

G. W. Meijnen en J. Kloprogge

## Samenvatting

In dit artikel wordt het Landelijk Evaluatie-Onderzoek van het Onderwijsvoorrangsbeleid (LEO-OVB) geanalyseerd (evaluatie van een evaluatie). Zowel de voorbereidingsfase als de uitvoering van het onderzoek worden beschreven. In het artikel wordt bovendien ingegaan op die onderwerpen die in de internationale literatuur als de meest relevante worden beschouwd in het kader van programma-evaluatie. Tot die onderwerpen behoren: designkeuze, de verhouding tussen interne en externe validiteit en de relaties met de omgeving van het evaluatieprogramma. De conclusie luidt dat zowel het ontwerp als de verbindingen met de context nog steeds beantwoorden aan recente inzichten omtrent opzet en inrichting van een extern evaluatie-onderzoek. Tevens wordt opgemerkt dat dit onderzoeksprogramma niet het totale Onderwijsvoorrangsbeleid evalueert.

## 1 Inleiding

In 1991 verscheen het Negende Jaarboek van de National Society for the Study of Education (McLaughlin & Phillips (Eds.), 1991). Een aantal gerenommeerde, deels bejaarde, evaluatie-deskundigen schetst daarin de ontwikkelingen in hun eigen denken over evaluatie-onderzoek gedurende de laatste vijftig jaar. De bijdragen concentreren zich in het bijzonder op de problematiek van programma-evaluatie. Opvallend is de convergentie in denken omtrent wat de belangrijkste onderwerpen zijn waar het programma-evaluatie betreft. Het betreft dan in het bijzonder (1) de notie dat programma-evaluatie wordt uitgevoerd binnen een maatschappelijk krachtenveld, (2) de spanning die bij de opzet van het programma optreedt in termen van interne en externe validiteit en, derhalve, (3) wat de meest geëigende designs zijn voor programma-evaluatie, (4) wat verwacht mag worden omtrent de impact van de verkre-

gen resultaten op de beleidsvoering en (5) de wetenschapsinterne ontwikkelingen in het evaluatie-onderzoek. In dit artikel zal het in uitvoering zijnde Landelijk Evaluatie-Onderzoek Onderwijsvoorrangsbeleid (LEO-OVB) mede in het perspectief van deze onderwerpen worden geanalyseerd.

Nadat zowel een beschrijving van de voorbereiding van het programma is gegeven als ook een beschrijving van het functioneren van 1986 tot 1992, worden deze periodes geëvalueerd waarbij deze onderwerpen onder meer aan de orde zullen komen. Ook deze activiteiten kunnen als een poging tot evaluatie aangemerkt worden (evaluatie van de evaluatie). Een van de punten van consensus onder de deskundigen is dat een externe instantie, dan wel een niet geëngageerde onderzoeker de eerst aangewezen is om deze taak op zich te nemen. Uit de rest van het betoog zal duidelijk worden dat de positie van de auteurs hier een hachelijke is, ze zijn beiden, hetzij als ontwerper, hetzij als programmacoördinator, betrokken bij het LEO-OVB.

## 2 De voorbereiding van het LEO-OVB

In 1983 verzoekt het Ministerie van Onderwijs en Wetenschappen een drietal onderzoekers van de Rijksuniversiteit Groningen (Hofstee, Koster, Meijnen) een plan te ontwerpen voor de evaluatie van het Onderwijsvoorrangsbeleid waarvan de contouren waren vastgelegd in een concept Onderwijsvoorrangsplan (definitieve versie: 1985) dat verscheen onder de verantwoordelijkheid van de toenmalige minister Van Kemenade en staatssecretaris Deetman. Een daarop gebaseerd wetsontwerp zou kort daarop bij de Tweede Kamer worden ingediend. In het toentertijd circulerende concept van de Onderwijsvoorrangswet (1984) wordt als centrale doelstelling van het beleid vermeld: "Het onderwijsvoorrangsbeleid is er op

gericht (...) *onderwijsachterstanden* van leerlingen ten gevolge van sociale, economische en culturele omstandigheden op te heffen of te verminderen". In de Memorie van Toelichting wordt onder 'onderwijsachterstand' verstaan; "(...) achterstand in ontwikkeling en in onderwijsleerprestaties in ruime zin".

De overheid tracht deze doelstelling te bereiken door (a) scholen met leerlingen uit achterstandssituaties extra te faciliteren (de zogenaamde gewichtenregeling) en door (b) faciliteiten ter beschikking te stellen aan *gebieden* met een concentratie van deze scholen.

Het verzoek aan de drie onderzoekers kan worden beschouwd als een hernieuwde poging van de overheid om een beleidsprogramma op dit terrein te evalueren. De eerste poging vond plaats in de periode 1977-1981. In dat laatste jaar verscheen namelijk het eindrapport van de evaluatie van het Stimuleringsbeleid, met min of meer soortgelijke doelstellingen, dat in 1974 was gestart (Mens & Van Calcar, 1981). Het overgrote deel van het verslag behelst beschrijvingen van activiteiten die in het veld plaatsvonden zowel op scholen als binnen betrokken welzijnsinstellingen, waarbij ook bestuurlijk-organisatorische aspecten niet uit het oog werden verloren. Van de oorspronkelijk geplande effectevaluatie werd zo goed als niets gerealiseerd. In evaluatiejargon: het onderzoeksverslag behelste bijna uitsluitend procesevaluatie. In het Ten Geleide bij het verslag somt de Begeleidingscommissie een aantal problemen op die de uitvoering van de evaluatie hebben bemoeilijkt en voorts kan men zinsneden van het volgende type lezen: "Een gedetailleerde bespreking, laat staan een beoordeling ligt niet op de weg van de begeleidingscommissie" en: "Wordt nu de inspanning en andere investeringen van alle betrokkenen in dit project (het evaluatie-onderzoek) gerechtvaardigd door het resultaat?"

Uit het Ten Geleide en de overige verslagen kan men opmaken dat de meeste problemen zijn voortgekomen uit de geringe consensus omtrent de te evalueren doelstellingen, door de discussie omtrent de vraag of sommige doelstellingen wel gemeten zouden kunnen worden en uit de variatie in visies op de vraag wiens belangen de evaluatie zou moeten dienen (scholen, wetenschap, overheid). De onderzoekers die hierover zo ook hun eigen soms

zeer geprononceerde, maar tevens uiteenlopende, meningen hadden, speelden zowel de rol van participant in de doelstellingendiscussie als uitvoerder van het evaluatie-onderzoek. Uit de verslagen wordt duidelijk dat de betrokken onderzoeksgroepen zowel van eigen programmadoelen als van eigen evaluatieopvattingen zijn uitgegaan. De Begeleidingscommissie is kennelijk niet in staat geweest er een geheel van te smeden. Dit verschijnsel had zich intussen ook bij de evaluatie van andere beleidsprogramma's voorgedaan (Scheerens, 1983). Het Ministerie van Onderwijs en Wetenschappen wilde dit probleem bij de evaluatie van het in gang te zetten Onderwijsvoor-rangsbeleid voorkomen door een evaluatieplan te laten ontwerpen door deskundigen die niet op voorhand ook de uitvoering ter hand zouden moeten nemen.

Gelet op de ruim geformuleerde doelstellingen door het beleid hield de keuze van de aan te zoeken deskundigen ook een sturing in van de specificering van de doelstellingen. De jaren zeventig werden immers gekenmerkt door het debat tussen voorstanders van de zogenaamde 'brede doelstellingen' voor het Stimulerings/Onderwijsvoor-rangsbeleid (cognitieve, sociaal-emotionele, expressieve) en de aanhangers van het beleid gericht op 'enge (cognitieve) doelstellingen' (schoolloopbanen, leerprestaties met betrekking tot de kernvakken). Dit debat was in 1983 nog niet geëindigd hoewel reeds een kentering merkbaar was ten faveure van de laatstgenoemde richting. De keuze van de deskundigen voor het ontwikkelen van het evaluatieplan was in dit kader niet oninteressant gegeven de beperkte middelen voor onderzoek. De te kiezen prioriteiten in het evaluatie-onderzoek zouden in de richting van de ene of de andere doelstellingsfilosofie moeten uitvallen. Het verzoek aan de drie Groningse wetenschappers was in dit bestek een duidelijke: alle drie hadden zich voorstander betoond van de accentuering van de cognitieve doelstellingen in het onderwijs.

De onderzoekers hebben het onderzoeksprogramma vervolgens geschreven na consultatie van een groot aantal deskundigen en na overleg met een interdepartementale werkgroep van ambtenaren, belast met de voorbereiding van de beleidsevaluatie van het onderwijsvoor-rangsplan.

Hofstee, Koster en Meijnen (1984) schrijven in een paragraaf voorafgaande aan de eigenlijke opzet van het onderzoek, dat bestudering van eerder verricht evaluatie-onderzoek naar landelijke innovatieprogramma's hen tot de volgende conclusies heeft gebracht (p. 9). "Ten eerste: het is noodzakelijk dat bij de start van een evaluatie-onderzoek de beleidsdoelstellingen van een vernieuwingsprogramma voldoende worden geëxpliciteerd. Ten tweede: na de vaststelling van een evaluatieplan in hoofdlijnen, dient de uitvoerende groep van onderzoekers daaraan relatief onafhankelijk uitvoering te kunnen geven en dient ook voor het overige voldaan te zijn aan minimale voorwaarden voor onafhankelijk onderzoek. Vierjaarlijkse bijstelling van het evaluatieplan lijkt een goede mogelijkheid om ook de voortgang van de evaluatie te beoordelen. Ten derde: de ervaringen met de evaluatie van grootscheepse onderwijsvernieuwingsprogramma's tonen aan dat variabelen op klasse- en leerlingniveau in het onderzoek worden verwaarloosd. Ten vierde: de uitkomsten van evaluatie-onderzoek dienen bij voorkeur een plaats te krijgen in de procedure voor de vaststelling van de onderwijsvoorarrangspunten".

Met betrekking tot de randvoorwaarden en uitvoering van het evaluatieonderzoek merken Hofstee e.a. ten slotte nog het volgende op. Zij wensen uitdrukkelijk onderscheid te maken tussen *beleidsevaluatie* en *extern evaluatie-onderzoek*. Beleidsevaluatie omvat meer dan effectonderzoek. Beleidsevaluatie dient bijvoorbeeld rekening te houden met gevestigde belangen, met maatschappelijke en politieke acceptatie van beleid, met werkgelegenheidsproblemen en dergelijke. Het voorgestelde onderzoek wordt door hen dan ook geen beleids-evaluatie genoemd maar een extern evaluatie-onderzoek. De beleidsevaluatie zou in hun ogen moeten worden overgelaten aan de volksvertegenwoordiging en het externe evaluatie-onderzoek aan een groep van onafhankelijke onderzoekers. Uiteraard zal tijdens de beleids-evaluatie gebruik gemaakt kunnen worden van de gegevens die het externe evaluatie-onderzoek heeft opgeleverd. Voorts kunnen de verkregen gegevens ook dienstig zijn voor andere functies (interne evaluatie van gebieden/scholen, ondersteuning). De roep om dienstbaarheid ten bate van deze functies zal echter

nooit mogen resulteren in een aantasting van het plan van onafhankelijke, externe, evaluatie. Wat de uitvoering betreft merken zij op dat de vaststelling van het evaluatieplan dient te geschieden door de Tweede Kamer en dat de uitvoering in handen dient te worden gegeven van een coördinator die belast wordt met de detaillering van de onderzoeksplannen, de organisatie van het onderzoek en de verslaggeving. Het Centraal Instituut voor Toetsontwikkeling (CITO) wordt als meest geschikte locatie gezien voor de gegevensverzameling.

Het voorstel van Hofstee e.a. (1984) benadrukt, naar verwachting, het belang van onderzoek naar de vraag of de door het beleid beoogde *effecten*, te weten het opheffen of verminderen van onderwijsachterstanden, optreden. Omdat elke school met een bepaald percentage achterstandsl leerlingen kan rekenen op extra faciliteiten, is een experimenteel design in strikte zin (experimentele groep vergelijken met een controlegroep) niet mogelijk. Wel mogelijk zou zijn een quasi-experimenteel design met een post hoc analyse van groepen scholen die vergelijkbaar zijn op sommige punten maar onvergelijkbaar op andere (die zeer relevant zijn voor de vraagstelling). De onderzoekers zijn echter van mening dat deze weg niet moet worden ingeslagen omdat zorgvuldige monitoring door middel van peiling van ontwikkeling van achterstanden meer tegemoet komt aan de informatiewensen van diverse actoren (parlement, departement, regering, publieke opinie) dan de informatie verkregen door een quasi-experimenteel design.

Gekozen wordt derhalve voor een zeer grootschalig cohortonderzoek waarbij leerlingen uit de groepen 2, 5 en 8 van de basisschool om de drie jaar worden getest op hun prestaties in rekenen en taal. Tevens wordt telkenmale een intelligentietest afgenomen. Door de testafname om de drie jaar te herhalen kunnen scholen ook met zichzelf in de tijd vergeleken worden. Het feit dat dezelfde leerlingen om de drie jaar worden hertest, heeft ook tot gevolg dat de leerwinst van leerlingen nauwkeurig in kaart kan worden gebracht. Een dergelijke opzet staat dus momentane vergelijkingen toe en vergelijkingen in de tijd (bijvoorbeeld: gebiedsscholen met niet-gebiedsscholen; scholen met veel versus scholen met weinig faciliteiten; leerjaar 4 op tijdstip Y vergeleken met

leerjaar 4 op tijdstip X; enz.)

Voorgesteld wordt om een steekproef te trekken van 1 op de 20 basisscholen. Aangezien op deze wijze het aantal scholen met veel achterstandsleerlingen in de steekproef beperkt zal zijn, wordt aanbevolen een substantieel extra aantal scholen van dit type aan dit cohort toe te voegen. De steekproef zou dan bestaan uit een representatieve selectie van scholen casu quo leerlingen (de referentiegroep), waaronder zich ook tal van gefaciliteerde scholen bevinden, aangevuld met scholen met veel leerlingen uit achterstandssituaties. Teneinde ook de effecten op langere termijn te kunnen traceren dienen ook scholen voor voortgezet onderwijs waar de leerlingen naar doorstromen, in de steekproef te worden opgenomen.

Effecten zouden moeten worden waargenomen in termen van (minder) verwijzing naar het speciaal onderwijs, (minder) doubleren, (hogere) toetsprestaties in de vakken taal en rekenen in de basisschool, (hoger) advies voortgezet onderwijs, (vaker) keuze voor hogere vormen van voortgezet onderwijs, (minder) doubleren in het voortgezet onderwijs, (geringere) afstroom, (hogere) opstroom en (hogere) toetsprestaties in de vakken Nederlands, wiskunde en Engels in het voortgezet onderwijs voor de leerlingen uit achterstandssituaties. Tevens wordt in de afname van intelligentietests voorzien. Deze kunnen als pure effectmaat dienst doen (heeft het beleid invloed op de intelligentie-ontwikkeling?) maar ook als referentiepunt: in hoeverre presteren leerlingen van een school beneden of boven hun intelligentieniveau ten gevolge van de door de school gekozen aanpak?

Teneinde te kunnen nagaan hoe beleidsmaatregelen door de overheid (extra faciliteiten) worden vertaald in maatregelen op school- en gebiedsniveau en hoe die zich verhouden tot de effecten, wordt implementatie-onderzoek voorgesteld. Variaties in implementatie op schoolniveau worden gerelateerd aan door de school behaalde resultaten. Nagegaan kan dan worden of sommige activiteiten op school- of gebiedsniveau effectiever zijn dan andere. Gegevens zullen derhalve moeten worden verzameld op landelijk-, gebieds- en schoolniveau. Ten slotte wordt in aanvulling op het groot-schalige cohortonderzoek een aantal deelstudies voorgesteld onder meer betrekking heb-

bend op de zogenaamde 'brede' cognitieve- en sociale vaardigheden.

Door het Ministerie van Onderwijs en Wetenschappen wordt aan een drietal andere deskundigen een reactie gevraagd op het door Hofstee e.a. geleverde voorstel. Beem, Pijl en Scheerens (1986) maken enige kritische kanttekeningen bij het plan voor het volgen van alle cohortleerlingen, inclusief verwezenen en doubleurs (kosten!), en bij het gekozen interval van drie jaar. Fundamenteel is hun kritiek echter waar zij het ontbreken van adequate controlegroepen aan de orde stellen. Zij pleiten toch voor een (quasi-)experimenteel design. Hofstee (1986) stelt in een reactie dat het omvormen van de referentiegroep tot een (niet-equivalente, quasi-)controlegroep leidt tot een ernstige verzwakking van het evaluatie-ontwerp. De belangrijkste functie van de evaluatie voor de politieke en publieke discussie, namelijk het in de tijd peilen van onderwijsachterstanden, opgevat als discrepantie in gemiddelden tussen OVB leerlingen (leerlingen met een gewicht hoger dan 1.00) en de landelijke referentiegroep, zou daardoor opgeofferd worden (zie ook: Van der Werf & Tesser, 1990). In het definitieve plan is de argumentatie van Hofstee e.a. met betrekking tot het peilen in de tijd geaccepteerd en is de kritiek van Beem e.a. omtrent de cohortstructuur serieus genomen.

In dezelfde periode vraagt het Ministerie aan Doddema-Winsemius en Hofstee (1985) een nadere uitwerking van de effectmaten waar het gaat om leerprestaties, intelligentie en effecten in het sociaal-emotionele domein. Met betrekking tot de laatste stellen de onderzoekers dat 'sociale redzaamheid' hier het ankerbegrip is en dat het meten van houdingen binnen dit domein niet wenselijk is vanwege het probleem van indoctrinatie en vanwege de 'voorwendbaarheid' van houdingen.

In de voorbereidingsperiode bleken de voorgestelde intelligentietests in het overlegcircuit bij sommige actoren op grote weerstand te stuiten. Inderdaad bleek uit een haalbaarheidsonderzoek (Hoeben & Schaveling, 1986) dat veel schoolhoofden tegen afname van intelligentietests waren uit vrees voor stigmatisering van groepen leerlingen. Zoals uit het nog te beschrijven uitvoeringsplan zal blijken, zijn na



ampele discussies over voor- en nadelen de intelligentietests toch in het evaluatieplan opgenomen.

Eveneens in deze periode is aan onderzoekers van het Instituut voor Toegepaste Sociale Wetenschappen (ITS) gevraagd een gedetailleerder voorstel omtrent het implementatiedeel van het evaluatie-onderzoek, te weten: onderzoek naar het gebruik van de faciliteiten op school- en gebiedsniveau, uit te werken (Jungbluth, Buis, Driessen & Mens, 1985). De eerder genoemde onderzoekers Beem e.a. hebben in hun reactie (1986) ook dit voorstel kritisch onder de loep genomen en alternatieven aangedragen. In een poging de verschillende opties te overbruggen is Jungbluth en Meijnen (1986) gevraagd een resumerend advies uit te brengen. De aanbevelingen uit dit advies zijn in 1986 door de latere coördinator omgezet in een uitvoeringsplan (Kloprogge, 1986).

Inmiddels werd ook gewerkt aan de organisatorische vormgeving. De coördinatie wordt in handen gelegd van het Instituut voor Onderzoek van het Onderwijs (SVO). Een coördinator wordt benoemd (Kloprogge) die de onderzoeksplannen en begrotingen zal opstellen en dient zorg te dragen voor de contacten met de omgeving van het programma. Behalve het opstellen van meerjaren- en jaarlijkse plannen behoort ook het opstellen van een jaaroverzicht, een samenvatting van en een reflectie op de onderzoeksresultaten tot zijn taak. Voorts wordt er een Projectgroep Evaluatie OVB bestaande uit onafhankelijke deskundigen, ambtenaren van betrokken departementen en vertegenwoordigers van CITO en SVO door het Ministerie van O & W geïnstalleerd (voorzitter: Lagerweij), die als intermediair fungeert tussen onderzoekers en beleidsmakers. Ten slotte wordt een externe SVO commissie in het leven geroepen bestaande uit vier onafhankelijke deskundigen (voorzitter: Leune), die alle onderzoeksvoorstellen dient te beoordelen en worden twee onderzoeksinstituten geselecteerd (RION en ITS) die op grond van hun expertise het cohortonderzoek zullen gaan uitvoeren.

### 3 De voorbereiding geëvalueerd

In het voorafgaande is de feitelijke gang van zaken beschreven met betrekking tot de voorbereiding van het Landelijk Evaluatie Onderzoek naar het Onderwijsvoorrangsbeleid (LEO-OVB). Terwille van de evaluatie van de voorbereiding is het wenselijk te refereren aan de literatuur over dat type evaluatie dat het meest verwant is met, of geëigend is voor, het in gang gezette beleid.

Het Onderwijsvoorrangsbeleid is een landelijk beleidsprogramma en het beoordelen van de opbrengst daarvan zou daarom het best omschreven kunnen worden met de term Beleids-evaluatie. Hoogerwerf (1992) omschrijft dat laatste als: "Beleids-evaluatie is het beoordelen van de waargenomen inhoud, processen of effecten van een beleid aan de hand van bepaalde criteria" (p. 9).

Een dergelijke evaluatie kan zowel ex ante als ex post plaatsvinden. In het eerste geval gaat het om een beoordeling van het voorgenomen beleid, ook wel analytische evaluatie genoemd. Rationaliteit en legitimiteit van doelstellingen alsmede de doel-middelen structuur van het beleidsprogramma worden aan beoordeling onderworpen in het licht van de beschikbare wetenschappelijke kennis.

Een expliciete ex ante evaluatie heeft niet plaatsgevonden, hoewel die in het uitvoeringsplan wel was opgenomen. Naar de reden daarvan kan men gissen. Wellicht wordt de behoefte daaraan door beleidsfunctionarissen het sterkst gevoeld als geheel nieuw beleid wordt overwogen. Dat was hier niet het geval: het OVB kan als een gewijzigde vorm van het Stimuleringsbeleid (1975-1984) worden beschouwd. Wellicht wordt ook de kritische functie die een analytische evaluatie kan hebben door beleidsfunctionarissen niet altijd op prijs gesteld.

Ex post evaluaties van beleidsprogramma's zoals het Onderwijsvoorrangsbeleid (een casus, met wellicht een vergelijkingsmogelijkheid in de tijd: het Stimuleringsbeleid) zijn er in vele soorten maar de volgende vragen zijn vrijwel altijd aan de orde (Hoogerwerf, 1992, p. 55, citaat):

"a. Welke zijn de doeleinden en instrumenten van het beleid? En welke evaluatiecrite-

ria kunnen aan de beleidsdoeleinden worden ontleend?

b. In hoeverre zijn de doeleinden van het beleid bereikt? En welke niet beoogde effecten treden op?

c. In hoeverre valt de mate van doelbereiking te verklaren uit het beleid als geheel en de gekozen instrumenten in het bijzonder?

d. Hoe verhouden zich de kosten (negatieve effecten) en de baten (positieve effecten) tot elkaar?

e. Uit welke andere factoren dan de beleidsinhoud vallen de mate van doelbereiking, effectiviteit en efficiëntie te verklaren? (...) Bij die andere factoren denkt men vooral aan het beleidsproces (met name de beleidsvoering), de beleidsorganisatie, het beleidsveld en de kenmerken van de afzonderlijke factoren, waaronder hun doelgerichtheid, informatie en macht."

Wanneer we het uitvoeringsplan (Kloprogge, 1986) als afronding van de voorbereiding van het evaluatieplan beschouwen en de dan voorliggende opzet confronteren met de door Hoogerwerf als gangbaar omschreven vragen inzake beleidsevaluatie, dan kunnen we het volgende vaststellen. Met betrekking tot vraag b wordt in het uitvoeringsplan wel aandacht geschonken aan het vraagstuk van de niet beoogde effecten maar worden daar geen expliciete hypothesen over geformuleerd. Dataverzameling specifiek hierop gericht is dan ook niet gepland. Voorts komen de vragen c en d niet of nauwelijks aan de orde. Er wordt met andere woorden in het evaluatieplan niet of nauwelijks voorzien in onderzoek anders dan naar de inzet van de middelen en de effecten op de beoogde doelen (leerprestaties, schoolcarrières). Zoals reeds aangegeven gaven Hofstee e.a. zelf al aan dat hun plan *geen* beleidsevaluatie inhoudt, beleidsevaluatie omvat in hun ogen meer en zou overgelaten moeten worden aan de volksvertegenwoordiging. De laatste zinsnede is meerduidig. Dat het primaat in de besluitvorming, en dus de finale beoordeling, over nationale beleidsvragen bij de politiek ligt en niet bij de wetenschap staat niet ter discussie. Men zou de zinsnede echter ook zo op kunnen vatten dat wat wetenschappelijk omtrent het OVB te evalueren valt wel zo ongeveer is opgeschreven in het door hen gepresenteerde plan. Dat laatste is echter niet het geval. Er valt wel degelijk we-

tenschappelijk verantwoord onderzoek te verrichten naar de inhoud, het krachtenveld, de organisatie en het proces van het beleidsprogramma (zie o.a. Karstanje, 1988; Verhoeff, 1992). En ook onderzoek naar niet beoogde effecten behoort zeker tot de mogelijkheden.

Wanneer we de voorbereiding evalueren, dienen we nog bij enkele zeer relevante methodische punten stil te staan. Als eerste is dat de discussie tussen Hofstee e.a. en Beem e.a. omtrent de wenselijkheid van adequate controlegroepen. In de evaluatieliteratuur is dit probleem bekend in termen van de spanning tussen externe en interne validiteit (Cronbach, 1983; Cook & Campbell, 1979). Anders gezegd: kiest men voor een opzet waarbij het functioneren van het programma in de context zo goed mogelijk wordt gedocumenteerd (externe validiteit) dan wel voor een ontwerp waarbij men het oorzakelijk verband tussen een activiteit en het effect binnen *dit* experiment wil aantonen (interne validiteit). De positie van Beem e.a. is te omschrijven als, in navolging van Cook en Campbell, het laten prevaleren van de interne validiteit terwijl Hofstee e.a. overeenkomstig Cronbach, de externe validiteit van de te vergaren gegevens benadrukken door te wijzen op de functie die de gegevens, verkregen door hun onderzoeksoptzet, in de politieke en publieke discussie zullen vervullen.

Het debat over de functie van controlegroepen in programma-evaluatie heeft de laatste tijd iets van zijn controverse verloren. In een recent overzichtsartikel omtrent de rol van quasi-experimenten in de evaluatie van beleidsprogramma's komt Cook (1991) tot de volgende conclusie: "In 1986, Campbell's work on quasi-experiments was still the most widely acknowledged method source among evaluators. However, our impression is that by then it had lost the hegemony it had once enjoyed. It is now among the methods to be used when descriptive causal questions can be justified as the centerpiece of an evaluation design. But it is definitely not the sole method of choice" (p. 141). Naar aanleiding van een presentatie van een paper van Kloprogge over LEO-OVB in Washington in 1991 constateerde Campbell (!) dan ook dat naar zijn oordeel de gekozen opzet voor de Nederlandse situatie de meest adequate was (mondelijke mededeling).

Een tweede relevant punt is het ontbreken van criteria ('standaarden') in het evaluatieplan. Nergens wordt aangegeven welke winst in leerprestaties of schoolcarrières leerlingen uit achterstandssituaties moeten boeken om van een succesvol beleid te kunnen spreken. Het onderwerp komt noch in het ontwerp van Hofstee e.a. noch in het uitvoeringsplan ter sprake. Uit de zinsnede dat het peilen van onderwijsachterstanden in de tijd de belangrijkste functie is van de externe evaluatie, zou men kunnen opmaken dat de auteurs kennelijk de effecten van dit specifieke beleid niet denken te kunnen schatten dan wel te kunnen isoleren en het derhalve ook geen zin heeft om een nauwkeurig omschreven standaard te formuleren.

Tot slot valt te concluderen dat de voorbereiding van het evaluatie-onderzoek vergelijkenderwijs, zeer grondig is geweest. Contra-expertises zijn verricht, uitwerkingsopdrachten zijn verleend, resumerende adviezen zijn gevraagd en er is een infrastructuur gecreëerd die de onafhankelijkheid en de professionele uitvoering van de evaluatie zo goed als mogelijk garanderen.

#### 4 De uitvoering: 1986-1992

Zoals al eerder is aangegeven, is het resultaat van de voorbereidingsfase neergelegd in een uitvoeringsplan (Kloprogge, 1986). Voorts is in een themanummer van dit tijdschrift getiteld: *Evolutie in Onderwijsevaluaties* door Van der Werf en Tesser (1990) al redelijk gedetailleerd verslag gedaan over de opzet en uitvoering tot halverwege 1989. Ook verschijnt er elk jaar een door de coördinator geschreven samenvattende rapportage (Kloprogge, 1988, 1989, 1990, 1991, 1992). Wij zullen ons daarom beperken tot een bespreking op hoofdlijnen die zich leent voor een kritische beschouwing.

Eveneens al in het voorgaande aangeduid kan LEO-OVB opgedeeld worden in twee hoofdstromen van onderzoek: (a) het cohortonderzoek en (b) het flankerend onderzoek. Voor het totaal is per jaar circa 1,5 miljoen gulden beschikbaar.

##### a. *Het cohortonderzoek*

Om de twee jaar worden bij de leerlingen uit de groepen 4, 6 en 8 van circa 700 basisscholen toetsen voor taal, rekenen en een intelligentietest afgenomen. Hiermee is gestart in het najaar van 1988. Aan het oorspronkelijke voornemen om in deze cyclus ook een test voor sociale redzaamheid op te nemen, kon geen uitvoering worden gegeven omdat de instrumentconstructie teveel problemen opleverde. Per leerjaar betreft het ongeveer 12000 leerlingen. Aansluitend op de dataverzameling bij de leerlingen worden door middel van vragenlijsten gegevens verzameld over de school als organisatie en over het curriculum. Met betrekking tot het laatste wordt een onderscheid gemaakt naar kenmerken die uit de literatuur over schooleffectiviteit belangrijk zijn gebleken en kenmerken die vanuit de OVB-literatuur van belang werden geacht. Voorts werd de ouders gevraagd een vragenlijst in te vullen omtrent relevante gezinskenmerken (opleiding, beroep, land van herkomst enz.).

Het onderzoek in het voortgezet onderwijs is gestart een jaar nadat de eerste golf leerlingen uit groep 8 de basisschool had verlaten, dus in 1989. Dit betrof een random steekproef van 50% van deze leerlingen. Zij werden weer opgespoord en de overgang naar het voortgezet onderwijs evenals hun prestaties (cijfers) werden in kaart gebracht. Een jaar later (1990) is hun onderwijspositie opnieuw geïnventariseerd. Inmiddels zijn de leerlingen van de groepen 4, 6 en 8 van de basisscholen zowel in 1990 als in 1992 weer getest en getoetst. Van een generatie leerlingen zijn dus nu de gegevens over de gehele basisschoolperiode verzameld. Voorts zijn in 1990 en 1992 opnieuw cohorten gestart in groep 4. Onderzoek naar verwezenen naar het Speciaal Onderwijs is eveneens gedaan binnen het cohort 1988. Evenals onderzoek naar schoolverlaters binnen het cohort in het voortgezet onderwijs.

##### b. *Het flankerend onderzoek*

In de loop der tijd is uit het LEO-OVB budget ook veel onderzoek verricht naar specifieke problemen het OVB betreffende. Elk jaar legt de coördinator van LEO-OVB een lijst met onderwerpen voor aan de Projectgroep Evaluatie OVB die naar zijn mening voor onderzoek in aanmerking komen. Voorts reserveert hij een

budget waarmee beloftevolle, plaatselijke initiatieven geëvalueerd kunnen worden. Ook deze initiatieven worden voorgelegd aan de Projectgroep. In totaal zijn binnen de categorie flankerend onderzoek 15 projecten uitgevoerd of in uitvoering. Deze worden niet alleen bij het RION of het ITS aanbesteed maar ook bij andere onderzoeksinstituten of vakgroepen van universiteiten.

Voorbeelden van flankerend onderzoek zijn: het functioneren van OVB-gebieden, de relatie tussen OVB en de basiseducatie, de aspiraties en percepties van buitenlandse ouders en hun kinderen, de evaluatie van specifieke projecten voor buitenlandse meisjes, zorgverbreding technisch lezen, een onderzoek naar de ondersteuning.

De onderzoeksvoorstellen voor zowel het cohortonderzoek als het flankerend onderzoek worden, zoals gepland, voorgelegd aan een externe beoordelingscommissie van SVO en aan de Projectgroep. De eerste beoordeelt met name de researchtechnische aspecten van het voorstel en de laatste de uitwerking van de probleemstelling. Jaarlijks worden de samenvattende rapportages van de coördinator aangeboden aan de betrokken ministeries en aan het parlement. Voorts worden de resultaten van de test- en toetsafnames teruggedoorgegeven aan de scholen.

De uitvoering van LEO-OVB loopt, in ruime zin, vanaf 1986 en in engere zin vanaf 1988 als we de start van het cohortonderzoek als uitgangspunt nemen, dus respectievelijk 6 en 4 jaar. Lang genoeg om er in evaluatief opzicht enige kanttekeningen bij te plaatsen.

## 5 Evaluatieve kanttekeningen bij de uitvoering 1986-1992

De omgeving van een vrij groot evaluatieprogramma als van het OVB, is in principe vrij uitgebreid. Met bijna alle instellingen en organisaties, die werkzaam zijn op het onderwerp van onderzoek, zijn er wel relaties. Voor het functioneren van het evaluatieprogramma zijn er echter drie categorieën uitermate relevant:

- a de opdrachtgever, casu quo het Ministerie van Onderwijs en Wetenschappen;
- b de wetenschap;

c de geëvalueerden.

In veel literatuur over evaluatie worden vooral de opdrachtgever en de wetenschap als relevante actoren gezien (JCSEE, 1981; Chelmsky, 1987). In de Nederlandse situatie met zijn unieke verzorgingsstructuur voor het onderwijs is het echter dringend aan te bevelen ook degenen die worden geëvalueerd tot het krachtenveld rond de evaluatie te rekenen.

Deze drie elementen in de omgeving van een evaluatieprogramma hebben enkele karakteristieken die voor de evaluatie van belang zijn:

- elk van de drie is geen eenheid op zich; er is een behoorlijke mate van heterogeniteit;
- elk van de drie heeft unieke inzichten, denkbeelden en belangen;
- er is sprake van relaties tussen deze drie omgevingsfactoren.

De heterogeniteit van de elementen in de omgeving van de evaluatie is wellicht het moeilijkst te hanteren vanuit een evaluatieprogramma. De evaluatoren moeten er rekening mee houden dat er voortdurend verschuivingen optreden zowel bij de opdrachtgever als in de wetenschap en in het veld van onderzoek.

Wat betreft de opdrachtgever is voor de evaluatie van het OVB steeds van belang geweest in hoeverre verschillende onderdelen van het Ministerie of individuele ambtenaren vonden dat de evaluatie ook hun werkterrein raakte. In de wetenschap zijn er uiteenlopende disciplines en wetenschapsopvattingen, die in de loop der tijd krachtiger of zwakker worden. Tijdens de uitvoering van de evaluatie kwam het onderzoek naar de effectiviteit van scholen en instructie sterk opzetten en raakte het kwalitatieve onderzoek deels uit de gratie. In het object van onderzoek treedt er ook voortdurend een wisseling op van machtsposities en aanspreekpunten. Toen het Ministerie van WVC zich uit het OVB tijdelijk terugtrok leidde dat bijvoorbeeld tot een verzwakking van de positie van de welzijnsinstellingen ten opzichte van de scholen. Met de komst van het beleid van Sociale Vernieuwing en met een hernieuwde aandacht van WVC voor het OVB hebben er zich in het veld opnieuw wisseling van machtsposities voorgedaan.

De eigenheid aan kennis, inzichten en belangen bij de drie elementen in de omgeving is



vanuit de organisatie van het programma minder moeilijk te hanteren. Aangezien de evaluatie zelf een wetenschappelijke activiteit is, behoort het volgen van de ontwikkeling van kennis en inzichten uit de wetenschap tot de reguliere taak van de evaluatie-onderzoekers. Ambtenaren op een departement beschikken vaak over informatie over te verwachten beleidsontwikkelingen, waar de evaluatie op kan of moet anticiperen. Degenen die worden geëvalueerd hebben veelal een gedetailleerd inzicht in de praktijk van het dagelijks werk, dat tot de evaluatie behoort. Dat kan zowel leiden tot nieuwe aandachtsvelden voor de evaluatie als bijdragen aan het bevestigen van de juistheid van de onderzoeksresultaten of het relativeren daarvan. In deze zin is de in de omgeving aanwezige informatie van groot belang om de kwaliteit van de evaluatie te verhogen.

Aan in de omgeving spelende belangen kan ook de evaluatie niet voorbij gaan. De meest geëigende oplossing daarvoor is bij de opzet van de evaluatie daarop te anticiperen zowel met betrekking tot de inhoud van het programma als met de structurering van de contacten tussen het programma en de omgeving. In het voorgaande is deze structuur al uiteengezet. Dat de evaluatie gedurende een aantal jaren vrij continu en ongestoord kon worden uitgevoerd, kan als een indicatie gelden voor de kwaliteit van de structurering van de omgeving van het programma.

Een goede structurering is echter geen panacee voor alle problemen. In een betrekkelijk kleinschalige omgeving als de Nederlandse, waar alle belangrijke actoren elkaar kennen en in allerlei circuits elkaar ontmoeten, zijn de drie elementen in de omgeving duidelijk vervlochten. Dit houdt in dat een departement als opdrachtgever zeer terughoudend zal zijn om evaluatieve activiteiten te financieren, die in wetenschappelijke kring omstreden zijn of waar in het veld van onderzoek grote bezwaren tegen bestaan. Omgekeerd is het moeilijk voor de onderzochten om bezwaar te maken tegen evaluatie-onderzoek, dat wordt gesteund vanuit het departement en waarvan wetenschappers het belang onderstrepen. De vervlechting die er bestaat tussen de actoren in de omgeving kan dus zowel de positie van de evaluatoren verzwakken als versterken. Het is in ieder geval van belang met deze vervlechting vanuit

managementoogpunt rekening te houden. Dit houdt onder meer in dat er heel snel moet worden gereageerd als er ergens tegen een programma-onderdeel verzet rijst. Bij het begin van het cohortonderzoek bleken er bijvoorbeeld in enkele scholen in een gemeente bezwaren te leven tegen het afnemen van toetsen, hoewel ze hadden ingestemd met de deelname aan het programma. Deze scholen zochten medestanders in de gemeente en daarbuiten, waaronder wetenschappers. Dergelijke ontwikkelingen kunnen snel de omvang aannemen van een onoplosbaar probleem. Door in een heel vroeg stadium contact op te nemen met de betrokken scholen, uitleg te geven over de feitelijke gang van zaken en voorzieningen te treffen op die punten waar de argumenten van de scholen ter zake bleken, konden de problemen worden opgelost.

Voor het zodanig structureren van de omgeving van het programma dat het een van tevoren gegarandeerd succes wordt, bieden wetenschappelijke kennis en onderzoekservaring weinig aanknopingspunten. Belangrijk is echter dat de voorgenomen onderzoeksactiviteiten als relevant en kwalitatief hoogwaardig worden beschouwd. Hierdoor wordt een basis gelegd op grond waarvan de interactie met de omgeving kan plaatshebben. Het blijft echter nodig in het onderzoek gemaakte keuzen expliciet en veelvuldig te verantwoorden, en het is van groot belang ervoor te zorgen dat voor dergelijke activiteiten tijd en geld beschikbaar is.

Na deze meer algemene opmerkingen zal in het kort worden ingegaan op de relatie van het programma met elk der drie categorieën.

#### a. *De opdrachtgever*

In de praktijk blijkt het Ministerie de onderzoeksvoorstellen van de Projectgroep ongewijzigd over te nemen, waarbij het belang van bepaalde onderdelen soms expliciet wordt onderstreept. Commentaar op de kwaliteit van de onderzoeksresultaten wordt van de zijde van het Ministerie, de verantwoordelijke bewindslieden en het parlement niet gehoord. De resultaten op zich hebben aanleiding gegeven zowel tot commentaar van parlementariërs als van beleidsvoornemens door staatssecretaris Wallage. De laatste heeft naar aanleiding van de bevinding dat meer dan 45% van de leerlingen tot de doelgroepen van het beleid behoort, beslo-

ten tot aanscherping van de criteria. Voorts heeft hij op grond van de peiling van de achterstand van groepen allochtone leerlingen door LEO-OVB een commissie onder voorzitterschap van Van Kemenade gevraagd alternatieven, dan wel wijzigingen, ten opzichte van het huidige beleid te formuleren. De genoemde commissie heeft bij de analyse van het probleem uitvoerig gebruik gemaakt van de gegevens van het evaluatieprogramma (Ministerie van Onderwijs en Wetenschappen, 1992a,b). Voorts wordt in de Toelichting op de Wet op het Onderwijsvoorrangsbeleid, die momenteel (januari 1993) nog in behandeling is bij de Eerste Kamer, veelvuldig verwezen naar de resultaten verkregen via LEO-OVB.

#### b. *De wetenschap*

De onderzoeksvoorstellen worden, zoals gezegd, voorgelegd aan een commissie van onafhankelijke deskundigen. Het komt maar sporadisch voor dat een eerste voorstel genade vindt in de ogen van de commissie. Voorstellen worden herschreven met in sommige gevallen als nog een negatief resultaat. Eindverslagen worden gepubliceerd in een afzonderlijke reeks na eerst te zijn becommentarieerd door de Projectgroep. Voorts verschijnen er bijdragen van de onderzoekers in wetenschappelijke tijdschriften, zowel nationaal als, sporadisch, internationaal en worden papers gepresenteerd op wetenschappelijke congressen. Gelet op de kwaliteit, omvang en longitudinale structuur van de datasets van de cohorten moet echter worden opgemerkt dat de wetenschappelijke potentie daarvan bij lange na niet wordt benut. Er gaat zoveel menskracht zitten in het verzamelen van de data en het globaal analyseren daarvan ten behoeve het beleid, dat tal van wetenschappelijke vragen die ook middels deze data geanalyseerd kunnen worden, niet aan bod komen (stabiliteit van prestaties van scholen, stabiliteit van prestaties van leerlingen, kwaliteit versus egaliserend vermogen van scholen, detailanalyses van specifieke groepen leerlingen enz.). Gelet op de kwaliteit en structuur van de dataset en gegeven de aanwezigheid van hoogwaardige analyse-expertise in Nederland zou ook in buitenlandse tijdschriften veel meer over het evaluatieprogramma gepubliceerd kunnen worden dan tot nu toe het geval is.

#### c. *De geëvalueerden*

Zoals gezegd, bestonden er bij de start van het onderzoek op veel scholen bezwaren tegen de afname van tests en toetsen. Na de eerste terugrapportage aan de scholen van zowel de toetsgegevens (gespecificeerd per school) als de samenvattende rapportages, ontstond echter een niet verwachte appreciatie in het veld. Tal van uitnodigingen bereikten de onderzoekers om de resultaten te komen toelichten dan wel om bijvoorbeeld analyses op gebiedsniveau te verrichten. Al spoedig bleek de vraag dusdanig aan te zwellen dat de inrichting van een 'transferpunt' noodzakelijk werd. Met instemming van het Ministerie van O en W is dat ook gebeurd en is aanvullende formatie ter beschikking gesteld. Spoedig daarna vroegen ook bevoegde gezagen van scholen buiten de steekproef of het toetsprogramma ook bij hun scholen, tegen betaling, kon worden afgenomen. Er was al spoedig sprake van olievlakwerking: in grote steden als Rotterdam en Den Haag worden momenteel alle scholen uit onderwijsvoorrangsgebieden onderworpen aan het toetsprogramma.

De bevoegde gezagen gebruiken deze gegevens voor interne evaluatiedoeleinden. Wat aanvankelijk een hindernis leek voor het evaluatieprogramma, verkeerde in haar grootste aantrekkingskracht: het rapporteren over effecten. Bijgevolg is ook de steekproefmortaliteit bijzonder gering en de medewerking van de scholen in de steekproef zeer groot. Een dergelijke wending is uiteraard niet te voorzien en zal in dit geval alles te maken hebben met de toenemende aandacht voor accountability in het onderwijs.

Resumerend: de relaties van het evaluatieprogramma met de belangrijkste actoren zijn zonder meer goed te noemen.

Ten slotte zullen we kort ingaan op een vraag die in veel handboeken over evaluatie aan de orde komt, namelijk die naar het relatieve belang van kwaliteit versus tijdigheid. Evaluaties die voor een opdrachtgever worden uitgevoerd, dienen meestal om beslissingen te nemen over voortzetting of bijstelling van door de opdrachtgever uitgevoerde of gefinancierde activiteiten. Het tijdig opleveren van rapportages behoort dan ook tot de standaarden die voor

evaluaties gelden. Chelimsky (1987) trekt uit haar ervaringen de les dat een adequaat design, waarbij de gevraagde resultaten tijdig kunnen worden opgeleverd, vaak de voorkeur verdient boven een optimaal design. Cronbach (1983) lijkt de factor tijd meer te zien als een van de vele randvoorwaarden die in een evaluatieprogramma zo goed mogelijk moeten worden verzoend. Scheerens (1983) vindt het wel van belang dat voornemens nauwgezet en volgens de tijdsplanning worden nagekomen als onderdeel van het streven dat projectmanagement van beleidsgerichte onderwijsresearch in het algemeen moet worden verbeterd.

In de evaluatie van het OVB heeft het werken volgens een strak tijdsschema een hoge prioriteit gekregen. Er hebben zich echter diverse situaties voorgedaan waarbij een afweging moest worden gemaakt tussen kwaliteit en tijdigheid. In het algemeen luidt de conclusie hier dat het is aan te bevelen de hoogste prioriteit toe te kennen aan kwaliteitsaspecten. Tijdigheid is rekbaarder dan kwaliteit. Harde termijnen waarop activiteiten worden gepland blijken achteraf toch vaak minder hard te zijn. Politieke besluitvorming kost de nodige tijd en verloopt in diverse fasen. Het al te strak vasthouden aan termijnen is daarom vaak niet zinvol. Wanneer gekozen wordt voor minder optimale designs of analyses, levert dit meestal moeilijk herstelbare problemen op. Tijdigheid blijft uiteraard belangrijk. Aantasting van de kwaliteit ter wille van de tijdsplanning lijkt echter zelden te lonen bij evaluatie van een beleidsprogramma dat zich over vele jaren uitstrekt.

## 6 Conclusies

Samenvattend kan worden gesteld dat het van belang is om bij de opzet van de evaluatie van een beleidsprogramma nauwkeurig na te gaan welke relevante actoren zich in de omgeving van het programma bevinden en welke eisen het omgaan met de omgeving aan het programma stelt. Het zorgvuldig opzetten van het onderzoek en het structureren van de omgeving van het programma kost tijd en geld maar ze kunnen de kwaliteit van de evaluatie sterk ten goede komen. Het zorgen voor hoogwaardige kwaliteit van het onderzoek blijft de eerste

prioriteit van een evaluatie van een beleidsprogramma en is ook voorwaarde om binnen een complexe omgeving te kunnen werken. Het werken binnen gestelde termijnen is voor een evaluatieprogramma eveneens van belang maar lijkt zelden voorrang te verdienen boven het kwaliteitsaspect. In het bijzonder geldt dat voor programma's als het OVB waar niet zozeer voor een eenmalige beleidsinzet wordt geopteerd maar voor beleid dat, weliswaar in meerdere vormen, decennia bestrijkt. Vanuit deze optiek is programma-evaluatie niet in eerste instantie de evaluatie van een in de tijd goed gedefinieerd beleidsproces maar de registratie van ontwikkelingen in de tijd ('rolling reform'). Een zorgvuldige monitoring van deze ontwikkelingen, in dit geval van prestaties en schoolcarrières van groepen leerlingen, gecombineerd met zo goed mogelijke analyses richting oorzaken, is dan te prefereren boven reeksen dieptestudies dan wel experimentele designs die een precieze tijdsbegrenzing vereisen.

Tot slot. Ook binnen de categorie 'flankerend onderzoek' treffen we geen onderzoek aan dat het beleidsproces in 'hogere sferen' als onderwerp heeft. Het feit dat de Onderwijsvoorrangswet na acht jaar nog steeds niet is aangenomen en de gevolgen daarvan, de wijzigingen in doelstellingen van het OVB in de loop der tijd, de werkgelegenheidsconsequenties van de maatregelen enz. zijn geen voorwerp van onderzoek geweest. In die zin is ook het LEO-OVB geen volledige beleidsevaluatie in de zin die Hoogerwerf (1992) er aan toekent. Prioriteiten in deze richting van de kant van de opdrachtgever zijn niet te verwachten. We zullen het hier moeten hebben van onderzoek gefinancierd door NWO of door een vakgroep van een universiteit.

## Literatuur

- Beem, A. L., Pijl, Y. J., Scheerens, J. (1986). *Methodologische overwegingen bij de voorbereiding van het evaluatieprogramma Onderwijsvoorrangsbepaald (OVB)*. Den Haag: SVO.
- Chelimsky, E. (1987). What have we learned about the politics of program evaluation? *Educational Evaluation and Policy Analysis*, 9(3), 199-213.

- Cook, Th. D. (1991). Clarifying the warrant for generalized causal inferences in quasi-experimentation. In M. W. McLaughlin & D. C. Phillips (Eds.), *Evaluation and Education: At Quarter Century*. (pp. 115-145). Chicago: The University of Chicago Press.
- Cook, Th. D., & Campbell D. T. (1979). *Quasi-experimentation: Design and Analysis Issues for Field Settings*. Chicago: Rand McNally.
- Cronbach, L. J. (1983). *Designing Evaluations of Educational and Social Programs*. San Francisco: Jossey-Bass Publishers.
- Doddema-Winsemius, H., & Hofstee, W. K. B. (1985). *Operationalisering van de effecten van het onderwijsvoorrangsbeleid (OVB)*. Den Haag: SVO.
- Hoeben, W. Th. J. G., & Schaveling, J. (1986). *Evaluatie van het onderwijsvoorrangsbeleid; Haalbaar?* Groningen: RION.
- Hofstee, W. K. B., Koster, K. B., Meijnen, G. W. (1984). *Programmavoorstel voor evaluatieonderzoek naar de effecten van het onderwijsvoorrangsbeleid*. Groningen: Rijksuniversiteit.
- Hofstee, W. K. B. (1986). *Reactie naar aanleiding van het eindrapport Methodologische overwegingen bij de voorbereiding van het evaluatieprogramma OVB*. Groningen: vakgroep persoonlijkheidspsychologie (interne notitie).
- Hoogerwerf, A. (1992). Naar een verbreding van de beleidsevaluatie. *Beleidswetenschap*, 6(1), 52-68.
- JCSEE, (1981). *Standards for Evaluation of Educational Programs, Projects and Materials*. New York: McGraw-Hill.
- Jungbluth, P., Buis, Th., Driessen, G., Mens, A. (1985). *De evaluatie van het onderwijsvoorrangsbeleid. Voorstellen voor onderzoek naar het gebruik van OVB-faciliteiten en naar de daarbij beoogde effecten*. Nijmegen: ITS.
- Jungbluth, P., & Meijnen, G. W. (1986). *Resumerend advies over de vormgeving van de evaluatie van het onderwijsvoorrangsbeleid*. Nijmegen: ITS.
- Karstanje, P. N. (1988). *Beleidsevaluatie bij controlerende onderwijsvernieuwing*. Amsterdam: SCO.
- Kloprogge, J. (1986). *Uitvoeringsplan evaluatie onderwijsvoorrangsbeleid 1987-1990*. Den Haag: SVO.
- Kloprogge, J. (1988). *En toen was er het OVB...? Notitie betreffende het onderwijsvoorrangsbeleid in 1987*. Den Haag: SVO.
- Kloprogge, J. (1989). *Voorrang voor de gebieden? Notitie over het onderwijsvoorrangsbeleid in 1988*. Den Haag: SVO.
- Kloprogge, J. (1990). *Naar beter onderwijsvoorrangsbeleid. Notitie over het onderwijsvoorrangsbeleid in 1989*. Den Haag: SVO.
- Kloprogge, J. (1991). *Het gewicht gewogen. Notitie over het onderwijsvoorrangsbeleid in 1990*. Den Haag: SVO.
- Kloprogge, J. (1992). *Een kompas voor de toekomst. Notitie over het onderwijsvoorrangsbeleid in 1991*. Den Haag: SVO.
- Mens, A., & Calcar, C. van. (1981). *Onderwijsstimulering op weg*. Nijmegen/Amsterdam: ITS/SCO.
- McLaughlin, M. W., & Phillips, D. C. (Eds.). (1991). *Evaluation and Education: At Quarter Century. Ninetieth yearbook of the National Society for the Study of Education*. Chicago: The University of Chicago Press.
- Ministerie van Onderwijs en Wetenschappen (1984). *Vaststelling van een wettelijk kader voor het voeren van een samenhangend onderwijsvoorrangsbeleid (onderwijsvoorrangswet)*. Den Haag: Staatsuitgeverij.
- Ministerie van Onderwijs en Wetenschappen (1985). *Onderwijsvoorrangsplan 1985-1989*. Den Haag: Staatsuitgeverij.
- Ministerie van Onderwijs en Wetenschappen (1992a). *Ceders in de tuin. Naar een nieuwe opzet van het onderwijsbeleid voor allochtone leerlingen. Deel I*. Den Haag: Sdu.
- Ministerie van Onderwijs en Wetenschappen (1992b). *Ceders in de tuin. Voorstudies. Deel II*. Den Haag: Sdu.
- Scheerens, J. (1983). *Evaluatie-onderzoek en beleid*. Den Haag: SVO.
- Verhoeff, C. C. (1992). *Netwerkgebonden beleidsprogramma's in het onderwijs. Een studie naar implementatie en doelrealisatie van veranderingen in scholenbestand en bekostiging*. De Lier: Academisch Boeken Centrum.
- Werf, M. P. C. van der, & Tesser, P. (1990). *De evaluatie van het onderwijsvoorrangsbeleid. Pedagogische Studiën*, 67, 127-139.

Manuscript aanvaard 11-1-1993



## Auteurs

**G.W. Meijnen** is hoogleraar onderwijskunde aan de Universiteit van Amsterdam en wetenschappelijk directeur van het SCO-Kohnstamm Instituut.

Adres: SCO-Kohnstamm Instituut, Grote Bickersstraat 72, 1013 KS, Amsterdam.

**J. Kloprogge** is coördinator van de landelijke evaluatie van het onderwijsvoorrangsbeleid bij SVO.

Adres: SVO, Sweelinckplein 14, 2517 GK, Den Haag.

## Abstract

### **The Evaluation of the Educational Priority Policy in The Netherlands Evaluated**

G. P. Meijnen & J. Kloprogge. *Pedagogische Studiën*, 1993, 70, 84-96.

This article presents the analysis of a national evaluation study regarding the Educational Priority Policy. Both, the preliminary phase and the implementation are described. Furthermore, the article focuses on those topics which are considered to be important in international literature, namely, internal versus external validity, design choices and interrelations between the research programme and the environment. It can be concluded that both design and the way the research programme is carried out, still meet recent views regarding external evaluation research. At the same time, it has been established that policy evaluation encompasses more than the present contents of the research programme.