

## 4 | Het klas- en schoolniveau

### 4.1 Inleiding

Op klas- en schoolniveau wordt, zo blijkt uit het vorige hoofdstuk, meer dan driekwart van de variantie in de beide toetscores gebonden door kenmerken van leerlingen en gezinnen. In dit hoofdstuk wil ik nagaan of er naast effecten van leerling- en gezinskenmerken ook effecten van docenten en scholen (inclusief contextuele kenmerken) zijn. Deze vraag probeer ik op verschillende manieren te beantwoorden. Allereerst wil ik nog een kenmerk op leerlingniveau toevoegen, namelijk het door een leerling gevolgde *schooltype*. Dit kenmerk heb ik in het vorige hoofdstuk niet opgenomen in de analyses, omdat het schooltype niet alleen een indicatie geeft van het prestatieniveau van een leerling. Het schooltype geeft ook de gelegenheid tot leren weer. Naarmate het schooltype hoger is, krijgen leerlingen moeilijkere leerstof aangeboden, waardoor bijvoorbeeld vwo-leerlingen meer gelegenheid tot leren hebben dan overige leerlingen. Ten tweede wil ik nagaan of er *random effecten* van leerling- en gezinskenmerken zijn. Onafhankelijke variabelen op het leerlingniveau kunnen immers op hogere niveaus differentiële effecten hebben. Met multiniveau analyse kan er in de derde plaats onderscheid worden gemaakt tussen *tussenschoolse* en *binnenschoolse regressie-effecten*. Hierdoor kan worden bekeken of leerling- en gezinskenmerken niet alleen op individueel maar ook op klas- of schoolniveau een effect hebben op prestaties.

Tot slot voeg ik *vaste effecten* van variabelen op *hoger niveau* toe. Hierbij besteed ik aandacht aan docent- en schoolkenmerken. Om de effecten van deze kenmerken te verklaren gebruik ik ook op klas- en schoolniveau de structuur van concentrische cirkels. Op deze niveaus bestaat de structuur uit drie cirkels: (C) contextuele kenmerken; (B) schoolkenmerken; (A) docentkenmerken.

De opzet van dit hoofdstuk is als volgt. In paragraaf 4.2 geef ik een overzicht van de invloed van docent- en schoolkenmerken. Op deze manier krijgt de structuur van concentrische cirkels nadere invulling. In paragraaf 4.3 behandel ik de respons, de operationalisatie en de beschrijvende statistieken van de docent- en schoolkenmerken. In de twee daarop volgende paragrafen ga ik in de op de effecten op klas- en schoolniveau. In paragraaf 4.4 voer ik drie stappen uit. Allereerst breid ik het spaarzame model uit hoofdstuk 3 uit met de variabele schooltype. In

de tweede en derde plaats onderzoek ik in die paragraaf of er random effecten en regressie-effecten tussen klassen en scholen zijn. In paragraaf 4.5 kijk ik naar de vaste effecten van docent- en schoolkenmerken.

In dit hoofdstuk is de algemene onderzoeksvraag: *In hoeverre en waarom zijn kenmerken op klas- en schoolniveau van invloed op de prestaties van leerlingen in het derde leerjaar van het voortgezet onderwijs?* Deze vraag kan worden gesplitst in vier delen:

1. In hoeverre en waarom heeft het schooltype effect op de prestaties van leerlingen in het derde leerjaar van het voortgezet onderwijs, gegeven de invloed van leerling- en gezinskenmerken?
2. Zijn de effecten van leerling- en gezinskenmerken vaste of random effecten?
3. Zijn de effecten van leerling- en gezinskenmerken regressie-effecten binnen of tussen klassen en scholen?
4. In hoeverre en waarom hebben docent- en schoolkenmerken effect op de prestaties van leerlingen in het derde leerjaar van voortgezet onderwijs, gegeven de invloed van leerling- en gezinskenmerken en het schooltype?

## 4.2 Belang van docent- en schoolkenmerken

In deze paragraaf bespreek ik de docent- en schoolkenmerken die van invloed kunnen zijn op de prestaties van leerlingen. Voor de keuze van deze kenmerken vormt het werk van Creemers (1991; 1994), Scheerens (1989; 1992) en Scheerens en Bosker (1997) het vertrekpunt. Hierbij maak ik een onderscheid tussen contextuele, school- en docentkenmerken.

Hoe ziet een effectieve school eruit? Wat voor activiteiten onderneemt een effectieve docent? Sinds de jaren zeventig worden deze vragen geregeld gesteld. In de loop van de jaren heeft dit tot talrijke opsommingen van kenmerken van effectieve scholen geleid. In hoofdstuk 1 heb ik de kenmerken die Edmonds (1979) als cruciaal beschouwde naar voren gebracht. Op dit werk zijn in latere jaren talrijke variaties ontstaan, waarbij het vijf-factorenmodel van Edmonds meestal met enkele kenmerken wordt uitgebreid (Brophy & Good, 1986; Sammons, Hillman & Mortimore, 1995; Scheerens & Bosker, 1997; Wilson & Corcoran, 1988). Op basis van dergelijke overzichten, kom ik tot een overzicht van verschillende kenmerken die mogelijk van belang zijn voor prestatieverschillen tussen scholen. Allereerst gaat het hierbij om drie contextuele kenmerken: (1) locatiegrootte; (2) leerlingencompositie; (3) denominatie. In de tweede plaats kunnen verscheidene schoolkenmerken van belang zijn: (4) nadruk op basisvaardigheden; (5) professioneel en onderwijskundig leiderschap; (6) eenheid in visie en doelen; (7) orde-

lijk schoolklimaat; (8) bewaking van leervorderingen; (9) samenwerking tussen school en thuis; (10) aard brugperiode; (11) mentoraatsysteem; (12) docent-leerling ratio; (13) toelatingscriteria. De eerste zes schoolkenmerken komen in verscheidene opsommingen voor. De laatste vier worden minder vaak genoemd. Tot de herhaaldelijk genoemde kenmerken van effectieve docenten behoren: (14) gelegenheid tot leren en effectieve leertijd en (15) gestructureerd onderwijs.

### **Contextuele kenmerken (cirkel C)**

De locatiegrootte, de leerlingencompositie en de denominatie zijn alledrie contextuele schoolkenmerken.<sup>i</sup> Deze kenmerken komen sterk overeen met de structurele kenmerken op leerlingniveau. Het zijn kenmerken die doorgaans als gegeven worden beschouwd. Voor deze schoolkenmerken kan de schoolleiding of het bestuur niet of nauwelijks beleid voeren.

Wat *locatiegrootte* betreft variëren scholen voor voortgezet onderwijs sterk. In het voortgezet onderwijs was in de jaren tachtig en negentig sprake van schaalvergroting (Peschar & Veenstra, 1999). Het aantal scholen ging van 2.472 in 1980 naar 803 in 1995. Die achteruitgang was het sterkst bij categorale scholen. Zo liep het aantal zelfstandige scholen voor voorbereidend beroepsonderwijs in vijftien jaar terug van 835 naar 82. Daarentegen groeide het aantal scholen met alle typen van voortgezet onderwijs, de zogeheten brede scholengemeenschappen. Het meest duidelijk wordt dat door te kijken naar het percentage leerlingen dat op een brede scholengemeenschap zit: in 1990 was dat 5 procent en in 1995 meer dan 50 procent. Bij deze schaalvergroting moet een duidelijke kanttekening worden gemaakt. Wat organisatie betreft zijn scholen de laatste jaren in omvang toegenomen, maar voor leerlingen is daar vaak weinig van te merken, want veel scholen bestaan uit verschillende locaties. Vandaar dat voor een effect op prestaties van leerlingen beter naar de locatiegrootte dan naar de schoolgrootte kan worden gekeken. De verwachting is dat de locatiegrootte negatief samenhangt met de prestaties van leerlingen. Een negatieve samenhang zou het gevolg kunnen zijn van de grootschaligheid op veel scholen, waarbij leerlingen mogelijk een ‘nummer’ worden en verloren gaan in de massa (Stoel, 1980; Hoffer & Coleman, 1990). In zijn column in de Volkskrant geeft Schuyt (1996) het als volgt weer: “Grote scholengemeenschappen verliezen het zicht op de leerlingen en de computer-gestuurde aanwezigheidslijsten geven wel een snelle uitdraai van spijbelende leerlingen, maar daarmee heb je ze nog niet terug in de klas. Persoonlijke binding helpt. (...) Wat in elk geval als kenmerk van grootschaligheid opvalt, is de geringere betrokkenheid van zeer velen, de sfeer van anonimiteit en onpersoonlijkheid, van aangemoedigde ontmoediging.”

Wat de leerlingencompositie betreft is met name een effect van het *percentage allochtone leerlingen* te verwachten. Scholen verschillen sterk in dit opzicht. Op de meeste scholen zitten niet of nauwelijks allochtone leerlingen, maar er zijn ook scholen die bijna volledig ‘zwart’ zijn.<sup>ii</sup> In hoofdstuk 2 heb ik laten zien dat deze segregatie sterk samenhangt met andere kenmerken van de leerlingencompositie. Scholen met veel allochtone leerlingen zijn doorgaans scholen waar ouders gemiddeld genomen een laag opleidingsniveau en leerlingen een laag instroomniveau hebben. Wanneer een leerling bijvoorbeeld op een locatie zit met veel leerlingen uit een bepaald sociaal milieu dan kan dit de keuze van vrienden beïnvloeden, wat van invloed kan zijn op de prestaties van leerlingen. Ook zou het kunnen zijn dat docenten op ‘zwarte’ scholen hun lessen aanpassen en lagere streefniveaus hanteren, omdat veel leerlingen taalproblemen hebben. Een dergelijke aanpassing kan een effect hebben op de prestaties van alle leerlingen.

De invloed van de *denominatie*, de levensbeschouwelijke richting, van de school is in hoofdstuk 1 reeds ter sprake gekomen. Sinds de jaren tachtig wordt er veelvuldig onderzoek gedaan naar de verschillen tussen openbaar en bijzonder onderwijs (Bryk, Lee & Holland, 1993; Coleman & Hoffer, 1987; Coleman, Hoffer & Kilgore, 1982). Deze onderzoekers concluderen dat katholieke scholen in de Verenigde Staten effectiever zijn dan andere scholen in de particuliere sector of scholen in de publieke sector. Dit succes van katholieke scholen kan grotendeels worden verklaard door verschillen in docent- en schoolkenmerken.

Met een aandeel van ongeveer 80 procent nemen bijzondere scholen in het Nederlandse voortgezet onderwijs een dominante positie in. In het bijzonder onderwijs zijn er twee grote religieuze denominaties, rooms-katholiek en protestant, die elk ongeveer een derde van het bestel vormen. Wanneer er rekening wordt gehouden met de effecten van kenmerken van leerlingen en gezinnen, is het de vraag of er prestatieverschillen tussen de denominaties zijn. De overheid bekostigt openbare en bijzondere scholen immers op gelijke wijze en geeft ook sturing aan beide sectoren. Het is dan ook minder gemakkelijk om een algemene uitspraak te doen over het succes van een bepaalde richting in het voortgezet onderwijs (Dijkstra, 1994; 1997).

### **Schoolkenmerken (cirkel B)**

In de bespreking van de invloed van schoolkenmerken komen negen kenmerken aan bod die in de literatuur aan effectieve scholen worden toegeschreven, te weten nadruk op basisvaardigheden, professioneel en onderwijskundig leiderschap, eenheid in visie en doelen, ordelijk schoolklimaat, bewaking van leervorderingen, samenwerking tussen school en thuis, aard brugperiode, mentoraatsysteem, docent-leerling ratio en toelatingscriteria. In tegenstelling tot de contextuele ken-

merken kunnen scholen op deze kenmerken beleid voeren. Dergelijke kenmerken fungeren dan ook mogelijk als inspiratiebron voor scholen die hun kwaliteit willen vergroten (Van Petegem, 1997).

Tot de vijf factoren die volgens Edmonds (1979) de prestaties van leerlingen beïnvloeden behoort de *nadruk op basisvaardigheden*. Voor het Nederlandse basisonderwijs kunnen taal, rekenen en schrijven tot de basisvaardigheden worden gerekend. Voor het voortgezet onderwijs gaat het dan met name om Nederlands, Engels en wiskunde. De betere prestaties van leerlingen op katholieke scholen verklaren Bryk, Lee en Holland (1993) door de nadruk op basisvaardigheden. Katholieke scholen in de Verenigde Staten, zo stellen zij, zijn meestal vrij klein. Door het geringe leerlingenaantal hebben deze scholen minder mogelijkheden om allerlei keuzevakken en richtingen in te voeren. Vrijwel alle leerlingen volgen een academisch curriculum waarin de basisvaardigheden centraal staan, wat gemiddeld genomen een positief effect heeft op hun prestaties. Daarentegen zijn publieke scholen vaak groot, waardoor ze tal van extra vakken en aparte richtingen kunnen aanbieden. Gekscherend vergelijken Powell, Farrar en Cohen (1985, p. 309) publieke scholen met winkelcentra waar leerlingen kunnen bepalen wat ze willen kopen. "The mall's central qualities - variety of offerings, choice among them, and neutrality about their value - have succeeded in holding most teenagers on terms they and their teachers can live with. The will to learn is perceived, in a deceptively sensible formulation, simply as the responsibility of students and their families." In tegenstelling tot de Verenigde Staten beschikt Nederland over een standaardcurriculum voor het voortgezet onderwijs met een grotendeels voorgeschreven lessentabel en een landelijk eindexamen. In het Nederlandse voortgezet onderwijs zijn er dus veel minder verschillen te verwachten in de mate waarin tijd wordt besteed aan de basisvaardigheden. Scheerens en Bosker (1997) stellen dan ook dat dit effectiviteitskenmerk typisch Amerikaans is.

Het professioneel en onderwijskundig leiderschap behoort ook tot de kenmerken die Edmonds in 1979 naar voren bracht. Onder een professioneel en onderwijskundig leiderschap wordt verstaan een directie die de structuur van en de relaties binnen de school bewaakt en die veel tijd en energie steekt in activiteiten om de school te verbeteren (Sammons, Hillman & Mortimore, 1995). Voor de structuur van de school is het belangrijk dat verschillende elementen van de school op elkaar aansluiten. De trefwoorden die Creemers (1991; 1994) hiervoor hanteert zijn consistentie, cohesie, constantie en controle. *Consistentie* houdt in dat de diverse componenten van het onderwijs op elkaar aansluiten, zoals de groeperingvorm, het onderwijsleermateriaal en het docentgedrag. Wanneer een leerling bij verschillende docenten en gedurende de gehele schooltijd op een zelfde manier leskrijgt dan is er sprake van *cohesie* en *constantie*. Door de leerresultaten

en het docentgedrag te evalueren kan de schoolleiding de effectieve instructie *controleren*. Deze trefwoorden sluiten aan bij een begrip dat Rutter e.a. (1979) hanteren. Verschillen tussen scholen hangen samen met *configuraties* van schoolkenmerken en niet zo zeer met specifieke schoolkenmerken. De schoolleiding moet dus nagaan of er succesvolle configuraties tot stand komen. Wat de omgang tussen docenten en directie betreft stellen Rutter e.a. dat het voor docenten belangrijk is om te merken dat ze deel uit maken van het beslissingsproces en dat ze worden betrokken bij beslissingen die voor hen van belang zijn. Het is volgens hen belangrijk dat er met docenten rekening wordt gehouden, maar wanneer nodig moeten de beslissingen hogerop worden genomen. In scholen die goed functioneren, zo vervolgen Rutter e.a., worden de belangrijke beslissingen genomen door de directie en niet door de docenten. Op deze manier schermt de directie het personeel af en vermijdt de leiding dat docenten te veel vergaderen. Wilson en Corcoran (1988) brengen naar voren dat een professionele leiding dient te beschikken over voldoende kennis van onderwijsonderzoek en van de praktijk in de school. Die kennis moet de directie inzetten om de school te verbeteren.

In tegenstelling tot het Angelsaksische onderzoek heeft het Nederlandse onderzoek naar een effect van het onderwijskundig leiderschap in het voortgezet onderwijs weinig empirische bevestiging opgeleverd (Scheerens & Bosker, 1997). Knuver (1993) meent dat de sterke rechtspositie van docenten een reden zou kunnen zijn voor het ontbreken van een effect van het leiderschap. De Nederlandse schoolleider heeft minder invloed op het aannemen of ontslaan van personeel dan hun collega's in andere landen, zoals een rector het treffend verwoordt (Drayer, 1997, p. 23). “Echt personeelsbeleid voeren kun je niet. (...) We zullen komend jaar tien leerkrachten moeten overplaatsen. En dat zijn de jonge mensen, want die zijn het laatst aangekomen. (...) Juist docenten die vitaal zijn, die nog vol *power* zitten, raak ik kwijt. Een man van zesenvijftig die graag weg wil, die is opgebrand, kan niet wegkomen. Absurd. De doorstroming is nihil. Eigen keuze heb ik niet. Als ik al een vacature heb, moet ik eerst wachtgelders aannemen. Zo kun je nooit onderwijskundige vernieuwingen doorvoeren.” Het is dus de vraag of directies van Nederlandse scholen voor voortgezet onderwijs op dit aspect de regie stevig in handen kunnen hebben.

Op scholen waar onder personeel een *eenheid in visie en doelen* leeft, zo komt uit onderzoek van Bryk, Lee & Holland (1993) naar voren, zijn de prestaties van leerlingen beter. Ook Chubb en Moe (1990, p. 91) benadrukken het belang van een gezamenlijke visie en doelen. Op effectieve scholen is volgens hen de collegialiteit onder docenten hoger. “(T)eachers in the best schools are especially likely to agree that their fellow teachers 'can be counted on anytime, anywhere'; that they all 'share beliefs in the central school mission'; and among other things, that the school is 'like

a big family'." In Nederland vermoeden verschillende wetenschappers, zie bijvoorbeeld Meijnen e.a. (1991), het belang van dit kenmerk voor prestaties van leerlingen, maar tot empirisch onderzoek heeft dit nog niet geleid.<sup>iii</sup> Overigens leiden niet alle onderzoeken naar eenheid in visie en doelen tot empirische ondersteuning. Phillips (1997, p. 635) vindt in haar onderzoek geen effect van wat zij communitarisme noemt. Zij noemt het onderzoek naar een kenmerk als eenheid in visie en doelen tijdsgevoelig. "In this postindustrial era of geographic mobility and rapidly changing family structures, many people believe that traditionally communal domains, such as neighborhoods and families, have become less stable and supportive than they once were. Communitarian theories suggest that teachers' disaffection and students' school failure arise from a lack of strong affective ties to both places and people; a void that communally organized schools may fill."

De geschiktheid van de omgeving voor het leren is een schoolkenmerk waaraan onderzoekers al sinds geruime tijd aandacht besteden. In onderzoek naar zowel schooleffectiviteit (Edmonds, 1979; Rutter e.a., 1979; Wilson & Corcoran, 1988) als verschillen tussen openbaar en bijzonder onderwijs (Bryk, Lee & Holland, 1993; Coleman & Hoffer, 1987; Coleman, Hoffer & Kilgore, 1982) komt een *ordelijk schoolklimaat* als belangrijk naar voren. Op de meeste effectieve scholen heerst volgens Sammons, Hillman en Mortimore (1995) voldoende orde om het lesgeven en leren ongestoord door te laten gaan. Een van de middelen die de directie kan aanwenden om het lesgeven zo veel mogelijk door te laten gaan is het invoeren van een invalregeling bij afwezigheid van een docent. Op sommige scholen blijken docenten op hun rooster een of twee uur te hebben staan waarop ze verplicht kunnen worden om in te vallen voor een collega (Veenstra, 1997). De ordelijkheid van de leeromgeving is mogelijk meer van belang voor leerlingen in lagere dan voor leerlingen in hogere vormen van voortgezet onderwijs. Wanneer de regels en afspraken op school niet al te straf zijn, gaan die leerlingen eerder over tot bijvoorbeeld spijbelen (vergelijk Rutter e.a., 1979).

De voortgang van leerlingen in prestaties dient volgens Edmonds (1979) niet alleen op klasniveau maar ook op schoolniveau te worden bewaakt. Om de *leervorderingen te bewaken* kunnen scholen een systeem opzetten waarin ze de voortgang registreren. In een dergelijk systeem houden scholen van iedere leerling gegevens bij over bijvoorbeeld de thuissituatie, de inzet, het gedrag, schoolverzuim, de proefwerkcijfers, de rapporten, de vakkenpakketkeuze en de wensen van vervolgopleiding. Van der Werf, Van der Tuin en Driessen (1995) wijzen erop dat een minderheid van de scholen zo uitgebreid over leerlingen rapporteren. In Scheerens en Bosker (1997) zijn geen aanwijzingen te vinden dat er voor dit kenmerk empirische steun is in onderzoek naar het Nederlandse voortgezet onderwijs.

De empirische onderbouwing van de invloed *samenwerking tussen de school en de ouders* van leerlingen op leerprestaties is volgens Scheerens (1992) vrij stevig. Die steun wordt echter verkregen door onderzoek in het basisonderwijs, waar het geregeld voorkomt dat ouders helpen bij bijvoorbeeld het lezen of buitenschoolse activiteiten. Sammons, Hillman en Mortimore (1995) noemen in hun overzichtsartikel dan ook alleen maar studies voor het basisonderwijs. In het voortgezet onderwijs blijken deze contacten tamelijk formeel (Veenstra, 1997). Een paar keer per jaar organiseert de school spreekuren voor ouders. Verder zijn er jaarlijks enkele thema-avonden om ouders voor te lichten over de keuze van een vakkenpakketten of vervolgopleidingen. Voor veel ouders zijn dat de weinige momenten waarop ze contact hebben met de school. Voor het voortgezet onderwijs is er voor het effect van dit schoolkenmerk weinig empirisch houvast (vergelijk Ho & Willms, 1996).

Wanneer er in het voortgezet onderwijs wordt gedifferentieerd, gaat het doorgaans om niveaudifferentiatie tussen klassen. Tussen scholen kan de mate van differentiatie, *de aard van de brugperiode*, variëren. Op sommige scholen worden leerlingen op basis van hun prestatieniveau in zo homogeen mogelijke prestatiegroepen geplaatst. Op andere scholen worden klassen juist heterogeen samengesteld. Oakes (1985) heeft een beschrijvend onderzoek gedaan naar de invloed van heterogeen en homogeen groeperen. In haar conclusie houdt zij een pleidooi voor heterogeen groeperen met als argument dat geen enkele groep van leerlingen volgens haar profiteert van het krijgen van les in een homogene groep. Daarentegen vinden Dar en Resh (1986) in hun onderzoek naar het voortgezet onderwijs in Israël wel een effect van de groeperingvorm. Ook zij zijn voorstander van heterogeen groeperen, want bij die vorm van groeperen is de achteruitgang in prestaties van goede leerlingen kleiner dan de vooruitgang van zwakke leerlingen.<sup>iv</sup> Gamoran (1992, p. 826) gaat in op verschillende aspecten van homogeen en heterogeen groeperen. Zijn conclusie is dat naarmate de klassensamenstelling van een school voor langere tijd vast ligt, de prestatieverschillen tussen klassen groter zijn en de leerprestaties voor alle klassen te zamen gemiddeld genomen lager zijn. “(T)he results for track immobility suggest that inflexible tracking systems have greater inequality along with lower average achievement, presumably reflecting especially poor conditions in the nonacademic tracks of such schools.” In Nederland heeft De Vries (1992) onderzoek gedaan naar de klassensamenstelling van scholen voor voortgezet onderwijs. Haar conclusie sluit aan bij die van Oakes. Op termijn blijkt er geen invloed van de klassensamenstelling op de onderwijspositie van leerlingen. Toch prefereert zij de heterogene groeperingvorm.

Scholen blijken behoorlijk te verschillen in de mate waarin ze aandacht besteden aan *leerlingbegeleiding* in het algemeen en het *mentoraatsysteem* in het bijzon-



der (Veenstra, 1998). Op minder dan 10 procent van de scholen blijkt het mentoraatsysteem uitgebreid opgezet, wat inhoudt dat zowel het aantal mentorlessen als de frequentie van overleg tussen mentoren hoog is. Op deze scholen hebben leerlingen in de eerste leerjaren in ieder geval één en vaak twee of drie mentoruren per week. Niet alleen voor de leerlingen maar ook voor de docenten is het mentoraat in dat systeem intensief, want op deze scholen overleggen mentoren doorgaans wekelijks. Daartegenover staat een grote groep van scholen, bijna viertiende van de scholen voor voortgezet onderwijs, waar het mentoraatsysteem beperkt is qua omvang. Op deze scholen hebben leerlingen meestal alleen in de brugklas een mentoruur en overleggen mentoren eens in de drie maanden. De overige scholen nemen een tussenpositie in. Naar de invloed van het mentoraatsysteem op prestaties van leerlingen is weinig onderzoek gedaan. Er zijn slechts drie studies bekend. Van Marwijk Kooy-von Baumhauer (1984) stelt dat scholen die een actief beleid voor begeleiding van leerlingen voeren goede examenresultaten behalen. Tesser (1986) vindt op het vwo een positieve correlatie tussen de hoeveelheid leerlingbegeleiding en het bereikte onderwijsniveau na vijf jaar in het voortgezet onderwijs. Van der Tuin (1993) laat voor scholengemeenschappen zien dat er een positieve bijdrage uitgaat van leerlingbegeleiding op de prestaties van leerlingen. Die positieve bijdrage zou kunnen komen doordat scholen voortijdig inspelen op mogelijke problemen van leerlingen. Tijdens de mentorlessen kan bijvoorbeeld aandacht worden besteed aan studie- en sociale vaardigheden.

Een grote school hoeft niet kil en onpersoonlijk te zijn. Sizer (1992) doet hiertoe de aanbeveling om de *docent-leerling ratio* te verkleinen. Hij stelt dat een verlaging van deze ratio mogelijk is zonder dat dit extra kosten met zich mee brengt. Zijn voorstel is om docenten te groeperen in clusters, en om elke docent meer dan een vak te laten geven. Vertaald naar de Nederlandse situatie, zouden geschiedenis, aardrijkskunde, maatschappijleer en economie te zamen een vak kunnen vormen, bijvoorbeeld onder de noemer 'mens en maatschappij'. Natuurkunde, scheikunde en biologie kunnen eveneens tot een geheel worden gesmeed. Ook bij de talen zou meer kunnen worden gecombineerd. Vaardigheden die bij alle talen van belang zijn, zoals spreekvaardigheid, zouden op een standaardmanier kunnen worden geoefend. Het ideaal van Sizer is dat docenten aan niet meer dan tachtig leerlingen lesgeven, zodat een persoonlijke band tussen docenten en leerlingen ook in het voortgezet onderwijs mogelijk wordt. Scholen waar docenten aan een dergelijk aantal leerlingen lesgeven, zijn er in Nederland nauwelijks. Het enige reguliere schooltype waar een dergelijke omvang wordt benaderd is het ivbo. Daar is de grootte van een klas maximaal vijftien leerlingen. In het ivbo geven docenten vaak verschillende vakken aan dezelfde groep leerlingen. De docent-leerlingratio is er dan ook laag. Wellicht dat een lage ratio

ook van belang is voor leerlingen van andere schooltype en dan met name vbo-leerlingen.

Door de selectiemechanismen die katholieke scholen in de Verenigde Staten toepassen behalen ze, zo beargumenteert Witte (1992), betere prestaties dan andere scholen. Zij hanteren volgens hem strengere *toelatingscriteria* dan publieke scholen en weren daardoor probleemrijke leerlingen. Katholieke scholen eisen van leerlingen niet alleen dat ze het niveau van de school aankunnen, maar ook dat de leerlingen binnen het schoolklimaat passen. In Nederland hanteren scholen als toelatingseis het advies van de basisschool. Sommige scholen willen daarnaast ook het prestatieniveau van een leerling weten. Mogelijk dat scholen op die manier strenger selecteren en daardoor een makkelijkere leerlingenpopulatie krijgen. Het prestatieniveau van een leerling en het advies van de basisschool hangen weliswaar sterk samen, maar deze samenhang is niet perfect. Zo krijgen allochtone leerlingen en leerlingen uit hogere sociale milieus hogere adviezen dan hun score op de Cito-toets aangeeft (Dronkers, 1997a).

### **Docentkenmerken (cirkel A)**

Tot slot bespreek ik twee kenmerken van effectieve docenten, namelijk gelegenheid tot leren en effectieve leertijd en gestructureerd onderwijs.<sup>v</sup>

De *gelegenheid tot leren* is een term die Carroll (1963) heeft ingevoerd. In een klassiek artikel werkt hij een model van de mate van leervorderingen uit. Dat model bestaat uit vijf elementen die zijn verdeeld over twee dimensies. De eerste dimensie is de daadwerkelijk bestede leertijd, die volgens Carroll gelijk is aan het minimum van de hoeveelheid tijd die nodig is om een taak onder ideale omstandigheden te beheersen (het leertempo van een leerling), de toegestane leertijd (de gelegenheid tot leren) of de tijd die leerlingen bereid zijn actief aan leren te besteden (het doorzettingsvermogen van een leerling). De benodigde leertijd vormt de tweede dimensie en is, zo vervolgt hij, gelijk aan het leertempo van leerlingen onder controle van de kwaliteit van de instructie en hun begripsvermogen. De mate van vooruitgang in prestaties is gelijk aan de quotiënt van de daadwerkelijk bestede en de benodigde leertijd (zie voor de formule: Carroll, 1962). Wanneer deze quotiënt groter is dan één, dan presteren leerlingen boven hun niveau en wanneer de quotiënt kleiner is dan één, dan presteren leerlingen onder hun niveau.

Van de genoemde kenmerken van het model zijn de gelegenheid tot leren, de kwaliteit van de instructie en het doorzettingsvermogen van de leerling mogelijk te beïnvloeden (Carroll, 1963). Volgens Carroll (1989) is het vooral belangrijk om er achter te komen hoe de gelegenheid tot leren kan worden vergroot. Om de toegestane leertijd te bepalen is het begrip *effectieve leertijd* van belang. De effectieve leertijd is de tijd die per lesuur resteert als onderbrekingen door orde- of

concentratieproblemen niet meetellen (vergelijk Leseman, Sijssling & De Vries, 1992). De effectieve leertijd is afhankelijk van (1) de ingeroosterde lestijd, (2) de orde en rust in de klas, (3) de prestatiemotivatie van leerlingen, (4) de taakgerichtheid en het planningsgedrag van docenten en (5) de extra leertijd door opgave en controle van huiswerk.

Om de effectieve leertijd te vergroten is het volgens Creemers (1991) belangrijk dat docenten leerlingen bij de les houden door het geven van *gestructureerd onderwijs*. Leerlingen leren volgens Veenman (1992) meer, geordender en efficiënter als de docent in het onderwijs die zaken aan de orde stelt die de leerling nog niet begrijpt of beheerst en die liggen in, wat de leerpsycholoog Vygotsky noemt, hun zone van naaste ontwikkeling. Het onderwijs moet ontwikkelend en niet volgend zijn. Hierbij is het volgens Brophy en Good (1986, p. 361) van belang dat docenten de activiteiten van leerlingen sturen. "Students achieve more in classes where they spend most of their time being taught or supervised by their teachers rather than working on their own." Leerlingen krijgen dan minder mogelijkheden om te verslapen tijdens de les. Gestructureerd onderwijs, dat ook wel bekend staat als directe instructie of sturend onderwijs, bestaat volgens Doyle (1985) uit een achttal aspecten.<sup>vi</sup> Allereerst is het van belang dat een docent de leerdoelen van een les duidelijk formuleert en precies bedenkt wat hij de leerlingen wil leren. Vervolgens is het nodig dat hij de leerstof opsplijst in voor de leerling overzienbare eenheden. De docent moet met kleine stappen door de stof gaan, waarbij hij deze stappen in een weloverwogen volgorde aanbiedt. In de derde plaats moet hij bij de aanvang van de les de leerlingen duidelijk laten weten wat de doelstellingen zijn. Daarna moet de docent eerder geleerde stof herhalen, opdat hij de benodigde voorkennis van leerlingen ophaalt. Veenman (1992) stelt dat een docent die bij het begin van zijn les duidelijk aangeeft wat de doelen zijn en die een samenvatting geeft van de voorafgaande leerstof, de leerlingen helpt om betekenis te geven aan de nieuwe leerstof. In de vijfde plaats moet de docent duidelijk uitleg geven. Hij moet heldere taal gebruiken, concrete voorbeelden geven, hoofdzaken samenvatten en geregeld vragen stellen. Als zesde punt geldt dat leerlingen voldoende tijd moeten krijgen om de stof te oefenen. Deze oefening dient eerst onder begeleiding en daarna zelfstandig te gebeuren. Het is hierbij van belang dat leerlingen weten dat docenten hun werk controleren en dat ze deze controle zo spoedig mogelijk uitvoeren. Op die manier merken leerlingen gelijk de invloed van het nakijken door de docent. Bij deze terugkoppeling moeten ze ook aangeven waarom iets goed of fout is. Een volgend aspect is de frequente evaluatie van de voortgang van leerlingen door de docent. De kennis, inzicht en vaardigheden van leerlingen moeten geregeld worden getoetst, omdat een docent er zo achter kan komen of de leerlingen de stof beheersen. Wanneer leerlingen de leerdoelen niet bereiken, moet hij daar daadwerkelijk iets aan

doen door bijvoorbeeld zijn lessen bij te sturen. Het laatste kenmerk van gestructureerd onderwijs is de periodieke herhaling van de leerstof met als doel dat de leerlingen de leerstof automatisch gaan beheersen. De kern van deze instructiemethode is volgens Rosenshine en Stevens (1986) de opeenvolging van de uitleg van de docent, het oefenen door de leerlingen en de corrigerende terugkoppeling door de docent. Op deze manier geven docenten richting aan het leerproces van leerlingen.

### **Besluit**

De empirische steun voor de invloed van schoolkenmerken op prestaties van leerlingen is doorgaans zwak of afwezig. Het is dan ook twijfelachtig of de genoemde kenmerken een effect hebben. Verscheidene kenmerken zijn eerder van kracht in het basis- dan in het voortgezet onderwijs. Ook geldt dat enkele in het buitenland gevonden verbanden, zoals de nadruk op basisvaardigheden en het professioneel en onderwijskundig leiderschap, in Nederland nauwelijks of niet worden teruggevonden. Verder geldt dat bij de empirische steun voor verschillende kenmerk kanttekeningen zijn te plaatsen. Wanneer onderzoekers een adequaat design gebruiken, een multiniveau analyse naar verschillen in netto prestaties, dan worden er minder sterke effecten gevonden dan in onderzoek op één niveau of in onderzoek naar verschillen in bruto prestaties (Scheerens & Bosker, 1997).

Voor de invloed van klassenkenmerken op prestaties van leerlingen is de empirische steun vrij stevig (Creemers, 1991). Effecten van gelegenheid tot leren en gestructureerd onderwijs worden geregeld aangetoond. De meeste van deze effecten zijn echter nog niet in onderzoek naar het voortgezet onderwijs in Nederland gevonden.

De verwachting is dat in de onderbouw van het voortgezet onderwijs in Nederland geldt dat de netto prestaties van scholen hoger zijn, wanneer: (1) de locatiegrootte kleiner is; (2) het percentage allochtone leerlingen lager is; (3) de eenheid in visie en doelen groter is; (4) het schoolklimaat ordelijker is; (5) het mentoraatsysteem intensiever is; (6) de docent-leerling ratio kleiner is; (7) de toelatingscriteria strenger zijn; (8) de gelegenheid tot leren groter is; (9) het onderwijs meer gestructureerd is. Verder wordt getoetst of er verschillen zijn in netto prestaties tussen scholen (10) van verschillende denominaties; (11) met verschillende vormen van brugperiode. In Nederland is het onduidelijk in welke richting deze effecten zullen optreden.

Deze effecten hoeven niet voor alle schooltypen hetzelfde te zijn. Voor het effect van bijvoorbeeld de ordelijkheid van het schoolklimaat of de docent-leerling ratio zou kunnen gelden dat het effect vooral voor het beroepsonderwijs opgang

doet. Over dergelijke verschillen in effecten tussen hogere en lagere vormen van voortgezet onderwijs is uit de literatuur overigens weinig bekend.

Van sommige kenmerken worden in het voortgezet onderwijs geen effecten verwacht, te weten de nadruk op basisvaardigheden, het professioneel en onderwijskundig leiderschap, de bewaking van leervorderingen en de samenwerking tussen school en thuis.<sup>vii</sup>

### 4.3 Meten van docent- en schoolkenmerken

In deze paragraaf ga ik allereerst in op de respons, de operationalisatie en ontbrekende waarden van kenmerken op klas- en schoolniveau. Vervolgens geef ik de beschrijvende statistieken.

#### Respons

In dit onderzoek maak ik gebruik van vragenlijsten voor de directie en voor docenten. Van de directieleden zijn voor het schooljaar 1993-'94 de gegevens beschikbaar van 142 van de 149 scholen (95 procent) die ik mijn analyses heb opgenomen. Voor 1995-'96 zijn de gegevens van 111 van de 149 scholen (74 procent) beschikbaar. Van 471 docenten van 124 scholen zijn gegevens beschikbaar. Aangezien er in dit onderzoek 453 klassen zijn en docenten voor Nederlands en wiskunde zijn ondervraagd, hadden er 906 docenten een vragenlijst kunnen inleveren. De respons bedroeg dus 52 procent. In 1995-'96 is informatie verkregen van 522 docenten van 135 scholen. Op docentniveau is de respons 58 procent.

#### Operationalisatie

De contextuele schoolkenmerken zijn afkomstig uit de directievragenlijst. Bij de *denominatie* onderscheid ik zeven typen: (1) openbaar; (2) algemeen bijzonder; (3) interconfessioneel (inclusief samenwerkingsscholen); (4) protestants-christelijk; (5) rooms-katholiek; (6) reformatorisch; (7) gereformeerd-vrijgemaakt. Verder neem ik het *percentage allochtone leerlingen*<sup>viii</sup> en de natuurlijke logaritme van de *locatiegrootte* in de analyses mee. In onderwijsonderzoek is het tamelijk gebruikelijk om de  $\log_{10}$  van de locatiegrootte te nemen, omdat het verband tussen de locatiegrootte en de netto prestaties zo het beste wordt weergegeven.

Tenzij anders vermeld zijn de schoolkenmerken ook afkomstig uit de directievragenlijst uit het schooljaar 1993-'94. Om de *eenheid in visie en doelen* te meten gebruik ik gegevens uit de docentenvragenlijst uit 1995-'96. Aan docenten zijn twaalf vragen gesteld over de eenheid binnen de school, zie Bijlage 6. Van deze vragen heb ik de standaarddeviatie gebruikt om de spreiding in de antwoorden van do-

centen te meten. Aan de docenten is in 1995-'96 gevraagd hoe lang ze al op hun school werken. Met die vraag kan de teamstabiliteit worden bepaald. De teamstabiliteit zegt mogelijk ook iets over de eenheid binnen een school. Voor de operationalisatie van een *ordelijke leeromgeving* gebruik ik gegevens over afspraken tussen de directie en docenten, zie Bijlage 6. De vragen vormen te zamen een schaal met een Cronbachs alfa van 0,74. Ook gebruik ik gegevens over de actie die de directie onderneemt als een les onverwacht uitvalt. Hierbij onderscheid ik drie categorieën: (1) huiswerk maken; (2) invalregeling; (3) lesuur vrij of een andere maatregel. Bij *de aard van de brugperiode* onderscheid ik drie vormen, te weten homogeen, dakpansgewijs en heterogeen. Homogeen is een klas, wanneer er alleen leerlingen van één schooltype bijeenzitten. Dakpansgewijs is een klas, wanneer er leerlingen van twee schooltypen bijeenzitten, bijvoorbeeld een vbo/mavo-klas, en heterogeen is een klas bij leerlingen van drie of meer schooltypen. Vervolgens is per school berekend hoeveel leerlingen er in elk van de groeperingvormen zitten. Sommige scholen hebben alle leerlingen op een zelfde wijze gegroepeerd. Andere scholen hanteren verschillende vormen naast elkaar. Deze scholen krijgen dan een score op meer dan één groeperingvorm. Bij het *mentoraatsysteem* (gemeten in 1995-'96) onderscheid ik vier categorieën: (1) smal-verticaal; (2) smal-horizontaal; (3) breed-verticaal; (4) breed-horizontaal. Scholen met een uitgebreid mentoraatsysteem noem ik breed-horizontaal. Dit zijn scholen waar in de eerste vier leerjaren gemiddeld meer dan één mentoruur per jaar staat ingeroosterd en waar mentoren vrijwel wekelijks vergaderen. Scholen met een beperkt mentoraatsysteem noem ik smal-verticaal (Veenstra, 1998). Op deze scholen hebben leerlingen alleen in de brugklas of helemaal geen mentoruur. Mentoren vergaderen op deze scholen uit eens per kwartaal of minder vaak. In de andere twee categorieën scoren scholen hoog op het ene en laag op het andere kenmerk. Smal-horizontale scholen zijn scholen waar een mentoruur niet al te vaak voorkomt, hoewel een mentorenoverleg maandelijks op het rooster staat. Het omgekeerde geldt voor breed-verticale scholen. Op deze scholen komt in de meeste leerjaren een mentoruur voor, maar mentoren overleggen gemiddeld niet veel vaker dan eens in de drie maanden. De *docent-leerling ratio* heb ik berekend door met gegevens uit 1995-'96 voor docenten na te gaan aan hoeveel leerlingen zij lesgeven en dat totaal te delen door het aantal uren dat ze wekelijks lesgeven. Als een docent met 24 lessen aan twaalf klassen van 25 leerlingen lesgeeft, dan is de ratio:  $12 * 25 / 24 = 12,5$ . Vervolgens heb ik per school het gemiddelde van deze ratio's berekend. Behalve het advies van de basisschool kunnen scholen het prestatieniveau van leerlingen als *extra toelatingscriterium* hanteren. Deze variabele is een dichotome variabele. Scholen hanteren dit extra criterium wel of niet.

Voor de operationalisatie van de docentkenmerken gebruik ik gegevens uit de vragenlijsten voor de docenten Nederlands en wiskunde die leerlingen in de brugklas hadden (afgenomen in schooljaar 1993-'94). Voor de operationalisatie van de *effectieve leertijd* gebruik ik een vraag over de tijdsindeling van docenten. De *effectieve leertijd* is het percentage van de totale lestijd dat aan instructie (uitleg van nieuwe leerstof en onder begeleiding oefenen van nieuwe stof), verwerking (individueel of in groepjes laten werken aan opdrachten of huiswerk) en controle (bespreken of overhoren van huiswerk of in de klas gemaakte opdrachten), maar niet aan ordehandhaving wordt besteed. Om na te gaan of docenten *gestructureerd onderwijst* geven heb ik geprobeerd verschillende vragen te combineren. Een docent geeft gestructureerd les als de les bestaat uit de opeenvolging van uitleg door de docent, oefenen door leerlingen en corrigerende terugkoppeling door de docent. In de docentenvragenlijsten zijn er verschillende schalen voor de verschillende onderdelen van lesgeven. In Bijlage 6 worden deze schalen vermeld. De schalen blijken betrouwbaar. De betrouwbaarheid is voor alle schalen hoger dan 0,80. De schalen blijken onderling ook vrij sterk te correleren. De correlaties variëren van 0,40 tot 0,77. Uit een nadere inspectie van de data valt echter de nodige twijfel bij deze docentgegevens te plaatsen, zie Tabel 4.1.

*Tabel 4.1: Enkele gegevens uit het bestand met docentkenmerken*

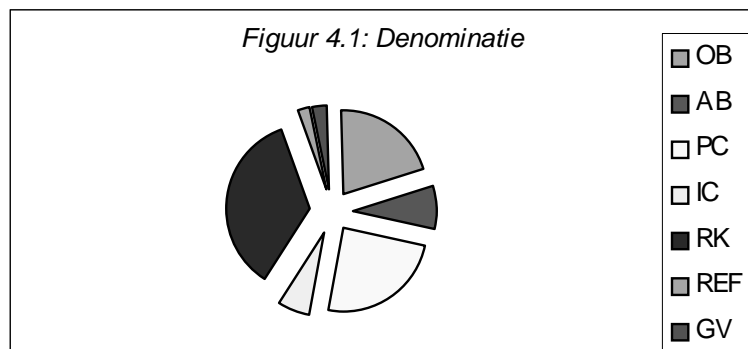
NUMMER	VR36A	VR36B	VR36C	VR36D	VR36E	VR36F
1	100	100	100	100	100	100
2	75	50	20	75	20	25
3	50	50	10	10	25	5
4	20	999	999	60	20	999
5	90	30	50	50	20	30
6	999	999	999	999	999	999
7	30	20	15	15	10	10
8	999	999	999	999	999	999
9	10	10	30	20	30	20
10	20	20	10	10	5	10

In Tabel 4.1 zijn voor tien docenten de antwoorden op de schaal 'Duidelijk en gestructureerd uitleggen' weergegeven. Het blijkt dat docenten de vragen op verschillende manieren benaderen. Sommige docenten hebben geen zin om deze vragen in te vullen en hebben talrijke ontbrekende waarden. In de tabel worden ontbrekende waarden met het getal 999 aangeduid. Docenten 6 en 8 hebben geen enkele vraag beantwoord. De andere docenten hebben één of meer vragen beant-

woord. Docent 1 legt naar eigen zeggen erg duidelijk en gestructureerd uit. Op alle vragen is het antwoord de maximale score van 100. Daartegenover staat docent 10 wiens antwoorden variëren van 5 tot 20. Tussen de scores van deze twee docenten in, bevinden zich de scores van bijvoorbeeld docent 2. Op sommige vragen scoort deze docent hoog en op andere vragen laag. Los van de ontbrekende waarden is er tot dusverre niet zo veel aan de hand met deze gegevens. Het probleem is echter dat voor sommige docenten de schaal van 0 tot 100 niet per vraag geldt, maar voor alle vragen gezamenlijk. De somscores van de zes vragen is voor docenten 4 en 7 gelijk aan 100. Voor docenten 9 en 10 is onduidelijk welke schaal ze hebben gehanteerd. Hebben ze per vraag een schaal van 0 tot 100 gehanteerd en scoren ze dus laag wat duidelijk en gestructureerd uitleggen betreft? Of hebben ze ook een schaal van 0 tot 100 voor alle vragen gezamenlijk gehanteerd en hebben ze een rekenfout gemaakt? Deze laatste mogelijkheid is reëel, omdat de omvang van de docentenvragenlijst zonder meer gevolgen moet hebben gehad voor de nauwkeurigheid in de antwoorden van docenten. De slotsom is dus dat de data betrouwbaar, maar niet valide zijn. Vandaar dat in het vervolg van dit onderzoek geen gebruik wordt gemaakt van docentgegevens. De vraag naar effecten van docentkenmerken kan daarom niet worden beantwoord.

### Ontbrekende waarden

Ontbrekende waarden komen in de vragenlijsten voor directieleden voor. Deze gegevens heb ik steeds zo goed mogelijk proberen aan te vullen. Een contextueel kenmerk als de denominatie van de school heb ik *opgezocht*. Andere gegevens heb ik vervangen met zogeheten *regressie-imputatie* (Huisman, 1999). Om een ontbrekende waarde bij het percentage allochtone leerlingen in het schooljaar 1993-'94 te vervangen, heb ik een regressieanalyse van hetzelfde percentage voor het schooljaar 1995-'96 op het percentage van twee jaar eerder verricht. De ontbrekende waarden heb ik vervolgens vervangen door de waarde die door de regressievergelijking wordt voorspeld. Als deze methode niet kon worden gehanteerd, heb ik gekozen voor *gemiddelde-score-imputatie*.





### Beschrijvende statistieken

Van de 149 scholen uit dit onderzoek is 21 procent openbaar (OB), 24 procent protestants-christelijk (PC) en 37 procent katholiek (RK). De overige scholen maken deel uit van een van de vier andere richtingen die ik onderscheid (AB: Algemeen bijzonder; IC: Interconfessioneel; REF: Reformatorisch; GV: Gerformeerd-vrijgemaakt). In Figuur 4.1 zijn de denominaties weergegeven met de richting van de wijzers van een klok.

Scholen variëren sterk wat locatiegrootte betreft. De gemiddelde grootte is 697 leerlingen, maar het bereik loopt van 124 tot 2.380 leerlingen. Het percentage allochtone leerlingen verschilt ook sterk van school tot school. Het gemiddelde is 7 procent, met als bereik 0 tot 87 procent (zie Tabel 4.2).

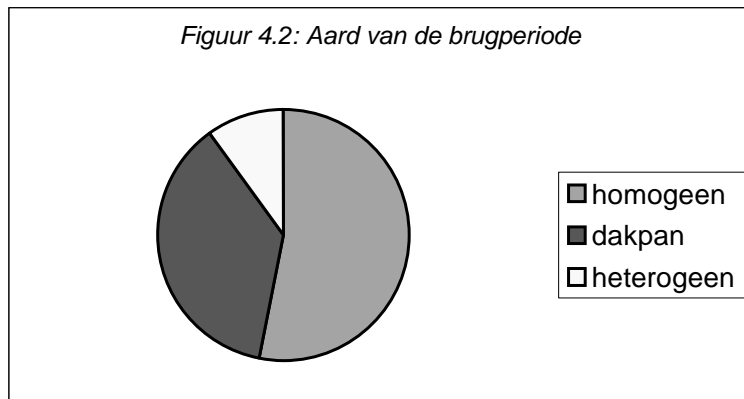
Tabel 4.2: Overzicht van schoolkenmerken

	GEM.	SD	MIN.	MAX.	N
<b>CIRKEL C: CONTEXTUELE KENMERKEN</b>					
Locatiegrootte	697	444	124	2.380	149
Percentage allochtone leerlingen	0,07	0,11	0,00	0,87	149
<b>CIRKEL B: SCHOOLKENMERKEN</b>					
Ordelijk schoolklimaat					
- afspraken over orderegels (4 items)	14,5	3,0	6,0	20,0	149
Eenheid in visie en doelen					
- standaarddeviatie (12 items)	9,8	0,9	5,7	12,4	149
Teamstabiliteit	15,0	3,6	7,2	25,3	149
Docent-leerling ratio	8,8	2,1	4,2	13,6	149

Directieleden scoren op de vier vragen over afspraken over orderegels gemiddeld 14,5, zie Tabel 4.2. Op scholen zijn er doorgaans wel afspraken vastgelegd over orderegels en docenten kunnen daarvan gemiddeld genomen alleen afwijken na overleg met de directie. Wat de maatregelen bij lesuitval betreft geven 83 van de 149 scholen aan te beschikken over een invalregeling, 21 scholen laten leerlingen tijdens het vrijgekomen uur huiswerk maken en 45 scholen laten dat lesuur vervallen.

De spreiding in de antwoorden van docenten over de eenheid binnen de school is 9,8. Op scholen waar deze spreiding gering is, zijn docenten het meer met elkaar eens en op scholen waar docenten het meer met elkaar oneens zijn is de spreiding groter. De Cronbachs alfa van de schaal is 0,74, zie Bijlage 6. Het aantal jaren dat docenten van een school gemiddeld genomen werkzaam zijn, varieert van 7,2 tot 25,3 jaar. Het gemiddelde bedraagt 15,0.

Wat de aard van de brugperiode betreft kunnen scholen verschillende scores tegelijkertijd hebben. Sommige scholen stellen alle klassen op een zelfde wijze samen. Het bereik op deze variabelen loopt dan ook van komt niet voor (0 procent) tot komt in alle klassen voor (100 procent). Gemiddeld zit 53 procent van de leerlingen in een homogene, 37 procent in een dakpansgewijze en 10 procent in een heterogene brugklas, zie Figuur 4.2.<sup>ix</sup>



Het mentoraatsysteem is op 54 scholen smal-verticaal, op 46 scholen smal-horizontaal, op 34 scholen breed-verticaal en op 15 scholen breed-horizontaal. De docent-leerling ratio is 8,8. De scores op deze variabele lopen uiteen van 4,2 tot 13,6. Uitgaande van 28 lessen zijn er dus scholen waar docenten gemiddeld genomen aan ongeveer 120 leerlingen lesgeven. Dit minimum is nog hoger dan wat Sizer (1992) ideaal acht, namelijk 80 leerlingen per docent. Er zijn echter ook scholen waar docenten aan ruim drie keer zoveel leerlingen lesgeven. Bij 28 lessen geeft de gemiddelde docent les aan 250 leerlingen. Tot slot wordt het prestatieniveau op 78 procent van de scholen als toelatingscriterium gehanteerd.

#### 4.4 Controles op klas- en schoolniveau

In deze paragraaf beantwoord ik de eerste drie deelvragen. Als eerste ga ik na in hoeverre het schooltype extra variantie verklaart in het spaarzame model met leerling- en gezinskenmerken uit het vorige hoofdstuk. Als tweede onderzoek ik of de effecten van leerling- en gezinskenmerken voor alle leerlingen gelijk zijn of dat er sprake is van differentiële effecten. Hierbij wordt gekeken naar random hellingen en intercept-helling covarianties. Als derde ga ik na of er regressie-effecten tussen klassen en scholen zijn.

#### 4.4.1 Belang van het schooltype

Door de toevoeging van het door een leerling gevolgde schooltype van leerlingen in leerjaar drie wordt extra informatie over leerlingen in de analyses opgenomen. Om scholen eerlijk met elkaar te kunnen vergelijken, is het van belang om zoveel mogelijk met verschillen op individueel niveau rekening te houden.<sup>x</sup> In hoofdstuk 3 heb ik talrijke leerling- en gezinskenmerken opgenomen, waaronder de eerdere prestaties van leerlingen. Die prestaties uit leerjaar één verklaren de prestaties uit leerjaar drie uiteraard goed. Naarmate brugklassers beter presteerden op de eerdere prestatietoetsen, scoren ze twee jaar later gemiddeld genomen hoger op Nederlands of wiskunde. Voor de wiskundeprestaties is de rekentoets de beste voorspeller. De verklaring van de wiskundeprestaties verbetert als de scores op de twee andere toetsen ook in de analyse worden opgenomen. Hoe meer er over het beginniveau van een leerling bekend is, hoe beter de verklaring. Door de opname van het schooltype dat leerlingen in het derde leerjaar volgen wordt de verklaring van hun scores aan het einde van dat leerjaar op Nederlands en wiskunde waarschijnlijk verbeterd.

Het schooltype is een ordinale variabele met zeven categorieën: (1) ivbo; (2) vbo; (3) vbo/mavo; (4) mavo; (5) havo; (6) havo/vwo; (7) vwo. Om het effect van schooltype weer te geven, is een model opgesteld met zes dummy's. In dit model vormen mavo-leerlingen de referentiegroep. De effecten van de overige schooltypen worden dus tegen die van mavo-leerlingen afgezet. Hierbij is de verwachting dat voor leerlingen op een schooltype hoger dan mavo geldt dat de regressiecoëfficiënt positief is. Voor leerlingen op een schooltype lager dan mavo is de verwachting dat de regressiecoëfficiënt negatief is. Verder is de verwachting dat de regressiecoëfficiënten dezelfde ordening vertonen als het schooltype. Het effect voor bijvoorbeeld vwo-leerlingen zal dan het meest positief zijn en het effect voor ivbo-leerlingen het meest negatief.

Tabel 4.3: Multiniveau analyse. Model met toevoeging van schooltype

	NEDERLANDS	WISKUNDE
Etniciteit (1=allochtoon)	-0,113 (0,026) ***	-0,113 (0,026) ***
Opleidingsniveau vader	0,011 (0,007)	0,011 (0,007)
Opleidingsniveau moeder	0,010 (0,006)	0,010 (0,006)
Aantal ouders (1=twee ouders)	<b>0,004 (0,024)</b>	<b>0,049 (0,020) *</b>
Sekse (1=jongens)	<b>-0,153 (0,019) ***</b>	<b>0,146 (0,016) ***</b>
Géén oudervragenlijst	-0,018 (0,018)	-0,018 (0,018)
Autoritatieve opvoedingsstijl	0,030 (0,006) ***	0,030 (0,006) ***
Praten met ouders	<b>0,006 (0,008)</b>	<b>-0,025 (0,007) ***</b>
Huiswerkcontrole	<b>-0,047 (0,008) ***</b>	<b>-0,020 (0,007) **</b>
Op straat met vrienden	-0,029 (0,006) ***	-0,029 (0,006) ***
Lezen en hobby's	<b>0,058 (0,009) ***</b>	<b>0,010 (0,008)</b>
Bijbaan	-0,010 (0,006)	-0,010 (0,006)
Uitgaan	<b>0,007 (0,009)</b>	<b>-0,028 (0,008) ***</b>
Zelfbeeld	0,040 (0,006) ***	0,040 (0,006) ***
Mening over klasgenoten	<b>0,029 (0,009) **</b>	<b>-0,005 (0,007)</b>
Mening over docenten en regels	0,033 (0,006) ***	0,033 (0,006) ***
Aantal dagen huiswerk	0,016 (0,007) *	0,016 (0,007) *
Non-verbale intelligentie	<b>0,015 (0,010)</b>	<b>0,092 (0,008) ***</b>
Prestatie taal	<b>0,146 (0,015) ***</b>	<b>0,032 (0,011) **</b>
Prestatie informatieverwerking	<b>0,196 (0,013) ***</b>	<b>0,110 (0,011) ***</b>
Prestatie rekenen	<b>0,033 (0,013) *</b>	<b>0,338 (0,013) ***</b>
Sekse * Taal	-0,037 (0,017) *	
Lezen en hobby's * Taal	0,025 (0,007) ***	
Sekse * Rekenen		-0,058 (0,014) ***
Schooltype (Referentiegroep: mavo)		
- ivbo	<b>-0,688 (0,052) ***</b>	<b>-1,014 (0,046) ***</b>
- vbo	<b>-0,459 (0,032) ***</b>	<b>-0,410 (0,029) ***</b>
- vbo/mavo	<b>-0,014 (0,074) ***</b>	<b>-0,028 (0,065) ***</b>
- havo	<b>0,327 (0,033) ***</b>	<b>0,179 (0,029) ***</b>
- havo/vwo	<b>0,507 (0,079) ***</b>	<b>0,343 (0,067) ***</b>
- vwo	<b>0,543 (0,037) ***</b>	<b>0,377 (0,033) ***</b>
Constante	0,141 (0,037)	-0,006 (0,034)

\*  $p < 0,05$ ; \*\*  $p < 0,01$ ; \*\*\*  $p < 0,001$ .

In Tabel 4.3 wordt een model weergegeven met de hoofd- en interactie-effecten uit het vorige hoofdstuk plus de effecten van de verschillende schooltypen. Met dit model kan dus worden nagegaan of het schooltype iets toevoegt aan de verklaring van verschillen in prestaties. Uit Tabel 4.3 blijkt dat bijvoorbeeld ivbo-leerlingen gemiddeld genomen ongeveer zeventiende standaarddeviatie lager op Nederlands en één standaarddeviatie lager op wiskunde scoren dan mavo-leerlingen (onder constant houding van de andere kenmerken). Vwo-leerlingen scoren het beste op beide toetsen. Rekening houdend met de overige kenmerken scoren ze gemiddeld meer dan een halve standaarddeviatie hoger op Nederlands en vierdiende standaarddeviatie hoger op wiskunde. Vbo/mavo-leerlingen verschillen wat netto prestaties betreft niet van mavo-leerlingen. Ook het onderscheid tussen havo/vwo- en vwo-leerlingen is gering.

Voor de verschillende typen geldt dat het effect voor ivbo-leerlingen negatiever is voor wiskunde dan voor Nederlands. Havo- en vwo-leerlingen scoren significant hoger op Nederlands dan op wiskunde. Uit de *simultane toets* (Goldstein, 1995<sup>2</sup>) voor de zes dummy's voor schooltype blijkt dat het contrast 66,80 is met zes vrijheidsgraden, wat overeenkomt met een significantieniveau kleiner dan 0,001. Hieruit kan worden opgemaakt dat het schooltype dus een verschillend effect heeft op Nederlands en wiskunde. In Tabel 4.3 zijn effecten die verschillend zijn voor beide vakken met vet aangegeven.

Door de opname van het schooltype zijn de effecten van het *opleidingsniveau van de vader en de moeder* niet meer significant. Deze effecten worden dus gemedieerd door het schooltype. Naarmate ouders hoger zijn opgeleid, presteren hun kinderen beter in leerjaar drie. Wanneer er rekening wordt gehouden met de prestaties uit de brugklas wordt dit effect minder sterk (hoofdstuk 3). Dit effect valt weg door de toevoeging van het schooltype aan het analysemodel. Er is dus geen direct effect meer van sociaal-economische status op prestaties van leerlingen. Het schooltype intermedieert het effect van sociaal milieu op prestaties.



Naarmate de sociaal-economische status hoger is, presteren leerlingen gemiddeld genomen hoger. De correlatie tussen de opleiding van de moeder, een van de indicatoren van sociaal-economische status, en de toetsscore op Nederlands is bijvoorbeeld 0,29. In een analyse met structurele kenmerken (Bijlage 3A, model 5) is het effect van de opleiding van de moeder op de prestaties op Nederlands 0,056. Dit effect is kleiner dan de correlatie, omdat de opleiding van de moeder matig

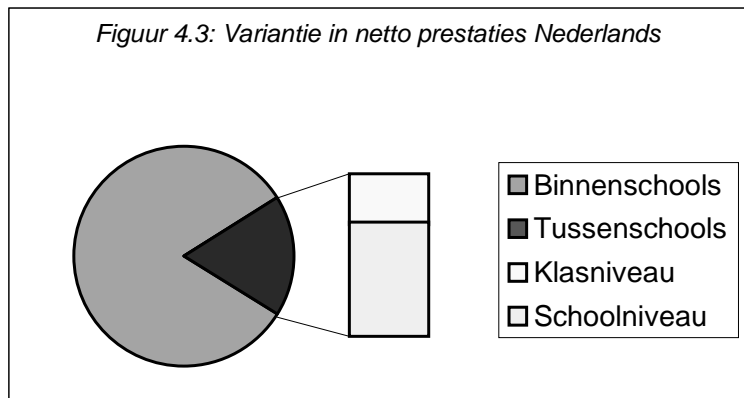
gecorrleerd is met de andere onafhankelijke variabelen, zoals het opleidingsniveau van de vader, het beroepsniveau van de kostwinner en de etniciteit. Vervolgens worden er allerlei culturele kenmerken aan het model toegevoegd, zoals de opvoedingsstijl van ouders of het leefpatroon van leerlingen. Deze variabelen zijn ook gecorrleerd met de opleiding van de moeder en verklaren een deel van het effect. Als vervolgens ook nog de eerdere prestaties van leerlingen worden toegevoegd (Tabel 3.5), dan is het directe effect van de opleiding van de moeder 0,029. Dit effect is klein, maar nog steeds relevant, zoals ik in paragraaf 3.4 naar voren heb gebracht. De vraag is vervolgens waarom heeft de sociaal-economische achtergrond nog steeds een effect op de prestaties van leerlingen. In Tabel 4.3 is het schooltype aan de analyses toegevoegd. Onder constanthouding van overige kenmerken (waaronder de eerdere prestaties) blijkt het schooltype in leerjaar drie een effect te hebben op de prestaties en door de opname van dit effect verdwijnt de invloed van de opleiding van de moeder. De parameterschatting is 0,010 en de standaardfout is 0,006 (Tabel 4.3). Het opleidingsniveau van de moeder is namelijk matig gecorrleerd aan het schooltype, zelfs als er rekening is gehouden met prestaties van leerlingen in leerjaar één.<sup>xi</sup> Naarmate het sociaal milieu hoger is, komen leerlingen niet alleen op hogere schooltype doordat ze beter presteren, maar waarschijnlijk ook doordat ze vaker hoger kiezen dan hun prestaties aangeven. Bij de overgang van basis- naar voortgezet onderwijs is bekend dat leerlingen uit hogere milieus allereerst een hoger advies krijgen dan hun prestaties aangeven en ten tweede hoger kiezen dan hun advies aangeeft (Dronkers, 1997). Na de brugklas is er op veel scholen nog een keuzemoment en ook dan geldt weer dat leerlingen uit hogere sociale milieus hoger kiezen dan hun prestaties aangeven. Op die hogere schooltypen krijgen ze vervolgens moeilijkere stof, waardoor ze in de loop van het voortgezet onderwijs meer leren dan leerlingen die naar een lagere schooltype gaan. Op de keuzemomenten in het onderwijs ontstaat een belangrijk deel van de ongelijkheid naar sociaal milieu.

Verder valt uit een vergelijking van Tabel 3.5 en Tabel 4.3 op dat ook de effecten van het al dan niet hebben van een *bijbaan* en van het aantal dagen dat een leerling *huiswerk* maakt niet langer significant is in een model met schooltype. Voor het maken van huiswerk geldt dus duidelijk dat dit vooral wordt bepaald door het schooltype van een leerling. Naarmate het schooltype hoger is, maken leerlingen gemiddeld genomen meer dagen in de week huiswerk. Ook geldt dat de effecten van de *eerdere prestaties* minder sterk worden. Uit deze veranderingen in de parameters blijkt dat het schooltype van leerlingen een indicatie is van hun prestatieniveau. De parameters voor eerdere prestaties nemen af. Het schooltype is dus een variabele die met talrijke andere kenmerken samenhangt. Voor onderzoek

naar verschillen in prestaties maakt het daarom sterk uit of er wel of niet rekening wordt gehouden met het schooltype.

### Variantiecomponenten en modelfit

Door de opname van het schooltype in de analyse worden de variantiecomponenten beduidend kleiner. Deze componenten bedragen voor Nederlands: op leerlingniveau 0,398; op klasniveau 0,027; op schoolniveau 0,059. De residuele intra-klasse correlatiecoëfficiënt tussen klassen en scholen is  $(0,027 + 0,059) / (0,397 + 0,027 + 0,059) = 0,18$  en tussen scholen  $0,059 / (0,027 + 0,059) = 0,69$ . In Figuur 4.3 is de netto variantie in prestaties Nederlands afgebeeld. De relatieve binnenschoolse variantie bedraagt 82 procent en de relatieve tussenschoolse variantie 18 procent. Deze tussenschoolse variantie betreft voor eenderde variantie tussen klassen en voor tweederde variantie tussen scholen.

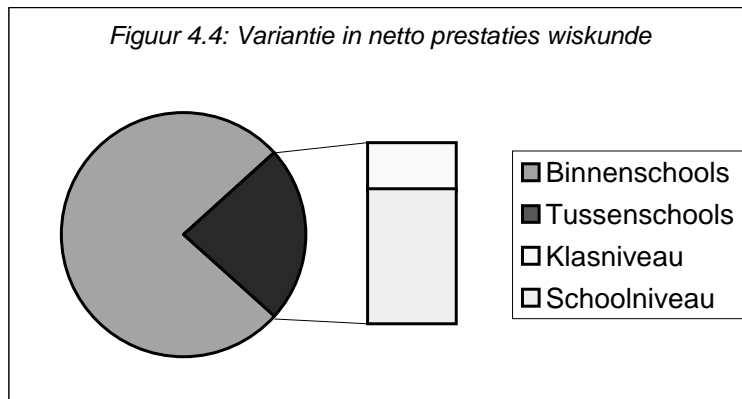


Voor wiskunde bedragen deze componenten op leerlingniveau 0,255; op klasniveau 0,021; op schoolniveau 0,057. Na toevoeging van schooltype bedraagt voor wiskunde de residuele variantie tussen klassen en scholen 23 procent, zie Figuur 4.4. Deze variantie tussen klassen en scholen valt op te delen in 27 procent op klas- en 73 procent op schoolniveau. Voor Nederlands en wiskunde wordt door het schooltype respectievelijk 5 en 7 procent extra variantie op individueel niveau verklaard. Deze gegevens worden in Tabel 4.7 samengevat onder het kopje 'schooltype'.

Met de opname van het schooltype is er op individueel niveau zo uitgebreid mogelijk gecontroleerd. De residuele intra-klasse correlatiecoëfficiënt voor verschillen tussen klassen en scholen is voor Nederlands gedaald van 0,51 voor het lege model tot 0,18 voor het model in Tabel 4.3. Voor wiskunde is er een afname van 0,65 naar 0,23. Als er rekening wordt gehouden met verschillen op indivi-

dueel niveau, dan is de variantie tussen klassen en scholen voor Nederlands 18 en voor wiskunde 23 procent.

Wat de verklaarde variantie betreft wordt op klas- en schoolniveau voor Nederlands 81 en voor wiskunde 88 procent van de variantie gebonden door kenmerken op individueel niveau. Zoals eerder geconstateerd, het overgrote deel van de verschillen tussen scholen is het gevolg van verschillen in leerlingpopulatie.



#### 4.4.2 Random effecten

Om na te gaan of klassen en scholen voor alle leerlingen even goed zijn, kan worden gekeken naar *differentiële effecten*. Vandaar dat ik op klas- en schoolniveau voor alle leerling- en gezinskenmerken uit hoofdstuk 3 ben nagegaan of er sprake is van *random hellingen* en *intercept-helling-covarianties*. Wanneer er sprake is van random hellingen dan is het effect van een bepaald kenmerk niet voor alle leerlingen in een klas of school gelijk. Uiteindelijk bleken er op klasniveau differentiële effecten te zijn voor vijf kenmerken, te weten etniciteit, sekse (en dan alleen voor het effect op wiskunde), het aantal dagen huiswerk, de non-verbale intelligentie (alleen voor wiskunde) en de score op de rekentoets (alleen voor wiskunde).

In hoeverre zijn deze random hellingen van belang? Voor etniciteit geldt dat de gemiddelde helling  $-0,114$  (vergelijk Tabel 4.3) en de random helling  $0,036$  (Tabel 4.4) is. De gemiddelde helling is het vaste effect van etniciteit op prestaties van leerlingen en de random helling is de variantie van dat effect. Wanneer de variantie  $0,036$  is, dan is de standaarddeviatie gelijk aan  $\sqrt{0,036} = 0,190$ . De gemiddelde helling  $\forall$  twee standaarddeviaties loopt van  $-0,494$  tot  $0,266$ . Het verschil tussen deze waarden is ongeveer driekwart standaarddeviatie.



Tabel 4.4: Multiniveau analyse. Random effecten

<b>SCHOOLNIVEAU</b>	<b>PAR. (S.E.)</b>
Variantie Nederlands	0,059 (0,009)
Variantie wiskunde	0,056 (0,008)
Covariantie Nederlands-wiskunde	0,032 (0,007)
<b>KLASNIVEAU</b>	
Variantie Nederlands	0,026 (0,005)
Variantie wiskunde	0,010 (0,003)
Covariantie Nederlands-wiskunde	0,005 (0,003)
Random helling van etniciteit	0,036 (0,015)
Random helling van sekse (wiskunde)	0,009 (0,004)
Random helling van aantal dagen huiswerk	0,004 (0,001)
Random helling van non-verbale intelligentie (wiskunde)	0,005 (0,002)
Random helling van rekenen (wiskunde)	0,007 (0,002)
Intercept-helling covariantie (wiskunde, rekenen)	-0,005 (0,002)
<b>LEERLINGNIVEAU</b>	
Variantie Nederlands	0,393 (0,007)
Variantie wiskunde	0,241 (0,005)
Covariantie Nederlands-wiskunde	0,041 (0,004)

**NOOT:** Rekening houdend met de vaste effecten uit Tabel 4.3

Het effect van etniciteit op prestaties van leerlingen is niet in alle klassen hetzelfde. In sommige klassen is dit effect sterker en in andere klassen zwakker, wat inhoudt dat voor allochtone leerlingen de ene klas dus gunstiger en de andere klas ongunstiger is. Waarom dit differentiële effect optreedt is vervolgens de vraag. Komt dit bijvoorbeeld door de manier van lesgeven door docenten? Op soortgelijke wijze kan het belang van de andere random hellingen worden bepaald. In Tabel 4.5 heb ik de intervallen weergegeven waarin de effecten van 95 procent van de scholen liggen.

Tabel 4.5 Voorspellingsinterval voor (random) effecten op klasniveau

	<b>GEM.</b>	<b>SD</b>	<b>95% VI</b>
Etniciteit	-0,114	0,190	-0,494; 0,266
Sekse (wiskunde)	0,146	0,094	-0,042; 0,334
Aantal dagen huiswerk	0,015	0,061	-0,107; 0,137
Non-verbale intelligentie (wiskunde)	0,092	0,072	-0,052; 0,236
Rekenen (wiskunde)	0,339	0,083	0,173; 0,505

Behalve random intercepts zijn er op klas- en schoolniveau ook intercept-helling covarianties berekend (Snijders & Bosker, 1999). Wat deze covarianties betreft heb ik allereerst gekeken naar *covarianties tussen het intercept van de afhankelijke variabelen en de helling van de soortgelijke toetsen* uit leerjaar één, dus taal met Nederlands, informatieverwerking met Nederlands, rekenen met wiskunde. Van deze covarianties blijkt alleen die tussen rekenen en wiskunde op klasniveau van belang. Met deze covariantie en de varianties van rekenen en wiskunde kan de correlatie tussen intercept en helling worden berekend. Deze correlatie is:  $-0,005 / \sqrt{(0,007 * 0,010)} = -0,60$ . Aangezien alle continue variabelen gestandaardiseerd zijn, correspondeert het intercept met de wiskundescore van een leerling met een gemiddelde score op de rekentoets in een klas met een gemiddelde rekentoetscore. De negatieve correlatie duidt erop dat in klassen met een gemiddeld hoog prestatieniveau (veelal havo- en vwo-klassen) het effect van de rekentoets op de wiskundescore minder sterk is. Het gemiddeld hoge niveau in leerjaar drie van havo- en vwo-klassen komt daarom meer door de goede prestaties van leerlingen die laag scoorden op de rekentoets uit leerjaar één dan door de goede prestaties van leerlingen die hoger scoorden op die toets. Voor (i)vbo-klassen geldt dat het gemiddeld lage niveau in leerjaar drie op wiskunde meer komt doordat de betere leerlingen in het beroepsonderwijs achterblijven bij hun verwachte score dan doordat de zwakkere leerlingen boven hun verwachte score presteren. De verschillen tussen leerlingen in wiskunde gaan dus per schooltype meer op elkaar lijken. In de hogere schooltypen presteren de ‘zwakkere leerlingen’ (met een voor havo- en vwo-leerlingen lage score op de toets uit leerjaar één) boven verwachting op wiskunde. In de lagere schooltypen presteren de ‘betere leerlingen’ beneden verwachting op wiskunde.

Vervolgens heb ik gekeken naar de *covarianties tussen het intercept van de afhankelijke variabelen en de helling van enkele structurele kenmerken*. Van de covarianties van het intercept van Nederlands of wiskunde met de helling van etniciteit, de opleidingsniveaus van vader en moeder en de sekse, blijkt geen van belang.

In totaal zijn er dus vijf random hellingen en één covariantie tussen intercept en helling gevonden. Voor al deze random effecten geldt dat ze differentiëren op klasniveau. Op schoolniveau bleken er geen differentiële effecten te zijn. De kanttekening hierbij is dat er voor dezelfde leerlingkenmerken wel random effecten op schoolniveau kunnen worden gevonden, wanneer de random effecten op klasniveau worden weggelaten. Door de opname van random effecten wordt de modelfit beter, zie Tabel 4.7. De deviance neemt af met 63,4. Bij zes vrijheidsgraden is deze afname significant op een niveau van 0,001. Door de opname van de random effecten veranderen de vaste effecten nauwelijks. Deze effecten worden daarom niet opnieuw weergegeven.

#### 4.4.3 Regressie-effecten tussen klassen of scholen

Bij een analyse met variabelen op één niveau kunnen er talrijke interpretatieproblemen ontstaan. Is een effect van de mening over docenten en schoolregels nu een effect binnen of tussen klassen en scholen? Gaat het bij dit kenmerk om verschillen tussen leerlingen of om verschillen tussen klassen en scholen? Om hier antwoord op te geven kunnen leerling- en gezinskenmerken worden geaggregeerd, waarna die geaggregeerde score als een aparte variabele in de analyses kan worden meegenomen.

Van de kenmerken op individueel niveau uit Tabel 4.3 heb ik op *klas- en schoolniveau* de geaggregeerde score genomen. Voor deze kenmerken geldt dat er in geen van de gevallen sprake is van regressie-effecten tussen klassen en scholen (zie paragraaf 2.3 voor het verschil tussen binnenschoolse en tussenschoolse effecten). Op scholen waar leerlingen bijvoorbeeld gemiddeld genomen positiever zijn over hun docenten en de schoolregels, presteren leerlingen dus niet beter dan op scholen waar leerlingen negatiever zijn. De mening over docenten en regels is een effect op leerlingniveau, dus binnen klassen en scholen. Ook voor de andere kenmerken blijken er geen regressie-effecten tussen klassen en scholen te zijn.

Uit nadere analyses blijkt dat er in een *model zonder schooltype* wel *regressie-effecten tussen klassen* zijn, zie Bijlage 4A. In een dergelijk model zijn er positieve effecten op Nederlands en wiskunde van het gemiddelde opleidingsniveau van de vader (cirkel vijf), de gemiddelde autoritatieve opvoedingsstijl (cirkel vier), het gemiddelde aantal dagen huiswerk, het gemiddelde taalniveau in leerjaar één, en het gemiddelde rekenniveau in leerjaar één (cirkel één). Wanneer er rekening wordt gehouden met het schooltype, verdwijnen deze contexteffecten.<sup>xii</sup> Deze effecten hangen dus duidelijk samen met dat kenmerk. Naarmate het schooltype hoger is, is het beginniveau ook hoger, maken leerlingen vaker huiswerk, voeden ouders hun kinderen meer autoritatief op en is het opleidingsniveau van de vader hoger.

Het gemiddelde niveau in een klas kan dus een invloed hebben op de prestaties van leerlingen, maar die invloed hangt samen met het schooltype waarin een leerling terecht komt. Wanneer in één analyse leerlingen van verschillende schooltypen worden samengenomen en er gekeken wordt naar regressie-effecten tussen klassen en scholen, moet er worden gecontroleerd voor het schooltype. Wanneer dat niet gebeurt kunnen er effecten van geaggregeerde kenmerken worden gevonden die eigenlijk aangeven in wat voor schooltype een leerling zit. Uit deze analyses komen geen effecten van de samenstelling van een klas of school naar voren.

## 4.5 Effectieve scholen

In de voorgaande paragraaf heb ik verschillende controles op klas- en schoolniveau uitgevoerd. Uit deze controles blijkt dat de verklaringen van de prestaties van leerlingen aanzienlijk verbeteren, wanneer rekening wordt gehouden met het schooltype van leerlingen. Op klasniveau blijken er verder talrijke differentiële effecten te zijn. Voor verschillende groepen leerlingen maakt het dus uit in welke klas ze zitten. Wanneer er rekening wordt gehouden met het schooltype, dan blijken er geen regressie-effecten tussen klassen of scholen te zijn. Op klas- en schoolniveau kan echter ook worden gedacht aan *vaste effecten* van schoolkenmerken. Ik heb daarom aan het model met vaste en random effecten van leerlingen en gezinskenmerken schoolkenmerken toegevoegd.

### Contextuele schoolkenmerken

Het percentage allochtone leerlingen, de locatiegrootte en de denominatie blijken invloed te hebben op de prestaties van leerlingen. In Tabel 4.6 worden deze effecten weergegeven. Aangezien het effect van deze kenmerken niet voor alle schooltypen hetzelfde hoeft te zijn, is ook gekeken of deze variabelen interacteren met het schooltype van leerlingen. Uit deze analyses per schooltype, zie Bijlage 4B en 4C, blijkt bijvoorbeeld dat er voor het effect van het percentage allochtone leerlingen op de toetscore Nederlands een onderscheid moet worden gemaakt tussen enerzijds leerlingen met een mavo-opleiding of lager en anderzijds havo/vwo-leerlingen.

Voor alle leerlingen geldt dat het percentage allochtone leerlingen een negatief effect heeft op de wiskundeprestaties. Een afwijking van één standaarddeviatie in het percentage allochtone leerlingen, vermindert de wiskundeprestaties met 0,03. Voor leerlingen die op de havo of hoger zitten geldt dat ze niet alleen op wiskunde maar ook op Nederlands slechter presteren naarmate het percentage allochtone leerlingen op school stijgt. Om het effect van havo/vwo-leerlingen te bekijken moeten het hoofd- en interactie-effect worden samengenomen. Een toename van één standaarddeviatie in het percentage allochtone leerlingen, leidt tot een verschil van  $-0,085$  ( $0,006 - 0,091$ ) op de toets Nederlands.

Het percentage allochtone leerlingen hangt samen met talrijke andere kenmerken van de leerlingencompositie. De prestaties van leerlingen nemen af naarmate er meer allochtone leerlingen op een school zitten. Zijn die prestaties lager, omdat het percentage allochtone leerlingen een meting is van ongemeten leerlingkenmerken? Of passen docenten op scholen met veel allochtone leerlingen hun lessen aan en verlagen ze hun standaarden? Deze vragen kunnen in dit onderzoek niet worden beantwoord. Het is bovendien ook onduidelijk waarom er voor ivbo-, vbo- en mavo-leerlingen geen effect op Nederlands is. Zijn er voor scholen met

veel allochtone leerlingen verschillen tussen hogere en lagere vormen van voortgezet onderwijs?

Voor leerlingen in het beroepsonderwijs maakt de locatiegrootte ( $\log_{10}$ ) een verschil uit voor de prestaties op beide vakken. Voor mavo-, havo- en vwo-leerlingen is er geen effect van de locatiegrootte. Een verschil van één standaarddeviatie leidt voor (i)vbo-scholen tot een verschil van 0,070 (-0,002 – 0,068) standaarddeviatie in de toetscores. Een standaarddeviatie verschil in de locatiegrootte blijkt ongeveer gelijk aan een verdubbeling van het leerlingenaantal, bijvoorbeeld een toename van 300 tot 600 of van 600 tot 1200. Bij een dergelijke verandering dalen de scores van leerlingen in het beroepsonderwijs met eenveertiende standaarddeviatie. De praktische betekenis van deze effecten is groot, omdat ze over vele leerlingen worden uitgespreid (Bosker & Scheerens, 1989).

Er blijkt ook een effect van de denominatie te zijn.<sup>xiii</sup> Hierbij zijn er geen verschillen naar vak (het contrast is 6,69 bij zes vrijheidsgraden). Wel blijkt er een apart effect te zijn op Nederlands voor gereformeerd-vrijgemaakte mavo-scholen. Vandaar dat het model voor alle schooltypen is uitgebreid met een interactieterm om deze relatie weer te geven.<sup>xiv</sup> In Tabel 4.6 worden de effecten van de denominatie gegeven.

Wanneer er rekening wordt gehouden met vaste en random effecten van kenmerken op individueel niveau en met de andere contextuele kenmerken, dan hebben algemeen bijzondere en protestants-christelijke scholen de hoogste scores. Scholen van deze denominatie verschillen echter niet significant van openbare scholen, de referentiegroep. De standaardfouten zijn immers (bijna) net zo groot als de parameterschattingen. De laagst scores zijn gemiddeld genomen van interconfessionele scholen. De verschillen tussen denominaties zijn maximaal 0,21 standaarddeviatie.

In eerder onderzoek werden doorgaans vier denominaties onderscheiden: openbaar, protestants-christelijk, rooms-katholiek en overig bijzonder. Reformatorische en gereformeerd-vrijgemaakte scholen werden dan ingedeeld bij protestants-christelijke of overig bijzondere scholen. Algemeen bijzondere en interconfessionele scholen vielen dan beide onder de noemer overig bijzonder. Uit Tabel 4.6 blijkt dat het verfijnen van de meting van denominatie door de opname van zeven in plaats van vier categorieën zinvol is. Algemeen bijzondere scholen scoren wat netto prestaties betreft significant hoger dan interconfessionele scholen. Ook tussen protestants-christelijke, reformatorische en gereformeerd-vrijgemaakte scholen zijn er verschillen.<sup>xv</sup>

Tabel 4.6: Multiniveau analyse. Spaarzaam model met toevoeging van contextuele en schoolkenmerken

	CONTRAST	O <sup>2</sup>	NEDERLANDS	WISKUNDE
<b>CONTEXTUELE KENMERKEN</b>				
% Allochtone leerlingen	0,03	<b>3,60</b>	<b>0,006 (0,014)}</b>	<b>-0,031 (0,015) *</b>
Havo/vwo * % All. lln			<b>-0,091 (0,053)}</b>	
Locatiegrootte (log <sub>10</sub> )	-0,03	1,22	-0,002 (0,023)} ***	-0,002 (0,023)} ***
(I)vbo * locatiegrootte	-0,04	1,52	-0,068 (0,019)} ***	-0,068 (0,019)} ***
Denominatie (Referentiegroep: openbaar)				
- algemeen bijzonder	0,12	1,56	0,040 (0,072)} *	0,040 (0,072)} *
- rooms-katholiek	0,01	0,07	-0,027 (0,047)} *	-0,027 (0,047)} *
- interconfessioneel	0,09	0,87	-0,171 (0,074)} *	-0,171 (0,074)} *
- protestants-christelijk	0,03	0,17	0,035 (0,051)} *	0,035 (0,051)} *
- reformatorisch	-0,21	2,23	-0,138 (0,116)} *	-0,138 (0,116)} *
- gereformeerd-vrijgemaakt	-0,01	0,08	0,015 (0,107)} *	0,015 (0,107)} *
- mavo * geref.-vrijgemaakt			0,199 (0,084)} *	
<b>SCHOOLKENMERKEN</b>				
Teamstabiliteit	-0,03	1,35	0,043 (0,019) *	0,043 (0,019) *
Maatregelen bij lesuitval (Referentiegroep: lesuitval)				
- invalregeling	0,16	<b>6,13</b>	<b>0,192 (0,058)}</b> ***	<b>0,031 (0,058)}</b>
- verplichte huiswerkles	0,28	<b>8,19</b>	<b>0,346 (0,087)}</b> ***	<b>0,035 (0,086)}</b>
- havo/vwo * invalregeling			<b>-0,119 (0,061)}</b> ***	
- havo/vwo * huiswerkles			<b>-0,244 (0,109)}</b> ***	
Aard brugperiode (Referentiegroep: dakpansgewijs)				
- homogeen	-0,14	2,15	0,066 (0,052)} ***	0,066 (0,052)} ***
- heterogeen	-0,19	<b>3,33</b>	<b>-0,213 (0,090)}</b> ***	<b>-0,001 (0,092)}</b>
- vwo * heterogeen			<b>0,414 (0,103)}</b> ***	
Mentoraatsysteem (Referentiegroep: smal/verticaal)				
- smal/horizontaal	0,00	0,02	-0,055 (0,044)} *	-0,055 (0,044)} *
- breed/verticaal	0,00	0,03	-0,115 (0,046)} *	-0,115 (0,046)} *
- breed/horizontaal	-0,14	2,70	-0,176 (0,067)} *	-0,176 (0,067)} ***
Docent-leerling ratio	0,01	0,05	0,039 (0,019)} ***	0,039 (0,019)} ***
(I)vbo * Doc.-ll. ratio	-0,05	1,35	-0,089 (0,023)} ***	-0,089 (0,023)} ***

NOOT: Op individueel niveau zijn de kenmerken uit Tabel 4.3 meegenomen.

\*  $p < 0,05$ ; \*\*  $p < 0,01$ ; \*\*\*  $p < 0,001$ .

### Schoolkenmerken

Een aantal schoolkenmerken heeft een effect op de prestaties van leerlingen. Onder constant houding van de leerling- en gezinskenmerken blijkt de teamstabiliteit een zwak positief effect te hebben op de leerprestaties op beide vakken, zie Tabel 4.6. Door één standaarddeviatie verschil in de teamstabiliteit, gaan de prestaties van leerlingen met 0,04 standaarddeviatie vooruit. Wanneer lesuitval wordt opgevangen door een invalregeling of een verplichte huiswerkles, dan presteren ivbo-, vbo- en mavo-leerlingen op Nederlands een vijfde tot eenderde standaarddeviatie hoger dan wanneer het uur komt te vervallen. Voor havo- en vwo-leerlingen gaat dit effect niet op. Er blijkt namelijk een interactie te zijn tussen de invalregeling en het schooltype.

Van de aard van de brugperiode gaat ook een effect uit op Nederlands. Het blijkt dat leerlingen in heterogene brugklassen het slechtst presteren op Nederlands. Dit effect geldt echter niet voor vwo-leerlingen, zie bijlage 4B. Deze leerlingen presteren niet slechter, wanneer leerlingen in de brugperiode met leerlingen uit meer dan drie schooltypen in een klas zitten. Behalve voor vwo-leerlingen wordt er dus een negatief effect gevonden van heterogeen groeperen op de toetscore Nederlands. Deze bevinding wijkt af van wat in eerder onderzoek is gevonden. Een verschil met dat eerdere onderzoek is echter dat er hier niet alleen een heterogene en homogene, maar ook een dakpansgewijze brugperiode wordt onderscheiden. Leerlingen in dakpansgewijze brugklassen verschillen niet significant in prestaties op beide vakken van leerlingen in een homogene brugklas.<sup>xvi</sup>

Wat de vormgeving van het mentoraatsysteem betreft valt het op dat hoe uitgebreider dat systeem is opgezet, hoe slechter de prestaties van leerlingen. Alle vormen scoren lager dan een smal-verticaal mentoraatsysteem, de referentiegroep. Voor Nederlands is het verschil tussen het breed-horizontale en het smal-verticale systeem significant. De invloed van een uitgebreid mentoraatsysteem is dus tegen de verwachting in negatief. Dit negatieve effect vraagt om vervolgonderzoek. In dat vervolgonderzoek zou achterhaald moeten worden hoe nuttig mentorlessen eigenlijk zijn. Ontwikkelen leerlingen meer schoolse of algemene vaardigheden tijdens die lessen? Sporen docenten vaker vroegtijdig problemen van leerlingen op? Kunnen docenten die problemen ook oplossen? Verder is het de vraag of oorzaak en gevolg zouden moeten worden omgedraaid: des te slechter de prestaties van leerlingen, des te uitgebreider de opzet van het mentoraatsysteem. Een bezwaar tegen deze redenering is dat het mentoraatsysteem een tamelijk stabiel kenmerk van een school is. Het systeem bestaat al voordat leerlingen de school binnenkomen. Voor het omdraaien van oorzaak en gevolg zou pleiten dat de instroom van scholen verschilt naar gelang het mentoraatsysteem. Vooralsnog zijn daar echter geen duidelijke aanwijzingen voor. Ook zou kunnen worden gekeken

naar effecten van mentoren in plaats van naar effecten van een mentoraatsysteem. Zijn bepaalde mentoren succesvoller dan anderen, en hebben risicoleerlingen de meeste baat bij een goede mentor?

Tot slot, blijkt de docent-leerling ratio geen effect te hebben dat significant is voor alle schooltypen. Voor (i)vbo-leerlingen geldt echter dat ze slechter presteren naarmate deze ratio hoger is. Hier geldt dat niet alleen ivbo- maar ook vbo-leerlingen baat hebben bij een structuur waarin ze niet al te veel verschillende docenten hebben. Uit Tabel 4.6 valt af te leiden dat een toename van één standaarddeviatie in deze ratio leidt tot een verschil van  $-0,050$  ( $0,039 - 0,089$ ) in de toetscores op beide vakken.

Geen effect is er van het prestatieniveau als extra toelatingscriterium op de prestaties van leerlingen. In Bijlage 4D valt er weliswaar een positief effect van dit kenmerk te bespeuren op de toetsscore Nederlands. Wanneer het effect van het prestatieniveau als extra criterium op de toetsscore op beide vakken wordt weggelaten, dan verslechtert het model echter niet significant ( $\chi^2 = 3,0$ ,  $df = 2$ ,  $p > 0,20$ ). De variatie in dit kenmerk blijkt ook niet zo groot. Van alle scholen hanteert 78 procent dat extra criterium. Als er wat verfijnder naar het toelatingsbeleid van scholen zou kunnen worden gekeken, dan zouden er misschien verschillen tussen scholen te ontdekken zijn. Zo verschillen zelfstandige gymnasia mogelijk van andere vwo-scholen door een onderscheid in toelatingsbeleid. Vooralsnog zijn er geen aanwijzingen dat een dergelijk effect optreedt.

Naarmate de spreiding in antwoorden tussen docenten op vragen over het schoolklimaat kleiner is en docenten dus meer eenheid in opvattingen vertonen, zijn de netto prestaties op een school niet hoger. Er is dus geen effect van eenheid in visie en doelen op de prestaties van leerlingen. Ook het kenmerk afspraken over orderegels blijkt geen effect te hebben op de prestaties op beide vakken. Uit de Bijlagen 4B, 4C en 4D kan worden opgemaakt dat er apart noch integraal effecten zijn van dit kenmerk.

In Tabel 4.6 zijn alle effecten op de leerprestaties weergegeven. In de tabel is ook aangegeven of effecten voor beide vakken gelijk of verschillend zijn. Uitgaande van een significantieniveau van 0,10 is de kritieke waarde in de chi-kwadraatverdeling 2,71. Vier effecten blijken dan te verschillen naar vak, te weten het percentage allochtone leerlingen, de maatregelen bij lesuitval (voor zowel de invalregeling als de verplichte huiswerkles) en de heterogene brugperiode.



Tabel 4.7: Multiniveau analyse. Variantie- en covariantiecomponenten en modelfit

	VARIANTIE NEDERLANDS	VARIANTIE WISKUNDE	COVARIANTIE	DEVIANCE	)
INDIVIDUEEL *	PAR. (S.E.)	PAR. (S.E.)	PAR. (S.E.)		
Leerlingniveau	0,420 (0,008)	0,270 (0,005)	0,064 (0,005)	22.984,1	
Klasniveau	0,051 (0,007)	0,050 (0,006)	0,030 (0,005)		
Schoolniveau	0,064 (0,011)	0,094 (0,015)	0,050 (0,010)		
<b>SCHOOLTYPE</b>					
Leerlingniveau	0,397 (0,007)	0,255 (0,005)	0,045 (0,004)	22.089,3	894,8
Klasniveau	0,027 (0,005)	0,021 (0,004)	0,008 (0,003)		(df=12)
Schoolniveau	0,059 (0,010)	0,057 (0,009)	0,033 (0,006)		p<0,001
<b>RANDOM EFFECTEN</b>					
Leerlingniveau	0,393 (0,007)	0,241 (0,005)	0,041 (0,004)	22.025,9	63,4
Klasniveau	0,026 (0,005)	0,010 (0,003)	0,005 (0,003)		(df=6)
Schoolniveau	0,059 (0,009)	0,056 (0,008)	0,032 (0,007)		p<0,001
<b>CONTEXTUELE KENMERKEN</b>					
Leerlingniveau	0,393 (0,007)	0,241 (0,005)	0,041 (0,004)	21.965,0	60,9
Klasniveau	0,023 (0,004)	0,010 (0,003)	0,004 (0,003)		(df=20)
Schoolniveau	0,047 (0,008)	0,046 (0,007)	0,025 (0,006)		p<0,001
<b>SCHOOLKENMERKEN</b>					
Leerlingniveau	0,392 (0,007)	0,242 (0,005)	0,041 (0,004)	21.879,3	85,7
Klasniveau	0,023 (0,004)	0,010 (0,003)	0,004 (0,003)		(df=29)
Schoolniveau	0,028 (0,006)	0,040 (0,007)	0,018 (0,005)		p<0,001
<b>SPAARZAAM MODEL</b>					
Leerlingniveau	0,392 (0,007)	0,241 (0,005)	0,041 (0,004)	21.897,8	18,5
Klasniveau	0,023 (0,004)	0,009 (0,003)	0,004 (0,003)		(df=21)
Schoolniveau	0,031 (0,006)	0,043 (0,007)	0,016 (0,005)		p>0,60

\* Spaarzaam model met interactie-effecten

In Tabel 4.7 heb ik de variantiecomponenten en de modelfit weergegeven. De resultaten hieruit zijn gedeeltelijk al besproken in de vorige paragraaf. Door de opname van het schooltype (4.4.1) en random effecten (4.4.2) neemt de deviance af. Door opname van de contextuele en de schoolkenmerken wordt de modelfit nog beter. De extra verklaarde variantie door deze kenmerken bedraagt op schoolniveau voor Nederlands 7,3 procent en voor wiskunde 2,9 procent.

In het spaarzame model is de verklaarde variantie voor Nederlands tot 58,0 op leerling-, 87,0 op klas- en 88,6 procent op schoolniveau. Voor wiskunde is de verklaarde variantie dan respectievelijk 74,0, 91,5 en 90,7 procent.

Op klas- en schoolniveau blijkt de structuur van concentrische cirkels geen meerwaarde te hebben. De effecten van contextuele kenmerken kunnen niet of nauwelijks worden verklaard door schoolkenmerken, zie Bijlage 4D. In dit hoofdstuk is dan ook geen extra inzicht verschaft in de relaties tussen deze kenmerken onderling. De aanname dat prestaties van leerlingen meer worden verklaard door schoolkenmerken dan door contextuele kenmerken wordt niet door de analyses ondersteund. Door de toevoeging van contextuele kenmerken gaat de deviance met 60,9 omlaag bij 20 vrijheidsgraden, zie Tabel 4.7. Door de toevoeging van de schoolkenmerken neemt deze maat met 85,7 af bij 29 vrijheidsgraden. Wanneer de afname in deviance wordt afgezet tegen het aantal vrijheidsgraden dan is de verbetering van het model door beide soorten kenmerken vergelijkbaar.

#### 4.6 Besluit

In *bruto prestaties* variëren scholen sterk. De leerlingenpopulatie verschilt immers sterk van school tot school en uit het vorige hoofdstuk blijkt dat kenmerken van leerlingen en gezinnen een sterke invloed op prestaties hebben. Het is dus geen wonder dat er zulke verschillen zijn. In *netto prestaties* zijn die verschillen veel minder groot. Wanneer dan geprobeerd wordt te ontrafelen waarom bepaalde scholen hogere netto prestaties behalen dan andere scholen dan blijkt dat een moeilijke opgave.

Eén van de belangrijkste uitkomsten uit dit hoofdstuk is dat, om de vraag naar verschillen in prestaties te beantwoorden, het van belang is om het *schooltype* in de analyses op te nemen. Het schooltype blijkt een zelfstandig effect te hebben op de leerprestaties en medieert bovendien andere effecten, zoals het effect van het opleidingsniveau van ouders of het aantal dagen waarop door leerlingen huiswerk wordt gemaakt.

Deze bevinding werpt een nieuw licht op onderzoek van Bosker (1990). Bosker heeft onderzocht op welke momenten in de schoolloopbaan ongelijkheid naar sociaal milieu ontstaat. Uit zijn analyses blijkt dat 68 procent van de verschillen naar sociaal-economische status in de voorschoolse periode of tijdens het basisonderwijs ontstaat. De overgang van basisschool naar voortgezet onderwijs heeft betrekking op 14 procent van de verschillen. Tijdens het voortgezet onderwijs ontstaat de resterende 18 procent aan ongelijkheid. Uit dit hoofdstuk blijkt dat die laatste vorm van ongelijkheid voor een belangrijk deel het gevolg is van de schoolkeuze die leerlingen maken na de brugklas. Net als bij de overgang van ba-

sis- naar voortgezet onderwijs kiezen leerlingen uit hogere sociale milieus bij de overgang na de brugklas hoger dan hun prestatieniveau aangeeft. In het voortgezet onderwijs zijn er dus geen directe effecten meer van de sociaal-economische achtergrond op de prestaties van leerlingen. Deze effecten worden volledig gemedieerd door andere leerling- en gezinskenmerken of door het schooltype. Ongelijkheid ontstaat in deze fase dus voor een belangrijk tijdens keuzemomenten.

Een tweede controle die in dit hoofdstuk is uitgevoerd heeft betrekking op *differentiële effecten*. Voor verschillende leerlingen blijkt het uit te maken in welke klas ze zitten. Het effect van etniciteit op prestaties is bijvoorbeeld niet in alle klassen hetzelfde. In sommige klassen is dit effect sterker en in andere klassen zwakker, wat inhoudt dat voor allochtone leerlingen de ene klas gunstiger en de andere klas ongunstiger is. Met multiniveau analyse kunnen dergelijke random effecten worden gemodelleerd. Vervolgens is het dan de vraag of deze random effecten op klasniveau kunnen worden verklaard door andere kenmerken. In dit onderzoek is het niet gelukt om die verklaring te geven, doordat er geen gebruik kon worden gemaakt van klaskenmerken.

Een ander voordeel van multiniveau analyse is dat er gekeken kan worden naar *regressie-effecten tussen klassen en scholen*. Dit soort effecten bleken er niet te zijn. Al zou het percentage allochtone leerlingen wel kunnen worden beschouwd als een tussenschools regressie-effect. Etniciteit blijkt niet alleen op individueel maar ook op schoolniveau een effect te hebben op prestaties van leerlingen. Naarmate er meer allochtone leerlingen op een school zitten presteren leerlingen (allochtoon en autochtoon) slechter. Van de compositie van de school gaat dus een effect uit. Hierbij moet echter worden opgemerkt dat in dit onderzoek het percentage allochtone leerlingen niet de op schoolniveau geaggregeerde score van etniciteit is, maar de door de schooladministratie opgegeven score. Stel, het percentage allochtone leerlingen op een vbo/mavo-school is 30 procent. Deze school heeft vier brugklassen, waarvan er twee in de steekproef komen. In deze twee klassen zitten vijftig leerlingen, waarvan tien allochtoon zijn en veertig autochtoon. De op schoolniveau geaggregeerde score van etniciteit is dus 0,20. Uit dit voorbeeld blijkt dat er een verschil is tussen de opgave door de schooladministratie en de op schoolniveau geaggregeerde score. Deze verschillen blijken in VOCL'93 groot. Van brede scholen zijn in het ene geval alleen maar vbo-klassen genomen en in het andere geval juist vwo-klassen. Om deze vertekening door de selectie van klassen tegen te gaan, wordt voor het meten van het percentage allochtone leerlingen de gegevens van de schooladministratie gebruikt.

*Vaste effecten* van contextuele en schoolkenmerken blijken er in enkele gevallen te zijn. Een positief effect is er van de teamstabiliteit. Verschillen tussen scholen blijken er ook te zijn naar het mentoraatsysteem. Een uitgebreid mentoraatsys-

teem blijkt negatief samen te hangen met de prestaties beide vakken. Dit verband blijkt tegen de verwachting in negatief. Verder blijken er verschillen te zijn naar denominatie. Gegeven de overige kenmerken, springen bij deze verschillen met name de achterstanden van leerlingen op interconfessionele scholen eruit. De verschillen tussen denominaties zijn overigens beduidend kleiner dan de verschillen tussen openbaar en bijzonder onderwijs in bijvoorbeeld de Verenigde Staten.

De bovengenoemde effecten gelden voor alle schooltypen, maar andere effecten gaan slechts voor een *deel van de schooltypen* of voor slechts *één van de vakken* op. Het percentage allochtone leerlingen heeft voor alle schooltypen een negatief effect op de wiskundeprestaties. Voor havo/vwo-leerlingen geldt dat het effect ook voor Nederlands negatief is. Een ander contextueel kenmerk dat een zwak negatief effect heeft is het effect van de locatiegrootte op de prestaties van leerlingen in het beroepsonderwijs. Ook voor de denominatie-effecten is er sprake van een interactie. Onder constant houding van de overige kenmerken, presteren mavo-leerlingen op gereformeerd-vrijgemaakte scholen beter op Nederlands dan de overige leerlingen.

Verschillen tussen vbo/mavo-scholen blijken er te zijn naar de maatregelen bij lesuitval. Op scholen die gebruik maken van een invalregeling of van verplichte huiswerklessen scoren vbo/mavo-leerlingen beter op Nederlands. De aard van de brugperiode maakt ook uit voor Nederlands. Wanneer leerlingen in de brugperiode heterogeen worden gegroepeerd presteren ze in leerjaar drie slechter op Nederlands. Dit effect blijkt voor leerlingen in alle schooltypen te gelden, met uitzondering van vwo-leerlingen. In de praktijk groeperen scholen leerlingen overigens steeds minder heterogeen. Meer dan de helft van de leerlingen zit in een homogene brugklas. In een heterogene brugklas zit één op de tien leerlingen. De overige leerlingen zitten in een tussenvorm, een dakpanklas.

De docent-leerling ratio heeft alleen een effect voor (i)vbo-leerlingen. Naarmate de ratio groter is, presteren deze leerlingen slechter op beide vakken. Er zijn dus enkele kenmerken die een effect hebben op de prestaties van leerlingen. Deze effecten gaan overwegend dezelfde richting op voor Nederlands en wiskunde, maar verschillen vaak in sterkte.

De multiniveau analyse zijn voor alle schooltypen tegelijkertijd verricht. Door het maken van interacties kunnen echter effecten worden opgenomen die specifiek voor een bepaald schooltype zijn. In het verleden hebben onderzoekers, bijvoorbeeld Van den Eeden, De Jong, Koopman en Roeleveld (1993), betoogd dat een onderzoek naar effectieve scholen per schooltype dient te geschieden. Hun belangrijkste argument is dat het al moeilijk is om op klas- en schoolniveau effecten

te vinden, wanneer er per schooltype wordt geanalyseerd, laat staan dat er een analyse wordt gedraaid voor alle schooltype tegelijkertijd. In dit hoofdstuk heb ik echter laten zien dat een dergelijke analyse wel degelijk kan worden verricht. Effecten die voor een deel van de schooltypen gelden, kunnen met een interactie-term in de analyse worden opgenomen. De meerwaarde van deze integrale analyse is dat het aantal onderzoekseenheden, met name op schoolniveau, niet wordt verkleind. Op die manier kunnen kleine effecten eerder worden opgespoord. Wanneer er afzonderlijk per vak en per schooltype wordt geanalyseerd (Bijlage 4B en 4C), dan heeft teamstabiliteit bijvoorbeeld in zeven van de acht gevallen een positief, maar niet een significant effect. Wanneer het effect van dit kenmerk voor alle schooltypen tegelijkertijd wordt bekeken, blijkt er wel een significant effect te bestaan.

Deze integrale analyses zijn overigens niet afwijkend van gangbaar onderzoek. In landen waar een ongedeelde voortgezet onderwijs is, zoals de Verenigde Staten, vinden analyses voor alle leerlingen tegelijk plaats. In Nederlands onderzoek naar het basisonderwijs is het ook vanzelfsprekend om analyses te verrichten met een steekproef van leerlingen uit het complete spectrum aan scholen, van elitescholen tot alloctonenscholen. Zelfs in het voortgezet onderwijs komen integrale analyses voor. Bosker (1990) heeft bijvoorbeeld gegevens van vbo- tot en met vwo-leerlingen gebruikt in een onderzoek met de positie op de leerjarenladder als afhankelijke variabele. In onderzoek met toetsscores als criteriumvariabele is het echter standaard geworden om analyses per schooltype te verrichten. In dit hoofdstuk is duidelijk gemaakt dat ook voor die analyses een integrale aanpak een meerwaarde heeft.

In onderzoek naar contextuele en schoolkenmerken blijkt de structuur van concentrische cirkels geen meerwaarde te hebben. De effecten van contextuele kenmerken kunnen niet of nauwelijks worden verklaard door schoolkenmerken. In dit hoofdstuk is dan ook geen extra inzicht verschaft in de relaties tussen deze kenmerken onderling. De aanname dat prestaties van leerlingen meer worden verklaard door schoolkenmerken dan door contextuele kenmerken wordt niet door de analyses ondersteund. De toevoeging van beide soorten kenmerken resulteert in een vergelijkbare verbetering van het analysemodel.

In dit hoofdstuk is naar voren gebracht dat de empirische onderbouwing van veel effectieve schoolkenmerken is gebaseerd op buitenlands onderzoek. In tegenstelling tot Amerikaanse scholen verschillen Nederlandse scholen bijvoorbeeld minder in de nadruk op basisvaardigheden. Vandaar dat er ook geen effect van dit kenmerk valt te vinden. Voor het schooleffectiviteitsonderzoek geldt dat de bevindingen uit onderzoek in een bepaald land niet zo maar kunnen worden gegene-

raliseerd naar andere landen. Buitenlands onderzoek naar effectieve scholen is uiteraard een goede leidraad voor Nederlands onderzoek op dat terrein. Wanneer het echter een keurslijf wordt, dan laten onderzoekers vragen liggen die specifiek voor ons onderwijssysteem gelden. Het is in dit opzicht opvallend dat er naar de aard van de brugperiode of het mentoraatsysteem nog weinig onderzoek is gedaan.

In de tweede plaats wordt in veel overzichten van effectiviteitsonderzoek geen onderscheid gemaakt naar onderwijsniveau. Er zijn bijvoorbeeld duidelijke verschillen te verwachten in effecten voor scholen voor basis- dan wel voortgezet onderwijs. Waar een effect van de samenwerking tussen school en thuis in het basisonderwijs mogelijk kan worden gevonden, is een dergelijk effect voor het voortgezet onderwijs minder waarschijnlijk.

In de derde plaats wordt in een deel van het effectiviteitsonderzoek onvoldoende gecontroleerd voor kenmerken op individueel niveau. Wanneer rekening wordt gehouden met leerling- en gezinskenmerken, zo blijkt uit dit onderzoek, blijft er weinig variantie op klas- en schoolniveau over. Als vervolgens ook nog wordt gecontroleerd voor het door de leerling bezochte schooltype, dan is er nog minder variantie op hogere niveaus. Verschillen tussen klassen en scholen worden voor meer dan 80 procent verklaard door kenmerken op individueel niveau. Het ontbreken van effecten van sommige schoolkenmerken kan dan ook het gevolg zijn van de beperkte invloed van scholen op prestaties van leerlingen. Volgens Van der Velden (1996) heeft ook onderzoek in het basisonderwijs niet kunnen aantonen dat de zogeheten effectiviteitskenmerken daadwerkelijk effectief zijn.

Toch kan het ontbreken van effecten van schoolkenmerken voor een deel ook komen door het onderzoeksdesign. Schriftelijke vragenlijsten zijn voor het meten van sommige kenmerken op klas- en schoolniveau ongeschikt. Zo kwam in dit hoofdstuk naar voren dat de gegevens uit de docentenvragenlijst weliswaar betrouwbaar, maar niet valide zijn. In het onderstaande wil ik kort ingaan op de vraag in hoeverre school- en docentkenmerken te meten zijn met vragenlijsten. Ook draag ik enkele oplossingen aan om de kwaliteit van de dataverzameling te vergroten.

Voor twee scholen heb ik de data die het CBS, GION en OCTO hebben verzameld vergeleken met gegevens uit veldwerk. Elders heb ik daarover gerapporteerd (Veenstra, 1997). Op grond van deze vergelijking wil ik enkele opmerkingen maken over de opzet van de schriftelijke vragenlijsten van VOCL'93. In de eerste plaats moet het voor directieleden duidelijk zijn of ze de vragenlijst alleen voor hun *locatie* of voor de *gehele school* moeten invullen. Uit de gegevens van VOCL'93 blijkt dat sommige directieleden het voor de locatie en andere het voor de school invullen.

Hierdoor kunnen de antwoorden op sommige vragen niet met elkaar worden vergeleken. Doorgaans is het denk ik het beste om een directielid vragen te stellen over één locatie, omdat locaties sterk van elkaar kunnen verschillen. Antwoorden die voor de hele school gelden zijn dan niet zinnig. Sommige gegevens, zoals de schoolgrootte of de omgang tussen de schooldirectie en docenten, zouden echter wel moeten worden verzameld.

Ten tweede kunnen verscheidene gegevens die nu via de vragenlijsten worden verzameld uit informatiebrochures, jaarverslagen of de toekomstige schoolgidsen worden gehaald. Dergelijk *schriftelijk materiaal* kan bij de scholen worden opgevraagd. Op die manier kunnen de vragenlijsten voor met name de directieleden wat minder omvangrijk worden. Ook kan op deze manier worden gezorgd dat er geen ontbrekende gegevens zijn voor bijvoorbeeld de locatiegrootte, het aantal lessen of de examenresultaten. Wanneer dergelijke gegevens jaarlijks worden bijgehouden komen storende fouten, zoals een verkeerde denominatie, waarschijnlijk minder vaak voor.

In de derde plaats merk ik op dat het voor de meeste gegevens over scholen en docenten nodig is om *herhaaldelijk metingen* te verrichten. Uit de vergelijking (Veenstra, 1997) blijkt bijvoorbeeld dat de ene school het nascholingsbeleid recent heeft gewijzigd en dat de aard van brugperiode op de andere school van dakpansgewijs is omgezet in heterogeen. Om een goed overzicht van de praktijk in het onderwijs te geven, is het dus nodig om dergelijke informatie om de paar jaar op te vragen. Een groot deel van de gegevens zal echter niet elke keer veranderen. Vandaar dat het misschien mogelijk is om alleen de eerste keer een uitgebreide vragenlijst naar de directie te sturen. Bij latere metingen kan de vragenlijst wellicht worden ingekort door alleen die vragen op te nemen die betrekking hebben op nieuwe ontwikkelingen. Wat de overige gegevens betreft zou het raadzaam kunnen zijn om aan de directie een uitdraai te geven van hun antwoorden op het eerste meetmoment met daarbij de vraag om na te gaan of bepaalde antwoorden inmiddels achterhaald zijn en dus moeten worden veranderd.

Ook voor de prestatie meting bij leerlingen is een herhaaldelijke meting van belang. Leerlingen in het voortgezet onderwijs wisselen vaak jaarlijks van docent. Het beste onderzoeksdesign is daarom waarschijnlijk een met een prestatie meting aan het begin en aan het einde van hetzelfde schooljaar. Het ideaal zou dan zijn om deze procedure enkele jaren achtereen uit te voeren. Op die manier kan worden nagegaan in hoeverre een leerling vooruit gaat in een bepaald leerjaar (en bij een bepaalde docent) en over verschillende leerjaren (en bij één of meer docenten).

Ten vierde is het nodig om de opzet van de vragenlijst voor docenten te veranderen. De huidige vragenlijsten zijn *te omvangrijk*. Verscheidene docenten hadden

daar geen goed woord voor over en maakten opmerkingen als 'De leukheid was omgekeerd evenredig met de tijd', 'Zonde van m'n tijd' of 'Ik werd steeds agressiever door de grote hoeveelheid vragen! Dus ik stop er nu ook mee!'. In paragraaf 4.3 bleken de vragen waarbij van docenten wordt verlangd om percentages of frequenties te schatten van hun activiteiten, niet valide te zijn. Docenten schudden deze gegevens, zo was mijn indruk tijdens het kwalitatieve onderzoek (Veenstra, 1997), zo uit de mouw. Ze geven altijd op dat ze tijd besteden aan het uitleggen van nieuwe stof, maar in de praktijk blijkt dat een aanzienlijk deel van de docenten een les als volgt opbouwt: ongeveer vijf minuten verloren tijd doordat leerlingen van lokaal moeten wisselen, vijftien minuten huiswerk bespreken, twintig minuten opgaven of huiswerk maken en vijf minuten orde houden, huiswerk opgeven en andere praktische zaken. Het uitleggen van nieuwe stof komt lang niet in elke les voor. Het lijkt er dan ook op rapportages die docenten over zichzelf geven twijfelachtig zijn. Voor het achterhalen van kenmerken van docenten zal moeten worden gekozen voor andere methoden van dataverzameling. Om dergelijke gegevens te achterhalen kan bijvoorbeeld gebruik worden gemaakt van observaties of logboeken, zie bijvoorbeeld De Jong en Westerhof (1998).

Een alternatief voor observaties of logboeken is mogelijk het ondervragen van leerlingen over een docent om zo het gemiddelde oordeel van leerlingen te bepalen (Brekelmans, 1989). Scheerens en Bosker (1997, p. 205) halen een studie van Grisay naar het onderwijs in Frankrijk aan, waar succesvol gebruik is gemaakt van de oordelen van leerlingen over een docent. "The fact that pupil judgments about likely effectiveness-enhancing conditions showed most of the significant process-output associations raises an interesting methodological point. Pupil judgments and observations on school organizational functioning and teaching can be seen as having two advantages: avoiding self-reports by teachers and head teachers, and measuring these process conditions at the level where the impact on actual performance is likely to be greatest." In tegenstelling tot experts of observatoren kunnen leerlingen hun mening baseren op vele lessen. Bovendien is de mening van leerlingen over docenten misschien ook wel relevanter. Docenten kunnen vinden dat ze bijvoorbeeld duidelijk uitleggen, maar die uitleg is pas echt duidelijk als leerlingen dat ook zo ervaren. In onderwijsonderzoek worden oordelen van leerlingen over docenten nog niet zo vaak gebruikt. De informatie die zo verkregen wordt is naar het schijnt echter geschikter dan die uit een vragenlijst voor docenten.

Een andere oplossing voor de problemen met de docentenvragenlijst is om kortere vragenlijsten bij meer docenten af te nemen. Op dit moment worden er alleen vragenlijsten afgenomen bij docenten Nederlands en wiskunde. Voor die vakken maken leerlingen prestatietoetsen en dus ligt het voor de hand om na te gaan of het



*onderwijsaanbod* in deze vakken en de *wijze van lesgeven* van de betreffende docenten van invloed is op de vooruitgang in prestaties van leerlingen. Vragen over het *functioneren van de school* kunnen echter ook aan andere docenten worden gesteld. Bij die vragen gaat het er niet zo zeer om wat de betreffende docent ervan vindt, maar om wat verschillende docenten gezamenlijk vinden en om de mate waarin hun antwoorden van elkaar verschillen.

Ten vijfde moet het mogelijk zijn om de vragenlijsten *computergestuurd* af te nemen. Tijdens het veldwerk bleek dat docenten het leuk vinden om op die manier een vragenlijst in te vullen. Deze manier van enquêteren heeft verscheidene voordelen. Het doorverwijzen in de vragenlijst kan automatisch gebeuren, waardoor hiermee geen fouten kunnen worden gemaakt. De vragenlijst kan zo ook sneller worden afgenomen. Verder zullen er minder ontbrekende gegevens voorkomen, omdat er pas een nieuwe vraag op het scherm verschijnt, wanneer de vorige vraag is beantwoord. Computergestuurd interviewen komt de laatste jaren steeds vaker voor. Het is goed om na te gaan hoe die methode ook in onderwijsonderzoek kan worden gebruikt.

## Eindnoten

<sup>i</sup> Een ander kenmerk dat als contextueel kan worden beschouwd is de hechtheid van de gemeenschap rondom de school. Wanneer er rekening wordt gehouden met effecten op individueel niveau zijn er geen aanwijzingen dat dit kenmerk van invloed is op prestatieverschillen tussen scholen (Veenstra & Dijkstra, 1998). Verder kan ook aan de bestuurlijke inrichting van scholen worden gedacht (Chubb & Moe, 1990; Hofman, 1993).

<sup>ii</sup> Voor het basisonderwijs stellen Jungbluth, Peetsma en Roeleveld (1996) dat er in Nederland sprake is van 'standenonderwijs'. Er zijn namelijk elitescholen (met veel autochtone leerlingen uit gezinnen waar de ouders hoger onderwijs hebben gevolgd), burgerscholen (met veel autochtone leerlingen uit gezinnen waar de ouders lager of middelbaar beroepsonderwijs hebben gevolgd) en alloctonenscholen (met veel allochtone leerlingen uit gezinnen waar de ouders hoogstens lager beroepsonderwijs hebben gevolgd). Uit deze laatste categorie distilleren zij nog een apart type waarop voornamelijk Marokkaanse en Turkse leerlingen zitten.

<sup>iii</sup> Een aanzet hiertoe is Heyl (1996). Heyl beschrijft de structuur, de intensiteit en de invloed van informele contacten tussen docenten in kleine mavo-scholen. Met prestaties van leerlingen worden deze informele contacten niet in verband gebracht.

<sup>iv</sup> Net als Oakes (1985) controleren zij onvoldoende voor andere kenmerken. Zie Peschar (1994) voor een bespreking van onderzoek naar de klassensamenstelling.

<sup>v</sup> Een kenmerk dat ook in deze cirkel zou passen is hoge verwachtingen van docenten over prestaties van leerlingen. Over de invloed van deze verwachtingen op de daadwerkelijke resultaten is vaak gerapporteerd (Brophy & Good, 1986; Edmonds, 1979; Madaus, Kellaghan, Rakow, & King, 1979; Rutter e.a., 1979; Wilson & Corcoran, 1988). Aan de basis hiervan ligt het onderzoek van Rosenthal en Jacobson (1968) op een basisschool met voornamelijk leerlingen uit lagere sociale milieus. De verwachtingen die een docent over een leerling heeft beïnvloedt, zo vermoeden zij, de houding van de docent tegenover de leerling. De verwachtingen werken op verscheidene manieren

door in het gedrag van docenten tegenover leerlingen, bijvoorbeeld door positieve of negatieve gevoelens, de hoeveelheid informatie en terugkoppeling die zij geven en het aantal keren dat ze een vraag stellen. Volgens Rosenthal en Jacobson is het cruciale hieraan dat docenten die hoge verwachtingen over hun leerlingen koesteren en dat ook naar alle leerlingen uitdragen, betere schoolprestaties bereiken dan docenten met lage verwachtingen. Als een docent zich eenmaal een negatief beeld van leerlingen heeft gevormd, is hij geneigd om dergelijke leerlingen minder aandacht te geven en minder moeilijke taken op te geven. Dit principe noemen zij het Pygmalion-effect en het effect dat dit veroorzaakt staat bekend als een zichzelf waarmakende voorspelling (self-fulfilling prophecy), een voorspelling die een reeks gebeurtenissen beïnvloedt waardoor de aanvankelijke verwachting uitkomt. Een ander kenmerk dat in deze cirkel zou passen is positieve bekrachtiging (Walberg, 1984). Door op antwoorden of huiswerk van leerlingen positief te reageren is het volgens Walberg voor een docent mogelijk om zijn betrokkenheid bij leerlingen te laten blijken, een prettige sfeer in de klas te scheppen, gewenst gedrag te stimuleren en ongewenst gedrag bij te sturen. Leerlingen gaan zich door deze bekrachtiging competentier voelen. Naar hoge verwachtingen en positieve bekrachtiging kan met gegevens uit VOCL'93 geen onderzoek worden gedaan.

<sup>vi</sup> Enigszins vergelijkbaar is de instructiemethode van Bloom die bekend staat als mastery learning (Block & Burns, 1976). Ook bij deze methode wordt gebruik gemaakt van duidelijke doelen, kleine eenheden in het curriculum, een weloverwogen volgorde in de leerstof en remediale hulp voor lacunes in de kennis of vaardigheden van leerlingen. Het belangrijkste verschil met gestructureerd onderwijs is volgens Bloom dat mastery learning het behalen van een standaardniveau voor alle leerlingen beoogt. Elke persoon kan een bepaald niveau bereiken, zo meent Bloom, mits de condities waaronder hij iets moet leren geschikt zijn. Voor het effect van deze methode is het empirisch houvast stevig.

<sup>vii</sup> In extra analyses blijken deze verwachtingen te worden bevestigd. Scholen in Nederland blijken niet of nauwelijks te verschillen in de mate waarin ze *basisvaardigheden benadrukken* (geoperationaliseerd als het aantal uren Nederlands en wiskunde dat leerlingen per week krijgen) en als er al verschillen zijn, vertaalt dat zich niet in de prestaties van leerlingen. Van het *professioneel en onderwijskundig leiderschap* gaat evenmin een effect uit op prestaties Nederlands en wiskunde. Net als voor effecten van docentkenmerken, zie paragraaf 4.3, geldt dat VOCL'93 waarschijnlijk niet zo geschikt is om deze effecten na te gaan. Met VOCL'93 kan alleen worden gekeken naar de tijdsindeling van de directie, waarbij vier categorieën worden onderscheiden: (1) onderwijsinhoudelijk, (2) administratief en financieel; (3) personeel; (4) interne en externe contacten (zie Bijlage 6 voor de oorspronkelijke vraag). In hoeverre de schoolleiding de structuur van en de relaties binnen de school bewaakt en tijd en energie steekt in activiteiten om de school te verbeteren, valt waarschijnlijk moeilijk te achterhalen met een schriftelijke vragenlijst. De *bewaking van leervorderingen* heeft ook geen effect. Over een leerlingvolgsysteem beschikt 87 procent van de scholen niet, 12 procent gedeeltelijk en 1 procent wel. Op dit kenmerk is er dus weinig variatie tussen scholen. Van de *samenwerking tussen school en thuis* gaat evenmin een effect uit. Ook voor dit kenmerk geldt dat er weinig variatie tussen scholen is. De band tussen school en ouders is in het voortgezet onderwijs formeel. Scholen organiseren voor ouders jaarlijks enkele ouder- en thema-avonden. De opkomst op die avonden is volgens de schooldirecties overwegend goed en verschilt tussen scholen niet altijd te sterk. Voor zover er verschillen zijn heeft dat geen invloed op de prestaties van leerlingen.

<sup>viii</sup> Het effect van het percentage allochtone leerlingen hoeft niet lineair te zijn. Om na te gaan of er een ander verband is, heb ik dit kenmerk gecategoriseerd in: (1) minder dan 6 procent; (2) 6 tot

en met 20 procent; (3) 21 tot en met 40 procent; (4) meer dan 40 procent. Vervolgens heb ik deze categorieën met drie dummy's gerepresenteerd. Het blijkt dat het gebruik van dummy's geen beter model oplevert.

<sup>ix</sup> De aard van de brugperiode hangt sterk samen met de duur van de brugperiode. In homogene brugklassen is de duur nul jaar, in heterogene brugklassen doorgaans één jaar en in dakpansgewijze brugklassen overwegend twee jaar. Als aparte variabele of in combinatie met de aard van de brugperiode, zo blijkt uit nadere analyses, voegt de brugduur niets toe aan de multiniveau analyse.

<sup>x</sup> Willms (1985) laat zien dat schooleffecten te groot uitvallen, wanneer te weinig wordt gecontroleerd.

<sup>xi</sup> De partiële correlatie tussen de opleiding van de moeder en het schooltype (waarbij gedaan is alsof het een metrische variabele is met een schaal van 0 voor ivbo tot 8 voor vwo) onder constanthouding van de scores op de prestatietoetsen taal, informatieverwerking en rekenen is 0,19.

<sup>xii</sup> In plaats van op klasniveau kunnen dezelfde vijf leerling- en gezinskenmerken ook op schoolniveau worden geaggregeerd. Er blijken dan soortgelijke regressie-effecten tussen scholen te zijn. Ook deze effecten verdwijnen zodra het schooltype aan de analyses wordt toegevoegd.

<sup>xiii</sup> Wanneer de denominatie wordt opgenomen, gaat de deviance omlaag met 18,1 bij zeven vrijheidsgraden ( $p < 0,05$ ).

<sup>xiv</sup> Dit effect was niet voorzien. Het is daarom verstandig om met andere datasets na te gaan of gereformeerd-vrijgemaakte mavo-scholen een aparte positie innemen.

<sup>xv</sup> In Bijlage 4B valt op dat er negatieve, significante effecten zijn van reformatorisch onderwijs op de prestaties Nederlands van vbo- en havo-leerlingen. In een analyse met alle schooltypen valt dit effect echter weg.

<sup>xvi</sup> De aard van de brugperiode is een variabele op schoolniveau. Met de gegevens van VOCL'93 hanteren Van der Werf, Lubbers en Kuyper (1999) een soortgelijke variabele op individueel niveau: het door leerlingen gevolgde *klatype*. Rekening houdend met enkele structurele kenmerken (sociaal-economische status en etniciteit), culturele kenmerken (prestatiemotivatie, schoolbeleving en eerdere prestaties) en het door leerlingen gevolgde schooltype, laten zij zien dat leerlingen die in een homogene brugklas hebben gezeten significant beter presteren op Nederlands dan leerlingen in een dakpansgewijze of heterogene brugklas. Voor wiskunde is er geen effect van het klastype. Daarentegen is het voor de onderwijspositie in leerjaar vijf, uitgedrukt in een score op de leerjarenladder, juist relatief minder gunstig om in een homogene brugklas te hebben gezeten. Van der Werf, Lubbers en Kuyper (1999) komen tot de slotsom dat een brugklas waarin een type is vertegenwoordigd dat hoger ligt dan het advies relatief gunstig is voor de onderwijspositie in het vijfde leerjaar. Hierbij geldt als kanttekening dat deze auteurs geen random effect hebben opgenomen voor de brugklas waarin de leerling zat. Zij onderscheiden ten onrechte alleen een individueel en schoolniveau.