

6 | Conclusie

6.1 Leerlingen – klassen - scholen

Met gegevens van ongeveer 7.000 leerlingen afkomstig uit 450 klassen van 150 scholen voor voortgezet onderwijs ben ik nagegaan in hoeverre verschillen in prestaties en vorderingen Nederlands en wiskunde kunnen worden verklaard door kenmerken van leerlingen, gezinnen, docenten en scholen. Om deze vraag te beantwoorden heb ik rekening gehouden met de geneste structuur van de gegevens. Hierbij is gebruik gemaakt van multiniveau analyse. De informatie komt uit het cohortonderzoek VOCL'93. Deze dataset bevat gegevens van leerlingen die in het schooljaar 1993-'94 op het ivbo, het vbo, de mavo, de havo of het vwo zaten. Voor de leerlingen die in de analyses zijn meegenomen geldt dat ze in het schooljaar 1995-'96 in leerjaar drie zaten. De analyses zijn dus gebaseerd op leerlingen die onvertraagd zijn doorgestroomd.

Een geneste structuur

Multiniveau analyse heeft verscheidene voordelen ten opzichte van analyses op één niveau, zoals een regressieanalyse met geaggregeerde of gededaggregeerde gegevens. Door rekening te houden met de hiërarchie van gegevens (leerlingen binnen klassen binnen scholen) zijn de uitkomsten correcter en beter interpreteerbaar dan die van een analyse op één niveau. Een ander voordeel van multiniveau analyse is dat er naar differentiële effecten kan worden gekeken. Zo blijkt het effect van etniciteit op prestaties niet in alle klassen hetzelfde (paragraaf 4.4.2). In sommige klassen is dit effect sterker en in andere klassen zwakker, wat inhoudt dat voor allochtone leerlingen de ene klas gemiddeld genomen gunstiger en de andere klas ongunstiger is. Een derde voordeel is dat er met multiniveau analyse een onderscheid kan worden gemaakt tussen binnenschoolse en tussenschoolse regressie. Na uitvoerige controle op individueel niveau bleken er in dit onderzoek geen regressie-effecten tussen klassen of scholen te zijn. Het percentage allochtone leerlingen kan overigens wel als een voorbeeld van een tussenschoolse regressie-effect worden beschouwd (zie paragraaf 4.5). Niet alleen op individueel maar ook op schoolniveau blijkt er namelijk een effect van etniciteit te zijn. Onder constanthouding van de kenmerken van leerlingen en gezinnen, scoren leerlingen op scholen met veel allochtonen slechter dan scholen met weinig allochtonen. Van de

compositie van de school gaat dus een effect uit. Verder geldt dat met multiniveau analyse scholen gemakkelijk te rangordenen zijn door gebruik te maken van zogeheten ‘posterior means’. In de multiniveau analyse in hoofdstuk 2 zijn met deze schattingen rapportcijfers van scholen verkregen.

Met het Trouw-onderzoek naar schoolprestaties heb ik geïllustreerd dat de schattingen van de netto prestaties van scholen sterk verschillen voor analyses op één dan wel twee niveaus. Voor meer dan de helft van de scholen geldt dat de rapportcijfers van scholen één of meer punten verschillen (op een schaal van vier tot acht) naar gelang de keuze van het aantal niveaus. Wanneer er in onderzoek geen rekenschap wordt gegeven van de aanwezigheid van een geneste structuur, dan is bij de interpretatie van de uitkomsten op zijn minst voorzichtigheid geboden. Succesmaten uit een analyse op één niveau, zoals de rapportcijfers uit het Trouw-onderzoek, geven dan ook zeker geen betrouwbaar beeld van de prestatieverschillen tussen scholen.

Doorgaans leiden analyses op één niveau tot fouten, wanneer de structuur genest is. Vandaar dat een onderzoeker er niet aan ontkomt om een multiniveau analyse te verrichten. In de discussie over openbare schoolgegevens van scholen komen Bosker en Scheerens (1999) tot dezelfde conclusie. Waar het om draait bij het bepalen van de door de school toegevoegde waarde (*posterior means*), zo stellen zij, is de vaststelling van de gemiddelde mate waarin individuele leerlingen boven of onder de statistische verwachting presteren op een school. Dat is iets anders dan de vaststelling of een schoolpopulatie als geheel verschilt van de verwachte score.

In hoofdstuk 2 heb ik ook laten zien dat de uitkomsten eveneens sterk verschillen tussen een analyse met bruto of netto prestaties. Een indicator voor de bruto prestatie van een school is bijvoorbeeld het gemiddelde niveau dat leerlingen op Nederlands en wiskunde hebben. Om op eerlijke wijze zichtbaar te maken ‘wat scholen er toe doen’ is het echter nodig om rekening te houden met de omstandigheden waarin scholen verkeren. Er moeten dus netto in plaats van bruto prestaties worden gehanteerd. De netto prestatie van een school is dan bijvoorbeeld het gemiddelde niveau van leerlingen op Nederlands en wiskunde, nadat er gecontroleerd is voor verschillen in de samenstelling van leerlingpopulaties. Met gegevens op één niveau, zoals het in Trouw-onderzoek, blijkt het slechts ten dele mogelijk om de netto prestatie van een school vast te stellen. Wanneer er echter gegevens op zowel individueel als schoolniveau worden verzameld, dan kan de netto prestatie van een school in een multiniveau analyse worden bepaald.

Uit de multiniveau analyses in hoofdstuk 3 en 4 blijkt dat in Nederland de bruto variantie tussen klassen en scholen hoog is. In dit onderzoek is deze variantie 51

procent voor Nederlands en 65 procent voor wiskunde. De variantie tussen klassen en scholen is zo hoog, omdat in het Nederlandse voortgezet onderwijs leerlingen worden opgedeeld naar schooltype. In *Education at a glance* (OECD, 1998) blijkt dat Nederland wat relatieve tussenschoolse variantie in het voortgezet onderwijs betreft internationaal koploper is. Landen die hier ook hoog op scores zijn Ierland, Duitsland, Zwitserland, Oostenrijk en het Vlaamse deel van België. In Korea en Japan is de bruto variantie tussen klassen en scholen daarentegen minder dan 5 procent. Deze maat zegt overigens niets over de ongelijkheid in prestaties binnen een land. Tussen de tussenschoolse variantie (een relatieve maat) en de algehele variatie in prestaties (een absolute maat) blijkt geen duidelijke relatie te bestaan (OECD, 1998).

Het individuele niveau: prestaties

Voor een verklaring van verschillen in prestaties zijn de gegevens op individueel niveau het belangrijkste. Het niveau waarop leerlingen presteren op Nederlands en wiskunde wordt bepaald door zowel structurele als culturele kenmerken. Zelfs als er rekening wordt gehouden met de eerdere prestaties van leerlingen, is het om na te gaan welke leerlingen een voorsprong of achterstand hebben van belang om te weten wat de *opleiding van hun ouders* is en of ze *autochtoon of allochtoon* zijn. Wanneer het effect van deze kenmerken in samenhang wordt bekeken, dan is er sprake van een sterk effect. Een autochtone leerling met ouders met een hogere beroepsopleiding komt bijvoorbeeld een kwart standaarddeviatie hoger uit dan een allochtone leerling met ouders met alleen basisonderwijs. Ter vergelijking, het verschil in prestaties tussen leerlingen die een niveau in schooltype van elkaar verschillen, bijvoorbeeld een mavo- en een havo-leerling, is ongeveer een halve standaarddeviatie. Het verschil tussen allochtonen en autochtonen valt niet weg, wanneer er rekening wordt gehouden met de kenmerken van de sociaal-economische status. Etniciteit heeft dus een extra effect op prestaties en dan niet alleen op Nederlands maar ook op wiskunde.

Onder constant houding van de overige kenmerken heeft de *sekse* eveneens een sterk effect op de prestaties van leerlingen. Meisjes blijken gemiddeld genomen beter te scoren op Nederlands en jongens op wiskunde.¹ Van het opgroeien in een *tweeoudergezin* gaat een positief effect uit op de wiskundeprestaties.

Naast effecten van structurele kenmerken zijn er ook effecten van culturele kenmerken. Van de interactie tussen ouders en leerlingen hebben de *huiswerkcontrole* en de *autoritatieve opvoedingsstijl* een behoorlijke samenhang met de prestaties van leerlingen. Leerlingen die niet door hun ouders worden aangespoord tot en gecontroleerd op het maken van huiswerk en die de helft van de beslissingen na overleg met hun ouders nemen (er wordt dus geregeld gebruik gemaakt van de au-

toritatieve opvoedingsstijl) scoren voor Nederlands 0,33 standaarddeviatie en voor wiskunde 0,18 standaarddeviatie hoger dan leerlingen die geregeld te maken hebben met huiswerkcontrole en die niet autoritatief worden opgevoed.

Van de algemene kenmerken van leerlingen heeft het *zelfbeeld* van leerlingen een sterk positief effect. Dit effect is net zo groot als het gezamenlijke effect van het opleidingsniveau van beide ouders. Ook het *leefpatroon* heeft een sterke invloed op de prestaties van leerlingen. Leerlingen die de helft van hun vrije tijd besteden aan uitgaan, een bijbaantje of op straat met vrienden rondhangen en die geen tijd besteden aan lezen of hobby's scoren een kwart standaarddeviatie lager op Nederlands en een vijfde standaarddeviatie lager op wiskunde dan leerlingen die 20 procent van hun vrije tijd besteden aan lezen of hobby's en die geen tijd besteden aan de drie andere aspecten.

De grootste effecten op de prestaties komen natuurlijk van de effecten van de *eerdere prestaties*. Wanneer een leerling bijvoorbeeld op alledrie de toetsen (taal, informatieverwerking en rekenen) één standaarddeviatie boven het gemiddelde scoort, presteert een leerling twee jaar later op havo-niveau. Bij twee standaarddeviaties boven het gemiddelde scoort die leerling in leerjaar drie op vwo-niveau. Een verschil in één standaarddeviatie resulteert in een verschil van ongeveer één schooltype. De eerdere prestaties zijn daarmee een goede voorspeller van de latere prestaties van leerlingen. Voor de andere kenmerken uit cirkel één geldt dat het aantal dagen huiswerk een behoorlijke invloed heeft. Wanneer een leerling in plaats van vijf maar twee dagen in de week huiswerk maakt resulteert dat in een afname van de prestaties op beide toetsen met eentiende standaarddeviatie.

Wanneer de verschillende cirkels met elkaar worden vergeleken, dan blijkt dat de prestaties van leerlingen meer worden verklaard door kenmerken van leerlingen dan van gezinnen. Uit het literatuuroverzicht komen veel meer kenmerken van leerlingen dan van de gezinsinteractie naar voren. Ook blijkt dat de verklaarde variantie veel sterker toeneemt door de opname van leerlingkenmerken (cirkel 1' en 2) dan door gezinskenmerken (cirkel 3 en 4). Voor leerlingkenmerken geldt verder dat de prestaties van leerlingen meer door schoolse dan door algemene kenmerken worden verklaard. Deze conclusie kan voor gezinskenmerken niet worden getrokken. Voor de verklaring van verschillen in prestaties blijken culturele kenmerken verder belangrijker dan structurele kenmerken. De structurele kenmerken binden minder dan 10 procent van de variantie. De culturele kenmerken verklaren vier à vijf keer zo veel variantie.

Het *schooltype* heeft ook een zelfstandig effect op de leerprestaties. Dit kenmerk geeft een indicatie van zowel het prestatieniveau van een leerling als de gelegenheid tot leren. Naarmate het schooltype hoger is, krijgen leerlingen een moeilijker leerstof aangeboden, waardoor bijvoorbeeld vwo-leerlingen meer ge-

legenheid tot leren hebben dan overige leerlingen. Het schooltype heeft niet alleen een extra effect. Het blijkt ook andere effecten te mediëren, zoals het effect van het opleidingsniveau van ouders of het aantal dagen waarop door leerlingen huiswerk wordt gemaakt.

Het individuele niveau: vorderingen

In het voorgaande stond de vraag centraal in hoeverre de prestaties van leerlingen op één bepaald tijdstip kunnen worden verklaard uit allerlei kenmerken van leerlingen en gezinnen. Om de netto prestaties van leerlingen te bepalen wordt hierbij rekening gehouden met de prestaties van leerlingen op een eerder tijdstip. De toetsscores in leerjaar drie zijn de afhankelijke variabelen en de eerdere prestaties in leerjaar één behoren tot de onafhankelijke variabelen. Een dergelijk model is standaard in het schooleffectiviteitsonderzoek en staat bekend als het *covariantie-analytisch model*.

Met het covariantie-analytisch model kunnen prestaties van leerlingen op een bepaald tijdstip worden verklaard. In deze verklaring gaat het om de vraag welke leerlingen een achterstand of voorsprong *hebben*. Vervolgens kan dan worden nagegaan welke kenmerken invloed hebben op de *leerprestaties*. Met dit model wordt dus naar invloeden op leerprestaties en niet naar invloeden op leervorderingen gekeken. Wanneer een onderzoek op dit model is gebaseerd, kan dan ook niet over effecten op vooruitgang in prestaties worden gesproken. Om naar vooruitgang in prestaties te kijken, moet worden gekozen voor een *variantie-analytisch model van verschillen in relatieve posities*. De afhankelijke variabelen hierbij zijn bijvoorbeeld de *verschilscores* tussen de gestandaardiseerde prestaties in leerjaar drie en één.

De uitkomsten van de covariantie-benadering en de verschilscore-benadering blijken op individueel niveau sterk van elkaar af te wijken. Structurele kenmerken (ethniciteit, opleidingsniveau van vader en moeder of aantal ouders) en een deel van de culturele kenmerken (autoritatieve opvoedingsstijl, tijd besteden aan lezen en hobby's, mening over klasgenoten en non-verbale intelligentie) hebben wel een effect op de toetsscores Nederlands en wiskunde, maar geen effect op de verschillen in relatieve positie voor beide vakken. Omgekeerd is er in de analyse van verschillen in relatieve posities juist een effect van een cultureel kenmerk als de prestatiemotivatie, terwijl dat effect in de covariantie-benadering ontbreekt. Behalve van deze leerling- en gezinskenmerken gaat er ook van het schooltype een zwakke invloed uit op de vorderingen van leerlingen.

Uit de analyses blijkt dat op de relatieve positie op beide vakken structurele kenmerken nauwelijks een effect hebben. Afwijkend van wat vaak wordt aangenomen, bijvoorbeeld Hacquebord (1989), wordt de achterstand van allochtone

leerlingen niet groter in het voortgezet onderwijs. Daarentegen hebben verscheidene culturele kenmerken wel een effect op de vorderingen Nederlands en wiskunde. De veranderingen in prestaties komen door kenmerken van leerlingen, zoals hun inzet voor school of hun leefpatroon.ⁱⁱ Vanuit meritocratisch oogpunt is dit een gunstige uitkomst. Kenmerken die gegeven zijn, zoals de sociaal-economische status en de etniciteit, hebben geen invloed op verschillen in relatieve posities, maar culturele kenmerken wel.

Toch moet de invloed van deze veranderbare kenmerken niet worden overschat. De culturele kenmerken verklaren namelijk weinig van de variantie in verschillen in relatieve posities. Individuele kenmerken hebben dus weinig invloed op leerwinst. Bij de intrede in het voortgezet onderwijs nemen leerlingen een bepaalde positie qua prestaties in. In die posities treden tijdens de eerste drie leerjaren van het voortgezet onderwijs alleen in beperkte mate veranderingen op. Ter indicatie, de correlatie tussen de scores op de toetsen uit leerjaar één en drie is 0,55 voor Nederlands en 0,74 voor wiskunde.

Het klas- en schoolniveau: prestaties en vorderingen

Op klas- en schoolniveau zijn de overeenkomsten tussen de covariantie- en verschillscore-benadering sterker. De kenmerken die van invloed zijn op de prestaties in leerjaar drie en op de vorderingen tussen leerjaar één en drie zijn in grote lijn vergelijkbaar.

Voor beide benaderingen geldt dat na controle voor kenmerken op individueel niveau scholen in geringe mate van elkaar verschillen. De netto variantie in prestaties (waarbij gecontroleerd wordt voor leerling- en gezinskenmerken en het door leerlingen gevolgde schooltype) tussen scholen is 12 procent voor Nederlands en 19 procent voor wiskunde. De netto variantie op schoolniveau in de verschillscore-benadering is 9 procent voor Nederlands en 12 procent voor wiskunde. In de verschillscore-benadering geldt ook voor de bruto variantie dat slechts een klein deel van de variantie te maken heeft met systematische verschillen tussen scholen. De relatieve tussenschoolse variantie is 6 en 13 procent voor respectievelijk Nederlands en wiskunde. In de covariantie-benadering zijn de percentages veel hoger, namelijk 29 en 38 procent (tussen klassen en scholen zelfs 51 en 65 procent). Leerprestaties hangen namelijk sterk samen met de samenstelling van scholen, terwijl leervorderingen veel minder door de compositie van een school worden beïnvloed. Verder geldt in beide benaderingen dat er meer relatieve tussenschoolse variantie is voor wiskunde dan voor Nederlands, wat komt doordat prestaties en vorderingen op wiskunde meer samenhangen met wat er op school gebeurt dan die op Nederlands.

Van de *contextuele kenmerken* blijken het percentage allochtone leerlingen en de locatiegrootte beide invloed te hebben op de *prestaties* van leerlingen. Voor leerlingen in alle schooltypen geldt dat het percentage allochtone leerlingen een negatief effect heeft op hun wiskunde-prestaties. Door een toename van een standaarddeviatie in het percentage allochtone leerlingen, gaat de score op de toets met 0,03 standaarddeviatie achteruit. Voor leerlingen die op de havo of het vwo zitten geldt dat ze niet alleen op wiskunde maar ook op Nederlands slechter presteren naarmate het percentage allochtone leerlingen op school stijgt. Voor elke standaarddeviatie verschil in het percentage allochtone leerlingen, verandert de score op de toets met 0,09 standaarddeviatie. Het percentage allochtone leerlingen hangt samen met talrijke andere kenmerken van de leerlingensamenstelling. Op zwarte scholen is het opleidingsniveau van ouders en het instroomniveau van leerlingen lager dan op witte scholen. In dit onderzoek blijkt het resultaat van dergelijke verschillen te zijn dat de prestaties van leerlingen lager zijn naarmate er meer allochtone leerlingen op een school zitten. Voor vervolgonderzoek zou het interessant zijn om na te gaan waarom het percentage allochtone leerlingen een negatief effect heeft op prestaties van leerlingen, en waarom de prestatie-achterstand voor havo/vwo-leerlingen niet alleen voor wiskunde maar ook voor Nederlands geldt. 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? En waarom zijn er verschillen in de effecten voor Nederlands en wiskunde?

Voor leerlingen in het beroepsonderwijs maakt de locatiegrootte een verschil uit voor de prestaties op beide vakken. Wanneer het aantal leerlingen op een zelfde locatie toeneemt van 300 tot 600 of van 600 tot 1200, dan leidt dit tot een afname in prestaties van 0,07 standaarddeviatie op beide vakken. Aangezien vele leerlingen de gevolgen merken van een dergelijk schooleffect, moet dit effect als relevant worden beschouwd. Voor mavo-, havo- en vwo-leerlingen is er geen effect van de locatiegrootte op de prestaties.

In onderzoek buiten Nederland worden geregeld verschillen in prestaties aangetoond tussen openbaar en bijzonder onderwijs. Die verschillen zijn er dan voor zowel bruto als netto prestaties. Uit dit onderzoek blijkt dat er in Nederland verschillen in netto prestaties tussen denominaties zijn. In tegenstelling tot bijvoorbeeld Amerikaans onderzoek blijken openbare scholen niet slechter te presteren dan bijzondere scholen. Binnen de bijzondere sector blijken er met name verschillen te zijn. Hierbij valt vooral de achterstand van leerlingen op interconfessionele scholen op.

Op de *vorderingen* van leerlingen blijkt er van het percentage allochtone leerlingen geen en van de locatiegrootte en de denominatie wel een effect uit te gaan.

Het percentage allochtone leerlingen is, zoals eerder aangegeven, wel gerelateerd aan prestaties van leerlingen in leerjaar drie van het voortgezet onderwijs. Deze relatie hangt waarschijnlijk vooral samen met de achterstand die leerlingen op scholen met veel allochtone leerlingen hadden bij aanvang van het voortgezet onderwijs en niet met de processen in de onderbouw van het voortgezet onderwijs. De achterstand van leerlingen op ‘zwarte scholen’ wordt namelijk niet groter tussen leerjaar één en drie.

In de covariantie-benadering heeft de locatiegrootte alleen een negatief effect op de prestaties van leerlingen in het beroepsonderwijs. In de verschilscore-benadering wordt echter een negatief effect voor alle leerlingen gevonden. Dit negatieve effect duidt erop dat de processen op scholen verschillen naar gelang de locatiegrootte. Een effect van de denominatie is ook aanwezig. Opnieuw zijn er geen significante verschillen tussen openbare, protestants-christelijke en rooms-katholieke scholen. Wel valt op dat interconfessionele scholen opnieuw negatief naar voren komen. Leerlingen op dit type scholen blijken een wiskunde-achterstand op te lopen.

Van de *schoolkenmerken* zijn er enkele significante effecten. De teamstabiliteit blijkt een zwak positief effect te hebben op de leerprestaties op beide vakken. Door elk jaar dat het schoolteam langer bijeen is, blijken de prestaties van leerlingen met eentwintigste standaarddeviatie vooruit te gaan. Op de leervorderingen heeft dit kenmerk geen effect.

Wanneer scholen lesuitval opvangen door een invalregeling of een verplichte huiswerkles, dan blijken vbo/mavo-leerlingen op Nederlands een vijfde tot een derde standaarddeviatie hoger te presteren dan wanneer het uur komt te vervallen. Een invalregeling blijkt ook op de vorderingen van vbo/mavo-leerlingen een effect te hebben, maar het effect van een verplichte huiswerkles blijkt in de verschilscore-benadering niet significant.

Met uitzondering van vwo-ers, scoren leerlingen in heterogene brugklassen in zowel de covariantie- als de verschilscore-benadering lager op Nederlands dan leerlingen in een dakpansgewijze of homogene brugklas. Wanneer leerlingen van drie of meer schooltypen bij elkaar zaten in het eerste leerjaar van het voortgezet onderwijs, dan hebben die leerlingen in het derde leerjaar een achterstand opgelopen voor Nederlands. In de praktijk groeperen scholen leerlingen overigens steeds minder heterogeen. Meer dan de helft van de leerlingen zit in een homogene brugklas. Deze leerlingen hebben hun keuze voor een schooltype dus al voor het voortgezet onderwijs moeten maken. In een heterogene brugklas zit één op de tien leerlingen. De overige leerlingen zitten in een tussenvorm, een dakpansgewijze brugklas (leerlingen van twee schooltypen in dezelfde klas).

Wat de vormgeving van het mentoraatsysteem betreft valt op dat het verband in beide benaderingen er als volgt uit ziet: hoe uitgebreider dat systeem is opgezet, hoe slechter de prestaties van leerlingen. Alle vormen scoren lager dan een smalverticaal mentoraatsysteem, de referentiegroep. Tot slot, blijkt dat (i)vbo-leerlingen slechter presteren naarmate de docent-leerling-ratio hoger is. Hier geldt dat vbo-leerlingen net als ivbo-leerlingen baat hebben bij een structuur waarin ze niet al te veel verschillende docenten hebben. In de verschilscore-benadering wordt dit effect alleen voor Nederlands gevonden. In het algemeen bleek in deze analyses dat het niet of nauwelijks mogelijk is om de invloed van contextuele kenmerken te verklaren door schoolkenmerken.

6.2 Wetenschappelijke relevantie

De conclusie dat leerling- en gezinskenmerken bepalend zijn voor verschillen in prestaties is in lijn met onderzoek van Coleman e.a. (1966) en Jencks e.a. (1972). In dertig jaar tijd is er in dit opzicht niet veel veranderd. Technisch gezien is er echter vooruitgang geboekt door de ontwikkeling van *multiniveau analyse*. Zoals aan het begin van dit hoofdstuk vermeld heeft multiniveau analyse verscheidene voordelen boven een één-niveau analyse van hiërarchische structuren.

In onderwijsonderzoek is het nog niet gebruikelijk om verscheidene afhankelijke variabelen simultaan in een *multivariaat model* op te nemen. In dit onderzoek zijn de analyses tegelijkertijd voor Nederlands en wiskunde verricht. Een voordeel hiervan is dat er kan worden getoetst of onafhankelijke variabelen een zelfde effect hebben op de twee afhankelijke variabelen. In de tweede plaats kunnen de correlaties tussen de afhankelijke variabelen op verschillende niveaus worden bepaald. Ook geldt dat leerlingen met een ontbrekende waarde op een van de afhankelijke variabelen in de analyse kunnen worden meegenomen, waardoor incomplete datasets efficiënter kunnen worden gebruikt. Een multivariaat multiniveau model heeft dus belangrijke voordelen boven de modellen die enkele decennia terug werden gehanteerd. Door de ontwikkeling van deze modellen kan er dus beter en uitgebreider onderzoek worden gedaan naar verschillen in leerprestaties en leervorderingen tussen leerlingen.

Door de ontwikkeling van multiniveau analyse is het goed mogelijk om naar de invloed van leerlingen, gezinnen, docenten en scholen in *samenhang* te kijken. Uit dit onderzoek blijkt dat het belangrijk is om op individueel niveau uitgebreid te controleren. In studies naar effectieve docenten en scholen blijkt die controle vaak beperkt. Met structurele kenmerken en eerdere prestaties van leerlingen houden effectiviteitsonderzoekers wel rekening, maar met culturele kenmerken niet.

Voor de verklaring van verschillen in prestaties tussen leerlingen zijn deze culturele kenmerken echter minstens zo belangrijk als structurele kenmerken. Voor de verklaring van toe- of afname in relatieve positie geldt zelfs dat structurele kenmerken niet of nauwelijks, maar culturele kenmerken wel van belang zijn. Om over effecten op klas- en schoolniveau uitspraak te doen moet daarom op individueel niveau met zowel structurele als culturele kenmerken rekening worden gehouden.

In onderzoek met verscheidene kenmerken op verschillende niveaus is het gevaar dat het onderzoek te sterk door de empirie wordt gestuurd. Met een sleepnet trekt de onderzoeker alle gegevens voort. Na verschillende analyses blijven er dan enkele kenmerken hangen en die zijn blijkbaar van invloed op de afhankelijke variabele. Om dat gevaar zo veel mogelijk te ontlopen is er in dit onderzoek een *structuur van concentrische cirkels* gehanteerd. Op grond van theoretische overwegingen zijn er kenmerken in deze structuur opgenomen. Met deze structuur zijn vervolgens cirkel voor cirkel analyses uitgevoerd. Met het gebruik van deze structuur kunnen de verschillende invloeden goed worden geordend en kan ook inzicht worden gegeven in de (causale) relaties tussen de kenmerken onderling. Zo blijken de effecten van kenmerken als de opleiding van de moeder of de etniciteit voor een belangrijk deel te verklaren door meer direct aan prestaties gerelateerde kenmerken. Beide effecten worden voor ongeveer drievijfde verklaard door culturele kenmerken. Door de toevoeging van het schooltype worden de effecten van de opleiding van de moeder en van de vader zelfs volledig verklaard. Voor de *verklaring* van verschillen in prestaties op individueel niveau lijkt de structuur van concentrische cirkels dan ook bruikbaar.

Op klas- en schoolniveau beantwoorden de analyses gedeeltelijk aan de verwachtingen vooraf. Op beide niveaus zijn er op voorhand verscheidene kenmerken die voor verschillen in prestaties in de onderbouw van het Nederlandse voortgezet onderwijs van belang kunnen worden geacht. Deze verwachtingen worden voor een deel empirisch bevestigd. Voor verscheidene kenmerken, zeker voor die van docenten, geldt echter dat de beoogde meetinstrumenten uit VOCL'93 niet geschikt bleken. Over de kenmerken van effectieve docenten kan hier dan ook niets worden gezegd.

In het multiniveau model is alleen maar rekening gehouden met de brugklas waarin de leerling zich bevond en niet met de klas waarin de leerling in het derde leerjaar zat. Technisch is het mogelijk om een geneste structuur op te stellen, waarbij leerlingen zich in een bepaalde klas op tijdstip 1 en een andere klas op tijdstip 2 bevinden. Daartoe kan gebruik worden gemaakt van cross-classificaties (Goldstein, 1995²). Praktisch lijkt het echter niet aantrekkelijk om die structuur voor alle 7.000 leerlingen uit dit onderzoek te gebruiken. Voor sommige scholen

geldt dat leerlingen na de brugklas uiteen zwermen over tal van klassen, waardoor het aantal leerlingen per klas drastisch verminderde. Voor andere scholen geldt dat leerlingen bij elkaar in de klas blijven, waardoor een onderscheid tussen de variantie van het eerste en het derde leerjaar niet kan worden gemaakt. Voor een deel van de scholen uit VOCL'93 is het wellicht mogelijk om een analyse met cross-classificaties uit te voeren.

In de analyses zijn (i)vbo-, mavo-, havo- en vwo-leerlingen gecombineerd meegenomen. Het blijkt technisch goed mogelijk om de analyses voor *alle schooltypen tegelijkertijd* te verrichten. Hiervan is het voordeel dat op klas- en schoolniveau het aantal onderzoekselementen niet onnodig wordt verkleind. Door het maken van interacties kunnen bovendien effecten worden opgenomen die specifiek voor een bepaald schooltype gelden. Een onderzoek naar kenmerken van effectieve docenten en scholen hoeft dus niet per schooltype te geschieden.

Tot slot geldt dat er met name op individueel niveau verschillen zijn tussen de uitkomsten van *de covariantie- en de verschillscore-benadering*. Een covariantie-analyse geeft geen antwoord op de vraag welke leerlingen leerwinst boeken. Alleen met een analyse van leerwinst of relatieve posities kan er worden gekeken naar de voor- of achteruitgang in prestaties van leerlingen. Uit een vergelijking van de twee analysemethoden blijkt dat ze elkaar goed aanvullen. Afhankelijk van de vraag moet een onderzoeker een van de methoden kiezen. Wanneer een onderzoeker geïnteresseerd is in welke leerlingen op een bepaald moment een voor-sprong of achterstand hebben in het onderwijs, dan is een covariantie-analyse gewenst. Wanneer een onderzoeker wil weten welke leerlingen voorop of achterop raken in het onderwijs, dan moet er worden gekozen voor een analyse van leerwinst of relatieve posities. In tegenstelling tot een onderzoek naar kenmerken van effectieve leerlingen en gezinnen, blijkt het voor een onderzoek naar kenmerken van effectieve scholen minder uit te maken welk model wordt gekozen.

Technisch en inhoudelijk is er de afgelopen decennia dus vooruitgang geboekt in het onderzoek naar verschillen in prestaties. Het antwoord op de vraag naar *verschillen in prestaties* kan daarom met meer kracht en met meer nuance worden gegeven. Voor een verklaring van verschillen in prestaties zijn kenmerken op individueel niveau cruciaal. Van deze kenmerken is de invloed van culturele kenmerken veel groter dan van structurele kenmerken. Van de culturele kenmerken is de invloed van de schoolse leerlingkenmerken het grootst, gevolgd door de algemene leerlingkenmerken en de gezinskenmerken.

Voor een verklaring van *verschillen in relatieve posities* geldt dat kenmerken op individueel niveau een veel minder goede verklaring vormen. Met uitzondering van de sekse, hebben structurele kenmerken geen invloed op verschillen in relatieve posities. Verscheidene culturele kenmerken hebben wel een invloed op die

verschillen, waarbij geldt dat de effecten van leerlingkenmerken sterker zijn dan die van gezinskenmerken. Tussen algemene en schoolse kenmerken is geen duidelijk verschil in de sterkte van effecten. Wat verklaarde variantie betreft valt op dat de kenmerken op individueel niveau nauwelijks variantie binden. Hieruit kan worden opgemaakt dat nauwelijks bekend is waarom bepaalde leerlingen in de onderbouw van het voortgezet onderwijs voorsprong boeken en anderen achterstand oplopen.

6.3 Maatschappelijke relevantie

Doelmatigheidsprikkel

Het ministerie van Onderwijs, Cultuur en Wetenschappen wil dat scholen zich publiekelijk gaan verantwoorden voor hun prestaties tegenover ouders en de overheid. Het ministerie beoogt op die manier ouders beter voor te lichten, de resultaten van scholen vergelijkbaar te maken en scholen te prikkelen tot betere prestaties.

In het voortgezet onderwijs heeft dit geleid tot de ontwikkeling van kwaliteitskaarten. Bosker en Scheerens (1999) beargumenteren dat het publiceren van opbrengstgegevens van scholen in de praktijk slechts een geringe invloed zal hebben. De rapportcijfers van Trouw suggereren, in tegenspraak met de uitkomsten uit onderwijsonderzoek, een enorme spreiding in kwaliteit tussen scholen. Deze spreiding is echter een artefact van de methode van het Trouw-onderzoek. Wanneer scores worden gestandaardiseerd met een zes als gemiddelde en een standaarddeviatie van één dan heeft automatisch 40 procent van de scholen een score tussen 5,5 en 6,5. Deze scholen krijgen dus een zes als cijfer. Van de overige scholen scoort ongeveer 30 procent lager dan een 5,5 en een zelfde percentage een 6,5 of hoger. Hoe gering de verschillen binnen een schooltype in werkelijkheid ook zijn, er zal op deze manier altijd blijken dat er verschillen tussen scholen zijn.

Voor het belang van openbare prestatiegegevens voor het schoolkeuzeprocess, het afleggen van rekenschap door scholen (*accountability*) en schoolverbetering zien Bosker en Scheerens allerlei haken en ogen.ⁱⁱⁱ De vraag is hoe belangrijk de overheid en ouders die prestatiegegevens vinden. Ook moet worden afgevraagd in hoeverre het schooleffectiviteitsonderzoek duidelijke aanknopingspunten biedt voor schoolverbetering. De empirische ondersteuning voor effecten van docent- en schoolkenmerken is immers beperkt. Voor schooldiagnose of zelfevaluatie is bovendien meer gedetailleerde informatie nodig. Bosker en Scheerens (1999) relativeren daarom de betekenis van het Trouw-onderzoek of de kwaliteitskaart. Ondanks deze relativeringen beschouwen zij openbare prestatiegegevens van scho-

len als een nuttige praktijk die bijdraagt aan de rationaliteit van keuzegedrag, de besturingsprocessen op scholen en het overheidsbeleid. De aandacht voor prestatie-indicatoren is volgens hen een belangrijke prikkel voor scholen om meer opbrengstgericht te gaan werken. Een dergelijke wijziging in de schoolcultuur, zo verwachten Bosker en Scheerens, leidt tot prestatieverbetering.

Kwaliteitskaart en Trouw-onderzoek

De kwaliteitskaart van de onderwijsinspectie vermeldt vanaf het schooljaar 1999-2000 niet alleen het uitstroomniveau (eindexamencijfers), maar ook het instroomniveau (advies voor voortgezet onderwijs). Door zowel het instroom- als uitstroomniveau weer te geven, is volgens de inspectie te zien wat leerlingen op een school leren en of een school meer of minder dan gemiddeld bijbrengt. De speciale omstandigheden waaronder een school werkt, hoeven daarom niet meer te worden meegenomen. Er wordt dus niet meer gecontroleerd voor het percentage allochtone leerlingen of de urbanisatiegraad. Het ministerie volgt hiermee ten dele de kritiek van de Onderwijsraad (1998) op. Het advies voor voortgezet onderwijs is namelijk een minder vage en onduidelijke maat dan het percentage allochtone leerlingen of de urbanisatiegraad. Toch is het advies voor voortgezet onderwijs niet objectief. Allochtone leerlingen en leerlingen uit hogere sociale milieus krijgen gemiddeld genomen hogere adviezen dan hun score op de Cito-toets aangeeft (Dronkers, 1997a). Ook zijn er verschillen in advies naar regio en denominatie (Dronkers, Van Erp, Robijns & Roeleveld, 1998). Om het rendement van een school te meten blijft een aanvangstoets daarom het ideaal. Pas als het instroomniveau op die manier wordt bepaald en er multiniveau gegevens beschikbaar zijn, kan worden nagegaan wat leerlingen op een school leren en wat de toegevoegde waarde van een school is.^{iv}

Van het Trouw-onderzoek zijn inmiddels twee nieuwe versies verschenen (Agerbeek, 1998; 1999b). De berekening van de beoordeling in de tweede editie van het Trouw-onderzoek leek sterk op het eerdere rapportcijfer. In de tweede editie liepen de beoordelingen niet van één tot en met tien, maar werden de mogelijke scores uitgedrukt in plussen en minnen. De beide beoordelingen zijn in elkaar om te rekenen: een zes als rapportcijfer is bijvoorbeeld een nul als beoordeling en een vier of lager als rapportcijfer komt overeen met een dubbele min. Voor de tweede editie van het Trouw-onderzoek heeft het dagblad haar analyses gebaseerd op betere gegevens van scholen. Trouw vermeldt dat scholen de inspectieformulieren nu nauwkeuriger en op een meer vergelijkbare wijze hebben ingevuld. Zo worden de uitkomsten van leerlingen die buiten het reguliere onderwijs vallen, zoals leerlingen in internationale schakelklassen, niet langer

meegerekend. Voor de cijfers geldt bovendien dat ze (op het percentage allochtone leerlingen na) beschikbaar zijn van de afzonderlijke locaties en niet alleen van de school.

In de versie van 1999 heeft Trouw de gegevens van de nieuwe kwaliteitskaarten gebruikt. Met deze gegevens heeft het dagblad wederom een beoordeling met plussen en minnen opgesteld. Deze beoordeling is gebaseerd op de gemiddelde inspectiebeoordeling van drie elementen, te weten het aantal leerlingen dat onvertraagd in de derde klas belandt, het aantal leerlingen dat zonder in de bovenbouw zitten te blijven hun diploma haalt en het gemiddelde eindexamencijfer. Daarbij is rekening gehouden met het advies van leerlingen voor voortgezet onderwijs (en voor het rendement in de bovenbouw ook met het percentage allochtone leerlingen en het percentage leerlingen met een tegemoetkoming in de studiekosten).

Voor de analyses in de twee volgende versies van het Trouw-onderzoek zijn nog steeds alleen maar gegevens op geaggregeerd niveau beschikbaar. De belangrijkste bezwaren uit hoofdstuk 2 zijn dan ook nog steeds van kracht. Van een deugdelijke opzet voor de publieke verantwoording van scholen kan dan ook nog niet worden gesproken.

Menigeen weet wel ongeveer dat voor verschillen in prestaties vooral kenmerken van leerlingen en gezinnen van belang zijn. Of veel mensen ook ervan uit gaan dat de invloed van scholen klein is, is de vraag. In de werkelijkheid verschillen scholen immers op talrijke punten van elkaar, bijvoorbeeld in de grootte van een locatie, de aard van de brugperiode en de inrichting van het mentoraat. Wanneer ouders voor de keuze van een school voor voortgezet onderwijs staan en de keuze hebben uit een aantal scholen, dan is de kans groot dat die scholen zich van elkaar onderscheiden. Dat onderscheid kan in de dagelijkse praktijk merkbaar zijn. Voor het eindresultaat, de prestaties van leerlingen of het rendement van de school, maakt het echter tamelijk weinig uit. Ouders hebben iets te kiezen, maar in dat keuzeproces gaat het er meer om dat de ene school *anders* is dan de andere en er minder om dat sommige scholen *beter* zijn.

De belangrijkste reden hiervoor is waarschijnlijk dat alle scholen hun programma richten op het niveau van het centraal schriftelijk eindexamen. Op die eindexamens bereiden scholen vwo-leerlingen zes jaar voor. Scholen kunnen niet variëren in de duur van dit traject. De enige variatie die mogelijk is, is dat op sommige vwo-scholen leerlingen vaker een extra examenvak nemen of dat op sommige vbo- en mavo-scholen meer leerlingen examen op het hoogste niveau afleggen. Verschillen tussen scholen van een zelfde schooltype maken dus weinig

uit voor de prestaties van leerlingen. Daarentegen zijn verschillen tussen schooltypen (hoofdstuk 4) wel van belang voor de prestaties van leerlingen.

Dit onderzoek werpt ook licht op de vraag naar ongelijkheid in het onderwijs. Hierbij blijkt dat deze vraag een genuanceerd antwoord verlangt. Wanneer wordt gekeken naar *verschillen in prestaties* valt op dat er grote verschillen zijn tussen leerlingen. In leerjaar 3 van het voortgezet onderwijs blijken allochtone leerlingen op Nederlands en wiskunde een achterstand te hebben op autochtone leerlingen. Behalve naar etniciteit zijn er ook duidelijke verschillen naar sociaal-economische status. Jongeren uit lagere sociale milieus scoren lager op de toetsen op beide vakken dan leerlingen uit hogere sociale milieus. Verder blijkt er ook sprake van sekse-ongelijkheid. Jongens presteren beter op wiskunde dan meisjes en, omgekeerd, scoren meisjes hoger op Nederlands dan jongens. Aan het einde van de leerplichtige leeftijd zijn er dus duidelijke verschillen op deze structurele kenmerken.

Wanneer echter naar *verschillen in vooruitgang* wordt gekeken, dan blijkt dat er geen invloed is van de kenmerken etniciteit en sociaal-economische status. Het enige structurele kenmerk dat een effect heeft op leervorderingen is de sekse. Meisjes blijken meer vooruit te gaan op Nederlands dan jongens. In leerjaar 3 van het voortgezet onderwijs blijken er aanzienlijke verschillen te zijn in prestaties, maar deze verschillen komen niet of nauwelijks tot stand door de vooruitgang die leerlingen in de onderbouw boeken. Ongelijkheid naar etniciteit en sociaal-economische status komt tot stand voordat leerlingen aan het voortgezet onderwijs beginnen en wordt niet vergroot tussen leerjaar één en drie, de periode van de basisvorming.^v In lijn met Dronkers (1997a) kan worden gesteld dat de verschillen die bij aanvang aanwezig zijn niet verder toenemen. Achterstandsbeleid voor leerlingen van een andere etnische afkomst of voor leerlingen uit lagere sociale milieus zal daarom gemiddeld genomen geen zin hebben in deze periode. Achterstandsbestrijding zou zich vooral op het leefpatroon van jongeren moeten richten. Van dergelijk beleid moeten geen grote wonderen worden verwacht. Culturele kenmerken, zoals het leefpatroon, hebben weliswaar een invloed op verschillen in relatieve posities, maar deze invloed is klein.

6.4 Naar vervolgonderzoek

In toekomstig onderzoek naar verschillen in prestaties acht ik een grotere *theoretische inbedding* wenselijk. In dit onderzoek heeft theorievorming niet centraal gestaan. Het hanteren van de structuur van concentrische cirkels en het vooraf op-

stellen van hypothesen is mijns inziens al een behoorlijke verbetering op exploratief onderzoek, waarbij allerlei verbanden achteraf worden verklaard.

Hoe moeilijk het is om voor onderzoek met een algemene vraagstelling een theorie op te zetten, blijkt uit het werk van Coleman. Uit een vergelijking van zijn publicaties over rationele-keuze theorie (Coleman, 1990) en onderwijsonderzoek (Coleman e.a., 1966; Coleman, Hoffer & Kilgore, 1982; Coleman & Hoffer, 1987) valt op dat zijn onderwijsonderzoek nauwelijks voortbouwt op zijn rationele-keuzetheorie. In de rapportages van Coleman over ongelijkheid in onderwijskansen of verschillen tussen openbaar en bijzonder onderwijs voert de empirie dan ook de boventoon over de theorie.

Voor meer specifieke vragen is een theoretische sturing eerder mogelijk. Hierbij kan de structuur van concentrische cirkels gebruikt worden als het algemene kader waarbinnen een specifieke vraag valt. Een voorbeeld hiervan is een onderzoek dat ik samen met Dijkstra heb uitgevoerd (Dijkstra & Veenstra, 1999; Veenstra & Dijkstra, 1998). In dat onderzoek ging primair de belangstelling uit naar de invloed van religieuze kenmerken van leerlingen, ouders en scholen op prestaties Nederlands, en fungeerden de kenmerken uit de structuur van concentrische cirkels als controlevariabelen.

In een meer theoretische benadering zou het uitgangspunt kunnen zijn dat het er voor prestaties van leerlingen om gaat waartoe ze in staat zijn (intelligentie, eerdere prestaties), wat ze er zelf aan willen doen (huiswerk maken, leefpatroon) en hoe ze worden beïnvloed in hun preferenties en intenties door hun zelfbeeld of door ouders, klasgenoten, docenten en de regels op school. De afwegingen die leerlingen maken en de beslissingen die ze nemen staan centraal, maar de structuur waarin ze zich bevinden kan beperkingen opleveren. Een dergelijke benadering zou passen binnen de *rationele-keuzetheorie*.

Rationele-keuzetheorie is gebruikt voor het verklaren waarom de schoolloopbanen van leerlingen verschillend verlopen na kritieke fasen. In een onderzoek naar de verschillende keuzemomenten in de schoolloopbaan van leerlingen heeft Boudon (1974) naar voren gebracht dat ouders bij elk van die keuzemomenten een afweging maken tussen kosten en baten. Deze afweging wordt volgens Boudon beïnvloed door het sociaal milieu van de ouders. Gambetta (1987) heeft een soortgelijke theorie gebruikt voor het verklaren van verschillen in voortijdig schoolverlaten. In de Nederlandse onderwijssociologie is vanuit dit gezichtspunt onder anderen onderzoek gedaan door De Vos (1986), die een model ontwikkelde om de invloed van de samenstelling van een schoolklas op leerprestaties te kunnen onderzoeken, Van Liere (1990), die onderzocht wat het advies voor voortgezet onderwijs is voor meer en minder regelcompetente leerlingen, en Maas (1992), die in groep vier van de basisschool de verwijzing van leerlingen naar het speciaal

onderwijs bestudeerde. Voor het verklaren van verschillen in prestaties is een dergelijke theorie voor zover bekend niet gehanteerd. Voor een algemene vraag naar verschillen in leerprestaties en leervorderingen tussen leerlingen zal het echter niet gemakkelijk zijn om een specifieke theorie op te stellen.

Behalve aan verdieping kan er in toekomstig onderzoek ook worden gedacht aan verlenging. Dit onderzoek richt zich op effecten op verschillen in prestaties over een periode van twee jaar. Het zou echter goed zijn om na te gaan wat de effecten op *langere termijn* zijn. In dit opzicht is het gunstig dat het CBS van de leerlingen uit VOCL'93 de loopbaan door het voltijdsonderwijs jaarlijks bijhoudt. Met de in hoofdstuk 2 besproken leerjarenladder als afhankelijke variabele kunnen er dus analyses worden verricht. Hierbij kan met de structuur van concentrische cirkels worden nagegaan wat de positie van leerlingen op de leerjarenladder na verloop van tijd is.^{vi}

Op dit onderzoek kan ook worden voortgebouwd door verbreding. Er kan namelijk naar meer dan invloeden op cognitieve uitkomsten worden gekeken. Nederlands en wiskunde mogen dan twee van de kernvakken van het voortgezet onderwijs zijn, toch zijn de prestaties die leerlingen daar op leveren niet de enige graadmeter voor het succes van leerlingen, docenten of scholen. Er zou bijvoorbeeld ook kunnen worden gekeken naar *andere uitkomsten*, zoals het probleemoplossend vermogen van leerlingen of de mate waarin ze beschikken over sociale en communicatieve vaardigheden (Trier & Peschar, 1995).^{vii} Voor de toetsen Nederlands en wiskunde geldt dat ze tamelijk traditioneel zijn. In de toetsen wordt bijvoorbeeld geen aandacht besteed aan leren leren, het gebruik van moderne media, of de mate waarin leerlingen kunnen samenwerken.

De verwachting is dat factoren die op leerlingniveau cognitieve uitkomsten beïnvloeden doorgaans ook van invloed zijn op non-cognitieve uitkomsten. Uit onderzoek van Grisay (1996) in Frankrijk blijkt dat non-cognitieve uitkomsten sterk worden beïnvloed door leerling- en gezinskenmerken. Verder blijkt de relatieve tussenschoolse variantie in non-cognitieve uitkomsten kleiner dan in cognitieve uitkomsten. Wat de vaardigheden en houdingen van leerlingen betreft zijn er voornamelijk verschillen binnen en nauwelijks verschillen tussen scholen. Hoewel eerdere bevindingen (Creemers, 1991) erop wijzen dat er geen sterke relatie bestaat tussen affectieve en cognitieve onderwijsuitkomsten, blijkt uit het onderzoek van Grisay dat schoolkenmerken die van invloed zijn op cognitieve opbrengsten in grote lijn ook van invloed zijn op affectieve opbrengsten.

Een andere mogelijkheid is om naar keuze- en selectiemomenten in het voortgezet onderwijs te kijken (Dekkers, Bosker & Meijnen, 1998). In dat kader kan worden gekeken naar de keuze van leerlingen voor een schooltype na de brugperiode, de keuze van een richting voor leerlingen in het beroepsonderwijs of van

een vakkenpakket voor leerlingen in het algemeen vormend onderwijs (Dekkers, 1996), en naar de selectie die optreedt door voortijdig schoolverlaten (Bosker & Hofman, 1994; Dekkers & De Wit, 1995; Gambetta, 1987). De vraag in welk schooltype leerlingen na de brugklas terechtkomen is bijvoorbeeld uiterst relevant. In hoofdstuk 4 bleek dat 37 procent van de leerlingen in een dakpansgewijze en 10 procent in een heterogene brugklas zitten. Voor deze leerlingen kan met logistische regressie worden nagegaan waarom sommige leerlingen doorstromen naar het hogere schooltype en andere naar het lagere schooltype. Voor onderzoek naar dergelijke keuzemomenten is de structuur van concentrische cirkels waarschijnlijk ook toepasbaar.

Ook verruiming behoort tot de mogelijkheden voor vervolgonderzoek. De structuur van concentrische cirkels zou ook in onderzoek met *andere datasets* gebruikt kunnen worden. Hierbij kan worden gedacht aan onderzoek in andere typen van onderwijs en aan onderzoek in andere landen.^{viii} Voor een validatie van deze structuur is het noodzakelijk om nieuwe gegevens te gebruiken (Freedman, 1991, p. 307). "Generally, replication and prediction of new results provide a harsher and more useful validation regime than statistical testing of many models on one data set. Fewer assumptions are needed, there is less chance of artifact, more kinds of variation can be explored, and alternative explanations can be ruled out. Indeed, taken to the extreme, developing a model by specification tests just comes back to curve fitting."

Wat de verdieping, verlenging, verbreding of verruiming ook mag inhouden, er zal rekening moeten worden gehouden met de technische vooruitgang die ik in paragraaf 6.2 heb beschreven. In onderwijsonderzoek zal daarom moeten worden gekozen voor multiniveau analyse.^{ix} Wanneer er verschillende afhankelijke variabelen beschikbaar zijn, kan het beste van de voordelen van een multivariaat model gebruik worden gemaakt. Om over effecten op hogere niveaus uitspraken te doen dient verder zo veel mogelijk rekening te worden gehouden met effecten op lagere niveaus. In onderzoek naar effectieve docenten en scholen, moet dus voor zowel structurele als culturele leerling- en gezinskenmerken worden gecontroleerd. Wanneer dergelijke kenmerken in samenhang worden bekeken, verdient het aanbeveling om bijvoorbeeld de structuur van concentrische cirkels te hanteren. Ook geldt dat als het om onderzoek in het voortgezet onderwijs gaat, het verstandig is om analyses voor alle schooltypen tegelijkertijd uit te voeren. Voor sociaal-wetenschappelijk onderzoek geldt tot slot dat er niet alleen naar invloeden op een toestand op een bepaald moment (leerprestaties), maar ook naar invloeden op verandering tussen twee momenten (leervorderingen) moet worden gekeken.

Eindnoten

ⁱ In Groot-Brittannië hebben meisjes sinds kort niet alleen in taal maar ook in wiskunde een voorsprong op jongens (Sammons, 1995). Uit dit onderzoek blijkt dat in Nederland meisjes nog geen algehele voorsprong hebben genomen.

ⁱⁱ Vergelijk het belang dat Harris (1998) aan vriendengroepen toeschrijft voor de ontwikkeling van de persoonlijkheid.

ⁱⁱⁱ Uit onderzoek van Trouw blijkt echter dat de rapportcijfers van Trouw een rol spelen bij de schoolkeuze (Agerbeek, 1999c). Voor het schooljaar 1998-'99 was de aanmelding op alle scholen gemiddeld genomen 9 procent hoger dan het jaar daarvoor. Op scholen met een rapportcijfer hoger dan een zeven steeg de aanmelding echter met 33 procent.

^{iv} Over openbare ranglijsten van scholen merkt Waslander (1999) op dat er in deze discussie niet alleen sprake is van een informatieprobleem (welke gegevens zijn nodig om de kwaliteit van scholen weer te geven?) maar ook van een kennisprobleem. Om een 'wijze' schoolkeuze te maken, moeten ouders ook beschikken over kennis van het onderwijssysteem, het functioneren van scholen, de aanmeldingsprocedures voor scholen, et cetera. Dit probleem, zo stelt zij, is niet op te lossen met meer informatie, maar vereist een meer omvattende aanpak van sociale ongelijkheid.

^v Van der Werf, Lubbers en Kuyper (1999) stellen dat sinds de invoering van de basisvorming de mate van ongelijkheid naar sociaal-economische achtergrond en etniciteit niet is afgenomen. Zij baseren zich hierbij op analyses met de covariantie-benadering. Met die benadering kunnen zij niet tot deze slotsom komen. De basisvorming moet worden beoordeeld op de vooruitgang van leerlingen tussen leerjaar één en drie en niet op verschillen die mogelijk in de periode voor het voortgezet onderwijs zijn ontstaan.

^{vi} Door cohorten lange tijd te volgen ontstaan mogelijkheden voor prachtige analyses. Befaamd is in dit opzicht het Malmö-cohort, waarin 1.544 kinderen die in 1938 tien jaar waren vijftig jaar lang zijn gevolgd. Met dit cohort hebben Husén (1969), Fägerlind (1975) en Tuijnman (1989) onder meer laten zien dat extra gevolgd onderwijs (her-, bij- of nascholing) op lange termijn invloed heeft op het bereikte beroepsniveau, het inkomen en de ervaren gevoelens van tevredenheid. Hierbij blijkt dat dergelijke effecten zich op 35-jarige leeftijd nog nauwelijks voordoen, maar op 56-jarige leeftijd bijzonder sterk zijn. Het effect van extra kwalificaties is onafhankelijk van andere kenmerken, zoals het sociaal milieu, de intelligentie en het initiële opleidingsniveau van leerlingen.

Een dergelijk cohort is uniek. In Nederland zijn er niet zulke langlopende cohorten, maar er zijn wel enkele cohorten die op een later tijdstip opnieuw zijn aangewend. Peschar (1975) heeft begin jaren zeventig opnieuw informatie verzameld bij personen die in het schooljaar 1958-'59 in de hoogste klas van het lager onderwijs zaten. Dronkers (1998b) heeft gebruik gemaakt van het gereactiveerde cohort van Matthijssen en Sonnemans (1958).

Voor onderzoek naar lange-termijneffecten zou het gunstig zijn, wanneer in Nederland de verschillende cohorten meer op elkaar aansluiten. In het basisonderwijs is er het PRIMA-cohort; in het voortgezet onderwijs is VOCL'93 het meest recente cohort en in het hoger onderwijs is er Verder Studeren. Tussen deze cohorten is echter geen overlap. Het zou naar mijn mening de moeite waard zijn, wanneer het Ministerie van Onderwijs, Cultuur en Wetenschappen en het Ministerie van Sociale Zaken en Werkgelegenheid zich hard zouden maken voor een lang lopend onderzoek naar verschillen in school- en beroepsloopbanen. Ook is het wenselijk om een data-infrastructuur te ontwikkelen waarbij verschillende bestanden aan elkaar kunnen worden gekoppeld via een landelijke basisregistratie (Allen, Heijke & Van der Velden, 1999). Door

koppeling van gegevens uit verschillende bronnen wordt tijd en geld bespaard, onnodige overlap vermeden, het onderzoeksveld ontlast, een ver gaande coördinatie tussen onderzoeken bevordert, en de informatiewaarde van verzamelde gegevens aanzienlijk verhoogd.

^{vii} In het verleden zijn talrijke non-cognitieve uitkomsten als afhankelijke variabele gehanteerd, zoals de mate van delinquentie (Rutter e.a., 1979), sociale vorming (Van Marwijk Kooy-von Baumhauer, 1984) en affectieve en sociale ontwikkeling (Bryk, Lee & Holland, 1993).

^{viii} Hierbij zou het ook zeker voorkeur verdienen om onderzoeksinstrumenten meer te standaardiseren. Replicatie wordt in onderwijsonderzoek vaak bemoeilijkt doordat er te veel verschillende instrumenten worden gebruikt.

^{ix} In toekomstig onderzoek moet ook zeker nog meer naar interacties tussen leerling-, klas- en schoolkenmerken worden gekeken. In dit onderzoek bleek dat sommige schoolkenmerken bijvoorbeeld wel van invloed zijn op leerlingen in het beroepsonderwijs, maar niet op leerlingen in het algemeen vormend onderwijs. Op soortgelijke wijze kan er ook worden gekeken naar interacties van sekse, etniciteit, sociaal-economische status of culturele leerling- en gezinskenmerken met kenmerken op klas- of schoolniveau.