

(KAFT)

**DE BEVORDERING VAN
SCHOOLEFFECTIVITEIT
IN HET BASISONDERWIJS**

Mogelijkheden tot "flankerend beleid" bij klassenverkleining

Jaap Scheerens

**DE BEVORDERING VAN
SCHOOLEFFECTIVITEIT
IN HET BASISONDERWIJS**

Mogelijkheden tot "flankerend beleid" bij klassenverkleining*)

Jaap Scheerens**)

*) Deze studie is uitgevoerd in opdracht van het Ministerie van OC&W, directie Primair Onderwijs.

**) De auteur is hoogleraar Onderwijsorganisatie en -management bij de Faculteit der Toegepaste Onderwijskunde van de Universiteit Twente en wetenschappelijk directeur van het OCTO.

Samenvatting en conclusies

Onderwijskwaliteit moet blijken uit de opbrengsten van het onderwijs. Er vanuit gaande dat een belangrijk deel van die opbrengsten in het geval van het basisonderwijs bestaat uit leerprestaties in basisvakken als rekenen, taal en lezen, is er een eenvoudige definitie van onderwijskwaliteit te geven: "de kwaliteit van het onderwijs blijkt uit het niveau van de leerprestaties, zoals bepaald met behulp van een curriculumonafhankelijke toets" (Hofstee, 1989, p. 42; vgl. ook De Groot, 1986).

Ook na een analyse vanuit een wat bredere optiek van "organisatorische effectiviteit" blijkt deze typering als de kern van de zaak te kunnen worden beschouwd, al blijft er over het begrip onderwijskwaliteit natuurlijk voldoende stof ter discussie. Onder meer ook over de vraag hoe curriculumonafhankelijk de toetsing zou moeten zijn.

Acceptatie van deze toegespitste definitie van onderwijskwaliteit maakt het onderzoek in de traditie van onderwijsproduktiviteit en schooleffectiviteit bijzonder relevant, als het gaat om de identificatie van manipuleerbare condities die de prestatieniveaus in de basisvakken kunnen verhogen. Dit laatste nu vormt de kern van de opdracht naar aanleiding waarvan deze studie is uitgevoerd, nl. "een overzicht en samenvatting samen te stellen van de literatuur uit binnen- en buitenland inzake factoren die een kwaliteitsverhogend effect hebben" (citaat uit de brief waarin de opdracht wordt verstrekt). Een verbijzondering hierbij is dat de opdrachtgever vooral geïnteresseerd is in variabelen die naast *groeps-grootteverkleining* van invloed zijn op de leerresultaten.

In het schooleffectiviteitsonderzoek, zoals dat zo ongeveer de laatste tien jaar wordt bedreven, heeft een integratie plaatsgevonden van economisch georiënteerd onderzoek naar "onderwijsproductiefuncties", onderzoek naar instructie-effectiviteit en onderzoek naar effectieve scholen, waarin bij het laatste type vooral een accent lag op variabelen op het niveau van de schoolorganisatie. Statistische technieken onder de noemer van "multi-niveau analyses" hebben een belangrijke bijdrage geleverd aan deze integratie.

In hoofdlijnen komt schooleffectiviteit neer op het zoeken van "procescondities" op school- en klasniveau die een positieve samenhang hebben met voor aanvangsniveau en andere achtergrondvariabelen van leerlingen gecorrigeerde leerprestaties. Het merendeel van het onderzoek heeft een survey-karakter en is correlatieel van aard. Verschillen in effectiviteit die worden geconstateerd zijn altijd relatieve verschillen: effectiviteit wordt dus *niet* bepaald door de vergelijking van een prestatieniveau met een absolute criteriumgebonden maatstaf.

Uit onderzoek naar de stabiliteit van schooleffecten in het basisonderwijs en de consistentie van effecten over deelsystemen binnen de school als geheel (vakken, leerjaren, parallelklassen) komt naar voren dat het concept een matige robuustheid heeft. Dat wil zeggen dat in globale zin wel gesproken kan worden van de effectiviteit van de school als geheel, en van effectiviteitsbevorderende kenmerken die generaliseerbaar zijn over deelsystemen binnen de school, maar dat voor fijnmaziger analyses beter partiële modellen (bijvoorbeeld per schoolvak) kunnen worden getoetst. In het voortgezet onderwijs worden er binnen scholen betrekkelijk grote verschillen tussen leerkrachten en vakken gevonden.

De marge waarbinnen prestatieverschillen tussen scholen in westerse landen kunnen worden verklaard zijn vrij smal. Door de factor school wordt "netto" zo'n 9% variantie in de leerprestaties gebonden. De correlaties met afzonderlijke effectiviteitsbevorderende factoren zijn over het algemeen laag. De set variabelen die gaandeweg naar voren is gekomen als de kennisbasis van effectiviteitsbevorderende condities verklaart wel een redelijk deel van de tussenschoolse variantie (meer dan 50%). Dat de verschillen tussen scholen klein zijn, betekent echter nog niet dat ze onbelangrijk zijn, vooral als het gaat om de achterstand van de ca. 10% van de scholen aan het "lage" eind van de verdeling.

In de hoofdtekst zijn verschillende overzichten gegeven van de voornaamste effectiviteitsbevorderende factoren op school- en klasniveau. Meestal wordt hierbij een driedeling gemaakt in a) materiële en financiële hulpbronnen, b) organisatie- en managementcondities en c) instructiecondities. Op hoofdlijnen samengevat zijn binnen deze driedeling de volgende condities van belang:

- a) *materiële en financiële hulpbronnen*: kosten per leerlingen blijken "ertoe te doen" als de verschillen niet te klein zijn. Hedges et al. (1994, p. 11) becijferen dat een stijging van zo'n 10% in uitgaven per leerling aanleiding zou geven tot een toename van .7 standaarddeviatie in prestaties. Opgemerkt moet worden dat een dergelijke uitkomst grote "overlap" heeft met conclusies in verband met effecten van klassenverkleining, omdat de personeelskosten nu eenmaal verreweg de grootste kostenpost in het onderwijs vormen.
- b) *organisatie- en managementcondities*: de condities op dit niveau, die er "toe doen", kunnen worden samengevat onder de noemers "*opbrengstgerichtheid*" en "*planmatigheid*". Deze worden gedragen door zogenoemd "onderwijskundig leiderschap", "een gedeelde visie binnen de staf die gekenmerkt wordt door positieve verwachtingen over de mogelijkheden van de leerlingen", "een sterke gerichtheid op de kernvakken", en een ordelijk klimaat. Een belangrijke component van planmatigheid is het *evaluerend vermogen* van de school; het toepassen en benutten van evaluatie en informatie over het functioneren van leerlingen en leerkrachten.
- c) *condities voor effectieve instructie*: op dit niveau is een tweedeling te maken tussen condities die te maken hebben met de *kwantiteit* en doelgerichtheid van de instructie, namelijk effectieve leertijd als resultante van goed klasmanagement en de passing van het aanbod bij de toetsing, en condities die te maken hebben met de *kwaliteit* van de instructie. Wat de kwaliteit van de instructie betreft heeft het effectiviteitsonderzoek uitgewezen dat vooral *gestructureerd* onderwijs effectief is. Door onderwijspsychologische ontwikkelingen, geïnspireerd door het constructivisme, is hierop een duidelijke nuancering aangebracht, waarin met name ook zelfstandig en ontdekkend leren worden benadrukt. Gemeenschappelijke kenmerken zijn echter *expliciteit* van de instructie in de zin dat er duidelijke doelen worden gesteld en er een adequate controle van de leerprocessen en leerresultaten plaatsvindt, inclusief diagnoses van leerproblemen en remedieering daarvan.

Als tweede hoofdaspect van de kwaliteit van de instructie kan "*adaptiviteit*" worden genoemd. Mits organisatorisch beheersbaar blijkt een

bepaalde differentiatie binnen klassenverband effectief te zijn. Voorbeelden van doorbreking van het jaarklassensysteem, zoals ook combinatieklassen, hebben meestal geen effecten (dus ook geen negatieve) op de cognitieve prestaties (vgl. Veenman, 1996).

De ecologische validiteit van deze verzameling van condities is aangetoond in schoolverbeteringsprojecten, zoals in Nederland het Rotterdamse KEA-project en het Landelijk Project Schoolverbetering.

Uit een nadere interpretatie van de set van condities aan de hand van conceptuele modellen en mogelijk achterliggende theoriegebonden mechanismen zijn drie conclusies te trekken:

- a) Het ligt het meest voor de hand om de interrelatie tussen schoolorganisatorische en instructiecondities als *faciliterend* te zien. Het weinige onderzoek op dit terrein heeft echter nog nauwelijks ondersteuning voor deze hypothese geboden.
- b) Een aanvullende interpretatie van tussen-niveau relaties is de gedachte van *subsidiariteit*: een hoger beïnvloedingsniveau moet niet proberen activiteiten over te nemen die ook op een lager niveau kunnen worden gedaan. Constructivistisch leren en de rol van de leerkracht als "facilitator" van leerprocessen passen in deze gedachte. Sturen op hoofdlijnen, door opbrengsten te controleren en processen vrij te laten, eveneens.
- c) Het cybernetisch principe van *evaluatie-feedback* en *bijsturing* heeft een prominente plaats binnen meerdere in aanmerking komende theoretische verklaringen van schooleffectiviteit; dit mechanisme wordt dan ook gezien als een basisconditie voor effectiviteitsbevordering. Andere condities kunnen hier dan op voortbouwen en bij aansluiten.

Het theoretisch interessantste principe, evaluatie-feedback en bijsturing, wordt tevens gezien als het beste beleidsrelevante antwoord als het gaat om een beleid gericht op kwaliteitsverbetering in het basisonderwijs (als beleid "op zich", of als flankerend beleid bij klassenverkleining). Hiertoe zou de ontwikkeling van een "eindtermen-valide" leerlingvolgsysteem, bestaand uit criteriumgeoriënteerde toetsen, gestimuleerd moeten worden,

alsmede de uitbreiding ervan tot een integraal schoolzelfevaluatie-instrumentarium voor basisscholen¹.

Een centraal vastgesteld stelsel van geconcretiseerde eindtermen en tussendoelen, waar de commissies Van Eindhoven en Kohnstamm voor pleiten, is geen absolute voorwaarde voor het beoogde zelfevaluatie-instrumentarium, maar zou de betekenis ervan wel eenduidiger maken.

Naast een dergelijke "breedtestrategie" zou de overheid als "dieptestrategie" een extra stimulans kunnen geven aan onderzoeks- en ontwikkelingsprojecten gericht op integrale schoolverbetering en experimenten met varianten van effectiviteitsbevorderende maatregelen, gevarieerde uitwerkingen van klassenverkleining inbegrepen.

¹ Een aanzet hiertoe wordt al gegeven in het SVO-project "Schoolzelfevaluatie", waarin is samengewerkt tussen CITO, SLO en OCTO.

Inhoudsopgave	Pagina
1. Inleiding; probleemstelling	1
2. Beschrijving van het onderzoeksterrein	5
2.1 Onderwijseffectiviteitsonderzoek: algemene typering en specifieke soorten	6
2.1.1 Onderwijsproductiefuncties	7
2.1.2 Instructie-effectiviteitsonderzoek	8
2.1.3 Schooleffectiviteitsonderzoek	9
2.2 Het antwoord op fundamentele vragen rondom het concept schooleffectiviteit	11
2.2.1 De grootte en betekenis van schooleffecten	11
2.2.2 De reikwijdte en generaliseerbaarheid van schooleffectsindices	12
3. De kennisbasis die het onderwijseffectiviteitsonderzoek tot nu toe heeft opgeleverd	14
3.1 De betekenis van de factoren die geacht worden te "werken"	14
3.2 De empirische evidentie voor de factoren die geacht worden te werken	17
3.3 Resultaten van Nederlands onderzoek	23
4. Interpretatie van de resultaten	28
4.1 Algemene conclusies	28
4.2 Modelmatige en theoriegebonden interpretatie	29
4.3 Directe instructie en constructivisme	31
5. Beleidsimplicaties	35
5.1 Evaluatie en feedback als kern van effectiviteitsbevordering	36
5.2 Voorschrijvend of voorwaardescheppend beleid?	37
Referenties	40

1. Inleiding; probleemstelling

De aanleiding voor deze studie is een verzoek van het Ministerie van OC&W om een overzicht samen te stellen van de literatuur uit binnen- en buitenland inzake factoren in het onderwijs die een kwaliteitsverhogend effect hebben. De context van deze vraag is het zogenoemde "flankerende beleid" dat te overwegen is in aanvulling op eventuele maatregelen ter verkleining van de klassen in het basisonderwijs. Dergelijk beleid wordt overwogen in verband met het advies van de Commissie Van Eindhoven over de klassengrootte in het basisonderwijs.

Het centrale begrip in de vraagstelling is de "kwaliteit" van het onderwijs. In onderwijskundige adviezen wordt het begrip "kwaliteit" nogal eens afgedaan als niet goed definieerbaar (vgl. het rapport Schaal en Kwaliteit in het Basisonderwijs uit 1990) dan wel "vaag" (citaat van In 't Veld, als auteur van de studie "Toekomst voor het Funderend Onderwijsbeleid" in de NRC van 3 december 1996).

Een dergelijke opvatting is zeer onbevredigend. Bevordering van de kwaliteit van het onderwijs moet worden beschouwd als het ultieme doel van onderwijsbeleid. Beleidsadviezen waarin de betekenis van kwaliteit in het midden wordt gelaten, lijken hun analyses dan ook bij voorbaat op drijfzand te plaatsen. Om dit laatste te voorkomen, zal de invulling die in deze studie wordt gegeven aan het begrip "kwaliteit" nauwkeurig worden omschreven.

Kwaliteit zou men, in de meest algemene zin, en vrij naar Van Dale, kunnen omschrijven als een hoedanigheid waaraan een positieve waardering wordt gegeven. Drie vragen moeten vervolgens gesteld worden om dit wat verder te concretiseren:

- a) kwaliteit van wat? (wat is het object waarvan de hoedanigheid gewaardeerd wordt);
- b) kwaliteit zoals beoordeeld door wie? (welke actor is gerechtigd het kwaliteitsoordeel te vellen);
- c) hoe is "een positieve waardering" nader te typeren? (op welke criteria beoordeelt men kwaliteit).

Ad a) In de onderhavige context wordt de school of onderwijsinstelling als geheel, dat wil zeggen de schoolorganisatie qua opbouw en functioneren, als primaire kwaliteitsdrager gezien.

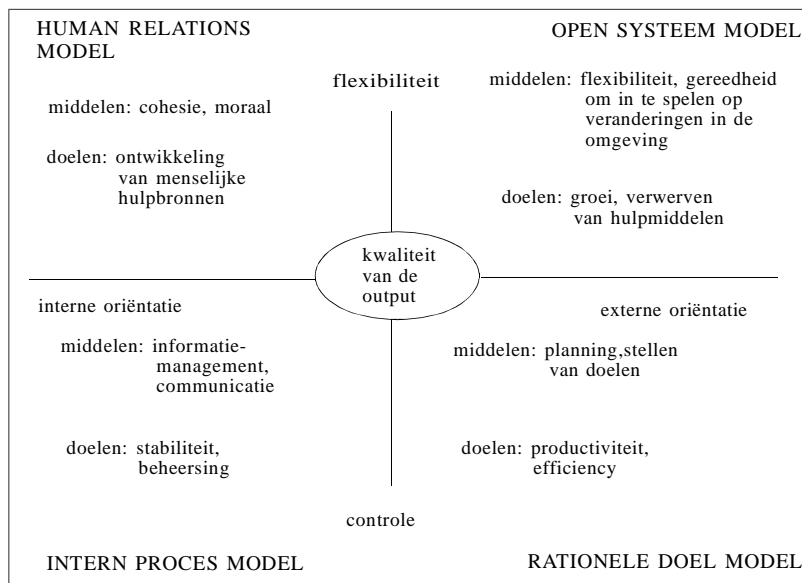
Ad b) De vraag "kwaliteit voor wie?" wordt doorgaans beantwoord door de directe afnemers van het produkt of de dienst in kwestie als de beoordelende actor te beschouwen. In het geval van scholen hebben we het dan over leerlingen en ouders, maar zeker ook over verder verwijderde afnemers zoals het vervolgonderwijs, werkgevers en de maatschappij in z'n volle breedte, zoals ook vertegenwoordigd door de rijksoverheid in het algemeen en de Inspectie van het Onderwijs in het bijzonder. Bij dit gebruik zal aangesloten worden.

Ad c) Een aangrijpingspunt om de vraag naar de waarderingsdimensies of -criteria van kwaliteit te beantwoorden is het gebruik van iets minder algemene typering van "positieve waardering", zoals deugdelijkheid, gezondheid of effectiviteit van organisaties.

Aan de keuze van zo'n kwaliteitssynoniem ligt vaak een bepaalde discipline invalshoek ten grondslag. De term deugdelijkheid past sterk in het denken van een Weberiaanse bureaucratie. "Gezondheid" sluit beter aan bij de organische systeembenadering, terwijl effectiviteit vooral in het economische denken past.

Hier zal gekozen worden voor een verdere invulling van organisatorische effectiviteit als kwaliteitssynoniem, omdat dit begrip in de literatuur het verst is uitgewerkt en een goed dekkende typering biedt van de hoofd-dimensies van (school)organisatorische kwaliteit.

Het effectiviteitsbegrip is gekoppeld aan de assumptie dat organisaties doelgericht zijn. De zich daarbij onmiddellijk aandienende volgende vraag is dan om *welke* doelen of welke soort doelen van de organisatie het gaat. In de organisatieliteratuur worden meerdere effectiviteitsperspectieven onderscheiden, die bepaald worden door een accentuering van verschillende soorten doelen en van daaraan ten grondslag liggende verschillende oriëntaties. Quinn en Rohrbaugh (1983) onderscheiden vier effectiviteitsperspectieven, die bepaald worden door twee onafhankelijke dimensies.



Figuur 1: *Typologie van effectiviteitsmodellen (naar Quinn en Rohrbaugh, 1983)*

De eerste dimensie heeft betrekking op de interne dan wel externe oriëntatie van de organisatie.

De tweede dimensie heeft te maken met de mate van structurering: oriëntatie op controle en beheersing enerzijds versus gerichtheid op flexibiliteit en verandering. Door deze twee dimensies met elkaar te kruisen, ontstaan vier kwadranten. Elk ervan representeert volgens Quinn en Rohrbaugh een effectiviteitsmodel. Dit is weergegeven in figuur 1.

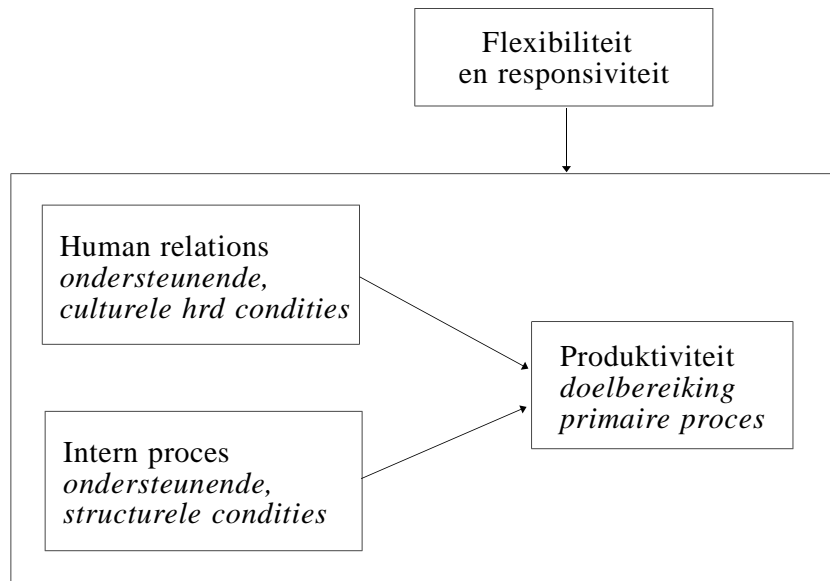
Voor een nadere bespreking van deze modellen zij verwezen naar Scheerens (1993).

Bij verdere analyse van deze vier modellen valt op dat zij onderling van nogal uiteenlopende aarde zijn.

In het "open-systeemmodel" wordt in feite de vraag centraal gesteld of de

organisatie is toegerust voor het op een juiste wijze kiezen van doelen. De oriëntatie is hier meer op "doing the right things" dan op "doing things right". Het model is het meest relevant in situaties waarin de organisatie in een dynamische omgeving verkeert. De oriëntatie op doelkeuze binnen dit model is in verband te brengen met het beeld van de lerende organisatie en met het concept van "dubbelslag leren", waarbij een organisatie de eigen doelen ter discussie stelt (vgl. Morgan, 1986, p. 88).

"Doing things right" is veeleer de oriëntatie van het rationele doel model, waarin een spaarzame en trefzekere keuze van middelen ter bereiking van de primaire produktiedoelen van de organisatie centraal staat. Het "formatief" evalueren van alternatieve middelen is een voorbeeld van "enkelslag leren", zoals dat in de conceptualisering van de lerende organisatie wordt behandeld. De beide modellen aan de linkerkant van figuur 1, het "human-relationsmodel" en het "interne-procesmodel" zijn eventueel op te vatten als "middelen" of ondersteunende condities ten opzichte van produktiviteit. Op deze wijze kunnen de vier modellen ook in onderlinge relatie worden geplaatst, zoals is gedaan in figuur 2. In deze figuur wordt het interne systeem van doelbereiking met ondersteunende culturele en structurele condities afhankelijk gezien van de externe oriëntatie (en de toerusting van de organisatie daartoe) bij de keuze van doelen.



Figuur 2: Nadere onderlinge relatering van effectiviteitsmodellen

Om het begrip "kwaliteit" hanteerbaar te maken, zijn inmiddels verschillende reducties toegepast: er is een aggregatieniveau gekozen: de school

(ook al kunnen daarbinnen sub-niveaus, klas en leerling, worden onderscheiden); externe "afnemers" zijn als belangrijkste referentiegroep voor het beantwoorden van de kwaliteitsvraag genoemd en effectiviteit, oftewel doelbereiking, is als centraal kwaliteitscriterium gekozen. Vervolgens is uit meerdere perspectieven op organisatorische effectiviteit het produktiviteitsgezichtspunt als het meest wezenlijke gekozen.

Verder toespitsing is mogelijk door aan te geven of alle mogelijke opbrengsten van het primaire proces van scholen als operationele criteria bij het meten van produktiviteit in aanmerking worden genomen of slechts een sub-set daarvan. In aanmerking komende operationele criteria zijn: rendement, in de zin van ononderbroken schoolcarrières van leerlingen, prestatieniveaus op centrale leergebieden (c.q. schoolvakken), meer algemene cognitieve vaardigheden en niet-cognitieve aspecten (sociale vaardigheden, normbesef e.d.). Hofstee (1989, p. 42) stelt in dit verband ondubbelzinnig dat "de kwaliteit van het onderwijs blijkt uit het niveau van de leerprestaties, zoals bepaald met behulp van een curriculumafhankelijke toets".

De ter beschikking staande onderzoeksgegevens in het kader van onderwijs-effectiviteitsonderzoek dwingen een verdere reductie tot een sub-set van operationele criteria af. Prestaties aan het einde van het basisonderwijs, dan wel het einde van de eerste fase voortgezet onderwijs, in centrale vakken zoals rekenen/wiskunde, taal en lezen zijn in het meeste effectiviteitsonderzoek gebruikt als effectcriteria. De hiermee gegeven beperking slaat uiteraard terug op de school- en klaskenmerken die uit dit onderzoek naar voren komen als effectiviteitsbevorderend. Het gaat hierbij dus om die factoren die een aantoonbaar effect hebben gehad op de resultaten in de desbetreffende leergebieden. Voor zover onderzocht is of scholen die het goed doen op deze cognitieve, leerstofgebonden criteria ook goed zijn op het niet-cognitieve vlak, is de voorlopige balans dat de resultaten geen eenduidig beeld geven (Scheerens, 1992).

De bovenstaande analyse van het begrip onderwijskwaliteit en de diverse reductiestappen die aan de orde zijn om van "kwaliteit" tot "resultaten van onderwijs-effectiviteitsonderzoek" te komen, geeft de beperking aan van

het kader van deze studie. De normatieve vraag hoe "erg" of "minder erg" deze beperking is, wordt overgelaten aan de lezer.

Vanuit de gedachte dat de desbetreffende leergebieden te zien zijn als de kern van het aanbod in het basisonderwijs wordt deze beperking acceptabel geacht. Tevens geldt dat - refererend aan de context van klassenverkleining van deze studie - de desbetreffende criteria ook zijn gehanteerd in het onderzoek naar de gevolgen van klassenverkleining.

Uit het advies van de cie. Van Eindhoven en daarbij gevoegde bijlagen komt naar voren dat er onder bepaalde condities een negatieve relatie tussen klassengrootte en prestaties wordt gevonden (hoe kleiner de klassen, hoe beter de prestaties).

De vraag naar de kennisbasis die relevant is met het oog op zogenoemd flankerend beleid bij klassenverkleining kan als volgt geconcretiseerd worden:

- a) welke factoren mediëren het effect van klassenverkleining, dat wil zeggen, welke effectiviteitsbevorderende factoren krijgen meer kans bij kleinere klassen?
- b) welke aanvullende factoren kunnen tezamen met klassenverkleining en de onder a bedoelde factoren een verdere effectiviteitsbevordering bewerkstelligen?

In de opbouw van dit verslag worden deze vragen aangepakt door middel van een samenvatting van de resultaten van onderwijseffectiviteitsonderzoek. Tevens wordt stilgestaan bij de interpretatie van deze resultaten. Tenslotte wordt ingegaan op de mogelijkheden om deze onderzoeksresultaten te benutten voor bepaalde beleidsmaatregelen.

2. Beschrijving van het onderzoeksterrein

De term *onderwijseffectiviteitsonderzoek* wordt wel gebruikt als een overkoepelend begrip voor onderzoek naar effectiviteitsbevorderende schoolkenmerken (schooleffectiviteitsonderzoek) en effectiviteitsbevorderende kenmerken van het onderwijs in de klas (instructie-effectiviteits-

onderzoek). Gezien het feit dat in veel recent *schooleffectiviteits*onderzoek tevens klassekenmerken worden meegenomen - en dus gewerkt wordt met gegevens op school- en klasniveau - is het onderscheid tussen "schooleffectiviteit" en "onderwijseffectiviteit" betrekkelijk arbitrair.

2.1 Onderwijseffectiviteitsonderzoek: algemene typering en specifieke soorten

Uit het onderzoek in de traditie van de gelijke kansenproblematiek, alsmede het onderzoek naar het effect van compensatieprogramma's in het onderwijs, is ondubbelzinnig gebleken dat verschillen in leerprestaties tussen leerlingen in belangrijke mate bepaald worden door verschillen in uitgangspositie van leerlingen, qua aanleg en milieu. Uit de onderzoeken van Coleman et al. (1966) en Jencks et al. (1972) bleek dat de invloed van in het onderwijs te manipuleren factoren hiermee vergeleken betrekkelijk klein was. Ruwweg wordt van de variantie in leerlingprestaties zo'n 40 à 50% gebonden door aanleg- en milieufactoren, 40% blijft onverklaard, terwijl de factor school zo'n 10 à 20% variantie bindt. In meer recente onderwijseffectiviteitsonderzoeken zijn de percentages door de factor school gebonden variantie vaak nog lager (vgl. bijvoorbeeld Hill, Rowe & Holmes-Smith, 1995).

Ondanks de betrekkelijk smalle marges waarbinnen het bezoeken van een bepaalde school en het daar gegeven onderwijs invloed heeft op de leerprestaties is hiermee toch het interessegebied van het onderwijseffectiviteitsonderzoek aangegeven. In de onderzoekspraktijk betekent dit dat gezocht wordt naar de "added value", oftewel de toegevoegde waarde van manipuleerbare school- en klassecondities, ongeacht de effecten van een al dan niet gunstige uitgangssituatie van de leerlingen. In onderzoektechnische termen wordt dit uitgedrukt als het bepalen van effecten, terwijl gecontroleerd wordt voor relevante achtergrondkenmerken van leerlingen. Globaal gezegd komt het onderwijseffectiviteitsonderzoek dus neer op het meten van leerprestaties op enig tijdstip in de schoolcarrière, het meten van achtergrondkenmerken en prestaties op een eerder tijdstip en het relateren van de voor achtergrondkenmerken en beginprestaties gecontro-

leerde prestaties aan school- en klaskenmerken waarvan verwacht wordt dat ze invloed hebben op de prestaties. Afhankelijk van de gehanteerde steekproefopzet is het daarbij mogelijk school-, klas- en/of leerkracht-effecten te isoleren, dat wil zeggen de vraag te beantwoorden of verschillen in prestaties kunnen worden toegeschreven aan verschillen tussen scholen, klassen en/of leerkrachten. Daarbij is dan in eerste instantie alleen duidelijk in hoeverre de factoren school, klas en leerkracht invloed hebben. Pas wanneer men in het onderzoek ook beschrijvende kenmerken van schoolorganisatie en

-management en instructieprocessen opneemt, is inzicht te krijgen in de vraag welke specifieke kenmerken er "iets toe doen".

Bij de interpretatie van de resultaten van dit soort onderzoek moet men er in de eerste plaats op bedacht zijn dat het steeds gaat om relatieve effecten, dat wil zeggen dat effecten worden vastgesteld op basis van onderlinge vergelijking en niet door te vergelijken met een vooraf vastgestelde standaard (bijvoorbeeld een bepaald scoringsniveau op een studietoets). In theorie is het dus mogelijk dat er in een land geen school-effectiviteit wordt geconstateerd, terwijl het onderwijspeil toch voldoende hoog is. Dit zou kunnen voorkomen wanneer er sprake is van een uiterst homogeen systeem, met nauwelijks verschillen tussen scholen.

In de tweede plaats is het onderzoek in feite gericht op de specificatie en empirische toetsing van causale modellen, waarbij het maar zelden mogelijk is dit volgens de daartoe gangbare experimentele methode te doen. Meestal is er slechts sprake van correlationeel onderzoek, waarbij causale interpretaties afhankelijk zijn van in de praktijk voorkomende variantie. Dit noopt tot de nodige terughoudendheid in de causale interpretatie; soms is bij dit onderzoek minder duidelijk wat de "oorzaken" en wat de "gevolgen" zijn, met name is dit het geval bij variabelen die gerelateerd zijn aan het concept "schoolcultuur", zoals prestatiegerichtheid, het hebben van positieve verwachtingen over de prestaties van leerlingen e.d.

Naast deze elementaire vragen over de aard en de interpretatie van onderwijseffectiviteitsonderzoek is er nog een categorie van fundamentele vragen over het concept school, c.q. onderwijseffectiviteit, welke door middel van empirisch onderzoek te beslechten zijn. Het gaat hier om de

vraag van de reikwijdte en "duurzaamheid" van schooleffectiviteit, waarbij bijvoorbeeld wordt nagegaan of een school die effectief is voor het vak rekenen dat ook is voor taal en of een school die het ene jaar effectief is, dat drie jaar later nog steeds is.

Deze vragen zullen in de volgende paragraaf kort aan de orde worden gesteld, waarbij de "stand van het weten" op dit terrein zal worden weergegeven. Ter nadere bepaling van het terrein van onderwijseffectiviteitsonderzoek wordt nu eerst een overzicht gegeven van de voornaamste soorten van dit onderzoek.

Voor een uitvoerige beschrijving van de historische ontwikkeling van het schooleffectiviteitsonderzoek zij verwezen naar Scheerens, 1992 (hfst. 3). Hier wordt volstaan met de bespreking van drie soorten onderzoek: economisch georiënteerd onderzoek naar "onderwijsproductiefuncties", instructie-effectiviteitsonderzoek en onderwijskundig onderzoek naar meer-niveau schooleffectiviteitsmodellen.

2.1.1 Onderwijsproductiefuncties

In dit economisch georiënteerde type effectiviteitsonderzoek ligt het accent op de impact van materiële inputs en "resources" op de onderwijsopbrengsten. In dit type onderzoek worden doorgaans dezelfde soort opbrengstgegevens, voor achtergrondkenmerken gecontroleerde prestaties in de basisvakken, gebruikt als in het onderwijskundig onderzoek (vgl. Monk, 1989, 1992). Scheerens en Bosker, 1997 (hfst. 2) vatten de kenmerken van het onderzoek naar onderwijsproductiefuncties als volgt samen:

- materiële en financiële hulpbronnen worden gekozen als de belangrijkste beïnvloedende variabelen;
- het bepalen van directe, in plaats van causaal gemedieerde effecten; Monk spreekt in deze van "fundamentally black-box formulations" (Monk, 1992, p. 309);
- gegevens die op één aggregatieniveau geanalyseerd worden, het leerling- dan wel het schoolniveau;
- het gebruik van een statisch, in tegenstelling tot een dynamisch (d.w.z.

in de tijd variërend) model.

Typische variabelen die in deze onderzoekstraditie gebruikt zijn, betreffen leerling/leerkracht-ratio, het salarisniveau van de leerkracht, ervaring van de leerkracht en de kosten per leerling. Op de nadere inhoud van deze variabelen en de resultaten van dit onderzoek wordt ingegaan in volgende paragrafen.

2.1.2 *Instructie-effectiviteitsonderzoek*

Oorspronkelijk leek het impliciete model van het instructie-effectiviteitsonderzoek, "educational productivity research" in de termen van Walberg (1984), sterk op het produktiefunctiemodel. Daarnaast heeft het instructie-effectiviteitsonderzoek wortels in onderzoek van het "onderwijzen" (research on teaching). Vanuit deze achtergronden werden gaandeweg modellen getoetst die een complexere causale structuur hadden, bijvoorbeeld door interacties tussen leerlingkenmerken en instructievormen te onderzoeken ("aptitude treatment interactions") en ook indirecte relaties (langere causale ketens) tussen omgevingskenmerken en instructievariabelen in beschouwing te nemen (vgl. Reynolds & Walberg, 1990).

Een elementair model bij het instructie-effectiviteitsonderzoek is het model van Carroll (1963) - zie ook Carroll (1989). Centrale elementen in dit model zijn:

- de daadwerkelijk bestede leertijd, die gezien wordt als resultante van:
 - de leermotivatie ("perseverance")
 - de roostertijd ("opportunity to learn");
- de benodigde leertijd, als resultante van:
 - het specifieke leertempo van de leerling ("aptitude")
 - de kwaliteit van het onderwijs ("quality of instruction")
 - het begripsvermogen van de leerling ("ability to understand instruction").

Leerprestaties zijn in het Carroll-model de resultante van vermogen en bereidheid om te leren (leerlingkenmerken dus) de onderwijstijd en de kwaliteit van de instructie. In latere modellen is vooral de factor "kwaliteit" van de instructie, in de zin van de kwaliteit van het *instructieproces*, nader gespecificeerd, bijvoorbeeld in het model van beheersingsleren

(Bloom, 1976) en het model van "direct instruction". Rosenshine (1983) en Doyle (1985) geven de kenmerken van "direct instruction" als volgt weer:

- 1) De leerdoelen worden duidelijk geformuleerd.
- 2) De voortgang door de leerstof is nauwgezet in taken uiteengelegd en op volgorde geplaatst.
- 3) De docent legt duidelijk uit wat de leerlingen moeten leren.
- 4) Leerlingen moeten aangetoond hebben bepaalde basisvaardigheden te beheersen voordat "hogere" en moeilijker taken worden aangeboden.
- 5) De docent checkt regelmatig de vorderingen van de leerlingen; wanneer er problemen en tekortkomingen worden gesignaleerd volgt daarop toegesneden remediëring.
- 6) De leerlingen krijgen ruimschoots de tijd om te oefenen, waarbij veelvuldig gebruik wordt gemaakt van "opstapjes" ("prompts") en feedback.
- 7) Gebruik van criterium-georiënteerde toetsen.

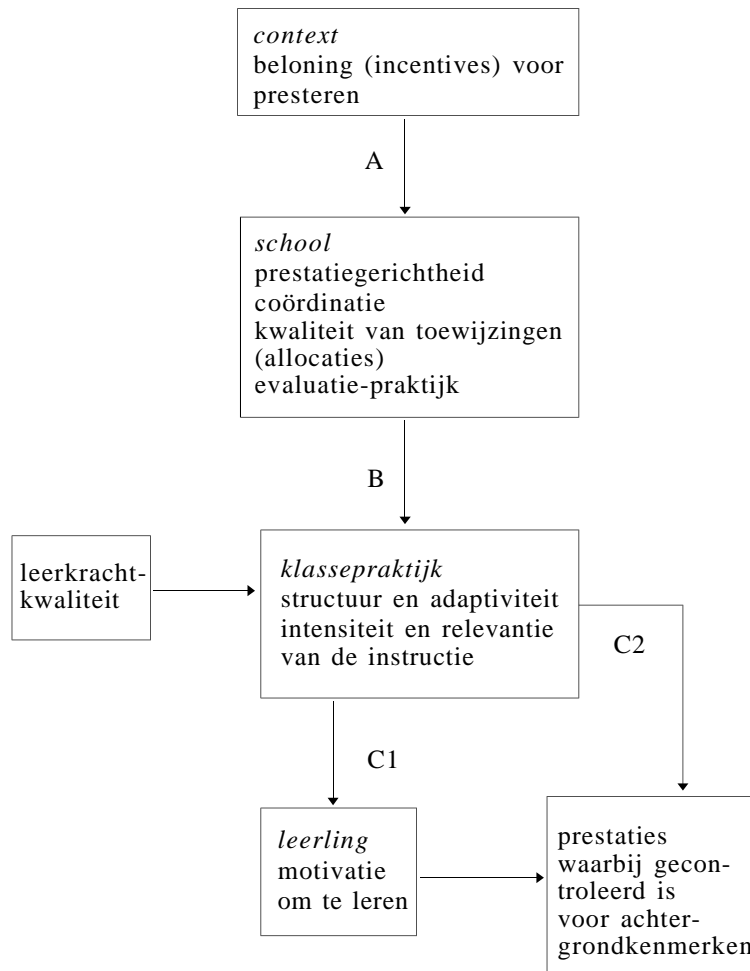
Creemers (1994) geeft een verdere uitsplitsing van instructie-effectiviteit door verschillende "dragers" toe te kennen aan kenmerken zoals genoemd in het "direct instruction" model, nl. het curriculum, groeperingsvormen en leerkrachtgedrag.

In hoeverre het paradigma van het constructivisme te rijmen is met de uitkomsten van instructie-effectiviteitsonderzoek, dan wel daar diametraal tegenover staat, zal in een volgende paragraaf worden nagegaan.

2.1.3 *Schooleffectiviteitsonderzoek*

Het moderne schooleffectiviteitsonderzoek integreert de beide eerder genoemde onderzoeksvormen en voegt er een derde categorie variabelen aan toe, namelijk schoolkenmerken die betrekking hebben op schoolcultuur, structurele aspecten van de schoolorganisatie en schoolmanagement. Incidenteel worden ook nog variabelen op het niveau van de omgeving van de school onderscheiden. Het desbetreffende onderzoek wordt geleid door integrale onderwijseffectiviteitsmodellen, vgl. Scheerens & Creemers (1989).

Een voorbeeld van een dergelijk model is dat van Bosker en Scheerens (1994).



Figuur 3: *Dimensies van schooleffectiviteit binnen een meer-niveau structuur.*

Scheerens & Bosker (1997) onderscheiden de volgende hoofdkenmerken van integrale schooleffectiviteitsmodellen:

- de beïnvloedende factoren die worden onderscheiden kunnen gecategoriseerd worden als inputs, processen en de context van het functioneren van scholen; het zijn in feite productiefunctiemodellen die ingebed zijn in organisatie- en omgevingsfactoren;
- er is sprake van een multi-niveau structuur van scholen binnen schoolomgevingen, klassen binnen scholen en leerlingen binnen klassen;
- in recente formuleringen worden complexe causale relaties gecombineerd met de multi-niveau structuren en worden niet-lineaire en dynamische aspecten onderscheiden (cf. Bosker & Scheerens, 1994).

Gegeven de meer-niveau structuur van het onderzoek en de daaraan ten grondslag liggende modellen, zijn hypothesen omtrent relaties tussen variabelen op verschillend niveau van centrale betekenis. De algemene gedachte is dat factoren op hoger aggregatieniveau, bv. kenmerken van schoolmanagement, condities op lager niveau, bv. kenmerken van effectieve instructie in klassen, faciliteren. Binnen deze algemene typering zijn echter uiteenlopende inhoudelijke interpretaties en formele specificaties te geven (Bosker & Scheerens, 1994). Op dit aspect zal verderop, bij de bespreking van de interpretatie van de onderzoeksresultaten, worden teruggekomen.

Het multi-level onderzoek geeft tevens uitsluitsel over de vraag vanuit welk niveau de beïnvloeding van leerprestaties het sterkst is.

2.2 Het antwoord op fundamentele vragen rondom het concept school-effectiviteit

De soort resultaten die worden opgeleverd door het onderwijseffectiviteitsonderzoek, zoals beschreven in het voorafgaande, corresponderen niet vanzelf met uitspraken als: school A is effectief en school B is dat niet. De robuustheid van de uitspraken over scholen op basis van het effectiviteitsonderzoek hangt af van twee zaken:

- a) de grootte van de verschillen tussen scholen;
- b) de reikwijdte van schooleffectsindices; waarbij het gaat om de stabiliteit en de generaliseerbaarheid van deze indices over vakken en leerjaren.

2.2.1 *De grootte en betekenis van schooleffecten*

Door Scheerens en Bosker (1997) wordt, op basis van een meta-analyse van 89 onderzoeken een netto effectgrootte voor de verschillen tussen scholen van .30 gevonden. De coëfficiënt van de effectgrootte is gebaseerd op een in meta-analyses gebruikelijke maat d , die gedefinieerd is als de wortel uit de intra-klassecorrelatie (dat is de ratio van de tussenschoolse en de totale variantie) gedeeld door $1 -$ de intra-klassecorrelatie). Het feit dat er gesproken wordt van een "netto" effectgrootte wil zeggen dat niet de "ruwe" schoolgemiddelden, maar de voor beginkenmerken van leerlingen gecorrigeerde gemiddelden tussen scholen gebruikt zijn bij de berekening van de intra-klassecorrelatie.

Afgemeten aan de maatstaven die Cohen (1969) aanlegt om kleine, gemiddelde en grote effecten te onderscheiden, gaat het hierbij om een klein tot gemiddeld effect.

Uitgedrukt in de in de schooleffectiviteitsliteratuur meer gebruikelijke aanduiding van het schooleffect als de proportie van de variantie tussen leerlingen in leerprestaties die gebonden wordt door de factor school is dit een waarde van 9%.

Bij de praktische interpretatie van de gevonden effectgroottes moet in aanmerking worden genomen dat er bij de schatting van het effect van een school sprake is van aanzienlijke standaardmeetfouten, hetgeen betekent dat de schatting van een gemiddelde schoolscore door tamelijk brede onzekerheidsmarges (betrouwbaarheidsintervallen) worden omgeven. Gegeven deze onzekerheidsmarges zijn alleen verschillen tussen scholen aan de uiteinden van de verdeling van gemiddelde scores significant. Zie ook Goldstein, 1996. Voor de praktijk die men in Engeland toepast, het plaatsen van scholen op ranglijsten, zogenoemde "league tables", is dit een uiterst storende conclusie, omdat men - zelfs bij het gebruik van "value-added" effectmaten - geen betrouwbaar onderscheid kan maken binnen het grote middengebied van scholen.

Wanneer men echter een vergelijking maakt tussen de 10% hoogst scorende en 10% laagst scorende scholen, dan zijn er wel degelijk belangrijke maatschappelijke implicaties. In Nederland komt het verschil tussen de 10% hoogst scorende en de 10% laagst scorende scholen neer

op een waarde van .65 van een standaarddeviatie. Dit verschil - gemeten aan het eind van de basisschool - correspondeert met een mavo/lbo-advies in de minst effectieve scholen, tegenover een havo/vwo-advies in de meest effectieve scholen (waarbij het dus gaat om een vergelijking van leerlingen die qua beginniveau niet verschillen).

Verder geldt dat men in feite een netto-schooleffect zou moeten vermenigvuldigen met het aantal leerlingen op een school; immers het maatschappelijk rendement van het bezocht hebben van een effectieve school komt in principe² ten goede aan alle leerlingen van die school.

Tenslotte is er bij de interpretaties van de betekenis van de verschillen tussen scholen nog de kanttekening te maken dat binnen een over het algemeen redelijk functionerend stelsel van basisscholen zoals in Nederland (vgl. de rapportage van de CEB, 1994) er kennelijk sprake is van een betrekkelijk smalle marge, waarbinnen spontaan in de praktijk voorkomende verschillen in beheersbare procescondities effect sorteren. Desalniettemin kunnen de marges belangrijk genoeg worden geacht om nader te bepalen waardoor ze veroorzaakt worden.

2.2.2 De reikwijdte en generaliseerbaarheid van schooleffectsindices

In het onderwijseffectiviteitsonderzoek worden in feite per schoolvak aparte causale modellen getoetst. Tegelijkertijd heeft het concept "school-effectiviteit" een bredere pretentie. De impliciete assumptie is dat een effectieve school niet alleen goede resultaten laat zien voor één schoolvak, maar voor het hele curriculum, of in ieder geval voor alle kernvakken. Tevens is de impliciete aanname dat een school niet alleen gedurende één schooljaar effectief is, maar dat blijvend is. En tenslotte is het de vraag of men de effectiviteit van een (basis)school uitsluitend moet afmeten aan de prestaties in het laatste leerjaar, dan wel tevens verwacht dat een effectieve school ook goed scoort op het niveau van, bijvoorbeeld, groep 4. Al deze vragen zijn in principe door middel van empirisch onderzoek te beantwoorden. Scheerens (1993) duidt dit soort onderzoek

² Aannemende dat de school niet differentieel effectief is, d.w.z. even effectief voor begaafde leerlingen als voor minder begaafde leerlingen.

aan als "funderend".

Op basis van een recent overzicht van Scheerens & Bosker (1997, hfst. 3) geeft tabel 1 een globaal overzicht.

Tabel 1: *Overzicht van stabiliteits- en consistentie-indices, naar Scheerens & Bosker, 1997, hfst. 3*

type consistentie/generaliseerbaarheid	gemiddelde correlatie
stabiliteit eindscore b.o. rekenen en taal (tijdsinterval 1 à 2 jaar)	$r = .70$ (range .34 - .87)
consistentie over leerjaren binnen scholen (rekenen en taal)	$r = .50$ (range .20 - .69)
consistentie over vakken (taal vs. rekenen)	$r = .70$ (range .59 - .83)

Scholen bleken verder stabiel in effectiviteit voor leerlingen met een lage socio-economische status dan voor leerlingen met een hogere socio-economische status.

In navolging van Scheerens en Bosker (ibid) zijn de volgende conclusies te formuleren inzake de reikwijdte en generaliseerbaarheid van school-effectsindices:

- 1) Bij relatief korte tijdsintervallen (1 à 2 jaar) zijn schooleffecten gemeten aan het eind van de opleiding betrekkelijk stabiel.
- 2) Er is een betrekkelijk lage consistentie in effecten wanneer verschillende leerjaren worden vergeleken.
- 3) De consistentie tussen vakken is hoger in het basisonderwijs dan in het voortgezet onderwijs.

Zowel de betrekkelijk lage consistentie tussen leerjaren in het basisonderwijs en de betrekkelijk lage consistentie tussen vakken in het voortgezet onderwijs wijzen op het belang van de invloed van de leerkracht. Scheerens en Bosker merken op dat de conclusie dat leerkrachteffecten sterker zijn dan schooleffecten vrijwel onontkoombaar is.

Deze beschouwing over fundamentele vragen betreffende het concept schooleffectiviteit, in de zin van vragen over de grootte van effecten en de reikwijdte van deze effecten voert tot de conclusie dat bij de interpretatie van de onderzoeksuitkomsten enige voorzichtigheid geboden is.

De gegevens over de grootte van schooleffecten laten zien dat een realistische verbeteringstarget eruit zou kunnen bestaan de 10 à 15% van de scholen aan het laagst scorende uiteinde van de verdeling op het niveau van de "gemiddelde" school te brengen. De uitkomsten van het onderzoek naar stabiliteit en consistentie van schooleffecten wijzen uit dat schooleffectiviteit niet zondermeer als een integraal kenmerk mag worden beschouwd, maar in niet onbelangrijke mate berust op sub-systemen binnen de school, zoals klassen en leerkrachten.

Deze constatering biedt steun voor de gedachte dat beleidsmaatregelen, gericht op kwaliteitsverbetering, zich niet tot het niveau van het schoolmanagement zouden moeten beperken, maar direct of indirect de leerkrachten moeten bereiken.

3. De kennisbasis die het onderwijseffectiviteitsonderzoek tot nu toe heeft opgeleverd

Uitgaande van de meest omvattende stroming in het onderwijseffectiviteitsonderzoek, het onderzoek naar integrale, multi-niveau schooleffectiviteitsmodellen, zijn er drie categorieën van beïnvloedende factoren te onderscheiden:

- materiële en financiële inputs;
- processen en structuren op schoolniveau, eventueel aangevuld met kenmerken van de omgeving van de school;
- processen op klasniveau.

In het overzicht van de beschikbare kennis op het terrein in kwestie zal eerst een schets worden gegeven van de inhoud en betekenis van de belangrijkste variabelen die binnen de genoemde drie categorieën zijn onderzocht. Vervolgens wordt achtereenvolgens een overzicht gegeven van de resultaten van enkele recente onderzoeksreviews, worden de resultaten van meta-analyses samengevat en wordt afzonderlijk stilgestaan bij het Nederlandse onderzoek op dit terrein.

3.1 De betekenis van de factoren die geacht worden te "werken"

De definitie van de "resource-input" factoren is minder gecompliceerd dan die van de meest onderzochte kenmerken op klas- en schoolniveau. De variabelen die in economisch georiënteerd effectiviteitsonderzoek de meeste aandacht hebben gekregen, zijn:

- leerling/leerkracht-ratio;
- het opleidingsniveau van de leerkracht (bijvoorbeeld het aantal jaren beroepsopleiding na de middelbare school);
- de ervaring van de leerkracht (bijvoorbeeld uitgedrukt in het aantal dienstjaren);
- het salaris van de leerkracht;
- de uitgaven per leerling; deze laatste variabele is in feite een variabele die mede is samengesteld op basis van sommige van de hierboven genoemde variabelen.

Eventueel is ook de variabele "beschikbare leertijd" als een resource-input te zien. De beschikbare tijd kan, variërend naar mate van verfijning worden getypeerd. De meest "bruto" maat is het aantal dagen (eventueel het aantal klokuren) dat een leerling geacht wordt op school te zijn. Vervolgens kan, liefst op schoolniveau, nagegaan worden hoeveel tijd er op het lesrooster is ingeruimd voor de diverse schoolvakken. Tenslotte is de tijd, die de leerlingen tijdens de lessen daadwerkelijk actief zijn, bepaald. De "time on task" op school moet eventueel aangevuld worden met de tijd die door de leerling aan huiswerk wordt besteed.

Een variabelencategorie die in het effectiviteitsonderzoek nog weinig aandacht heeft gekregen is de verdeling van leerkrachten over groepen leerlingen. Zo stelt Monk (1992) voor om "moeilijke" groepen leerlingen te matchen met extra goede leerkrachten. De variatie van de klassegrootte, al naar gelang de sociaal-economische status of leeftijd van de leerlingen, past ook in deze variabelencategorie.

De variabelen op schoolniveau die onderzocht zijn in het schooleffectiviteitsonderzoek worden meestal onder tamelijk algemene labels, zoals "een opbrengstgericht klimaat" of "onderwijskundig leiderschap" samengevat. Uit nadere analyse (Scheerens & Bosker, 1997, hfst. 4) blijkt dat er vele

facetten aan deze algemene concepten zitten. Deze complexiteit brengt met zich mee dat er in verschillende onderzoeken nogal uiteenlopende operationalisaties van de factoren in kwestie zijn gekozen.

In de onderstaande tabel, ontleend aan Scheerens & Bosker (1997), wordt een overzicht gegeven van de componenten die binnen de belangrijkste algemene factoren op schoolniveau worden onderscheiden.

Tabel 2: Componenten van effectiviteitsgerichte factoren

Factoren	Componenten
Prestatiegerichtheid, hoge verwachtingen	<ul style="list-style-type: none"> • het centraal stellen van basisvaardigheden • hoge verwachtingen (schoolniveau) • hoge verwachtingen (klasniveau) • het vastleggen van leerlingresultaten
Onderwijskundig leiderschap	<ul style="list-style-type: none"> • algemene leiderschapsvaardigheden • informatievoorziening • (participatie in) besluitvorming • coördinatie • tijdsbesteding aan algemeen en onderwijskundig leiderschap • het volgen van de vooruitgang van (groepen) leerlingen • het begeleiden van leerkrachten en bewaken van de kwaliteit • het initiëren en faciliteren van professionalisering
Samenwerking en overleg	<ul style="list-style-type: none"> • vorm en frequentie van formeel en niet-formeel overleg • inhoud van de samenwerking • tevredenheid van samenwerking • belang van samenwerking • (andere) indicatoren van succesvolle samenwerking
Curriculumkwaliteit/opportunity to learn	<ul style="list-style-type: none"> • wijze waarop accenten binnen het curriculum bepaald worden • keuze van methoden • omgang met methoden • opportunity to learn • tevredenheid m.b.t. het curriculum
Schoolklimaat	<p><i>a) ordelijk klimaat</i></p> <ul style="list-style-type: none"> • belang van ordelijk klimaat • regels en procedures • straffen en belonen • omgang met verzuim, lesuitval en voortijdig schoolverlaten • (het sturen van) goed gedrag van leerlingen • tevredenheid m.b.t. een ordelijk schoolklimaat <p><i>b) klimaat in termen van effectiviteitsoriëntatie en goede interne relaties</i></p> <ul style="list-style-type: none"> • prioriteiten ten aanzien van een effectiviteitsverhogend schoolklimaat • percepties ten aanzien van effectiviteitsverhogende condities • relatie leerlingen onderling • relatie leerkracht-leerlingen • relatie leerkrachten onderling • relatie leerkrachten-schoolleiding • medeverantwoordelijkheid van leerlingen • waardering leerkracht voor functie en taken • waardering voor het werk wat betreft faciliteiten, arbeidsvoorwaarden en promotiemogelijkheden, werkbelasting en algemene tevredenheid • middelen en gebouw(en)
Evaluerend vermogen	<ul style="list-style-type: none"> • het belang van evalueren • het evalueren van leerlingresultaten • het gebruik van een leerlingvolgsysteem • het vastleggen van leerlingresultaten • evaluatie van het functioneren van de school • het gebruik van evaluatieresultaten • tevredenheid m.b.t. evaluatie-activiteiten
Betrokkenheid van ouders	<ul style="list-style-type: none"> • nadruk op ouderbetrokkenheid in schoolbeleid • contacten met ouders • tevredenheid m.b.t. betrokkenheid van ouders
Effectieve leertijd	<ul style="list-style-type: none"> • belang van effectieve leertijd • afwezigheid en te laat komen • onderwijstijd op schoolniveau • onderwijstijd op klasniveau • klasmanagement • huiswerk

De tabel is gebaseerd op een analyse van een dertigtal onderzoeksinstrumenten die gebruikt zijn in schooleffectiviteitsonderzoek en schoolzelf-evaluatiesystemen. Uit de tabel is op te maken dat de factoren niet onderling uitsluitend zijn. Er is sprake van een zekere overlap, waarbij planmatigheid, prestatiegerichtheid en ordelijkheid meerdere keren voorkomen, als aandachtspunten bij klimaat, curriculum, leiderschap en opbrengstgerichtheid.

Op het niveau van de omgeving van de school worden soms classificatievariabelen gebruikt als urbaniciteit en regionale setting. Meer inhoudelijk relevante variabelen zijn eisen en standaarden die vanuit hoger bestuurlijke niveaus worden gesteld (bijvoorbeeld het districtsniveau in de Verenigde Staten), of kenmerken van het functioneren van schoolbesturen (vgl. Hofman, 1993).

Op het klasniveau zijn er twee hoofdcategorieën van variabelen: variabelen die een indicatie geven van de *hoeveelheid en de doelgerichtheid van het onderwijs* enerzijds en variabelen die zijn samen te vatten onder de noemer *instructiekwaliteit* anderzijds.

Onder de eerste categorie vallen de variabelen effectieve leertijd "time on task" en de aansluiting tussen het onderwijsaanbod en de toetsing daarvan: "opportunity to learn". Bosker en Scheerens (1994) duiden een samengestelde indicator, gebaseerd op beide componenten, aan met de term "targetedness of instruction".

De instructiekenmerken die uit het instructie-effectiviteitsonderzoek naar voren komen als indicatief voor instructiekwaliteit zijn in het algemeen samen te vatten onder de noemer van een gestructureerde aanpak. Zie de eerdere typering van "directe instructie" in par. 2.1.2.

Deelaspecten van gestructureerd onderwijs die in het onderzoek vooral aandacht gekregen hebben, zijn: het geven van feedback aan leerlingen, reinforcement, samen leren en aspecten van differentiatie en adaptief onderwijs. Binnen deze laatste categorie is zowel homogeen als heterogeen groeperen tussen en binnen klassen onderzocht (vgl. Slavin, 1996).

3.2 De empirische evidentie voor de factoren die geacht worden te werken

Op het terrein van de resource-input factoren is de meta-analyse van Hanushek uit 1986 nog steeds de belangrijkste bron. De aan hem ontleende overzichtstabel drukt vooral de zwakheid van de empirische steun voor deze variabelen uit. In een heranalyse van het gegevensbestand van Hanushek komen Hedges, Laine & Greenwald, op basis van statistische toetsing, tot iets positievere conclusies. Met name vonden zij een effect van "uitgaven per leerling" dat van voldoende praktische betekenis werd geacht (zijnde een correlatie van .20 met leerprestaties).

Tabel 3: *Een overzicht van 147 input-output studies. Een + duidt aan een positieve correlatie, terwijl een - een negatieve correlatie tussen input en output variabelen weergeeft. Gebaseerd op Hanushek, 1986, p. 1161.*

<i>Input</i>	<i>statistisch significant</i>			<i>statistisch insignificant</i>				
	<i>Aantal studies</i>	<i>+</i>	<i>-</i>	<i>insig.</i>	<i>totaal</i>	<i>+</i>	<i>-</i>	<i>Onbekende tekens</i>
Leerkracht/leerling ratio	112	9	14	89	25	43	21	
Leerkracht training	106	6	5	95	26	32	37	
Leerkracht ervaring	209	33	7	69	32	22	125	
Salaris leerkrachten	60	9	1	50	15	11	24	
Uitgaven per leerling	65	13	3	49	25	13	11	

Effectiviteitsbevorderende schoolkenmerken zijn tot nu toe voornamelijk in meer kwalitatieve reviews van de onderzoeksliteratuur beschreven. Een eerste meta-analyse op dit terrein wordt beschreven in Scheerens & Bosker (1997, hfst. 6), zie ook Witziers & Bosker (1997).

Tabel 4: *Effectiviteitsbevorderende schoolcondities in drie review studies (cursief in de kolom van de Cotton studie wijst naar sub-categorieën).*

<i>Levine & Lezotte, 1990</i>	<i>Sammons, Hillman & Mortimore, 1995</i>	<i>Cotton, 1995</i>
Opbrengstgericht klimaat	Gedeelde visie en doelen Een op leren gerichte omgeving Positieve beloning	Planning en leerdoelen Curriculumplanning en ontwikkeling
Accent op centrale cognitieve vaardigheden	Accent op onderwijs en leren	Schoolbreed accent op leren
Toetsing van leer-vorderingen	Voortgangstoetsing	Evaluatie op district-, school- en klasniveau
Praktijkgeörienteerde nascholing	Een lerende organisatie	<i>professionele ontwikkeling</i> collegiaal leren
Uitstekend leiderschap	Professioneel leiderschap	Schoolmanagement en organisatie Leiderschap en schoolverbetering Leiderschap en planning
Indringende betrokkenheid van de ouders	Samenwerking tussen ouders en scholen	Betrokkenheid van de ouders
Hoge verwachtingen	Hoge verwachtingen Rechten en verantwoordelijkheden van leerlingen	Interacties tussen leerkracht en leerlingen Interacties tussen district en school Gelijkheid Speciale programma's

In overzichtstabel 4 worden de belangrijkste effectiviteitsbevorderende schoolkenmerken samengevat, zoals gepresenteerd in drie research reviews: Levine & Lezotte, 1990; Sammons, Hillman & Mortimore, 1995 en Cotton, 1995.

Wat opvalt is dat er een sterke overeenkomst bestaat in de kernfactoren die elk van de drie reviews onderscheidt. Een meer impressionistische interpretatie van de schoolkenmerken in kwestie zou kunnen zijn dat de basisingrediënten voor een effectiviteitsbevorderende aanpak op schoolniveau bestaan uit: opbrengstgerichtheid, coördinatie, een leergerichte oriëntatie bij alle betrokkenen (ook leerkrachten en schoolleiders) en toepassing van het cybernetisch principe van evalueren en feedback, met de daarmee verbonden cognitieve (leren) en conatieve implicaties (belonen, motiveren).

Op het vlak van de instructie-effectiviteit geeft tabel 5, ontleend aan Creemers, 1994, een gedetailleerd overzicht. Wel moet worden aangetekend dat deze tabel niet gebaseerd is op een meta-analyse, maar alleen een impressie geeft van de bewijskracht voor een scala van instructievariabelen.

Tabel 5: *Overzicht van empirische evidentie voor de kenmerken van effectieve instructie, ontleend aan Creemers, 1994, p. 9*

<i>Kenmerken</i>	<i>Sterk empirisch bewijs</i>	<i>Gemiddeld empirisch bewijs</i>	<i>Plausibel</i>
Curriculum		x	
Groeperingsvormen	x		
Leerkrachtgedrag	x		
<i>Curriculum</i>			
Expliciteit en ordening van doelen en inhouden	x		
Structuur en duidelijkheid van de leerstof		x	
Voorstructurering (advance organizers)	x		
Evaluatie	x		
Feedback	x		
Correctieve instructie			x
<i>Groeperingsvormen</i>			
Mastery leren	x		
Niveau-groepen		x	
Samenwerkend leren			x
Gedifferentieerd materiaal			x
Evaluatie	x		
Feedback		x	
Correctieve instructie		x	

<i>Leerkrachtgedrag</i>			
Klasmanagement/ordelijke en rustige sfeer	x		
Huiswerk		x	
Hoge verwachtingen		x	
Duidelijke doelen		x	
Beperkte set van doelen		x	
Nadruk op basisvaardigheden			x
Nadruk op cognitief leren en transfer		x	
Structurering van de inhoud		x	
Ordening van doelen en inhoud	x		
Voorstructurering (advance organizers)		x	
Reeds aanwezige kennis		x	
Duidelijkheid van de presentatie	x		
Stellen van vragen		x	
Direct oefeningen aanbieden	x		
Evaluatie		x	
Feedback			x
Correctieve instructie			

Kwantitatieve research syntheses van Walberg (1984), Fraser et al. (1987) en Slavin (1996) laten zien dat er relatief belangrijke effecten kunnen worden aangetoond voor reinforcement en feedback en minder grote effecten voor samenwerkend leren en differentiatie binnen klasseverband. Voor een meer gedetailleerde beschrijving zij verwezen naar Scheerens & Bosker, 1997, hfst. 6.

In de onderstaande tabel, ontleend aan Scheerens en Bosker, 1997, worden de resultaten van reviews en meta-analyses nog eens samengevat. Tevens zijn in de tabel resultaten opgenomen die gebaseerd zijn op twee grootschalige internationale onderzoeken, de IEA Tweede Wiskunde Studie en de IEA Reading Literacy Study. De coëfficiënten (correlaties) die vermeld worden bij de research syntheses betreffende resource-input factoren zijn gebaseerd op de meta-analyse van Hedges et al. (1994); de kolom van correlatie-coëfficiënten betreffende de schoolorganisatiefactoren berust op de eigen meta-analyse, terwijl de coëfficiënten vermeld bij "instructiecondities" zijn overgenomen uit de meta-analyses van Fraser et al. (1987) en de "best evidence syntheses" van Slavin (1996) en, voor wat betreft "opportunity to learn" gebaseerd op de eigen meta-analyses.

Tabel 6: *Overzicht van de empirische evidentie op het terrein van school-effectiviteit, gebaseerd op kwalitatieve reviews, internationaal vergelijkende onderzoeken en meta-analyses. De getallen zijn correlaties, waarvan de betekenis als volgt geïnterpreteerd kan worden (vgl. Cohen, 1988): 0.10 klein, 0.30 gemiddeld, 0.50 groot. Het plusteken betekent een positieve invloed; n.s. betekent: niet significant. Ontleend aan Scheerens & Bosker, 1997.*

	<i>kwalitatieve reviews</i>	<i>internationale studies</i>	<i>meta-analyses</i>
<i>variabelen die betrekking hebben op materiële inputs</i>			
leerling-leerkracht ratio		-.03	.02
opleiding leerkrachten		.00	-.03
ervaring van de leerkrachten			.04
salaris van de leerkrachten			-.07 ¹⁾
kosten per leerling			.20 ²⁾
<i>Schoolorganisatiefactoren</i>			
opbrengstgericht klimaat	+		
accent op prestaties in basisvaardigheden	+	.02	.14
onderwijskundig leiderschap	+	.04	.05
monitoring/evaluatie	+	.00	.15
samenwerking/consensus	+	-.02	.02
betrokkenheid ouders	+	.08	.13
professionele ontwikkeling leerkrachten	+		
hoge verwachtingen	+	.20	
ordelijk klimaat	+	.04	.11
<i>Instructieconditie</i>			
gelegenheid om te leren	+	.15	.09
leertijd/huiswerk	+	.00/-.01	.19/.06
gestructureerd onderwijs	+	(n.s.)	.11
aspecten van gestructureerd onderwijs:		-.01 (n.s.)	(n.s.)
- samenwerkend leren			
- feedback			.27
- beloning			.48
differentiatie/adaptieve instructie			.58
			.22

¹⁾ Uitgaande van een standaarddeviatie van \$ 5000 voor het leerkrachtsalaris

²⁾ Uitgaande van een standaarddeviatie van \$ 100 voor kosten per leerling

De overzichtstabel laat zien dat de effecten van de afzonderlijke *resource-input* variabelen nihil zijn, maar dat de "container-factor" uitgaven per leerling een effect laat zien dat ligt tussen "klein" en "redelijk", wanneer men Cohen's (1988) standaarden hanteert. In de meta-analyses blijken de schoolorganisatiefactoren gemiddeld rond de .10 te correleren met prestaties, hetgeen neerkomt op als "klein" te interpreteren effecten. Opvallend is dat de effectschattingen gebaseerd op de twee internationale studies beduidend kleiner zijn, terwijl in de kwalitatieve reviews wordt aangenomen dat deze factoren zondermeer een positieve invloed hebben.

De discrepantie tussen de uitkomsten van de kwalitatieve reviews en de kwantitatieve analyses roept vragen op. Een te gemakkelijke interpretatie zou, mijns inziens, zijn de resultaten van de kwalitatieve reviews in het licht van de "hard evidence" van de meta-analyses als een overschatting af te doen. Enerzijds moet de "hardheid" van de meta-analyses niet overdreven worden en anderzijds is het betekenisvol dat onderzoekers uit verschillende landen een duidelijke gelijkkluidende mening hebben over effectiviteitsbevorderende schoolkenmerken. Ook ervaringen met andere bij het onderwijs betrokken professionals, zoals inspecteurs en nationale afgevaardigden bij het OECD-indicatorenproject, wijzen erop dat er brede overeenstemming bestaat over het belang van deze factoren. Hierbij kan overigens in het midden worden gelaten of men hierbij sterk in instrumentele termen denkt: goed voor de bevordering van leerprestaties, dan wel of er een meer intrinsieke waarde wordt toegekend aan een onderwijsinrichting die door de factoren in kwestie gekenmerkt wordt. In dit laatste geval lijkt er eerder sprake van het opvatten van de factoren in kwestie als voorbeelden van "educational good practice" (zie ook Scheerens & Ten Brummelhuis, 1996).

Op het instructieniveau worden kleine effecten gevonden voor de variabelen "opportunity to learn" en "time on task" en gemiddelde tot grote effecten voor aspecten van gestructureerd onderwijs, met name feedback en reinforcement.

3.3 Resultaten van Nederlands onderzoek

Bij de meta-analyses die zijn samengevat in de vorige paragraaf, nl. in tabel 6, voor zover het gaat om de schoolorganisatiefactoren, is nagegaan in hoeverre Nederlandse onderzoeken afwijkende resultaten te zien geven. Dit bleek alleen het geval te zijn voor de factor onderwijskundig leiderschap. Het (zeer kleine) positieve effect dat in andere landen gevonden werd, bleek niet in het Nederlandse onderzoek te vinden. Het ontbreken van een effect van onderwijskundig leiderschap is ook gedocumenteerd in verschillende publicaties van Van de Grift (bijv. Van de Grift, 1990). Uit een recent onderzoek zou echter wel een positief effect van onderwijskundig leiderschap gebleken zijn, te verklaren uit een toegenomen manifestatie van dit soort gedrag bij Nederlandse schoolleiders in het basisonderwijs (Van de Grift, persoonlijke communicatie).

In de onderstaande tabel, ontleend aan Scheerens en Creemers, 1996, p. 187, wordt een nog somberder beeld gegeven van de empirische evidentie van het "schooleffectiviteitsmodel". Slechts in een fractie van het totaal aan Nederlandse onderwijseffectiviteitsonderzoeken blijken de factoren die verondersteld worden om effectiviteitsbevorderend te zijn, inderdaad een significant positief effect te hebben. Wel geldt dat in het onderzoek in het basisonderwijs dat, tot nu toe, het beste aan in aanmerking komende methodologische eisen voldoet, nl. het onderzoek van Brandsma (1993), nog relatief de meeste positieve verbanden zijn gevonden.

Tabel 7: *Aantallen Nederlandse onderzoeken waarin bepaalde condities significant samenhangen met leeropbrengsten, na controle voor achtergrondkenmerken van leerlingen.*

	Primair onderwijs		Secundair onderwijs	
	positief verband	negatief verband	positief verband	negatief verband
Gestructureerd onderwijs/ feedback	5		1	
Leerkrachtering	3	1		1
Onderwijskundig leiderschap		2	1	
Ordelijk klimaat	2		3	1
Leerlingevaluatie	5		0	
Differentiatie		2	0	
Klassikaal onderwijs	3		0	
Prestatiegerichtheid	4		4	
Stabiliteit team/samenwerking		3		3
Leertijd/huiswerk	4		4	
Andere variabelen	16		8	
Gemiddelde tussenschoolse variantie		9		13,5
Aantal onderzoeken		29		13

Opmerking: niet alle variabelen werden gemeten in elk onderzoek.

Uit een recent internationaal vergelijkend onderzoek, uitgevoerd in het kader van het OECD/CERI onderwijsindicatorenproject, blijkt overigens dat Nederland op zes van de zeven procesindicatoren, die sterk geïnspireerd zijn op de modelvorming betreffende schooleffectiviteit, boven het landengemiddelde scoort.

Landenprofiel Nederland (uit Scheerens & Ten Brummelhuis, 1996)

(Hier country profile the Netherlands - OECD School Survey Report p. 68)

De enige indicator waarbij Nederland onder het gemiddelde scoort is onderwijskundig leiderschap.

Naast onderzoek van het survey-type, waarop de voorafgaande uitkomsten voornamelijk zijn gebaseerd, vormt de evaluatie van zogenoemde schoolverbeteringsprojecten, die geïnspireerd zijn op de assumpties en het kennisbestand inzake school- en instructie-effectiviteit, een tweede relevante onderzoeksbenadering. Voor een overzicht van evaluaties van voornamelijk Amerikaanse schoolverbeteringsprogramma's zij verwezen naar Scheerens (1992, p. 56-58). Hoewel er regelmatig positieve resultaten werden gerapporteerd, bleek de kwaliteit van het onderzoek in een aantal gevallen minder overtuigend. Meer recente Amerikaanse schoolverbeteringsprogramma's die positieve resultaten hebben laten zien zijn het programma "Success for All" van Slavin (1996) en het Barclay-Calvert programma (Stringfield, Bedinger & Herman, 1995). Kenmerkend voor deze laatste programma's zijn een sterk (voor)gestructureerde en planma-

tige aanpak met veel nadruk op diagnostiek en onmiddellijke remediëring (voor een nadere beschrijving van deze projecten zij verwezen naar Scheerens & Bosker, 1997).

Ook in Nederland zijn er de laatste jaren schoolverbeteringsprojecten voor het basisonderwijs ontwikkeld, waarin de inzichten die het school- en instructie-effectiviteitsonderzoek hebben opgeleverd richtinggevend zijn geweest.

Het eerste project in kwestie is het Landelijk Project Schoolverbetering, waaraan werd deelgenomen door 28 scholen. De doelstellingen van dit project luiden als volgt:

- 1) Het realiseren van een resultaatgeoriënteerd schoolbeleid, waarin vooraf concrete streefdoelen op het terrein van technisch lezen, op groeps- en schoolniveau, zijn vastgesteld.
- 2) Het in samenhang met 1 vergroten van het evaluerend vermogen van de school (c.q. het gebruik van een leerlingvolgsysteem en het maken van een sterkte/zwakte analyse van het onderwijsaanbod).

Naast deze doelen, die betrekking hebben op het functioneren van de school als geheel, werden op het leerkrachtniveau de volgende subdoelen geformuleerd:

- 1) Het invoeren van een systeem van planmatig werken bij het technisch leesonderwijs.
- 2) Deskundigheidsbevordering ten aanzien van vaardigheden ten behoeve van het optimaliseren van de leer- en instructietijd, en vaardigheden op het terrein van directe instructie.
- 3) Deskundigheidsbevordering ten aanzien van de kwaliteit van vakspecifieke instructie, namelijk het werken volgens de principes van de klanksynthese opbouwmethodiek bij technisch lezen.

(bron: Houtveen, Booy, De Jong & Van de Grift, 1996, p. 8)

Het project werd, over een looptijd van 3 schooljaren, geëvalueerd volgens een quasi-experimenteel onderzoeksontwerp. Uit het evaluatieonderzoek bleek dat de leerkrachten van de projectscholen duidelijk vooruit waren gegaan op de verschillende onderdelen van het wenselijk geachte instructiegedrag. Op de gebruikte prestatietoets voor technisch

lezen scoorden de leerlingen van de projectscholen gemiddeld significant beter dan de leerlingen uit de controlegroep. Deze leerwinst, geconstateerd op het niveau van groep 3, bleek overigens in groep 4 te zijn verdwenen (Houtveen et al., 1996).

Inmiddels zijn vervolgpiloten gestart op het terrein van rekenen en wiskunde en begrijpend lezen (Booy, Houtveen en Overmars, 1996; Houtveen, Booy en Overmars, 1996). In deze projecten wordt dezelfde aanpak gekozen als in het Landelijk Project Schoolverbetering.

Een tweede recent schoolverbeteringsproject is het Rotterdamse KEA-project (Kleinschalig Experiment Achterstandsbestrijding). Dit project is erop gericht om de leerprestaties in basisscholen met een hoog percentage leerlingen uit kansarme milieus op het terrein van rekenen/wiskunde en Nederlandse taal te verhogen.

Als inhoudelijke rationale voor het project wordt het min of meer klassieke vijf-factoren model van schooleffectiviteit gebruikt:

- sterk onderwijskundig leiderschap;
- hoge verwachtingen bij leerkrachten m.b.t. leerresultaten;
- nadruk op basisvaardigheden;
- veilig en ordelijk schoolklimaat;
- frequente evaluatie van leerlingprestaties.

(bron: Hogendijk en Wolfgram, 1995)

Bij de nadere uitwerking blijkt toepassing van de principes van "directe instructie", gekoppeld aan doelmatig klasmanagement en vergroting van de effectieve leertijd, en frequente, ook diagnostisch georiënteerde, evaluatie de kern van het programma.

In de praktische aanpak wordt voortgebouwd op bestaande methodes die worden versterkt door het verder voorstructureren van de leerstof in zogenoemde leerlijnen en een directe vorm van leerkrachtbegeleiding (klasseconsultatie) door schoolbegeleiders.

De eerste resultaten van het programma zijn positief, zowel wat betreft de gewenste verbetering van het leerkrachtgedrag als de verhoging van de leerprestaties (Hogendijk en Wolfgram, 1995).

De uitkomsten van dit soort onderzoek zijn niet alleen relevant omdat er een toetsing plaatsvindt van belangrijke elementen van school- en instructie-effectiviteitsmodellen, maar vooral ook, omdat er een praktische toepasbare uitwerking is gegeven aan deze modellen. Het is daarom van belang de belangrijkste kenmerken van de zojuist besproken Nederlandse schoolverbeteringsprojecten nog eens te resumeren.

In de eerste plaats is er sprake van een integratie van factoren die naar voren zijn gekomen uit enerzijds het schooleffectiviteitsonderzoek en anderzijds het instructie-effectiviteitsonderzoek. Instructiekenmerken als een gestructureerde aanpak met veel nadruk op frequente toetsing met het oog op diagnosticering en remediëring vormen de kern, terwijl "opbrengstgerichtheid" de voornaamste ondersteunende conditie op schoolniveau is.

In de tweede plaats zijn de projecten, qua innovatiebenadering, duidelijk inhoudelijk voorgestructureerd en dus niet direct voorbeeld van de in innovatiekringen populaire "bottom up" benadering. De inhoudelijke voorstructurering heeft zowel betrekking op de articulatie van leerinhouden, via leerlijnen en toetsen, als op de aanbiedingswijze (directe instructie).

In de derde plaats wordt in de desbetreffende projecten gebruik gemaakt van de mogelijkheden van schoolbegeleidingsdiensten, enerzijds door middel van cursussen aan leerkrachten en directe begeleiding in de classesituatie, anderzijds door middel van vakinhoudelijke aanvullingen op lesmethoden.

Tot slot moet worden opgemerkt dat bij de interpretatie van de resultaten van dit soort evaluaties van schoolverbeteringsprogramma's meestal geen rekenschap wordt gegeven van de invloed van het zogenoemde Hawthorne-effect (de prestaties verbeteren alleen al door de extra aandacht die projectscholen krijgen). Tevens wordt het kostenaspect van de schoolverbetering meestal over het hoofd gezien. Het KEA-project lijkt hierop een gunstige uitzondering, omdat in een persbericht wordt aangegeven dat de succesvolle benadering van het project "zeventig gulden extra per leerling per jaar" kost (Didaktief, februari 1996).

4. Interpretatie van de resultaten

4.1 Algemene conclusies

Zoals de vorige paragraaf laat zien, is het niet eenvoudig om een algemene en eenduidige conclusie te trekken over de factoren die "werken" in het (basis)onderwijs. Als de beste poging hiertoe moeten de gegevens, gebaseerd op review-artikelen, internationaal vergelijkende onderzoeken en meta-analyses, zoals samengevat in tabel 6 worden beschouwd. Bij de interpretatie van de gegevens uit tabel 6 moet worden aangetekend dat het merendeel van de afzonderlijke onderzoeken in kwestie uit survey-onderzoek bestond, waarbij correlatieve verbanden zijn onderzocht die gebaseerd zijn op "spontaan" in de onderwijspraktijk voorkomende variantie. Dit in tegenstelling tot (quasi-)experimenteel onderzoek, waarbij bepaalde school- en instructiekenmerken doelgericht gestimuleerd worden. De evaluatie van schoolverbeteringsprogramma's is een voorbeeld van dit laatste type onderzoek. De effecten die in dit soort onderzoek worden gemeten lijken, op basis van een casuïstische impressie (meta-analyses zijn momenteel niet voorhanden), groter dan bij het survey-onderzoek (zie bv. Slavin, 1996). Daar moet echter weer bij worden aangetekend dat deze effecten voor een deel aan het zogenoemde Hawthorne-effect moeten worden toegeschreven. Bovendien levert de evaluatie van een schoolverbeteringsprogramma alleen informatie over de werking van het programma als geheel, en niet over de invloed van afzonderlijke effectiviteitsbevorderende school- en klaskenmerken. Al dit soort nuances in acht nemend, kunnen er desalniettemin algemene conclusies worden getrokken:

- 1) Er blijkt een brede consensus te zijn gegroeid over zo'n 10 à 15 school- en instructiekenmerken, waarvan wordt aangenomen dat ze effectiviteitsbevorderend zijn op het terrein van de basisvakken (rekenen en taal). Bij die consensus gaat het om een algemene typering en kan er bij de operationalisering van de desbetreffende kenmerken nog grote variëteit zijn. De factoren in kwestie zijn in het voorafgaande besproken en bijvoorbeeld weergegeven in de tabellen 4 en 6. Tezamen verklaren deze factoren een behoorlijk deel van de "netto" tussenschoolse variantie (50 à 60%).

- 2) Gezien de betrekkelijk geringe variantie tussen scholen in "netto"-effecten (zo'n 10% van de variantie in leerlingprestaties), lijkt, in Westerse maatschappijen, de marge om met gerichte manipulatie van de onder 1 bedoelde factoren de effectiviteit te verhogen betrekkelijk gering. Eerder is in dit verband opgemerkt dat inspanningen vooral gericht zouden kunnen worden op het "bijtrekken" van de 10% laagst scorende scholen; verder is opgemerkt dat er enige aanwijzing is voor de veronderstelling dat de desbetreffende factoren vooral werken bij de minder kansrijke groepen leerlingen. Hoewel dit soort beschouwingen, gebaseerd op het schooleffectiviteitsonderzoek, nuttig zijn om bescheidenheid ten aanzien van de maakbaarheid van het onderwijs in acht te nemen, moet wel gewaarschuwd worden voor een interpretatiefout. In het effectiviteitsonderzoek wordt, zoals eerder gesteld, steeds met onderling vergelijkende maatstaven, in tegenstelling tot absolute maatstaven, gewerkt. Kortom, wat men ook doet, er zullen altijd 10% laagst scorende scholen zijn.

Het verdient daarom overweging om condities te creëren die een meer inhoudelijke vaststelling van het niveau van een school mogelijk maken. Dit is mogelijk door middel van een stelsel van doelstellingsvalide, zogenoemde criteriumgeoriënteerde toetsen. (Zie in dit verband ook het advies van de Commissie Kohnstamm).

- 3) Wanneer de drie onderscheiden hoofdcategorieën van manipuleerbare factoren onderling worden vergeleken, dan blijkt in ieder geval dat de impact van verbetering van instructiecondities op klasniveau groter is dan die van verbetering van de condities van management en organisatie op schoolniveau. Toch laten de resultaten van de meta-analyses zien dat de effecten van schoolorganisatiecondities weliswaar klein, maar toch niet betekenisloos zijn. Voor de materiële input-condities geldt dat er in het algemeen behoorlijk forse verbeteringen (c.q. hogere investeringen) nodig zijn om effecten te sorteren. Dit laatste principe is uitvoerig gedocumenteerd in het rapport van de cie. Van Eindhoven over de klassengrootte.

In de volgende sub-paragraaf zal kort worden stilgestaan bij veronderstelde interrelaties in de werking van condities op klas- en schoolniveau.

- 4) De resultaten van evaluatie van schoolverbeteringsprogramma's die geïnspireerd zijn op de uitkomsten van school- en instructie-effectiviteitsonderzoek laten in ieder geval zien dat deze resultaten vertaalbaar zijn in concrete maatregelen die in praktijk kunnen worden gebracht. In zekere zin is dit, gezien het praktisch nabije karakter van het effectiviteitsonderzoek, ook niet verwonderlijk. Recente ervaringen (zie o.m. de Nederlandse projecten die in par. 3.3 besproken zijn) laten zien dat goede resultaten worden bereikt met programma's die een gestructureerde instructie-aanpak integreren met ondersteunende maatregelen op schoolniveau.

4.2 Modelmatige en theoriegebonden interpretatie

Tot dusver heeft in de samenvatting van de onderzoeksliteratuur nogal het accent gelegen op de opsomming van factoren die "werken", c.q. een positieve correlatie hebben met voor aanvangsniveau gecontroleerde leerlingprestaties. Daarbij zijn verschillende categorieën van factoren onderscheiden, resource-inputs, schoolorganisatiecondities en instructiecondities. In het meer recente schooleffectiviteitsonderzoek worden "integrale" conceptuele modellen als uitgangspunt genomen. Multi-niveau analysetechnieken maken het mogelijk om de invloed van de beïnvloedende factoren die op school- en klasniveau worden onderscheiden simultaan te schatten. Op deze wijze kan dus onder meer worden vastgesteld dat instructiecondities op klasniveau over het algemeen een grotere impact hebben dan condities op schoolniveau.

Bij een verdere uitwerking van dit soort integrale schooleffectiviteitsmodellen (een voorbeeld ervan is weergegeven in figuur 3 in par. 2.1.3) wordt een nadere invulling gegeven aan de aard van de beïnvloeding van leerprestaties door factoren die op verschillend aggregatieniveau zijn gedefinieerd. Met name is hierbij de vraag aan de orde hoe de causale relaties tussen de verschillende aggregatieniveaus eruit zien, d.w.z. inhoudelijk geïnterpreteerd moeten worden en in statistische modellen geformaliseerd kunnen worden. De algemene aanname hierbij is dat condities op lager niveau vergemakkelijkt worden door condities op hoger niveau. Dus bijvoorbeeld condities voor effectieve instructie die gestimu-

leerd worden door onderwijskundig leiderschap en een opbrengstgerichte onderwijsvisie, die gedeeld wordt door de staf. Bosker en Scheerens (1994) laten zien dat er meerdere varianten van deze algemene conceptie bestaan, zowel conceptueel-inhoudelijk als in formeel-statistische zin.

Alternatieve invullingen van facilitering vanuit het organisatieniveau zijn afscherming ("buffering"), verschaffen van materiële randvoorwaarden, directe en meer indirecte stimulering. In de statistische modellering komen deze varianten onder meer terug in het onderscheid tussen modellen met alleen hoofdeffecten en modellen met interactietermen.

Er is nog onvoldoende empirisch onderzoek gedaan om een goed beeld te krijgen van de houdbaarheid van de algemene faciliteringshypothese, en de verbijzonderingen ervan. De handvol onderzoeken die beschikbaar zijn (o.m. Bosker & Scheerens, 1994; Hofman, 1993; Hill, Rowe & Holmes-Smith, 1995) bieden betrekkelijk weinig steun voor de hypothese. Een mogelijke verklaring hiervoor is dat schooleffectiviteitsonderzoeken zelden kunnen worden opgezet vanuit conceptuele modellen, maar meestal als bijproduct van evaluatie-onderzoek moeten worden uitgevoerd. (Zie ook het overzichtsartikel van Creemers en Reezigt, 1996).

Overigens bieden positieve uitkomsten van evaluatie-onderzoek naar "geïntegreerde" (d.w.z. zowel op school- als instructieniveau gerichte) schoolverbeteringsprogramma's enige steun voor de faciliteringshypothese, ook al is het bij dit soort onderzoek niet goed mogelijk om effecten precies te lokaliseren.

Op meer theoretische gronden wordt overigens een meer *indirecte* werking van hoger-niveau facilitering het meest effectief geacht. De basisgedachte hierbij is dat leer- en instructieprocessen primair als constructies van de directe actoren moeten worden gezien en dat inmenging vanuit hogere bestuurlijke of organisatieniveaus in ieder geval voldoende ruimte moet bieden om deze processen tot hun recht te laten komen. Scheerens en Bosker (1997) brengen dit inzicht in verband met het subsidiariteitsprincipe uit de politicologie. In de discussie over de mate van structurering dan wel autonomie in leer- en instructieprocessen, die in de volgende sub-paragraaf aan de orde komt, komt dit vraagstuk van structurering en ruimte bieden voor eigen construerende processen opnieuw aan de orde

(vgl. Scheerens, 1994; Scheerens & Bosker, 1997, hfst. 9).

Een vraag die verder voert dan die naar de specifieke interpretatie van tussen-niveau relaties in conceptuele modellen van schooleffectiviteit is die naar de achterliggende verklarende mechanismen. Waarom "werken" de factoren die uit het effectiviteitsonderzoek naar voren zijn gekomen? Is deze werking terug te voeren op algemene principes die in de theorievorming over sociale verschijnselen in ruimere zin beschikbaar zijn?

Scheerens (1995) behandelt in dit verband vijf mechanismen die voorkomen in de economische, organisatiekundige en bestuurskundige theorie:

- proactieve, rationele planning (planningstheorie, theorie van de bureaucratie);
- het op één lijn brengen van individuele preferenties en organisatiedoelen door middel van het marktmechanisme (public choice theorie);
- retro-actieve sturing op basis van het cybernetisch principe (planningstheorie, systeemleer, theorie van de lerende organisatie);
- "passing" en onderlinge consistentie van organisatie- en situationele kenmerken (contingentietheorie);
- zelforganisatie (systeemdynamica, chaos-theorie).

Het voert in dit verband te ver om een samenvatting van deze analyse te geven (zie ook Scheerens & Bosker, 1997). Een belangrijke conclusie is dat het evaluatiemechanisme een centrale plaats heeft in vier van deze vijf algemene theoretische principes. Versterking van het evaluatiemechanisme op schoolniveau, waarin de vaststelling van de progressie in leerresultaten centraal staat, is te rijmen met het eerder genoemde subsidiariteitsprincipe, waarin betrekkelijke vrijheid wordt gegeven voor de invulling van processen, maar op uitkomsten wordt gecontroleerd.

4.3 Directe instructie en constructivisme

In de traditie van het schooleffectiviteitsonderzoek en het onderzoek van "educational productivity" is er sprake van een duidelijke accentuering van een gestructureerde aanpak van het onderwijsproces. Zie de kenmerken van "direct instruction" die geciteerd zijn in par. 2.1.2 en de kenmerken die worden onderscheiden door Creemers en zijn geciteerd in tabel 5.

Binnen de onderwijspsychologie en het onderzoek naar onderwijzen (research on teaching) is het laatste decennium een nieuw paradigma voor leren en instructie ontwikkeld, waarin andere accenten worden gelegd dan in het "direct instruction" model.

Globaal gezien komt deze accentverlegging neer op de volgende hoofdpunten:

- meer nadruk op "leren" en leeractiviteiten dan op instructie en aanbiedingsvormen;
- de stelling dat leren eerder een proces van actieve "constructie" is, dan een meer passieve absorptie van kennis;
- in verband met het voorafgaande punt, aandacht voor de kennis die de leerling al heeft aan het begin van een onderwijsleersituatie, aansluiting bij de ervaringskennis van de leerlingen, inbedding van leertaken in praktijksituaties en, in het bijzonder, veel aandacht voor het leerproces zelf en denkstrategieën (meta-cognities) die de leerlingen hierbij hanteren;
- accentuering van samenwerken en sociale aspecten van leren.

In de volgende tabel, ontleend aan Scheerens (1994), worden "klassieke" en op het "constructivisme" gebaseerde instructiemodellen met elkaar vergeleken.

Tabel 8: *Vergelijking van traditionele en constructivistische instructie-modellen (naar Scheerens, 1994, p. 30)*

<i>Traditionele instructie</i>	<i>Instructie geïnspireerd door het constructivisme</i>
nadruk op basisvaardigheden (kennis als doel)	neiging om hogere cognitieve vaardigheden centraal te stellen (kennis als middel)
accent op leerstof en kennisoverdracht	nadruk op het leerproces
een gestructureerde aanpak - vooraf gespecificeerde doelstellingen - kleine stapjes - frequent stellen van vragen en geven van feedback - beloning door een hoog percentage leerstofbeheersing	ontdekkend leren
abstracte, generaliseerbare kennis	leren van gevallen
gestandaardiseerde studietoetsen	alternatieve evaluatiemethoden, zoals overzichten van uitgevoerde werkstukken

In verschillende overzichtsartikelen (Haller, Child & Walberg, 1988; Wang, Haertel & Walberg, 1993; Veenman, 1992 en De Jong & Van Joling, 1997) worden positieve resultaten van op constructivisme geïnspireerde instructiepraktijken gemeld. De vraag doet zich voor wat er nu beter werkt: een gestructureerde "direct instruction" benadering, of een benadering waarin juist het zelf ontdekken van de leerling centraal staat? Het antwoord op dit vraag is niet een simpel ja of nee, en dat heeft verschillende achtergronden. Om te beginnen is het onderscheid "direct instruction" versus "op constructivisme geïnspireerde instructie" wel erg grof. De Jong (1992) laat zien dat het hierbij in feite gaat om twee polen van een continuüm met verschillende tussenposities. De Jong onderscheidt

de volgende varianten, die liggen tussen "leerlinggestuurd onderwijs" en "leraargestuurd onderwijs":

- 1) zelfontdekkend leren (discovery learning)
- 2) geleid ontdekken leren (guided discovery)
- 3) observerend leren (observational learning)
- 4) geleid participerend leren (guided participation)
- 5) context-onafhankelijk leren (boeken en cursussen)
- 6) directe instructie (direct explanation, teacher directed)
- 7) dialogisch onderwijzen (dyadic instruction)

(De Jong, 1992, zoals geciteerd door Veenman, 1992, p. 242)

In de tweede plaats is er niet alleen sprake van verschillende instructievormen in de zin van verschillende *middelen*, maar ook, tot op zekere hoogte, van onderwijs met verschillende *doelen*. Praktisch "ingebede" kennis (situated cognition), leer- en probleemoplossingsstrategieën (metacognition) en "transfer" krijgen als *onderwijsdoelstellingen* een zwaarder accent bij een op het constructivisme geïnspireerde onderwijspraktijk dan bij een meer traditionele onderwijspraktijk (basisvaardigheden en basis-kennis). Uiteraard leidt dit ook tot verschillende eisen ten aanzien van de toetsing en evaluatie van de leerprestaties.

In de derde plaats speelt een rol dat in het onderzoek naar effectieve instructie en "educational productivity" een sterk accent op klasmanagement heeft gelegen. De kern hiervan is dat de leerkracht het onderwijs zodanig organiseert dat de effectieve leertijd voor alle leerlingen zo groot mogelijk is. Brophy (1996) beargumenteert dat ook voor meer ontdekkend en zelfstandig leren "classroom management", in de zin van het optimaliseren van leertijd, van groot belang is. Overigens stellen de principes van zelfstandig en ontdekken leren aanvullende eisen aan het klasmanagement, zowel wat betreft mate van individualisering, samenwerken in groepjes en de creatie van "rijke leeromgevingen", waarbij de computer een belangrijke rol speelt. "Structurering", tenslotte, heeft ook in een op het constructivisme georiënteerde instructiepraktijk een plaats, zij het dan niet alleen gericht op de leerstof, maar ook op de structurering van leerprocessen.

In de Nederlandse onderwijspraktijk zijn aan het constructivisme gerelateerde benaderingen toegepast in de methode van het realistisch rekenen en de ideeën betreffende het "studiehuis" in de tweede fase van het voortgezet onderwijs. In de diverse adviezen van de voormalige Advies Raad voor het Onderwijs (ARO) komt dit gedachtengoed ook duidelijk naar voren.

De gedachte dat het misschien minder zinvol is om de meer recente onderwijspsychologische ideeën als diametraal tegenover "directe instructie" te plaatsen, komt naar voren uit bijdragen als die van De Jong (1992), Veenman (1992), De Jong en Van Jolingen (1997) en Boekaerts en Simons (1993). Ook in op de praktijk gerichte benaderingen wordt in feite gedemonstreerd dat integratie, dan wel wederzijdse aanpassing van de meer extreme posities mogelijk is. Dit komt bijvoorbeeld naar voren uit het project "kwaliteitsbewaking rekenen en wiskunde" (Houtveen et al., 1996) en uit de leergang "effectieve instructie" welke ontwikkeld is voor studenten van PABO's door Veenman, Leenders, Meyers en Sanders (1993). De items van de Directe Instructie Schaal, ontwikkeld in het kader van het laatstgenoemde project, illustreren een instructiepraktijk waarin elementen van "klassieke" directe instructie en aspecten van zelfstandig leren gezamenlijk voor komen.

Subschaal Presentatie

Dagelijkse terugblik

Activeren voorkennis

Controleren voorkennis

Aangeven lesdoelstelling

Helder taalgebruik

Het geven van handelingswijzers

Onderwijzen in kleine stapjes

Vaardigheid voordoen

Gebruik concrete voorbeelden

Geven van samenvattingen

Gebruik van de DI-structuur

Subschaal Verwerking

(geleid en zelfstandig)

Controle op begrip/inzicht

Circuleren door de klas

Duidelijke opdrachten

Stellen van veel vragen

Wederzijds onderwijzen

Leerlingen beginnen meteen aan hun opdrachten

Leerlingen helpen elkaar in kleine groepen

Verwijzing naar nieuwe toepassingssituaties

Leerlingen verantwoordelijk stellen voor hun werk

In het handboek van de instructiepraktijk van Boekaerts en Simons (1993) vindt eveneens een afwisseling van meer structurerende en "autonome" momenten plaats.

Desalniettemin is het continuüm van direct instruction naar discovery learning een belangrijke basis voor verder onderzoek. Hierbij is vooral ook van belang voor welke leeftijds categorieën en subcategorieën van leerlingen een bepaalde positie op het continuüm optimaal is. Hierbij wordt vooralsnog aangenomen dat structurerende des te belangrijker is, naarmate leerlingen jonger en minder bevoorrecht qua aanleg en milieu zijn.

Voor de onderwijspraktijk en (eventueel) het onderwijsbeleid zijn verder vooral de volgende punten van belang:

- 1) De mate van accentuering van situationeel gebonden "praktische" kennis en algemene cognitieve vaardigheden is een kwestie van onderwijsdoelstellingen en dus van onderwijspolitiek en pas in tweede instantie een kwestie van instrumentatie (het vinden van effectieve middelen en methoden).
- 2) De ecologische validiteit van veel bevindingen betreffende op het constructivisme geïnspireerde instructiepraktijk is nog "fragwürdig" (Veenman, 1992).
Evenals dat geldt voor adaptieve instructie is er een kloof tussen

wenselijkheid en haalbaarheid.

- 3) Op een wat algemener niveau zijn de gemeenschappelijke kenmerken van *directe instructie* en *discovery learning* wellicht van meer primair belang voor beleidsvoerders. Die gemeenschappelijke kenmerken kunnen worden samengevat onder de noemer "expliciete instructie" (Veenman, 1992). Dat wil zeggen dat de leerkracht helder maakt wat de doelen van de instructie zijn en hoe wordt vastgesteld of die doelen ook door alle leerlingen bereikt zijn. Op instructieniveau gaat het hierbij om zaken die men op schoolniveau aanduidt als *opbrengstgerichtheid* en *evaluerend vermogen*.
- 4) De belangrijkste media om goede posities te kiezen op het continuüm van directe instructie naar discovery learning zijn de leergangen, oftewel onderwijsmethoden, de repertoires van leerkrachten, zoals zij die verwerven in initiële en nascholing, en de toetssystemen. Informatietechnologie heeft betekenis op elk punt van het continuüm, van drill and practice bij gestructureerde instructie tot simulaties in het kader van discovery learning (vgl. De Jong & Van Jolingen, 1997).

5. Beleidsimplicaties

Dit overzicht van onderzoeksuitkomsten op het terrein van school- en instructie-effectiviteit is uitgevoerd om keuzes ten aanzien van flankerend beleid bij een eventuele beleidsmaatregel tot klassenverkleining in het basisonderwijs te vergemakkelijken. In de inleiding is opgemerkt dat de desbetreffende onderzoeksliteratuur relevant is voor twee vragen:

- a) welke factoren mediëren het effect van klassenverkleining, dat wil zeggen, welke effectiviteitsbevorderende factoren krijgen meer kans bij kleinere klassen?
- b) welke aanvullende factoren kunnen tezamen met klassenverkleining en de onder a bedoelde factoren een verdere effectiviteitsbevorderende werking hebben?

ad a) Specifiek onderzoek naar "mediërende" factoren die betere prestaties bij kleinere klassen kunnen verklaren is schaars (vgl. Bosker, 1997). Uit de meta-analyses van Glass et al. (1982) kwam naar

voren dat de *tevredenheid over het werk* bij leerkrachten in kleinere klassen evenals het *instructiegedrag* beter waren dan in grotere klassen. Bosker (1997) stelt in dit verband verder dat veel auteurs erop wijzen dat "leerkrachten in pedagogisch-didactisch opzicht waarschijnlijk onvoldoende gebruik maken van een klassenverkleining".

Wanneer we de factoren bekijken die uit het school- en met name het instructie-effectiviteitsonderzoek naar voren zijn gekomen, dan is een redenering te geven over de manier waarop klassenverkleining indirect, dus via een aantal meer rechtstreeks op de leerprestaties inwerkende factoren, effectiviteitsbevorderend kan werken. De meest plausibele redenering is dat klassenverkleining in algemene zin de opgave van het klasmanagement verkleint, waardoor de effectieve leertijd voor individuele leerlingen kan worden vergroot. Verder moet onder ogen gezien worden dat overige instructie-effectiviteitsbevorderende factoren, zoals directe instructie en meer leerlinggerichte instructie, het volgen van leervorderingen, feedback, diagnose en remediëring in de context van adaptief onderwijs de opgave voor het klasmanagement juist groter maken. Verlichting door klassenverkleining maakt dus de kansen op realisering van de desbetreffende instructiepraktijken groter.

- ad b) Zoals gesteld, blijkt dat de onderwijspraktijk de mogelijkheden die klassenverkleining in potentie biedt niet "vanzelf" verzilvert. Op verschillende manieren zou geprobeerd kunnen worden om deze mogelijkheden wèl te benutten. Het effectiviteitsonderzoek biedt enig houvast wat betreft de inhoudelijke accenten die bij maatregelen om de school- en instructiepraktijk ten aanzien van het presteren op de basisvakken te vergroten. Na alles wat in het voorafgaande hierover gezegd is, resten wat dit betreft twee vragen: 1) welke effectiviteitsbevorderende factoren vormen het beste aangrijpingspunt? en 2) welke maatregelen als "algemeen beleid gericht op kwaliteitsverbetering" dan wel als "flankerend beleid" bij klassenverkleining komen in aanmerking?

5.1 Evaluatie en feedback als kern van effectiviteitsbevordering

Beschouwingen over hoe het onderwijs verbeterd kan worden leiden al gauw tot een waslijst van wenselijkheden: beter opgeleide leerkrachten, professionelere schoolleiders, aangepaste gebouwen, een betere didactische praktijk en een ruimere en betere toepassing van communicatie- en informatietechnologie. Gezien de complexiteit van de materie is dit ook begrijpelijk. Het effectiviteitsonderzoek heeft deze complexiteit weten terug te brengen tot een beperkte set van factoren, die in verschillende modaliteiten van het onderwijs een betekenis hebben: in het gedrag van schoolleiders, de opdeling en presentatie van leerstof, in ideeën over effectief instructiegedrag en klasmanagement, groeperingsvormen, het schoolklimaat en kwaliteitsbewaking. Schoolverbeteringsprojecten, ook Nederlandse, laten zien dat deze inzichten ook in praktijkprogramma's zijn uit te werken. Integrale schoolverbeteringsprojecten, zoals het Landelijk Project Schoolverbetering en het KEA-project, zijn daarom zeker een antwoord op de vraag welk flankerend beleid het effect van klassenverkleining kan vergroten.

Zowel vanuit een meer wetenschappelijke belangstelling als met het oog op een zo spaarzaam mogelijke toepassing van beleid is het echter interessant om de vraag te stellen wat nu de kern van de zaak is bij kwaliteitsverbetering in de zin van een grotere effectiviteit in de basisvakken van het primaire onderwijs. De presentatie van onderzoeksresultaten en met name de meer theoriegebonden interpretatie hiervan in het voorafgaande wijzen op de evaluatie- en feedback-functie als het kernmechanisme voor effectiviteitsverbetering. Een nauwgezet volgen van de vorderingen van leerlingen door de hele schoolloopbaan vormt hierbij de basis. De resultaten vormen het uitgangspunt voor diagnose en remediëring. "Opbrengstgericht schoolbeleid" en "onderwijskundig leiderschap" krijgen een concrete invulling wanneer er gewerkt kan worden op basis van concrete gegevens. Hetzelfde geldt voor de zo wenselijk geachte communicatie en samenwerking tussen leerkrachten, om de nodige continuïteit en consistentie te creëren.

Uiteraard stelt dit hoge eisen aan het evaluatie-instrumentarium. Wanneer

hogere cognitieve vaardigheden en meta-cognitie meer accent krijgen als onderwijsdoelen moet de evaluatiepraktijk hierbij worden aangepast. Het beleidsvoerende potentieel van het kerninstrumentarium, bijvoorbeeld een leerlingvolgsysteem, wordt verhoogt wanneer er een aanvullend instrumentarium wordt aangereikt dat ook het onderwijsaanbod en andere procesindicatoren omvat.

Voorwaarde voor het tot z'n recht laten komen van een evaluatiesysteem is dat er sprake is van een duidelijk schoolconcept, in de zin van concrete eindtermen en tussendoelen. Op de vraag of dit moet worden gezien als een opgave voor iedere afzonderlijke school, of als een overheidstaak wordt in de volgende sub-paragraaf ingegaan. Een tweede voorwaarde is ondersteuning en professionalisering van leerkrachten en schoolleiding bij het gebruik van een integraal evaluatiesysteem.

5.2 Voorschrijvend of voorwaardescheppend beleid?

Verbetering van de evaluatiefunctie van scholen kan zowel worden gezien vanuit het perspectief van de zelfverbetering van autonome individuele scholen als vanuit het perspectief van een overheidsbeleid, waarin output-controle wordt gezien in het verlengde van de toetsing van het onderwijs aan de eisen van deugdelijkheid. In deze is een continuüm te schetsen dat loopt van maximale decentralisatie tot centralisatie:

- a) scholen zijn geheel vrij bij de interpretatie van eindtermen en de mate waarin en de wijze waarop zij vorm geven aan (zelf)evaluatie;
- b) scholen kiezen eigen evaluatieprocedures bij een van overheidswege voorgeschreven normatief kader van eindtermen en tussendoelen (vgl. het advies van de cie. Kohnstamm om voor het eind van de onderbouw van de basisschool concrete tussendoelen te ontwikkelen en het advies van de cie. Van Eindhoven, waarin gepleit wordt voor "structuurbiedende programmering voor de onderbouw");
- c) gegeven het kader van eindtermen en concrete tussendoelen wordt de ontwikkeling gestimuleerd van een hierop afgestemd zelfevaluatie-instrumentarium, dat naast een leerlingvolgsysteem aanbodsindicatoren en procesindicatoren omvat; scholen zijn geheel vrij om dit instrumentarium al dan niet aan te schaffen en te gebruiken;

- d) gegeven eindtermen en tussendoelen, wordt een hierop afgestemd integraal instrumentarium ontwikkeld dat scholen verplicht zijn te hanteren en waarvan zij rekenschap afleggen in rapportages aan de overheid en de lokale gemeenschap.

Bij de hierboven opgesomde varianten geldt overigens dat zij de inspectiefunctie onvermeld laten. Gezien het sterke accent op procesevaluatie in de werkwijze van de Inspectie van het Onderwijs, vormt een output-georiënteerd (zelf)evaluatie-instrumentarium, geënt op eindtermen en tussendoelen een essentiële aanvulling. Bij een centralistische procesevaluatie is het gevaar van een bureaucratisering en "goal displacement" levensgroot aanwezig.

Naast politieke en beleidsmatige overwegingen die pleiten voor het nemen van overheidsverantwoordelijkheid voor het vaststellen van de eindtermen en tussendoelen voor het basisonderwijs zijn er nog twee aanvullende argumenten voor een dergelijke beleidsmaatregel te noemen.

Operationalisatie van eindtermen maakt de specificatie van te bereiken beheersingsniveaus voor de basisvakken mogelijk. Een hierop toegesneden toetsinstrumentarium maakt een betere kwaliteitsbeheersing mogelijk dan het gebruik van allerlei relatieve maatstaven, zoals die bijvoorbeeld ook in het schooleffectiviteitsonderzoek gebruikelijk zijn.

Een tweede argument is de waarschijnlijkheid dat, zo de overheid geen tussendoelen en concretisering van eindtermen vaststelt, dit wel op een andere manier zal gebeuren, namelijk via leerlingvolgsystemen. Een verkorte route naar de door de commissies Van Eindhoven en Kohnstamm bepleite specificatie van eindtermen en tussendoelen is dan ook een check op de doelstellingsvaliditeit van het CITO-leerlingvolgsysteem.

Beleid dat bijdraagt tot een verhoging van de kwaliteit van het basisonderwijs, in de zin van een vergrote effectiviteit op het vlak van de basisvakken (rekenen/wiskunde, lezen en taal) is gebaat bij een concretisering van de eindtermen en tussendoelen op deze gebieden. Daarnaast is voor een breedte-strategie voor effectiviteitsbevordering - steeds in de hierboven bedoelde specifieke betekenis van doelbereiking op het terrein

van de basisvakken - vooral een versterking van de outputgeoriënteerde evaluatiefunctie van belang. Hiertoe is voorwaardescheppend beleid bij het tot stand brengen van een geïntegreerd instrumentarium voor schoolzelfevaluatie voldoende.

Voor overige aspecten van effectiviteitsbevordering wordt geen specifiek overheidsbeleid noodzakelijk geacht. Gegeven de bestaande onderwijsverzorgingsvoorzieningen en de mogelijkheden voor toegepast onderwijskundig onderzoek in Nederland, mag worden aangenomen dat in principe binnen deze verbanden voldoende mogelijkheden zijn om bij te dragen aan effectiviteitsverhoging. Wel zou de overheid een extra stimulans kunnen bieden aan een geïntegreerd onderzoeks- en ontwikkelingsprogramma op de desbetreffende gebieden.

Aandachtsgebieden hierbij - naast verbetering van de evaluatie en feedbackfunctie - zijn: instructiebenaderingen op een continuüm van directe instructie tot ontdekkend leren, haalbare vormen van differentiatie binnen klasseverband, al dan niet gepaard gaand met een gedeeltelijke doorbreking van het leerstof-jaarklassensysteem en resultaatgeoriënteerd schoolleiderschap. Deze aandachtsgebieden zouden langs diverse wegen kunnen worden aangepakt: de initiële en nascholing van docenten, leergangontwikkeling, integratie schoolverbeteringsprojecten en experimenten met varianten van klassenverkleining (zie de voorstellen van Bosker en Meijnen hiertoe) en vormen van op effectiviteitsbevorderend gericht flankerend beleid. Van belang bij het uitproberen van varianten is het mede betrekken van de kostenfactor, door bijvoorbeeld een kostenneutrale vergelijking te maken tussen een bepaalde mate van klassenverkleining en een schoolverbeteringsaanpak.

Referenties

Bloom, B.S. (1976). *Human Characteristics and School Learning*. New York: McGraw-Hill.

Boekaerts, M., & Simons, P.R.J. (1993). *Leren en Instructie. Psychologie van de Leerling en het Leerproces*. Assen: Dekker & Van de Vegt.

Booy, N., Houtveen, A.A.M., & Overmars, B.A.M. (1996). *Evaluatieproject kwaliteitsverbetering begrijpend lezen: het eerste projectjaar*.

Bosker, R.J. (1997). Het einde van de klassenstrijd? Achtergronden bij de discussie over klassengrootte in het basisonderwijs. *Pedagogische Studiën* (in druk).

Bosker, R.J., & Scheerens, J. (1994). Alternative models of school effectiveness put to the test. In R.J. Bosker, B.P.M. Creemers & J. Scheerens (eds.), *Conceptual and Methodological Advances in Educational Effectiveness Research*, pp. 159-180. Special issue of the *International Journal of Educational Research*, 21(2).

Brandsma, H.P. (1993). *Basisschoolkenmerken en de Kwaliteit van het Onderwijs*. Groningen: RION.

Brophy, J. (1996). *Classroom management as socializing students into clearly articulated roles*. Paper presented at AERA, New York.

Carroll, J.B. (1963). A model of school learning. *Teachers College Record*, 64, 722-733.

Carroll, J.B. (1989). The Carroll Model, a 25-year retrospective and prospective view. *Educational Researcher*, 18, 26-31.

Cohen, J. (1969). *Statistical Power Analysis for the Behavioral Sciences*. (2nd ed.) Hillsdale, NJ: Lawrence Erlbaum Ass.

Cohen, D.K. (1988). Teaching practice ... Plus ça change ... In Ph. Jackson (ed.), *Contributing to Educational Change: Perspectives on Research and Practice*. Berkeley, CA.: McCutchan.

Coleman, J.S., Campbell, E.Q., Hobson, C.F., McPartland, J., Mood, A.M., Weifeld, F.D., & York, R.L. (1966). *Equality of Educational Opportunity*. Washington: US Government Printing Office.

Commissie Evaluatie Basisonderwijs (1994). *Zicht op kwaliteit, evaluatie van het basisonderwijs. Eindrapport*. Den Haag: SDU.

Commissie Indicatiestelling Onderwijsachterstanden (1996). *Zo onvoorspelbaar als het leven zelf*. Zoetermeer: Ministerie van OC&W.

Commissie Kwalitatieve Aspecten van Groepsgrootte in het Basisonderwijs (1996). *Klassenverkleining*. Zoetermeer: Ministerie van OC&W.

Cotton, K. (1995). *Effective schooling practices: a research synthesis*. 1995 Update. School Improvement Research Series. Northwest Regional Educational Laboratory.

Creemers, B.P.M. (1994). *The Effective Classroom*. London: Cassell.

Creemers, B.P.M, & Reezigt, R.J. (1996). School level conditions affecting the effectiveness of instruction. *School Effectiveness and School Improvement*, 7(3), 197-228.

Doyle, W. (1985). Effective secondary classroom practices. In M.J. Kyle (ed.), *Reaching for Excellence: An Effective Schools Sourcebook*. Washington, DC: US Government Printing Office.

Fraser, B.J., Walberg, H.J., Welch, W.W., & Hattie, J.A. (1987). *Syntheses of Educational Productivity Research*. Special issue of the *International Journal of Educational Research*, 11(2).

Glass, G.V., Cahen, L.S., Smith, M.L., & Filby, N.N. (1982). *School Class Size*. London: Sage Publishers.

Goldstein, H. (1996). *Methodological aspects of school effectiveness research*. Unpublished manuscript. London: Institute of Education.

Grift, W. van de (1990). Educational leadership and academic achievement in elementary education. *School Effectiveness and School Improvement*, 1(1), 26-40.

Groot, A.D. de (1986). *Begrip van evalueren*. Den Haag: VUGA.

Haller, E.P., Child, D.A., & Walberg, H.J. (1988). Can Comprehension be Taught? A quantitative synthesis of 'metacognitive' studies. *Educational Researcher*, 17(9), 5-8.

Hanushek, E.A. (1986). The economics of schooling: production and efficiency in public schools. *Journal of Economic Literature*, 24, 1141-1177.

Hedges, L.V., Laine, R.D., & Greenwald, R. (1994). Does money matter? A meta-analysis of studies of the effects of differential school inputs on student outcomes. *Educational Researcher*, 23, 3, 5-14.

Hill, P.W., Rowe, K.J., & Holmes-Smith, P. (1995). *Factors Affecting Students' Educational Progress: Multilevel Modelling of Educational Effectiveness*. Paper presented at the 8th International Congress for School Effectiveness and Improvement, Leeuwarden, the Netherlands, January 3-6, 1995.

Hofman, R.H. (1993). *Effectief Schoolbestuur*. Groningen: RION, proefschrift.

Hofstee, W.K.B. (1989). Naar een zuiver wetenschappelijke opvatting van evaluatie-onderzoek. In J. Scheerens (ed.), *Evaluatie: om de kwaliteit van het onderwijs*. Lisse: Swets & Zeitlinger.

Hogendijk & Wolfgram, P. (1995). *KEA halverwege*. Projectverslag schooljaar 1994/1995. Rotterdam: CED.

Houtveen, A.A.M., Booy, N., Jong, R. de, & Grift, W. van de (1996). *Effecten van adaptief onderwijs*. Evaluatie van het Landelijk Project Schoolverbetering.

Houtveen, A.A.M., Booy, N., & Overmars, B.A.M. (1996). *Versterking van het reken- en wiskundeonderwijs*. Vakgroep Onderwijskunde, Universiteit Utrecht.

Jencks, C., Smith, M.S., Ackland, H., Bane, M.J., Cohen, D., Grintlis, H., Heynes, B., & Michelson, S. (1972). *Inequality*. New York: Basic Books.

Jong, F.P.C.M. de (1992). *Zelfstandig leren: Regulatie van het leerproces en leren reguleren: een procesbenadering*. Tilburg: Katholieke Universiteit Brabant, proefschrift.

Jong, T. de, & Joling, W.R. van (1996). *Discovery learning with computer simulations of conceptual domains*. Manuscript University of Twente.

Levine, D.U., & Lezotte, L.W. (1990). *Unusually Effective Schools: A Review and Analysis of Research and Practice*. Madison, Wis.: National Center for Effective Schools Research and Development.

Monk, D.H. (1989) The education production function: it's evolving role in policy analysis. *Educational Evaluation and Policy Analysis*, 11, 1, 31-45.

Monk, D.H. (1992). *Microeconomics of School Productions*. Paper for the Economics of Education Section of the International Encyclopedia of Education.

Morgan, G. (1986). *Images of Organization*. London: Sage.

Projectgroep Schaalvergroting Basisonderwijs (1990). *Schaal en kwaliteit in het basisonderwijs*. Zoetermeer: Ministerie van O&W, p. 34-36.

Quinn, R.E., & Rohrbaugh, J. (1983). Spatial model of effectiveness criteria towards a competing values approach to organizational analysis. *Management Science*, 29, 363-377.

Reynolds, A.J., & Walberg, H.J. (1990). *A Structural Model of Educational Productivity*. Northern Illinois University (unpublished manuscript).

Rosenshine, B.V. (1983). Teaching functions in instructional programs. *Elementary School Journal*, 3, 335-351.

Sammons, P., Hillman, J., & Mortimore, P. (1995). *Key Characteristics of Effective Schools: a Review of School Effectiveness Research*. London: OFSTED.

Scheerens, J. (1992). *Effective schooling. Research, Theory and Practice*. London: Cassell.

Scheerens, J. (1993). Basic school effectiveness research: items for a research agenda. *School Effectiveness and School Improvement*, 4(1), 17-36.

Scheerens, J. (1994). The school-level context of instructional effectiveness: a comparison between school effectiveness and restructuring models. *Tijdschrift voor Onderwijsresearch*, 19(1), 26-38.

Scheerens, J. (1995). *School Effectiveness as a Research Discipline*. Paper presented at the ICSEI Congress, Leeuwarden, the Netherlands, January 1995.

Scheerens, J., & Brummelhuis, A.C.A. ten (1996). *Process Indicators on the Functioning of Schools: Results from an International Survey*. New York: AERA paper.

Scheerens, J., & Bosker, R.J. (1997). *The Foundations of Educational Effectiveness*. Oxford: Elsevier Science Publ.

Scheerens, J., & Creemers, B.P.M. (1989). Conceptualizing school effectiveness. In *Developments in School Effectiveness Research*. Special issue of the *International Journal of Educational Research*, 13(7), 691-706.

Scheerens, J., & Creemers, B.P.M. (1996). School effectiveness in the Netherlands; the modest influence of a research programme. *School Effectiveness and School Improvement*, 7, 181-195.

Slavin, R.E. (1996). *Success for all*. Lisse: Swets & Zeitlinger.

Stringfield, S.C., Bedinger, S., & Herman, R. (1995). *Implementing a Private School Program in an Inner-City Public School: Processes, Effects, and Implications from a Four Year Evaluation*. Paper presented at the ICSEI Conference, Leeuwarden, the Netherlands.

Veenman, S. (1992). Effectieve instructie volgens het directe instructie-model. *Pedagogische Studiën*, 69, 242-269.

Veenman, S. (1996). Effects of Multigrade and Multi-Age Classes Reconsidered. *Review of Educational Research*, 3.

Veenman, S., Leenders, Y., & Sanders, M. (1993). Leren lesgeven met het directe instructiemodel. *Pedagogische Studiën*, 70, 2-16

Walberg, H.J. (1984). Improving the productivity of American schools. *Educational Leadership*, 41, 19-27.

Wang, M.C., Haertel, G.D., & Walberg, H.J. (1993). Toward a knowledge base for school learning. *Review of Educational Research*, 63(3), 249-294.

Witziers, B., & Bosker, R.J. (1997). *A meta-analysis on the effects of presumed school effectiveness enhancing factors*. Paper presented at the International Congress on School Effectiveness and Improvement, Memphis, USA, 1997.