

# De effectiviteit van onderwijs bij 17 Rotterdamse stimuleringscholen\*

C. J. VERMEULEN

*Project Onderwijs en Sociaal Milieu,  
Rotterdam*

## Samenvatting

*De laatste tijd is er veel aandacht in de wereld van het onderwijsonderzoek voor de relatie tussen bepaalde schoolkenmerken (met name het schoolklimaat) en de effectiviteit van het gegeven onderwijs. Voor de invloed van een vijftal schoolkenmerken is zo regelmatig ondersteuning gevonden dat men al spreekt van het 5-kenmerken model van effectieve scholen. Het merendeel van het onderzoek waaruit dit model naar voren is gekomen is van Amerikaanse origine. Aangezien er nogal wat verschillen bestaan tussen het Amerikaanse en het Nederlandse onderwijssysteem, leek het zinvol om binnen de Nederlandse situatie (i.c. een aantal Rotterdamse stimuleringscholen) na te gaan of daar ook een dergelijke relatie kon worden aangetoond. Uit het onderzoek blijkt dat de scholen inderdaad onderling verschillen op de 5 schoolkenmerken, maar een relatie met de leerprestaties kon maar voor één kenmerk (veilige en ordelijke schoolomgeving) worden aangetoond. Gezien deze resultaten worden er in dit onderzoek vraagtekens gezet bij de validiteit, in ieder geval in de Nederlandse situatie, van het gebruikte, Amerikaanse, meetinstrument.*

## 1 Inleiding

Op het terrein van het onderzoek naar onderwijs is er de laatste tijd een verhoogde belangstelling voor het rendement van het onderwijs. Meer specifiek gaat het, vooral in Amerikaans

onderzoek, om de vraag hoe het komt dat sommige scholen aanmerkelijk effectiever zijn in het aanleren van vooral de cognitieve vaardigheden (lezen, taal en rekenen) dan andere, vergelijkbare scholen. Onderzocht wordt welke eigenschappen van scholen deze verschillen kunnen verklaren om zo inzicht te krijgen in de factoren waarop het beleid zich moet richten om tot verbetering van het onderwijs te komen. Langzamerhand is op basis van verschillende onderzoeken (o.a. Edmonds, 1979; Schweitzer, 1984) een consensus ontstaan over een 5-tal kenmerken waarop meer en minder effectieve scholen zich van elkaar onderscheiden. Het gaat daarbij om de volgende kenmerken:

1. een sterk op het onderwijs gericht leiderschap van het teamleider;
2. nadruk op het aanleren van de basisvaardigheden;
3. een ordelijke en veilige schoolomgeving;
4. hoge verwachtingen ten aanzien van de prestaties van de leerlingen;
5. het regelmatig controleren van de vorderingen van de leerlingen.

De genoemde kenmerken hebben dus vooral betrekking op het school- en leerklimaat, waarbinnen het onderwijsproces zich afspeelt.

## 2 Probleemstelling

Dit onderzoek is opgezet om de vraag te beantwoorden of er binnen de onderzoeksgroep, bestaande uit 22 Rotterdamse stimuleringscholen, duidelijke verschillen bestaan in de mate van effectiviteit van het gegeven onderwijs en of deze verschillen zijn terug te vinden in de leerprestaties van de leerlingen. Meer algemeen kan de doelstelling omschreven worden als het empirisch toetsen van het zgn. 5-kenmerken model van effectieve scholen.

\*Onderzoek verricht in het kader van de subsidieregeling van de Stichting voor Onderzoek van het Onderwijs (SVO) voor aspirant onderzoekers

### 3 *Overzicht van de theorie*

De huidige hausse in onderzoek naar de effectiviteit van scholen kan gezien worden als een reactie op het jarenlang domineren van de visie die inhield dat de school er weinig toe deed. Deze opvatting is in hoofdzaak gegrondvest door de onderzoeksrapporten Equality of educational opportunity, van Coleman (1966) en Inequality: A reassessment of the effect of family and schooling in America, van Jencks (1972). In beide rapporten werd geconstateerd dat de verschillen die konden worden vastgesteld tussen de leerprestaties van leerlingen uit minderheidsgroepen en overige leerlingen geen verband hielden met de gemeten schoolkenmerken. Het ging daarbij om schoolkenmerken als: faciliteiten, geldelijke middelen, kwalificaties van leerkrachten e.d. Deze rapporten suggereerden dat wat kinderen op school leren voornamelijk afhankelijk is van hun gezins- en sociale achtergrond en niet van variatie in wat de school hun biedt. Ook de in die tijd verschenen evaluatieverslagen van de onderwijsstimuleringsprogramma's Head Start en Title I toonden het geringe effect aan dat de programma's hadden op het verkleinen van de kloof in prestaties tussen geprivilegeerde en niet-geprivilegeerde leerlingen. Het gevolg van deze consistente reeks van onderzoeksconclusies was dat er ernstige twijfel rees over de algemene overtuiging dat de school een belangrijke rol kon en moest spelen in de verbetering van de leerprestaties van kinderen in achterstandssituaties. Dat daarbij de conclusies ten aanzien van de onderzochte schoolkenmerken werden generaliseerd naar alle schoolkenmerken werd over het hoofd gezien, terwijl ook andere methodologische beperkingen van deze studies weinig of geen aandacht kregen. Toch bleef de opvatting van Coleman c.s. niet onbestreden: gedeeltelijk op basis van methodologische kritiek en gedeeltelijk op basis van het intuïtieve gevoel dat scholen wel verschil moesten maken, kwam in de zeventiger jaren onderzoek op gang dat het ongelijk van deze richting probeerde aan te tonen. Dit onderzoek richtte zich niet zozeer op de eenvoudig te identificeren en te manipuleren variabelen zoals faciliteiten e.d. maar op meer complexe variabelen als schoolklimaat, de wijze van leidinggeven, programma-coördinatie en interactie binnen het schoolteam. Met andere

woorden, de aandacht verschoof van statische naar procesvariabelen. Terwijl statische variabelen eenvoudig te meten en te beïnvloeden zijn, bieden procesvariabelen een grotere mogelijkheid tot het begrijpen van de verschillen in schooleffectiviteit, maar geven minder snel aanwijzingen over concrete beleidsmaatregelen. Voorbeelden van dit soort onderzoeken zijn die van Wellisch (1978), Brookover (1979) en Schweitzer (1984) in de Verenigde Staten, Rutter (1979) in Engeland en Meijnen (1984) en Van der Wolf (1984) in Nederland. In al deze onderzoeken komen veelal dezelfde elementen naar voren die in belangrijke mate bepalend zijn voor de effectiviteit van een school. Veel blijkt af te hangen van de opstelling van de teamleider en het team van onderwijzers. Vooral de teamleider krijgt een grote rol toebedeeld in het tot stand brengen en handhaven van een stimulerende leeromgeving. Belangrijk voor effectiviteit blijkt ook te zijn het bestaan van hoge verwachtingen over de leerprestaties aan de kant van de onderwijsgeevenden. Maar ook de overige elementen uit het 5-kenmerken model komen in de genoemde onderzoeken in meer of mindere mate terug, waardoor het gerechtvaardigd lijkt om ook in de Nederlandse situatie het 5-kenmerken model te gebruiken om onderzoek te doen naar de effectiviteit van het onderwijs.

Voor een meer uitgebreide bespreking van deze onderzoeken wordt verwezen naar Van der Grift (1985) en Fokkema (1986).

### 4 *De onderzoeksopzet*

#### 4.1 *Metten van de effectiviteit van scholen*

In dit onderzoek is de mate waarin de onderzochte scholen beantwoorden aan de 5 veronderstelde kenmerken van een effectieve school bepaald met het meetinstrument uit het onderzoek van Schweitzer. Daarom zal eerst wat nader worden ingegaan op dit onderzoek. Schweitzer heeft onderzocht of met het inmiddels bekende 5-kenmerken model van effectieve scholen een significante proportie van de tussen-school variantie in leerprestaties kon worden verklaard. Het onderzoek werd verricht in een steekproef van 16 lagere scholen gevestigd in het centrum van een Amerikaanse stad. De populatie van de scholen bestond vrijwel uitsluitend uit zwarte leerlingen uit de

lage inkomensgroepen. De schoolkenmerken werden gemeten met de antwoorden van 456 leerkrachten en 16 teamleiders op een vragenlijst bestaande uit 49 Likert-type items, die ontworpen waren om de 5 kenmerken van effectieve scholen te meten. De items bestonden uit uitspraken, betrekking hebbend op het onderwijsleerklimaat, waarvan de leerkrachten en de teamleiders op een schaal die liep van 'waar' via 'grotendeels waar', 'voor de helft waar', 'enigszins waar' tot 'niet waar', moesten aangeven in welke mate ze deze uitspraken op hun school van toepassing achten. Ter illustratie van het soort uitspraken waaruit de vragenlijst is opgebouwd zal uit elk van de 5 kenmerk-schalen een uitspraak worden weergegeven.

*Kenmerk 1:* De teamleider bespreekt met de leerkrachten op welke wijze het onderwijsprogramma didactisch het best kan worden uitgevoerd

*Kenmerk 2:* Het hele schoolteam is zich duidelijk bewust van zijn verantwoordelijkheid om de leerlingen de basisvaardigheden aan te leren

*Kenmerk 3:* Zowel leerkrachten, niet-onderwijzend personeel als ouders voelen zich verantwoordelijk voor het handhaven van de orde op school

*Kenmerk 4:* Welk percentage van de leerlingen op Uw school is volgens U in staat om de leerdoelen van hun jaar te halen? (Antwoordmogelijkheden:  $\leq 29\%$ , 30 - 49 %, 50 - 69 %, 70 - 89 %,  $\geq 90\%$ ).

*Kenmerk 5:* Middelen om leerlingen te beoordelen (zoals programmatoetsen, voortgangslijsten e.d.) worden regelmatig gebruikt om aan de leerlingen te vertellen op welke onderdelen ze achterblijven en om aangepaste lessen samen te stellen.

In zijn onderzoek vond Schweitzer een duidelijk verband (vooral voor de leerkrachten) tussen de effectiviteitskenmerken en de leerprestaties. De conclusie van Schweitzer was dan ook dat met dit meetinstrument de 5 kenmerken betrouwbaar bepaald kunnen worden in een homogene steekproef van scholen.

Om de volgende redenen is ervoor gekozen om in dit onderzoek de effectiviteitskenmerken te bepalen met het meetinstrument van Schweitzer:

- In de eerste plaats omdat dit meetinstrument in minstens één onderzoek heeft bewezen een valide meting van de effectiviteitskenmerken mogelijk te maken.

- In de tweede plaats omdat dit is gebeurd met een onderzoek in een steekproef die verwant is met de onderhavige onderzoeksgroep. Het gaat immers in beide onderzoekingen om scholen gelegen in de oude binnenstad, bevolkt door kinderen uit kansarme groepen.

- In de derde plaats omdat het verrichten van een onderzoek met een reeds eerder gebruikt meetinstrument tot voordeel heeft dat het onderzoek tevens functioneert als replicatie-onderzoek.

Als de tot nu toe gevonden kenmerken van effectieve scholen generaliseerbaar zijn, wat immers geclaimd wordt, dan moet het mogelijk zijn om op basis van de genoemde kenmerken ook in deze onderzoeksgroep te kunnen differentiëren tussen effectieve en minder effectieve scholen.

Nadat de vragenlijst uit het onderzoek van Schweitzer vertaald was, is de volgorde van de vragen veranderd aangezien in de oorspronkelijke per kenmerk geordende vragenlijst de vragen over het leiderschap van de teamleider een te overheersende indruk konden maken.

#### 4.2 Meten van de leerprestaties

De effectiviteit van scholen wordt binnen het 'effectieve school'-paradigma gedefinieerd in termen van schoolprestaties op het gebied van de basisvaardigheden. De schoolprestaties worden gemeten met behulp van gestandaardiseerde toetsen die in de hoogste klas van de lagere school worden afgenomen (in dit onderzoek de eindtoets basisonderwijs van het CITO), aangezien men er van uit gaat dat daarmee de uiteindelijke effecten van 6 jaar lagere school kunnen worden bepaald, en daarmee dus ook de effecten van het gedurende die tijd verblijven in een bepaald schoolklimaat. Daarbij wordt verondersteld dat het schoolklimaat een stabiel schoolkenmerk is. Hoewel het klasseklimaat onder invloed van de verschillende leerkrachten van jaar tot jaar wisselt, speelt alles zich af binnen een specifiek sociaal systeem. Een sociaal systeem heeft een stabiel karakter als de personen die er deel van uit maken (hier: de leden van het schoolteam), dit over langere tijd doen en als nieuwe medewerkers mede geselecteerd worden op hun geschiktheid tot integratie in het systeem. Om deze redenen is het gerechtvaardigd om op basis van een eenmalige meting een uitspraak te doen over het schoolklimaat.

Ter voorbereiding van de afname is met alle teamleiders een gesprek gevoerd om de bedoeling van het onderzoek uit te leggen en om toestemming voor het afnemen van de vragenlijst te verkrijgen. Vrijwel alle teamleiders wilden aan het onderzoek meewerken, al maakten enkelen de toestemming afhankelijk van de goedkeuring van het hele schoolteam. Bij de afname van de vragenlijst is er op toe gezien dat de leden van het schoolteam zo veel mogelijk de kans kregen om de vragenlijst zonder beïnvloeding van collega's in te vullen. Voor het invullen van de vragenlijst kwamen alle leden van het schoolteam in aanmerking die zich bezig hielden met de basisvaardigheden, aangezien de effectiviteit van scholen vooral daaraan wordt afgemeten. Dat betekent dat b.v. onderwijzers op het gebied van het tweede-taalonderricht, handenarbeid en gymnastiek niet in het onderzoek betrokken zijn. Voor één school was het feit dat niet het hele schoolteam werd ondervraagd aanleiding om hun medewerking aan het onderzoek in te trekken. Van 4 scholen konden wel de leerkrachten maar niet de teamleider bereid gevonden worden, zodat uiteindelijk 17 scholen overbleven waarbij van zowel teamleider als leerkrachten gegevens bekend zijn. Op deze 17 scholen waren 87 leerkrachten bereid om de vragenlijst in te vullen. De afname van de vragenlijst heeft plaatsgevonden in de periode van 13 tot 28 juni 1985.

## 6 *De resultaten*

Om de antwoorden op de vragenlijst te kunnen verwerken zijn aan de antwoordcategoricën van de items waarden (van 0 tot 8) toegekend, zodanig dat voor alle items geldt dat naarmate de score hoger is er meer is voldaan

aan de voorwaarden voor effectiviteit. Voor elke item is per school de gemiddelde score van de leerkrachten berekend. Dit betekent dat per school voor elk effectiviteits-item een score bekend is van de teamleider en de gemiddelde score van de leerkrachten.

### 6.1 *De betrouwbaarheid van de vragenlijst*

De interne consistentie van de vragenlijst is bepaald door het berekenen van de betrouwbaarheidscoëfficiënt Cronbachs alpha. Daarbij gaat het om de vraag in welke mate de items die geacht worden een bepaald begrip (hier een effectiviteitskenmerk) te meten, samen een schaal vormen en dus daadwerkelijk allen metingen zijn van hetzelfde begrip. De betrouwbaarheid is bepaald op schoolniveau ( $n=17$ ), dus op basis van de gemiddelde itemscores van de leerkrachten en de itemscores van de teamleiders. De betrouwbaarheid van de meting van elk van de 5 kenmerken is te vinden in Tabel 1 (zie paragraaf 1 voor de inhoud van de kenmerken).

Uit Tabel 1 blijkt dat de betrouwbaarheid van de leerkrachtsschalen beduidend hoger ligt dan die van de teamleiderschalen. Gaan we uit van een grens van .70 (zie Fox: *The research proces of evaluation*, p. 362) als voldoende voor interne consistentie dan voldoen 3 van de 5 kenmerkschalen bij de leerkrachten en 1 van de 5 schalen bij de teamleiders. Dit betekent dat voor wat de leerkrachten betreft er gewerkt kan worden met de schalen 1, 3 en 4, en voor de teamleiders met schaal 3. Het is niet zinvol om aan de resultaten verkregen met de overige schalen, die voor de volledigheid wel zullen worden vermeld, enige waarde toe te kennen. Vergelijken we deze gegevens met die zoals Schweitzer heeft gevonden dan blijkt dat de interne consistentie van het meetinstrument in zijn onderzoek aanmerkelijk hoger uitvalt. Bij zijn onderzoek had maar 1 van de 10 schalen een Cronbachs alpha die lager was

Tabel 1 *Betrouwbaarheid van de meting van de effectieve schoolkenmerken*

kenmerken	leerkrachten	leiders	aantal items
kenmerk 1	.90 (.95)	.56 (.73)	18
kenmerk 2	.64 (.85)	.25 (.78)	6
kenmerk 3	.89 (.95)	.70 (.69)	10
kenmerk 4	.76 (.93)	.13 (.73)	7
kenmerk 5	.55 (.85)	.05 (.77)	8

(tussen haakjes de alpha-coëfficiënten uit het onderzoek van Schweitzer)

Tabel 2 *Verdelingskarakteristieken van de kenmerkscores op schoolniveau (n=17)*

		gem.	s.d.	min.	max.	kurt.	skew.
kenmerk 1	TL	3.49	0.79	1.50	4.78	1.25	-0.86
	LK*	3.63	1.36	2.09	7.00	1.44	1.38
kenmerk 2	TL	5.81	0.80	4.30	6.80	0.42	-0.98
	LK	5.79	0.54	4.52	6.83	1.19	-0.45
kenmerk 3	TL*	5.55	1.03	4.09	7.60	-0.80	0.43
	LK*	5.27	1.30	3.15	7.20	-1.24	-0.43
kenmerk 4	TL	4.61	0.66	3.43	6.00	0.77	0.05
	LK*	4.56	0.65	3.05	5.54	0.17	-0.73
kenmerk 5	TL	3.11	3.95	1.50	4.75	-3.55	0.03
	LK	4.12	1.27	2.42	7.43	2.21	1.38

(\* schaal met voldoende hoge betrouwbaarheid)  
(TL = teamleider, LK = leerkrachten)

dan .70, terwijl zijn onderzoek op vrijwel hetzelfde aantal scholen is uitgevoerd (n = 16). De vertaalde versie van het meetinstrument dat Schweitzer gebruikte, blijkt dus in een andere sociale omgeving een lagere interne consistentie te bezitten. Dit opvallende verschil kan zijn veroorzaakt door de vertaling en/of door het niet aansluiten van de vragenlijst bij de Nederlandse situatie en mentaliteit.

#### 6.2 *Bepaling van de scores op de 5 effectiviteitskenmerken*

De scores op de 5 kenmerken zijn bepaald door van de schalen die de operationalisatie vormen van deze kenmerken de items te sommeren. Bij de leerkrachten gaat het daarbij om de sommering van hun per school gemiddelde itemscores, terwijl voor de teamleiders direct de itemscores konden worden opgeteld. De schalen waarmee de kenmerken zijn gemeten bestaan uit respectievelijk 18, 6, 10, 7 en 8 items. Uiteindelijk heeft dus elke school op de 5 kenmerkschalen een score van de teamleider en een score die de gemiddelde opinie van de leerkrachten weergeeft. Om bij het ontbreken van antwoorden op bepaalde items (weet niet/geen mening, of helemaal niet ingevuld) toch een schaal score te kunnen bepalen is een gewogen berekening over de wel aanwezige scores toegepast, indien minder dan 25 procent ontbrak van de items, waaruit een schaal score is opgebouwd. Daarbij bleek dat, met uitzondering van 1 school waarbij voor de teamleider geen score op factor 5 kon worden bepaald, alle schaal scores konden worden berekend. Door de gewogen correctie voor het optreden van ontbrekende gegevens kan de

score van de kenmerken net als die van de items variëren van 0 tot 8. In Tabel 2 zijn de verdelingskarakteristieken van de schaal scores voor leerkrachten en teamleiders weergegeven.

De vergelijking tussen de schaal scores van teamleiders en leerkrachten is alleen goed mogelijk voor het kenmerk 'veilige en ordelijke schoolomgeving', omdat voor dit kenmerk beide schalen een voldoende hoge interne consistentie te zien gaven. Daarbij blijkt dat er slechts een gering verschil bestaat tussen de gemiddelde beoordeling van de leerkrachten en het oordeel van de teamleider (gemeten over alle scholen). De theoretische vereisten voor effectiviteit worden door de betrokken scholen het best benaderd voor kenmerk 2 (nadruk op de basisvaardigheden) en voor kenmerk 3 (een veilige en ordelijke omgeving). Toch blijven de scholen ook op deze kenmerken nog 2 tot 2.5 schaaldelen van het maximum verwijderd. Op het punt van het onderwijskundig leiderschap van de teamleider (kenmerk 1), bleken de betrokken scholen, zowel volgens de leerkrachten als de teamleiders zelf, het minst te voldoen aan het profiel van een effectieve school.

Met behulp van variantie-analyse is voor de leerkrachtscores bepaald of de schoolgemiddelden op de 5 kenmerken significant van elkaar verschillen. Uit de berekeningen blijkt dat voor de kenmerken 1, 3 en 4 er sprake is van een variantie in de schoolgemiddelden die significant groter is dan de variantie in schaal scores binnen de scholen (kenmerk 1:  $F = 6.55$ ,  $p = .001$ , kenmerk 3:  $F = 9.06$ ,  $p = .001$ , ken-

Tabel 3 *Onderlinge samenhangen tussen de kenmerken (n = 17)*

	Leerkrachten				
	kenmerk 1	kenmerk 2	kenmerk 3	kenmerk 4	kenmerk 5
kenmerk 1	0.17	0.48*	0.48*	0.28	0.75*
kenmerk 2	-0.29	0.08	0.31	0.21	0.77*
kenmerk 3	-0.32	0.36*	0.55*	0.17	0.42*
kenmerk 4	-0.05	-0.05	-0.18	-0.11	0.36*
kenmerk 5	-0.03	-0.05	0.08	-0.07	0.25

Teamleiders

(\* = significant op .05 niveau)

merk 4:  $F = 2.64$ ,  $p = .006$ ). Op deze kenmerken bestaan er dus tussen de scholen duidelijke verschillen, wat het mogelijk en zinvol maakt om de relatie met de leerprestaties te onderzoeken.

### 6.3 *Onderlinge samenhangen tussen de kenmerken*

Om na te gaan welke onderlinge samenhangen er bestaan tussen de scores van teamleiders en leerkrachten op de effectiviteitskenmerken zijn de onderlinge correlaties berekend. De resulterende correlatiematrix is te vinden in Tabel 3. De bovenste driehoek van de tabel geeft de onderlinge samenhang tussen de schaa scores gebaseerd op de antwoorden van de leerkrachten, de onderste driehoek geeft de onderlinge samenhang tussen de schaa scores bij de teamleiders, terwijl op de diagonaal de samenhangen te vinden zijn tussen schaa scores van leerkrachten en teamleiders.

Uit Tabel 3 blijkt dat alleen ten aanzien van de beoordeling van de schoolomgeving er substantiële overeenstemming bestaat tussen leerkrachten en teamleiders. De overige samenhangen tussen de leerkracht- en teamleideropinions zijn niet significant. Dit kan zowel veroorzaakt worden door het bestaan van een aanzienlijke discrepantie tussen teamleiders en leerkrachten in de beoordeling van hun school, als door onvoldoende interne consistentie van de gebruikte schalen. In het onderzoek van Schweitzer werden sterkere verbanden tussen de schaa scores van leerkrachten en teamleiders aangetroffen (deze bedroegen voor de 5 kenmerken resp. .24, .30,

.53, .69 en .56). De onderlinge correlatie tussen de schaa scores van de leerkrachten is in 6 van de 10 gevallen significant terwijl dit bij de teamleiders maar in 1 van de 10 gevallen voorkomt. Een hoge onderlinge correlatie duidt erop dat leerkrachten of teamleiders die op een kenmerk tot de laagst of hoogst scorende behoren dit ook op de andere kenmerken doen, m.a.w. ze zijn, afgemeten aan het effectiviteitscriterium, consequent in hun beoordeling.

### 6.4 *De leerprestaties*

Tot nu toe is uit het onderzoek gebleken dat er een vrij grote variatie is in de gemiddelde score van de scholen op de effectiviteitskenmerken, zowel voor de leerkrachten als voor de teamleiders. Op zich heeft deze constatering geen betekenis als niet kan worden aangetoond dat er een verband bestaat tussen de mate waarin een school volgens het gebruikte meetinstrument effectief is en de gemiddelde leerprestaties op die school, m.a.w. als het meetinstrument empirisch gevalideerd kan worden.

De leerprestaties zijn in dit onderzoek bepaald aan de hand van de resultaten die de 6 klas leerlingen hebben behaald op de eindtoets basisonderwijs 1985 van het CITO. Deze toets is op de betrokken scholen afgenomen op 5 en 7 maart 1985. De eindtoets bestaat uit 180 opgaven waarvan 60 betrekking hebben op taal, 60 op rekenen en 60 op informatieverwerking. Met behulp van variantie-analyse is nagegaan of de scholen significant van elkaar verschillende scores op de eindtoets behaalden. Daarbij bleek dat, hoewel de be-

trokken scholen zich landelijk gezien in een klein gebied van de verdeling bevinden, er binnen de groep toch voldoende spreiding in schoolgemiddelde is om een verband met de verschillen in effectiviteit mogelijk te maken ( $F = 2.16, p = .006$ ).

Een onderlinge vergelijking van scholen kan vertekend worden door verschillen in instroom-niveau van de leerlingen. Hierdoor kan een school die lagere gemiddelde leerprestaties behaalt dan een andere school, toch de meer effectieve school zijn omdat ze voor een grotere toename in leerprestaties heeft gezorgd. Als, zoals voor de onderzochte groep, deze gegevens niet bekend zijn dan kan dit probleem ondervangen worden door de toetscores over meerdere jaren te bekijken. Als daarbij blijkt dat er consequent hoog en consequent laag scorende scholen zijn, dan zijn de leerprestaties blijkbaar onafhankelijk van mogelijk fluctuaties in het instroom-niveau. Van de betrokken scholen zijn de gegevens van de CITO eindtoets basisonderwijs bekend van de jaren 1983 tot 1985. Het blijkt echter niet mogelijk om binnen deze groep een rangorde aan te brengen van consequent hoog en laag scorende scholen. Door het, landelijk gezien, smalle gebied binnen de verdeling van de toetsscores, waarin de betrokken scholen zich bevinden, hebben relatief kleine veranderingen in de gemiddeld behaalde toetsscore, al gauw een grote plaatswijziging in de rangorde tot gevolg. Daarom wordt in dit onderzoek vastgehouden aan de rangorde, zoals bepaald in hetzelfde jaar als waarin de effectiviteitsvragenlijst is afgenomen, waarbij niet kan worden uitgesloten dat de rangorde wordt vertekend door verschillen in instroomniveau.

### 6.5 Samenhang tussen effectiviteitscore en leerprestaties op schoolniveau

De correlatiecoëfficiënten die de sterkte van de samenhang tussen enerzijds de schaalscores van leerkrachten en teamleiders, en anderzijds de schoolgemiddelden op de eindtoets basisonderwijs weergeven zijn te vinden in Tabel 4.

Er moet geconcludeerd worden dat in dit onderzoek veel minder ondersteuning voor het 5-kenmerken model is gevonden dan in het onderzoek van Schweitzer. Alle samenhangen zijn beduidend zwakker, waarbij voor de leerkrachten alleen het effectiviteitskenmerk 3 (een veilige en ordelijke omgeving) significant met de leerprestaties samenhangt, terwijl dit bij de teamleiders geldt voor de effectiviteitskenmerken 1 en 4. De richting van de samenhang is in deze gevallen echter omgekeerd aan de verwachting. Hieraan kan echter geen conclusie worden verbonden, aangezien de betrouwbaarheid van de meting van deze kenmerken onvoldoende was. Uiteindelijk blijkt op maar 1 van de 10 verwachte samenhangen een significante samenhang in de veronderstelde richting te worden aangetroffen.

Dit resultaat kan beïnvloed zijn door het verschil dat bestaat in homogeniteit tussen deze onderzoekspopulatie en die van Schweitzer. De betrokken stimuleringscholen zijn wel homogeen ten aanzien van de sociale klasse, maar niet ten aanzien van de etnische achtergrond van de kinderen. Uit het onderzoek van Van Esch (1982) is bekend dat allochtone kinderen gemiddeld lager scoren op de eindtoets basisonderwijs van het CITO dan autochtone kinderen. Van Esch heeft 1271 allochtone kinderen verdeeld over 342 lagere scholen onderzocht. Daarbij bleek dat de Nederlandse kinderen gemiddeld bijna 130 van de 180 opgaven goed maakten (= 71.7%),

Tabel 4 Samenhang tussen effectiviteitskenmerken en leerlingprestaties ( $n = 17$ )

	leerkr.	gemiddelde CITO-toetsscore		
		dit onderzoek leider	onderzoek Schweitzer leider	
kenmerk 1	0.05	-0.56*	0.58	0.09
kenmerk 2	0.12	-0.19	0.79	0.44
kenmerk 3	0.34*	0.21	0.59	0.28
kenmerk 4	0.21	-0.45*	0.79	0.71
kenmerk 5	0.23	0.13	0.68	0.21

(\* = significant op .05 niveau, van Schweitzer-onderzoek geen significantieniveau bekend)

tegen 100 van de 180 voor de allochtone kinderen (= 53.6%). Om de invloed van het percentage allochtone kinderen op het schoolgemiddelde op de CITO-toets na te gaan is voor de 6e klas de ratio allochtone kinderen bepaald. Deze ratio bleek te variëren tussen 7.1 en 72.2 %. De samenhang tussen het ratio allochtone kinderen en de eindtoets basisonderwijs blijkt  $-.41^*$  te bedragen en tussen de ratio en de toetsonderdelen taal, rekenen en informatieverwerking resp.  $-.46^*$ ,  $.03$  en  $-.66^*$  (\* = significant op .05 niveau). Uit deze gegevens blijkt dat het gedeelte van de klas dat uit allochtone kinderen bestaat een duidelijke invloed heeft op de gemiddelde toetsscore, met uitzondering van het onderdeel rekenen, waarschijnlijk omdat dit onderdeel het minst taal-afhankelijk is. De ratio kan dus van invloed zijn op het verband tussen de effectiviteitskenmerken en de leerlingprestaties. Of dit het geval is is nagegaan door het berekenen van de partiële correlatie tussen de mate waarin de school voldoet aan de effectiviteitscriteria en de leerlingprestaties onder controle van de ratio allochtone leerlingen. De gevonden partiële correlatiecoëfficiënten zijn te vinden in Tabel 5.

Het blijkt dat de invloed van de ratio allochtone leerlingen op de samenhang tussen effectiviteit en leerprestaties gering is. De meeste verbanden worden iets sterker, waarbij voor de teamleiders het verband tussen effectiviteitskenmerk 3 en de leerprestaties nu ook significant is. De conclusie moet dus blijven dat er zeer weinig ondersteuning is gevonden voor het 5-kenmerken model in dit onderzoek.

## 7 Samenvatting en conclusie

In dit onderzoek is nagegaan, met behulp van

Tabel 5 Samenhang tussen mate van effectiviteit en leerlingprestaties onder controle van de ratio allochtone leerlingen ( $n = 17$ )

	gemiddelde score eindtoets			
	direct	leerkrachten partieel	direct	teamleiders partieel
kenmerk 1	0.05	0.08	$-0.56^*$	$-0.49^*$
kenmerk 2	0.12	0.24	$-0.19$	$-0.15$
kenmerk 3	$0.34^*$	$0.40^*$	0.21	$0.44^*$
kenmerk 4	0.21	0.21	$-0.45^*$	$-0.58^*$
kenmerk 5	0.23	0.25	0.13	0.28

(\* = significant op .05 niveau)

een vragenlijst overgenomen uit een Amerikaans onderzoek, of er een verband kan worden aangetoond tussen de mate waarin een school voldoet aan de theoretische vereisten van een effectieve school en de daadwerkelijke optredende gevolgen van effectief onderwijs, zoals die hun weerslag vinden in de resultaten van de leerlingen op een genormeerde, landelijk afgenomen, toets. Voor het onderzoek zijn 22 Rotterdamse stimuleringscholen benaderd, waarvan er 17 overbleven waar zowel de leerkrachten als de teamleiders konden worden ondervraagd. Het onderzoek is daarom gebaseerd op de gegevens van 17 teamleiders en 87 leerkrachten. De effectiviteit van de scholen is uitgedrukt in een score van de teamleider en een gemiddelde score van de leerkrachten op de 5 kenmerken, die met het gekozen meetinstrument zijn bepaald. Bij analyse van de gegevens bleek dat bij 2 van de 5 leerkrachtsschalen sprake was van een onvoldoende interne consistentie, terwijl dit bij de teamleiders voor 4 van de 5 schalen het geval was. Hierdoor is het maar in beperkte mate mogelijk om op basis van dit onderzoek uitspraken over de validiteit van het 5-kenmerken model te doen. Wat wel geconcludeerd kan worden is dat er duidelijk verschillen bestaan tussen de betrokken scholen in hun score op de 5 kenmerken, d.w.z. in de mate waarin op de scholen een schoolklimaat heerst dat bevorderlijk geacht wordt te zijn voor de leerprestaties. De betrokken scholen blijken het meest te beantwoorden aan de effectiviteitsnormen voor de kenmerken 'nadruk op de basisvaardigheden' en 'veilige en ordelijke omgeving'. Er kan echter maar voor één kenmerk (veilige en ordelijke omgeving) een verband worden aangetoond tussen deze indeling van de scholen op het effectiviteitscri-

terium en de op de scholen behaalde leerprestaties, zoals gemeten met de eindtoets basisonderwijs van het CITO. Dit resultaat bleek onafhankelijk te zijn van het percentage allochtone kinderen in de klas.

Zoals blijkt is in dit onderzoek met het vertaalde meetinstrument uit het onderzoek van Schweitzer veel minder ondersteuning gevonden voor het 5-kenmerken model dan op basis van het oorspronkelijke onderzoek verwacht mocht worden. Deze constatering kan tot twee gevolgtrekkingen aanleiding geven:

- of de variatie in effectiviteitskenmerken die werd aangetroffen tussen de scholen heeft daadwerkelijk geen relatie met de leerprestaties, zodat in dit onderzoek is gebleken dat het schoolklimaat veel minder invloed heeft op de leerprestaties dan de laatste tijd wordt aangenomen;

- of de effectiviteit en/of de leerprestaties zijn niet goed gemeten.

Voorlopig lijkt de eerste conclusie, gezien de resultaten van andere onderzoekingen, niet aannemelijk, zodat voor een oorzaak voor het niet terug vinden van de verwachte verbanden eerder bij de meting gezocht moet worden. Daarbij kan worden aangenomen dat de leerprestaties voldoende betrouwbaar gemeten zijn, gezien de status die de CITO-toets heeft o.a. bij het doorverwijzen naar het voortgezet onderwijs. Dan blijft er over te concluderen dat het gebruikte meetinstrument, in ieder geval in de onderhavige situatie, niet geschikt is. Deze conclusie wordt ondersteund door de ervaringen opgedaan bij het afnemen van de vragenlijst. Daarbij bleek dat de vragen die betrekking hadden op het leiderschap van de teamleider weerstanden oproepen. De in deze vragen impliciet veronderstelde hiërarchische verhouding bleek op veel scholen niet te bestaan en zou waarschijnlijk ook niet door de leerkrachten getolereerd worden. Er werd door de ondervraagden regelmatig benadrukt dat men zich als team verantwoordelijk voelde voor het onderwijs en dat daarin aan de teamleider geen aparte taak werd toegekend. Ook uit de reacties bij de onderwijsgevenden op de vragen die betrekking hadden op het kenmerk 'gebruik van beoordelingsinstrumenten' bleek dat de achterliggende gedachtengang weinig weerklink vond bij de onderzochte onderwijsgevenden.

Geconcludeerd moet daarom worden:

- dat er binnen de onderzochte groep stimule-

ringsscholen verschil bestaat in de mate waarin voldaan wordt aan het effectiviteitscriterium 'een veilige en ordelijke omgeving', dat positief samenhangt met de variatie in leerprestaties op die scholen;

- dat ten aanzien van de andere effectiviteitscriteria een dergelijk verband niet kon worden aangetoond;

- dat daarbij aannemelijk is dat het gebruikte meetinstrument niet het meest geschikt is gebleken om binnen deze onderzoeksgroep de verschillen in de, voor leerprestaties bevorderlijk geachte, kenmerken van het schoolklimaat te bepalen.

### Literatuur

- Brookover, W., C. Beady, P. Flood, P. Schweitzer, J. Wisenbaker, *School social systems and student achievement. School can make a difference*. New York: Praeger Publishers, 1979.
- Coleman, J. S., E. Campbell, C. Hobson, J. McPortland, A. Mood, F. Weinfeld & R. York, *Equality of educational opportunity*. Washington D.C.: U.S. Government Printing Office, 1966.
- Edmonds, R. R., *Effective schools for the urban poor. Educational Leadership*, 1979, 37, 15-24.
- Esch, W. van, *Toetsprestaties en doorstroomadviezen van allochtone leerlingen in de zesde klas van de lagere school*. Nijmegen: I.T.S., 1982.
- Fokkema, E., *Schoolmanagement en de effectiviteit van scholen*. In: F. J. van der Krogt, *Schoolleiding en Management*. Lisse: Swets & Zeitlinger, 1986.
- Fox, D. J., *The research proces in education*. New York: Holt, Rinehart and Winston, 1969.
- Grift, W. van der, *Onderwijsleerklimaat en leerlingprestaties. Pedagogische Studiën*, 1985, 62, 401-413.
- Jencks, C. S., M. Smith, H. Acland, M. Bane, D. Cohen, H. Gintis, B. Hyns & S. Michelson, *Inequality: A reassessment of the effect of family and schooling in America*. New York: Basic Books, 1972.
- Madaus, G., P. Airasian, T. Kellaghan, *School effectiveness. A reassessment of the evidence*. New York: McGraw-Hill, 1980.
- Meijnen, G. W., *Van zes tot twaalf. Een longitudinaal onderzoek naar de milieu- en schooleffecten van loopbanen in het lagere onderwijs*. Harlingen: Flevodruk, 1984.
- Rutter, M., B. Maugham, P. Mortimer & J. Duston, *Fifteen thousand hours, secondary schools and their effects on children*. London: Open books, 1979.
- Schweitzer, J. H., *Characteristics of effective schools*. Paper presented at the annual meeting of

the American Educational Research Association.  
New Orleans, april 1984.

Wellish, J. B., A. MacQueen, R. Carriere, G. Duck,  
Management and organisation in succesful  
schools. *Sociology of education*, 1978, nr 51, 211-  
226.

Wolf, J. C. van der, *Schooluitval. Een empirisch on-  
derzoek naar de samenhangen tussen schoolinterne  
factoren en schooluitval in het reguliere onderwijs*.  
Lisse: Swets & Zeitlinger, 1984.

### *Curriculum vitae*

*C. J. Vermeulen* (1949) studeerde sociologie aan de  
Vrije Universiteit te Amsterdam met als specialisatie  
methoden en technieken van onderzoek. Hij was  
van januari 1984 tot september 1986 verbonden aan  
het Project Onderwijs en Sociaal Milieu te Rot-  
terdam. Hij was daarvoor werkzaam bij het Insti-  
tuut voor Milieuvraagstukken van de Vrije Univer-  
siteit en de afdeling Kijk- en Luisteronderzoek van  
de N.O.S. te Hilversum.

*Adres:* Overtoom 226-1, 1054 HZ Amsterdam.

*Manuscript aanvaard* 25-6-'86

### **Summary**

Vermeulen, C. J. 'The effectiveness of 17 elementary schools in Rotterdam.' *Pedagogische Studiën*, 1987, 64, 49-58.

In this study the validity of the so-called 5-factor model of effective schools was tested. To this end a group of 17 elementary schools in Rotterdam was studied. The 5 school characteristics were measured by the responses of 17 principals and 87 teachers to a questionnaire containing 49 Likert-type items. School achievement was measured by the Cito achievement test administered in the sixth grade. The internal consistency of the measures used was determined by computing Cronbach's coefficient alpha. The reliability of 2 of the 5 teacher scales and 4 of the 5 principal scales proved to be insufficient. When the correlations between the remaining factors and school achievement were computed, only the characteristic 'a safe and orderly environment' accounted for a significant proportion of in between school variance in achievement. The conclusion therefore must be that little has been found in this study to corroborate the 5-factor model.