderen. Retrieved April 18, 2012, from http://fac.ppw.kuleuven.be/o_en_o/COE/pirls/steekproef.html
- Finn, J. D., & Achilles, C. M. (1990). Answers and questions about class size: A statewide experiment. *American Educational Research Journal*, 27(3), 557-577.
- Fox, J. (1997). *Applied regression analysis, linear models, and related methods* (pp. 59-60). California: Sage Publications.
- Foy, P., & Kennedy, A. M. (2008). *PIRLS 2006 user guide for the international database. Evaluation. TIMSS & PIRLS International Study Center, Lynch School of Education, Boston College.*
- Glass, G. V., Cahen, L. S., Smith, M. L., & Filby, N. N. (1982). *School class size: Research and policy* (pp. 1-160).
- Glass, G. V., & Smith, M. L. (1978). Meta-analysis of research on the relationship of class-size and achievement. The class size and instruction project.
- Hancock, T. M. (1996). Effects of class size on college student achievement. *College Student Journal*, 30(4), 479-81.
- Hanushek, E. A. (1997). Assessing the effects of school resources on student performance: An update. *Educational Evaluation and Policy Analysis*, 19(2), 141-64.
- Hoxby, C. M. (2000). The effects of class size on student achievement: New evidence from population variation. *Quarterly Journal of economics*, 115(4), 1239-1285. MIT Press.
- International Association for the Evaluation of Educational Achievement. (2012). About PIRLS 2006. *TIMSS & PIRLS International Study Center, Lynch School of Education, Boston College.* Retrieved January 16, 2012, from <http://timss.bc.edu/pirls2006/about.html>
- Kabinet Vlaams minister van Onderwijs en Vorming. (2007). Eerste reactie van de minister op de PISA 2006-resultaten voor Vlaanderen. Retrieved February 7, 2012, from <http://www.ond.vlaanderen.be/nieuws/archief/2007/2007p/1204-PISA.htm>
- Keil, J., & Partell, P. J. (1997). The effect of class size on student performance and retention at Binghamton University. *Office of Budget and Institutional Research.*
- Krueger, A. (1999). Experimental estimates of education production functions. *The Quarterly Journal of Economics*, 114(2), 497-532.
- Linnakyla, P., Malin, A., & Taube, K. (2004). Factors behind low reading literacy achievement. *Scandinavian Journal of Educational Research*, 48(3), 231-249.
- Mishel, L., & Rothstein, R. (2002). *The class size debate. Education* (pp. 1-102). Washington, D.C.: Economy Policy Institute.
- Mol, S., & Bus, A. (2011). Lezen loont een leven lang. *Levende Talen Tijdschrift*, 3.
- Molnar, A., Smith, P., Zahorik, J., & Palmer, A. (1999). Evaluating the SAGE program: A pilot program in targeted pupil-teacher reduction in Wisconsin. *Educational Evaluation and Policy Analysis*, 21(2), 165-177.

- Mosteller, F. (1995). The Tennessee study of class size in the early school grades. *The future of children*, 5(2), 113–127. JSTOR.
- Nye, B., Hedges, L., & Konstantopoulos, S. (1999). The long-term effects of small classes: A five-year follow-up of the Tennessee class size experiment. *Educational Evaluation and Policy Analysis*, 21(2), 127-142.
- Nye, B., Hedges, L. V., & Konstantopoulos, S. (2000). The effects of small classes on academic achievement: The results of the Tennessee class size experiment. *American Educational Research Journal*, 37(1), 123–151. Sage Publications.
- Porwell, P. J. (1978). *Class size: A summary of research*. Educational Research Service (pp. 68-69). Arlington VA: Educational Research Service.
- Purves, A. C., & Elley, W. B. (1994). The role of the home and student differences in reading performance. In W.B. Elley (Ed.), *The IEA study of reading literacy: Achievement and instruction in thirty-two school systems*. Oxford, UK: Pergamon Press.
- Rice, J. M. (1902). Educational research: A test in arithmetic. *The Forum*, 281-297.
- Siegfried, J. J., & Kennedy, P. E. (1995). Does pedagogy vary with class size in introductory economics? *The American Economic Review*, 85(2), 347-351.
- Tope, R. E., Groom, E., & Beeson, M. F. (1924). Size of class and school efficiency. *Journal of Educational Research*, 9, 126-132.
- Van Damme, J., Liu, H., Vanhee, L., & Pustjens, H. (2010). Longitudinal studies at the country level as a new approach to educational effectiveness: Explaining change in reading achievement (PIRLS) by change in age, socio-economic status and class size. *Effective Education*, 2(1), 53–84. Taylor & Francis.
- Wester, F., Renckstorf, K., & Scheepers, P. (2006). *Onderzoekstypen in de communicatiewetenschap* (Tweede dru., p. 237). Alphen aan den Rijn: Kluwer.
- Woessmann, L., & West, M. (2006). Class-size effects in school systems around the world: Evidence from between-grade variation in TIMSS. *European Economic Review*, 50(3), 695–736. Elsevier.

8 Bijlagen

De output uit het statistische programma R wordt hier als bijlage toegevoegd. Hieronder een overzicht van de bijgevoegde modellen:

- 8.1 Analyse van het nulmodel level 2= school
- 8.2 Analyse van het nulmodel level 2= klas
- 8.3 Analyse van het RImodel
- 8.4 Analyse van het model met leerlingkenmerken toegevoegd: geslacht, leeftijd, socio-economische status, cultureel kapitaal, thuistaal
- 8.5 Analyse van het model met klaskenmerken toegevoegd: geslacht en ervaring van de leerkracht, gemiddeld leesniveau, gemiddelde socio-economische status, gemiddelde thuistaal
- 8.6 Analyse van het model met schoolkenmerken toegevoegd: schoolgrootte, urbanisatiegraad
- 8.7 Analyse van het model met alle kenmerken samen
- 8.8 Analyse om een curvilineair verband te testen

8.1 Analyse van het nulmodel level 2= school

```
> Nulmodelschool<-lmer(OVERALLREAD_z~1+(1|IDSCHOOL),data=Data,REML=FALSE)
> summary(Nulmodelschool)
Linear mixed model fit by maximum likelihood
Formula: OVERALLREAD_z ~ 1 + (1 | IDSCHOOL)
Data: Data
  AIC   BIC logLik deviance REMLdev
9548 9566  -4771    9542    9547
Random effects:
Groups   Name              Variance Std.Dev.
IDSCHOOL (Intercept) 0.039295 0.19823
Residual                0.931201 0.96499
Number of obs: 3415, groups: IDSCHOOL, 131

Fixed effects:
              Estimate Std. Error t value
(Intercept)  0.05137    0.02452   2.095
> pvals.fnc(Nulmodelschool)
$fixed
  Estimate MCMCmean HPD95lower HPD95upper  pMCMC Pr(>|t|)
1  0.0514   0.0519    0.0057    0.0989 0.0328  0.0363

$random
  Groups   Name              Std.Dev. MCMCmedian MCMCmean HPD95lower HPD95upper
1 IDSCHOOL (Intercept)    0.1982    0.1844   0.1849    0.1411    0.2332
2 Residual                0.9650    0.9669   0.9669    0.9434    0.9900

> ICCschool<- 0.039295/(0.039295+ 0.931201)
> ICCschool
[1] 0.04048961
```

8.2 Analyse van het nulmodel level 2= klas

```
> Nulmodelklas<-lmer(OVERALLREAD_z~1+(1>IDCLASS),data=Data,REML=FALSE)
> summary(Nulmodelklas)
Linear mixed model fit by maximum likelihood
Formula: OVERALLREAD_z ~ 1 + (1 | IDCLASS)
Data: Data
  AIC   BIC logLik deviance REMLdev
9556 9574  -4775    9550    9556
Random effects:
Groups   Name              Variance Std.Dev.
IDCLASS  (Intercept) 0.042808 0.20690
Residual                0.927104 0.96286
Number of obs: 3415, groups: IDCLASS, 223

Fixed effects:
              Estimate Std. Error t value
(Intercept)  0.05826    0.02186   2.665

> pvals.fnc(Nulmodelklas)
$fixed
  Estimate MCMCmean HPD95lower HPD95upper  pMCMC Pr(>|t|)
1  0.0583   0.0596    0.019    0.1004 0.0058  0.0077

$random
  Groups   Name              Std.Dev. MCMCmedian MCMCmean HPD95lower HPD95upper
1 IDCLASS  (Intercept)    0.2069    0.1815   0.1814    0.1348    0.2273
2 Residual                0.9629    0.9666   0.9668    0.9441    0.9912
```

```
> ICCKlas<- 0.042808/(0.042808+ 0.927104)
> ICCKlas
[1] 0.04413596
```

8.3 Analyse van het RImodel

```
> RImodel<-lmer(OVERALLREAD_z~KLGROOTTE_z+(1|IDCLASS),data=Data,REML=FALSE)
> summary(RImodel)
Linear mixed model fit by maximum likelihood
Formula: OVERALLREAD_z ~ KLGROOTTE_z + (1 | IDCLASS)
Data: Data
   AIC   BIC logLik deviance REMLdev
9558 9582 -4775    9550    9562
Random effects:
 Groups      Name      Variance Std.Dev.
IDCLASS (Intercept) 0.042809 0.20690
Residual                0.927103 0.96286
Number of obs: 3415, groups: IDCLASS, 223

Fixed effects:
              Estimate Std. Error t value
(Intercept) 0.0582701  0.0219263  2.658
KLGROOTTE_z 0.0001308  0.0217037  0.006

Correlation of Fixed Effects:
      (Intr)
KLGROOTTE_z 0.078
> pvals.fnc(RImodel)
$fixed
              Estimate MCMCmean HPD95lower HPD95upper  pMCMC Pr(>|t|)
(Intercept)  0.0583  0.0591  0.0190  0.1008 0.0054  0.0079
KLGROOTTE_z  0.0001 -0.0003 -0.0413  0.0400 0.9924  0.9952

$random
  Groups      Name Std.Dev. MCMCmedian MCMCmean HPD95lower HPD95upper
1 IDCLASS (Intercept)  0.2069  0.1824  0.1825  0.1348  0.2281
2 Residual                0.9629  0.9665  0.9667  0.9420  0.9900
```

8.4 Analyse van het model met leerlingkenmerken toegevoegd: geslacht, leeftijd, SES, cultureel kapitaal, thuistaal

```
> RImodel2<-
lmer(OVERALLREAD_z~KLGROOTTE_z+Meisje+AGE_z+SES+Elfboeken+Zestwintigboeken+Honde
rdeenboeken+Meerdanboeken+Nederlands+(1|IDCLASS),data=Data,REML=FALSE)
> summary(RImodel2)
Linear mixed model fit by maximum likelihood
Formula: OVERALLREAD_z ~ KLGROOTTE_z + Meisje + AGE_z + SES + Elfboeken +
Zestwintigboeken + Honderdeenboeken + Meerdanboeken + Nederlands + (1 |
IDCLASS)
Data: Data

   AIC   BIC logLik deviance REMLdev
9078 9151 -4527    9054    9104
Random effects:
 Groups      Name      Variance Std.Dev.
IDCLASS (Intercept) 0.024772 0.15739
Residual                0.809409 0.89967
Number of obs: 3415, groups: IDCLASS, 223
```

Fixed effects:

	Estimate	Std. Error	t value
(Intercept)	-0.54865	0.08494	-6.459
KLGROOTTE_z	-0.01681	0.01893	-0.888
Meisje	0.11694	0.03114	3.755
AGE_z	-0.08732	0.01736	-5.031
SES	0.26633	0.01763	15.102
Elfboeken	0.05813	0.05507	1.056
Zestwintigboeken	0.09944	0.05000	1.989
Honderdeenboeken	0.24422	0.06018	4.058
Meerdanboeken	0.25616	0.06038	4.243
Nederlands	0.44197	0.07585	5.827

Correlation of Fixed Effects:

	(Intr)	KLGR00	Meisje	AGE_z	SES	Elfbkn	Zstwnt	Hndrdrn	Mrdnbk
KLGROOTTE_z	0.009								
Meisje	-0.174	0.011							
AGE_z	-0.092	-0.009	0.031						
SES	0.129	-0.025	0.027	0.196					
Elfboeken	-0.349	-0.013	-0.036	-0.035	-0.075				
Zestwntgbkn	-0.401	-0.041	-0.039	-0.018	-0.202	0.631			
Honderdnbkn	-0.332	-0.024	-0.035	-0.014	-0.259	0.533	0.624		
Meerdanbokn	-0.373	-0.044	-0.049	-0.016	-0.363	0.538	0.643	0.573	
Nederlands	-0.841	0.024	0.021	0.133	-0.008	-0.021	-0.024	-0.036	-0.001

> pvals.fnc(RImodel2)

\$fixed

	Estimate	MCMCmean	HPD95lower	HPD95upper	pMCMC	Pr(> t)
(Intercept)	-0.5486	-0.5538	-0.7242	-0.3905	0.0001	0.0000
KLGROOTTE_z	-0.0168	-0.0167	-0.0538	0.0177	0.3644	0.3745
Meisje	0.1169	0.1165	0.0557	0.1773	0.0002	0.0002
AGE_z	-0.0873	-0.0873	-0.1220	-0.0537	0.0001	0.0000
SES	0.2663	0.2666	0.2328	0.3020	0.0001	0.0000
Elfboeken	0.0581	0.0597	-0.0441	0.1684	0.2744	0.2912
Zestwintigboeken	0.0994	0.0999	-0.0037	0.1954	0.0502	0.0468
Honderdeenboeken	0.2442	0.2442	0.1259	0.3603	0.0001	0.0001
Meerdanboeken	0.2562	0.2549	0.1396	0.3742	0.0001	0.0000
Nederlands	0.4420	0.4478	0.2992	0.5967	0.0001	0.0000

\$random

Groups	Name	Std.Dev.	MCMCmedian	MCMCmean	HPD95lower	HPD95upper
1	IDCLASS (Intercept)	0.1574	0.1361	0.1356	0.0912	0.1845
2	Residual	0.8997	0.9038	0.9038	0.8810	0.9252

8.5 Analyse van het model met klaskenmerken toegevoegd: geslacht en ervaring van de leerkracht, gemiddeld leesniveau, gemiddelde SES, gemiddelde thuistaal

```
> RImodel3<-
lmer(OVERALLREAD_z~KLGROOTTE_z+Vrouw+ERVARING_z+OVERALLREAD_gem+SES_gem+Nederlands_perc+(1|IDCLASS),data=Data,REML=FALSE)
> summary(RImodel3)
Linear mixed model fit by maximum likelihood
Formula: OVERALLREAD_z ~ KLGROOTTE_z + Vrouw + ERVARING_z + OVERALLREAD_gem + SES_gem + Nederlands_perc + (1 | IDCLASS)
Data: Data
   AIC   BIC logLik deviance REMLdev
 9245 9301 -4614   9227   9261
Random effects:
Groups   Name      Variance Std.Dev.
IDCLASS (Intercept) 0.00000  0.00000
Residual              0.87295  0.93432
Number of obs: 3415, groups: IDCLASS, 223
```


Fixed effects:

	Estimate	Std. Error	t value
(Intercept)	0.1453892	0.1853889	0.784
KLGRROOTTE_z	0.0143445	0.0166318	0.862
Vrouw	0.0152577	0.0395889	0.385
ERVARING_z	-0.0007337	0.0168149	-0.044
OVERALLREAD_gem	0.9532132	0.0593228	16.068
SES_gem	-0.0121257	0.0425302	-0.285
Nederlands_perc	-0.1143391	0.1939699	-0.589

Correlation of Fixed Effects:

	(Intr)	KLGRROO	Vrouw	ERVARI	OVERAL	SES_gm
KLGRROOTTE_z	-0.069					
Vrouw	-0.149	0.144				
ERVARING_z	-0.073	0.033	0.289			
OVERALLREAD	0.342	0.098	0.036	-0.043		
SES_gem	-0.145	-0.152	0.011	0.089	-0.512	
Ndrlns_prc	-0.983	0.046	-0.014	0.027	-0.358	0.149

> pvals.fnc(RImodel3)
\$fixed

	Estimate	MCMCmean	HPD95lower	HPD95upper	pMCMC	Pr(> t)
(Intercept)	0.1454	0.1482	-0.2038	0.5357	0.4274	0.4330
KLGRROOTTE_z	0.0143	0.0144	-0.0181	0.0475	0.3890	0.3885
Vrouw	0.0153	0.0155	-0.0629	0.0925	0.7062	0.7000
ERVARING_z	-0.0007	-0.0009	-0.0336	0.0325	0.9584	0.9652
OVERALLREAD_gem	0.9532	0.9551	0.8417	1.0685	0.0001	0.0000
SES_gem	-0.0121	-0.0126	-0.0935	0.0723	0.7726	0.7756
Nederlands_perc	-0.1143	-0.1177	-0.4927	0.2812	0.5474	0.5556

\$random

Groups	Name	Std.Dev.	MCMCmedian	MCMCmean	HPD95lower	HPD95upper
1	IDCLASS (Intercept)	0.0000	0.0014	0.0075	0.0000	0.0292
2	Residual	0.9343	0.9353	0.9354	0.9144	0.9586

8.6 Analyse van het model met schoolkenmerken toegevoegd: schoolgrootte, urbanisatiegraad

```
> RImodel4<-  
lmer(OVERALLREAD_z~KLGRROOTTE_z+SCHGROOTTE_z+Urban+Suburban+(1|IDCLASS),data=Data,  
REML=FALSE)
```

```
> summary(RImodel4)
```

Linear mixed model fit by maximum likelihood
Formula: OVERALLREAD_z ~ KLGRROOTTE_z + SCHGROOTTE_z + Urban + Suburban + (1 | IDCLASS)

Data: Data

AIC	BIC	logLik	deviance	REMLdev
9558	9601	-4772	9544	9570

Random effects:

Groups	Name	Variance	Std.Dev.
IDCLASS	(Intercept)	0.041351	0.20335
	Residual	0.926389	0.96249

Number of obs: 3415, groups: IDCLASS, 223

Fixed effects:

	Estimate	Std. Error	t value
(Intercept)	0.10528	0.03376	3.119
KLGRROOTTE_z	0.01206	0.02216	0.544
SCHGROOTTE_z	0.03642	0.02159	1.687
Urban	-0.07663	0.05884	-1.302
Suburban	-0.08421	0.05012	-1.680

```

Correlation of Fixed Effects:
      (Intr) KLGROO SCHGRO Urban
KLGROOTTE_z  0.205
SCHGROOTTE_  0.066  0.127
Urban        -0.577 -0.145  0.023
Suburban     -0.688 -0.192 -0.137  0.393
> pvals.fnc(RImodel4)
$fixed
      Estimate MCMCmean HPD95lower HPD95upper  pMCMC Pr(>|t|)
(Intercept)  0.1053  0.1053  0.0438  0.1718 0.0004  0.0018
KLGROOTTE_z  0.0121  0.0114  -0.0304  0.0531 0.5870  0.5862
SCHGROOTTE_z  0.0364  0.0367  -0.0059  0.0767 0.0772  0.0918
Urban        -0.0766  -0.0729  -0.1856  0.0359 0.1966  0.1929
Suburban     -0.0842  -0.0833  -0.1765  0.0091 0.0820  0.0930

$random
      Groups      Name Std.Dev. MCMCmedian MCMCmean HPD95lower HPD95upper
1  IDCLASS (Intercept)  0.2033  0.1815  0.1814  0.1346  0.2289
2  Residual              0.9625  0.9663  0.9664  0.9425  0.9906

```

8.7 Analyse van het model met alle kenmerken samen

```

> RImodel5<-
lmer(OVERALLREAD_z~KLGROOTTE_z+Meisje+AGE_z+SES+Elfboeken+Zestwintigboeken+Honderdeen
boeken+Meerdanboeken+Nederlands+Vrouw+ERVARING_z+OVERALLREAD_gem+SES_gem+Nederlands_
perc+SCHGROOTTE_z+Urban+Suburban+(1|IDCLASS),data=Data,REML=FALSE)

> summary(RImodel5)
Linear mixed model fit by maximum likelihood
Formula: OVERALLREAD_z ~ KLGROOTTE_z + Meisje + AGE_z + SES + Elfboeken +
Zestwintigboeken + Honderdeenboeken + Meerdanboeken + Nederlands + Vrouw +
ERVARING_z + OVERALLREAD_gem + SES_gem + Nederlands_perc + SCHGROOTTE_z +
Urban + Suburban + (1 | IDCLASS)
Data: Data
AIC   BIC logLik deviance REMLdev
8815 8938 -4388   8775   8865
Random effects:
Groups      Name      Variance Std.Dev.
IDCLASS (Intercept) 0.00000  0.00000
Residual              0.76472  0.87448
Number of obs: 3415, groups: IDCLASS, 223

Fixed effects:
      Estimate Std. Error t value
(Intercept)  -0.0986540  0.1937749  -0.509
KLGROOTTE_z   0.0153661  0.0159338   0.964
Meisje        0.1110645  0.0300342   3.698
AGE_z        -0.0816304  0.0166891  -4.891
SES           0.2652503  0.0179801  14.752
Elfboeken     0.0346029  0.0530652   0.652
Zestwintigboeken 0.0873046  0.0481210   1.814
Honderdeenboeken 0.2381277  0.0579846   4.107
Meerdanboeken 0.2688938  0.0583791   4.606
Nederlands    0.3592231  0.0772389   4.651
Vrouw         0.0074571  0.0372093   0.200
ERVARING_z    0.0009359  0.0160068   0.058
OVERALLREAD_gem 0.9502554  0.0558427  17.017
SES_gem       -0.3521731  0.0433406  -8.126
Nederlands_perc -0.3990934  0.2012900  -1.983
SCHGROOTTE_z  0.0093482  0.0151963   0.615
Urban         -0.0115671  0.0438670  -0.264
Suburban      -0.0031180  0.0354781  -0.088

```

Correlation of Fixed Effects:

```
(Intr) KLGROO Meisje AGE_z SES Elfbkn Zstwnt Hndrdrn Mrdnbk
KLGROOTTE_z -0.001
Meisje -0.065 0.003
AGE_z -0.017 -0.024 0.033
SES 0.055 0.006 0.023 0.175
Elfboeken -0.185 -0.018 -0.036 -0.033 -0.072
Zestwntgbkn -0.209 -0.046 -0.039 -0.016 -0.187 0.632
Honderdnbkkn -0.161 -0.017 -0.034 -0.013 -0.237 0.531 0.623
Meerdanbokkn -0.201 -0.033 -0.049 -0.015 -0.324 0.536 0.643 0.575
Nederlands -0.052 -0.001 0.022 0.127 -0.001 -0.026 -0.033 -0.045 -0.022
Vrouw -0.148 0.140 -0.031 -0.020 0.006 0.003 -0.019 -0.015 -0.019
ERVARING_z -0.125 0.011 -0.001 -0.013 0.000 -0.012 -0.013 -0.003 -0.014
OVERALLREAD 0.303 0.079 -0.003 0.022 -0.001 -0.045 -0.030 -0.020 -0.005
SES_gem -0.100 -0.131 0.015 0.018 -0.316 0.023 -0.001 -0.023 -0.059
Ndrlns_prc -0.889 0.003 -0.006 -0.021 -0.010 0.040 0.047 0.031 0.059
SCHGROOTTE_ -0.015 0.132 -0.030 -0.021 -0.016 0.016 0.022 0.043 0.025
Urban -0.343 -0.109 -0.010 -0.022 0.020 0.010 0.011 -0.028 -0.054
Suburban -0.252 -0.166 0.018 0.018 0.012 0.006 -0.002 -0.024 -0.012
```

```
Ndrlns Vrouw ERVARI OVERAL SES_gem Ndrln_ SCHGRO Urban
```

```
KLGROOTTE_z
Meisje
AGE_z
SES
Elfboeken
Zestwntgbkn
Honderdnbkkn
Meerdanbokkn
Nederlands
Vrouw -0.015
ERVARING_z 0.002 0.295
OVERALLREAD 0.002 0.032 -0.040
SES_gem 0.012 0.008 0.075 -0.471
Ndrlns_prc -0.326 0.009 0.069 -0.318 0.104
SCHGROOTTE_ 0.005 0.042 0.041 -0.080 0.007 0.012
Urban 0.005 0.067 0.156 0.005 -0.046 0.266 0.020
Suburban -0.005 0.022 0.136 0.053 -0.003 0.167 -0.139 0.434
```

```
> pvals.fnc(RImodel5)
```

```
$fixed
```

	Estimate	MCMCmean	HPD95lower	HPD95upper	pMCMC	Pr(> t)
(Intercept)	-0.0987	-0.0999	-0.4701	0.2824	0.5920	0.6107
KLGROOTTE_z	0.0154	0.0152	-0.0157	0.0464	0.3370	0.3349
Meisje	0.1111	0.1101	0.0538	0.1708	0.0004	0.0002
AGE_z	-0.0816	-0.0817	-0.1146	-0.0488	0.0001	0.0000
SES	0.2653	0.2649	0.2292	0.2999	0.0001	0.0000
Elfboeken	0.0346	0.0344	-0.0708	0.1388	0.5076	0.5144
Zestwintigboeken	0.0873	0.0876	-0.0076	0.1834	0.0736	0.0697
Honderdeenboeken	0.2381	0.2391	0.1246	0.3483	0.0001	0.0000
Meerdanboeken	0.2689	0.2699	0.1583	0.3882	0.0001	0.0000
Nederlands	0.3592	0.3597	0.2123	0.5156	0.0001	0.0000
Vrouw	0.0075	0.0073	-0.0649	0.0801	0.8514	0.8412
ERVARING_z	0.0009	0.0011	-0.0299	0.0330	0.9382	0.9534
OVERALLREAD_gem	0.9503	0.9502	0.8372	1.0567	0.0001	0.0000
SES_gem	-0.3522	-0.3519	-0.4354	-0.2623	0.0001	0.0000
Nederlands_perc	-0.3991	-0.3989	-0.8065	-0.0171	0.0438	0.0475
SCHGROOTTE_z	0.0093	0.0094	-0.0215	0.0378	0.5350	0.5385
Urban	-0.0116	-0.0114	-0.0974	0.0734	0.7918	0.7920
Suburban	-0.0031	-0.0022	-0.0724	0.0678	0.9538	0.9300

```
$random
```

Groups	Name	Std.Dev.	MCMCmedian	MCMCmean	HPD95lower	HPD95upper
1	IDCLASS (Intercept)	0.0000	0.0011	0.0071	0.0000	0.0278
2	Residual	0.8745	0.8767	0.8770	0.8569	0.8991

8.8 Analyse om een curvilineair verband te testen

```

> Dataset$OVERALLREAD_z<-(Dataset$OVERALLREAD-
mean(Dataset$OVERALLREAD,na.rm=TRUE))/sd(Dataset$OVERALLREAD,na.rm=TRUE)
> Dataset$KLGROOTTE_z<-(Dataset$KLGROOTTE-
mean(Dataset$KLGROOTTE,na.rm=TRUE))/sd(Dataset$KLGROOTTE,na.rm=TRUE)
> Data<-na.omit(Dataset[,c("IDCLASS","KLGROOTTE_z","OVERALLREAD_z")])
> Data$Kl2<-(Data$KLGROOTTE_z*Data$KLGROOTTE_z)
> summary(Data)
  IDCLASS      KLGROOTTE_z      OVERALLREAD_z      Kl2
Min.   :   101  Min.   :-4.240e+00  Min.   :-4.306058  Min.   : 0.00327
1st Qu.:  5204  1st Qu.:-5.492e-01  1st Qu.:-0.627317  1st Qu.: 0.09193
Median :  9803  Median :-5.718e-02  Median : 0.026346  Median : 0.46360
Mean   : 41212  Mean   :-1.117e-16  Mean   :-0.001558  Mean   : 0.99977
3rd Qu.:101201 3rd Qu.: 6.809e-01  3rd Qu.: 0.657279  3rd Qu.: 1.37574
Max.   :214502  Max.   : 2.649e+00  Max.   : 4.101924  Max.   :17.97358
> nrow(Data)
[1] 4410
> library(lme4)
> library(languageR)
> Modelcurv<-
lmer(OVERALLREAD_z~KLGROOTTE_z+Kl2+(1|IDCLASS),data=Data,REML=FALSE)
> summary(Modelcurv)
Linear mixed model fit by maximum likelihood
Formula: OVERALLREAD_z ~ KLGROOTTE_z + Kl2 + (1 | IDCLASS)
Data: Data
      AIC      BIC logLik deviance REMLdev
12385 12417  -6188   12375   12393
Random effects:
Groups   Name          Variance Std.Dev.
IDCLASS (Intercept) 0.08028  0.28334
Residual                0.92096  0.95966
Number of obs: 4410, groups: IDCLASS, 233

Fixed effects:
              Estimate Std. Error t value
(Intercept)  0.015616   0.028067   0.556
KLGROOTTE_z -0.001424   0.024292  -0.059
Kl2          -0.018679   0.014309  -1.305

Correlation of Fixed Effects:
      (Intr) KLGROO
KLGROOTTE_z -0.069
Kl2          -0.519  0.326
> pvals.fnc(Modelcurv)
$fixed
              Estimate MCMCmean HPD95lower HPD95upper  pMCMC Pr(>|t|)
(Intercept)  0.0156  0.0158   -0.0349   0.0683 0.5450  0.5780
KLGROOTTE_z -0.0014 -0.0016   -0.0449   0.0446 0.9406  0.9533
Kl2          -0.0187 -0.0188   -0.0457   0.0075 0.1726  0.1918

$random
  Groups      Name Std.Dev. MCMCmedian MCMCmean HPD95lower HPD95upper
1 IDCLASS (Intercept)  0.2833  0.2572  0.2577  0.2193  0.2953
2 Residual                0.9597  0.9632  0.9633  0.9426  0.9832

> ICC<-0.08028/(0.08028+0.92096)
> ICC
[1] 0.08018058

> Data$Voorspeld<-0.015616+(-0.001424)*Data$KLGROOTTE_z+(-0.018679)*Data$Kl2
> plot(Data$KLGROOTTE_z,Data$Voorspeld,main="Voorspelde
leesscores",xlab="Klasgrootte (z-score)",ylab="Voorspelling",ylim=c(-2,2))

```