

Formatieve curriculumevaluatie, een heroverweging van de CURVO-strategie*

B. P. M. CREEMERS

RION Instituut voor Onderwijsonderzoek,
Groningen

Samenvatting

Gebaseerd op ervaringen in een aantal curriculumontwikkelingsprojecten en de evaluatie van de onderwijsvernieuwing in Nederland, wordt een herwaardering van de CURVO-strategie gepresenteerd. De voornaamste conclusies zijn:

- *Er bestaat slechts een relatief verschil tussen formatieve en summatieve evaluatie. Beide vormen van evaluatie zijn gebaseerd op oordelen over de effectiviteit van het curriculum en zijn erop gericht de relatie tussen doelen en middelen, oorzaken en effecten na te gaan.*
- *Voor een goede evaluatie is het noodzakelijk vooraf criteria en kenmerken van het curriculum nauwkeurig te formuleren. Daarbij kan een onderscheid gemaakt worden tussen standaard en specifieke kenmerken van een curriculum. Op de standaard kenmerken zijn de verschillende curricula vergelijkbaar, specifieke kenmerken verwijzen naar de extra's van een bepaald curriculum.*
- *In een project kunnen ontwikkeling en evaluatie plaatsvinden. Met het oog op de onafhankelijkheid en de kwaliteit van het curriculum wordt de voorkeur gegeven aan externe evaluatie.*

1 Inleiding

Het CURVO-project startte indertijd als een voorstudie op het terrein van curriculumontwikkeling en -evaluatie. De bedoeling ervan was na te gaan van welke theorieën en empi-

risch onderzoek gebruik kon worden gemaakt bij de ontwikkeling van curricula. In de voorstudiefase bleek dat er geen empirisch-gevalideerde theorie voorhanden was op het terrein van curriculumontwikkeling en -evaluatie (zie Creemers & Weeda, 1974; Creemers e.a., 1975). Besloten werd een dergelijke strategie te ontwikkelen op basis van beschikbare empirische gegevens en theorieën, en deze strategie vervolgens in een aantal ontwikkelingsprojecten te beproeven.

Aangezien de strategie bruikbaar zou moeten zijn in een groot aantal uiteenlopende ontwikkelingssituaties was zij inhoudelijk leeg. Dat betekent dat het ging om een formele strategie die aangaf welke stappen in het ontwikkelingsproces genomen zouden moeten worden, welke informatie verzameld zou moeten worden en welke gegevens daarvoor nodig waren. Er werd vanuit gegaan dat vakinhoudelijke kennis en informatie over de specifieke situatie waarin het curriculum zou moeten functioneren, in de ontwikkelingsgroep aanwezig waren.

Wanneer onderdelen van het curriculum klaar waren, dan zouden deze onderwerp zijn van evaluatieonderzoek. In een dergelijk evaluatieonderzoek zouden de onderdelen van het curriculum op hun waarde beproefd worden ten behoeve van de verbetering ervan. Deze evaluatie van het curriculum – tijdens de ontwikkeling ervan gericht op het nemen van beslissingen ten behoeve van de verbetering – werd volgens de gebruikelijke terminologie formatieve curriculumevaluatie genoemd. Ook voor de curriculumevaluatie werd een strategie ontwikkeld waarbij gedacht werd een oplossing te vinden voor een aantal problemen bij curriculumevaluatie, zoals bepaling van de criteria voor evaluatie, relatie tussen evaluatoren en ontwikkelaars en de verhouding tussen evaluatie tijdens de ontwikkeling van een curriculum en de summatieve evaluatie wanneer het curriculum voltooid is (Creemers & Meijer, 1978). Met de CURVO-strategie zijn inmiddels ervaringen opgedaan, die zijn neergelegd in een aantal ontwikke-

*Deze tekst is gebaseerd op een bijdrage voor de werkgroep evaluatieresearch van de Rijksuniversiteit Groningen van K. Meijer en B. P. M. Creemers

lingsverslagen (Beumer, Schuit & Terwel, 1982; Herb & Treffers, 1981; Terwel, 1980) en in een aantal wetenschappelijke publikaties, (Hoeben, 1981; Terwel, 1982, 1984). Verder is een aantal evaluaties uitgevoerd in het kader van de onderwijsvernieuwing (Scheerens, 1983; Creemers & Terlouw, 1984). Achteraf kan op basis van deze ervaringen een heroverweging plaatsvinden op een aantal belangrijke elementen van de CURVO-strategie, namelijk:

- het onderscheid tussen formatieve en summatieve evaluatie, gerelateerd aan het formele karakter van de evaluatiestrategie;
- het onderscheid in standaard en onderscheidende kenmerken van een curriculum en het probleem van de evaluatiecriteria, eveneens gerelateerd aan de formele voorstellen hieromtrent in de CURVO-strategie;
- de keuze van de CURVO-strategie voor interne evaluatie.

2 *Formatieve en summatieve curriculumevaluatie*

Wanneer men het onderscheid tussen formatieve en summatieve evaluatie hanteert, dan is formatieve evaluatie gericht op het nemen van beslissingen in het kader van de ontwikkeling van het curriculum. Summatieve evaluatie is gericht op een waardering van het ontwikkelde curriculum na afloop van het ontwikkelingsproces. Bij summatieve evaluatie gaat het dan vaak om de vergelijking van het curriculum ten opzichte van bestaande alternatieven. In beide gevallen gaat het om de vaststelling van de waarde van het curriculum. Het verschil is vooral temporeel van aard: tijdens het ontwikkelingsproces of na afloop ervan. Formatieve en summatieve evaluatie gaan vaak in elkaar over.

Over het belang van summatieve evaluatie is men het veelal eens, omdat een adequate definitie van objecten, kenmerken en criteria mogelijk is. Want tijdens de duur van het ontwikkelingsproces en de onderwijsvernieuwing zijn de kenmerken die bij de evaluatie betrokken moeten worden gespecificeerd en kunnen ze helder en duidelijk gedefinieerd worden, mede aan de hand van het ontwikkelde curriculum. Men gaat er daarbij vanuit dat tijdens de ontwikkeling de standaard- en

onderscheidende kenmerken en de criteria waarop de beoordeling moet plaatsvinden wel duidelijk zullen worden. Tijdens het ontwikkelingsproces verandert meestal een aantal doelen. De doelen worden meer realistisch in relatie tot de mogelijkheden van wat er in het onderwijs aan de gang is en worden bovendien meer geconfronteerd met traditionele onderwijsdoelen die gehandhaafd moeten blijven. Dat kan een wijziging in de onderscheidende kenmerken inhouden. Aan het eind van het ontwikkelingsproces, wanneer het produkt klaar is, kan door evaluatieonderzoek een vergelijking tussen het nieuwe curriculum en traditionele curricula worden gemaakt, of eventueel met mogelijke alternatieven aan de hand van een duidelijk set van standaard- en onderscheidende kenmerken en criteria waaraan het curriculum moet voldoen.

Summatieve evaluatie wordt opgevat als een vorm van evaluatie die een antwoord geeft op de vraag of een curriculum al of niet moet worden ingevoerd. Ook bij een invoeringsbeslissing staan nog veel alternatieven open. Daarom zou summatieve evaluatie op zo'n wijze moeten worden ingevoerd, dat het mogelijk is, wanneer bepaalde criteria niet worden bereikt, aan te geven wat daarvan de mogelijke oorzaken zijn, zoals bijvoorbeeld het ontwikkelde materiaal of de wijze van implementatie van het materiaal. Dan komt niet alleen informatie beschikbaar voor de uiteindelijke beslissing over al of niet invoeren maar ook voor de wijze waarop, of voor verbetering van het betreffende curriculum. Dat betekent dat een model moet worden opgesteld voor summatief evaluatieonderzoek, waarin de mogelijke variabelen in de onderwijsleersituatie in hun onderlinge relatie zijn opgenomen. Een dergelijk model kan in voorafgaand formatief evaluatieonderzoek ontwikkeld, uitgewerkt en uitgeprobeerd worden en vervolgens economisch worden gehanteerd in het summatief evaluatieonderzoek. Dit betekent opnieuw dat er slechts een temporeel verschil tussen formatief en summatief onderzoek bestaat.

Summatieve evaluatie, zoals hierboven omschreven, schijnt bedreigend te zijn voor curriculumontwikkelaars, want er valt een uiteindelijke beslissing over de waarde van het nieuwe curriculum.

Formatieve evaluatie wordt als minder bedreigend gezien; het staat dichterbij de ont-

wikkelaars en de overige betrokkenen. Dat formatieve evaluatie als minder bedreigend wordt ervaren, komt doordat formatieve evaluatie zich veelal beperkt tot het geven van beschrijvingen van het ontwikkelingsproces en de ontwikkelde producten van het proces zonder tot waarderende uitspraken te komen. Wanneer er al effectuering plaatsvindt, gebeurt dat in formatieve evaluatie meestal uitsluitend op interne criteria, te weten: de inhoud, (be)doel(ing)en van het curriculum, en niet op basis van externe maatstaven als schoolvorderingen, toetsen e.d., die wel vaak doorslaggevend zijn bij summatieve evaluatie. Met dergelijke gegevens is echter geen oordeel te geven over de waarde van het curriculum, maar slechts over de mate van dekking van, c.q. overeenkomst tussen inhouden van het curriculum en de evaluatie-instrumenten. Dat betekent dat het niet gaat om een vergelijking tussen formatieve en summatieve evaluatie, maar tussen summatieve evaluatie en geen evaluatie. Formatieve evaluatie dient in alle opzichten gelijk te zijn aan summatieve evaluatie, behalve dat ze plaatsvindt in de loop van het ontwikkelingsproces. Beide evaluatievormen verschaffen beoordelende informatie over ontwikkelde produktie door standaard- en onderscheidende kenmerken te relateren aan bepaalde duidelijk omschreven criteria.

Aangezien het bij formatieve evaluatie gaat om informatie die moet leiden tot verbeterde curriculumprodukten, is het van groot belang dat de informatie die verzameld wordt, geïnterpreteerd kan worden ofwel in termen van diverse alternatieven of in relatie tot een meer op basis van een theoretisch inzicht ontwikkeld model. Bij de ontwikkelaars, maar zeker bij de evaluatoren, moet inzicht bestaan in de mogelijke oorzaken, c.q. onafhankelijke variabelen, die tot bepaalde resultaten (leerprestaties) kunnen leiden. Kan bij de summatieve evaluatie soms volstaan worden met het constateren dat bepaalde resultaten bereikt of niet bereikt worden, wanneer het gaat om de verbetering van curricula, dan zal ook via evaluatieonderzoek inzicht verschaft moeten worden in de mogelijke oorza(a)k(en) voor het niet bereiken. Formatieve evaluatie vereist een goede theoretische kennis en een mogelijkheid om aan evaluatievraagstellingen een mogelijke theoretische onderbouwing te geven. Verder dient de methodisch-technische uitvoering van het onderzoek gegarandeerd te

zijn. Een aantal modellen, vaak voortbouwend op Bloom, Carroll, Harnischfeger en Wiley, Lundgren en Husèn (zie Creemers, 1985), kan bij opzet en uitvoering van evaluatieonderzoek hulp verlenen, doordat deze modellen de variabelen van belang bij de evaluatie van curricula bevatten. Essentieel in deze modellen is dat ze verschillen in leerprestaties willen verklaren. Uit het instructieonderzoek is duidelijk geworden dat een aantal (o.a. leerkracht) variabelen samenhangt met effectieve leertijd en via deze weer met het uiteindelijke leerresultaat. Alleen tijd-voor-instructie is echter nog onvoldoende om effectieve leerresultaten te bereiken. Daarvoor is nodig dat de beschikbare leertijd zo adequaat mogelijk wordt gevuld met duidelijke, heldere instructies die gestructureerd aan de leerling worden aangeboden met een maximum aan oefentijd. Dit betekent dat leerpakketten geëvalueerd moeten worden op de mate waarin zij én effectief én efficiënt gebruik maken van de beschikbare leertijd van de leerlingen door de structurering van het leerproces. Uit onderzoek naar de variabelen die op schoolniveau belangrijk zijn, blijkt dat maximalisering van de leertijd en de effectieve en efficiënte besteding ervan bevorderd worden door een stimulerende, het leren motiverende schoolorganisatie, waarbij controle van het werk van leerkrachten en leerlingen plaatsvindt en er een gezamenlijke verantwoordelijkheid is voor het bereiken van leerresultaten. Op deze wijze wordt de implementatie van het curriculum op een specifieke manier opgevat, namelijk dat intern in de school een zodanige structuur ontstaat dat (via controle, begeleiding e.d.) leerkrachten het curriculum, dat effectief is gebleken, op de juiste wijze hanteren.

Bovenstaande verhandeling over de relatie tussen formatieve en summatieve evaluatie betekent een sterke nuancering op het formele karakter van de CURVO-strategie; immers de overeenkomst tussen formatieve en summatieve evaluatie is erin gelegen, dat beide uitgaan van een variabelenmodel dat inhoudelijk de belangrijke elementen van de onderwijsleersituatie benoemt en in een onderling verband plaatst. De fundering in theorie van formatieve en summatieve evaluatie, waarvoor in het bovenstaande een betoog is gevoerd, laat de onvolkomenheid zien van een formele, inhoudelijke lege evaluatiestrategie. Curriculum-evaluatie, onderwijs-evaluatie in het alge-

meen vereist een inhoudelijke theoretische fundering.

3 *Evaluatiecriteria, standaard- en onderscheidende kenmerken*

3.1 *Standaard-, onderscheidende kenmerken en condities*

De vaststelling van de centrale evaluatievragen, zoals de kenmerken van het curriculum, die in de evaluatie betrokken zullen worden, verloopt niet altijd probleemloos.

- De doelen van het curriculumvernieuingsproject zijn nogal eens in vage termen geformuleerd;
- belanghebbenden bij de evaluatie verschillen soms van mening over de te evalueren objecten. Dit geldt met name wanneer het om politiek gevoelige vernieuwingsprojecten gaat (vgl. de discussie rondom de Middenschoolevaluatie);
- er kunnen verschillen van mening voorkomen over de 'breedte' van de evaluatieopzet. Wordt alleen de doelmatigheid van het curriculum nagegaan of worden zowel proces- als produktiekenmerken onderzocht?

Om een overzicht te krijgen van de curriculumaspecten die voor de evaluatie relevant zijn, is in de CURVO-strategie voorgesteld te differentiëren tussen de onderscheidende en de standaardkenmerken van een curriculum.

Onderscheidende kenmerken van een curriculum zijn die aspecten van een curriculum waarin het afwijkend is van andere curricula op het betreffende vakgebied.

Standaardkenmerken van een curriculum zijn die aspecten waaraan alle curricula op het betreffende vakgebied dienen te voldoen. Het gaat dus om de overlap die tussen de curricula op een bepaald gebied bestaan. Zowel bij vernieuwde als bij standaardkenmerken kan het gaan om de doelen, de werkvormen of om de hoeveelheid tijd die voor een bepaald vak beschikbaar zijn.

Bij de evaluatie zal ook gelet moeten worden op de condities waaronder het curriculum moet kunnen functioneren. Leerlingen (bijv. voorkennis), leerkrachten (bijv. didactische vaardigheden) en schoolorganisatie (bijv. aanwezigheid van practicumlokalen) zullen aan bepaalde condities moeten voldoen wil een curriculum uitvoerbaar en doelmatig

kunnen zijn. In de conceptversie van de CURVO-strategie werd een duidelijk onderscheid gemaakt tussen conditie-, standaard- en onderscheidende kenmerken. In de definitieve versie kwam het onderscheidende nog wel voor, maar niet in de vorm van handlingsaanwijzingen. De projecten die met de conceptversie werkten, hebben met deze drieling goed te functioneren en hanteerbaar te zijn bij het formuleren van evaluatievragen (De Kok-Damave, 1977; Meijer, 1977; Terwel, 1980). Bij de projecten die met de definitieve versie werkten was dit niet het geval. Overigens waren de aanwijzingen die gegeven werden voor het identificeren van de kenmerken en condities weinig concreet. Voor het opsporen van standaardkenmerken werd als aanwijzing gegeven: ga na op welke punten u kritiek kunt verwachten op het curriculum, neem dit soort vragen in het evaluatieplan op. Voor vernieuwende kenmerken en condities zou het curriculum geanalyseerd moeten worden. Aansluitend op hetgeen in de voorgaande paragraaf is opgemerkt, kan gesteld worden dat een sterke relatering nodig is aan de onderwijstheorie waarop het curriculum gebaseerd is.

Ieder curriculum is gebouwd op veronderstellingen omtrent doel-middel-relaties; hoe impliciet of onuitgewerkt ook. Het zal deze theorie moeten zijn die bij het opstellen van een evaluatieplan, het verzamelen en het interpreteren van de gegevens een belangrijke rol speelt. Bij het - vooraf - tentatief expliciteren van de samenhang tussen onderwijsactiviteiten, te bereiken leereffecten en de invloed van omstandigheden daarop kan de evaluator gebruik maken van vakinhoudelijke of meer algemene theorieën over leren en onderwijzen. Alhoewel een ontwikkelingsgroep in principe altijd in staat zou moeten zijn tot het, desnoods achteraf, expliciteren van haar theoretische opvattingen, zijn er toch situaties denkbaar waarin dit de nodige problemen zal opleveren. Wanneer dit het geval is kunnen de meer algemene theorieën over onderwijzen en leren enig soelaas bieden. Ze geven ontwikkelaar en evaluator zicht op de factoren die mogelijk relevant zijn.

3.2 *De evaluatiecriteria*

Het vaststellen van de criteria is van belang voor de interpretatie van gegevens uit evaluatieonderzoek en het geven van een waardeoor-

deel. Het, met name vooraf, vastleggen van criteria is weinig 'populair', omdat de handlingsmarges na het bekend worden van de resultaten uiterst beperkt worden. Deze onwil, of huiver, geldt zowel tijdens de eerste formatieve evaluatieronde van een curriculum ('We weten nog niet precies wat haalbaar is') als bij politiek gevoelig onderzoek van onderwijsvernieuingsprocessen (vgl. Middenschool-discussie en de Innovatie-Project-Amsterdamdebatten). Bij ontbreken van duidelijke productcriteria ziet men dan een verschuiving in de richting van implementatieonderzoek optreden (vgl. Scheerens, 1983).

In de CURVO-strategie wordt gesteld dat de criteria vooraf geformuleerd dienen te worden. Dit geldt voor de eisen waaraan zowel het proces als het produkt dienen te voldoen.

In de praktijk van het CURVO-onderzoek (Meijer, 1977a, 1977b; Terwel, 1984, 42) is gebleken dat het vooraf formuleren tot onoverkomelijke barrières heeft geleid en vaak niet gericht is. Toch is het vooraf formuleren van criteria noodzakelijk. Theoretisch gezien zijn criteria nodig, omdat zonder criteria evalueren verwordt tot een 'meten zonder meetlat'. Wanneer niet aangegeven wordt op welke onderdelen men met welke maten zal oordelen, krijgt het lezen van een evaluatierapport iets van 'het grabbelen uit een data-ton': alleen de in de eigen opvattingen passende gegevens worden behouden (vgl. Creemers & Terlouw, 1984, 140).

Het standpunt dat criteria geformuleerd dienen te worden, wijkt af van de conclusie die Terwel (1984) op basis van uitvoeringsproblemen heeft getrokken. Deze auteur stelt dat de CURVO-evaluatiestrategie te typeren valt als een anticiperend-discrepantie model. Anticiperend, omdat de eisen waaraan het curriculum moet voldoen vooraf geformuleerd moeten worden; discrepantie, omdat de feitelijke uitvoering vergeleken wordt met in theorie beoogde, waarbij 'afwijkingen van het model worden gezien als, eventueel door de werkelijkheid opgedrongen, maar altijd te betreuren discrepanties' (Terwel, 1984, 33).

Erkend wordt dat het discrepantie denken voordelen heeft, maar de nadelen wegen toch zwaarder. Het is moeilijk uitvoerbaar; de criteria worden niet serieus geformuleerd; er is weinig oog voor niet-beoogde processen en resultaten. Terwel pleit er dan ook voor 'het

getrouwheidsperspectief enigszins af te zwakken ten gunste van een minder voorgestructureerde, een opener en flexibeler aanpak. De eerste evaluatieronde zou veel meer de functie kunnen krijgen van het genereren van vragen, criteria enz. in plaats van het a priori stellen hiervan' (ibid, 50). Deze nieuwe opzet wordt soms omschreven als het 'process' of 'mutual adaptation' 'perspective'.

Ofschoon Terwels voorstellen tot het afzwakken van allerlei eisen zich lijken te beperken tot de eerste evaluatieronde zijn er toch principieel theoretische bezwaren tegen de gevolgde gedachtengang te maken. Het is onmogelijk om zonder criteria conclusies te formuleren. Uitkomsten van een evaluatieonderzoek zullen altijd geïnterpreteerd, dus in verband met criteria gebracht moeten worden. Er zullen hoe dan ook impliciet of expliciet, vooraf of achteraf criteria gehanteerd worden.

Een aanpak hierbij zou kunnen zijn dat de criteria eerst voorgelegd worden nadat de eerste fase van het ontwikkelingswerk is afgerond. Overigens is alle evaluatie in de kern gebaseerd op het anticiperende-discrepantie denken. Uitsluitend door te weten wat er verwacht wordt, is het mogelijk zinvolle afwijkingen op te sporen.

Het vooraf formuleren van criteria heeft ook een aantal praktische voordelen: voorkomen wordt dat uitsluitend op procescriteria gelet wordt;

juist door het niet wegmoffelen van de criteria worden ze tot een onderwerp van debat.

Wellicht wordt het daarmee voor beslissers moeilijker om een stellingname te ontlopen;

ten slotte geldt dat uit onderzoek is gebleken dat binnen projecten waar met een strakke(re) evaluatieopzet gewerkt werd de invloed van de evaluatieuitkomsten op de besluitvorming groter was dan in projecten met een minder strakke opzet (vgl. Van den Berg & Hoeben, 1983).

4 *Evalueren en beslissen, interne of externe evaluatie*

Sinds 1978 is een groot aantal studies verschenen die de oorzaken van de problematische verhouding tussen evaluator en beslisser trachtten bloot te leggen. Daarbij werden di-

verse mogelijkheden geopperd: soms triviaal, zoals het feit dat evaluatie-informatie vaak te laat komt, of dat er geen duidelijke vraagstellingen zijn ten aanzien van evaluatieonderzoek. Soms werd vastgesteld dat er een heel ander paradigma in evaluatieonderzoek gehanteerd zou moeten worden dan in het traditionele evaluatieonderzoek dat min of meer toch van dezelfde opvattingen uitgaat als het empirische onderzoek (Van de Vall, 1980). Een aantal onderzoeken dat in Nederland is verricht naar de relatie tussen besluitvorming en evaluatieonderzoek, de opzet en uitvoering ervan (Van den Berg & Hoeben, 1983 en Van Oyen e.a., 1984) wijst in de richting dat het niet nodig is van heel andere paradigma's uit te gaan, maar dat het min of meer traditioneel empirische onderzoek een adequaat model is voor evaluatieonderzoek en tot duidelijke conclusies kan leiden voor de besluitvormers. Er lijkt dan ook weinig reden om om te zien naar andere, meer communicatieve modellen binnen evaluatieonderzoek ter oplossing van het probleem van de relatie tussen evaluator en beslisser. Overigens zal de evaluator zich er mee moeten verzoenen dat het evaluatieonderzoek niet altijd leidt tot die ene beslissing die hij wellicht geïndiceerd acht. Evaluatieonderzoek kan vaststellen dat een criterium via dit produkt wel of niet bereikt wordt, echter niet hoe het, in het geval het criterium niet bereikt wordt, wel bereikt kan worden. Wel kunnen eventueel op grond van het theoretische model mogelijkheden worden aange-reikt.

Voor de beslissers kunnen er andere mogelijkheden zijn op grond waarvan toch tot andere dan door evaluatieonderzoek aange-reikte mogelijkheden kan worden besloten. Uit een onderzoek van Meijer blijkt dat in vele gevallen de resultaten van evaluatieonderzoek wel in de overwegingen van de beslissers zijn opgenomen. Dat betekent dus dat ze een referentiepunt hebben gevormd, maar dat soms de beslissing toch anders uitvalt. Overigens zal dit in een rationeel besluitvormingsproces, wanneer kenmerken en criteria van te voren duidelijk gedefinieerd zijn, in mindere mate voorkomen. Zeker bij summatief evaluatieonderzoek, maar ook wanneer de in paragraaf 2 aangeduide modellen ontwikkeld zijn, minder bij formatief evaluatieonderzoek.

Wanneer een duidelijke en heldere vraagstelling, een goede analyse ervan, duidelijke

kenmerken en criteria van evaluatieonderzoek, het verwerven van betrouwbare en valide objectieve informatie het grootste nut hebben voor het uiteindelijke besluitvormingsproces, dan dient de organisatie van het evaluatieonderzoek zodanig te zijn geregeld, dat de evaluator in staat is, mits hij aan de voorwaarden voldoet die in de vorige paragrafen zijn beschreven, tot een dergelijke opzet en uitvoering van evaluatieonderzoek te geraken.

In verband hiermee vindt vaak discussie plaats over de interne en externe evaluator. Hierbij wordt vaak het standpunt ingenomen dat in formatieve evaluatie een interne evaluator toelaatbaar is, vaak zelfs gecombineerd met de functie van ontwikkelaar, maar dat bij de summatieve evaluatie extern evaluatieonderzoek nodig is. Dit standpunt is discutabel. Wanneer formatieve evaluatie duidelijk gericht moet zijn op de verbetering van het te ontwikkelen produkt, dan zou een nauwe relatie tussen ontwikkeling en evaluatie aantrekkelijk zijn. Wanneer van meet af aan de concurrentie met het bestaande curriculum plaatsvindt, dan zou juist een formatieve evaluator extern gesitueerd moeten zijn, en een zo groot mogelijke afstand ten opzichte van de ontwikkelaar moeten bewaren. Dezelfde redenering volgend bij summatieve evaluatie zou, aangezien het produkt gereed is, de evaluator ook intern kunnen zijn, omdat concurrerende beïnvloeding bijna onmogelijk is.

Op grond van een aantal teleurstellende ervaringen in het verleden ten aanzien van de (formatieve) evaluatie bij onderwijsvernieuwing (middenschool, basisonderwijs) en curriculumvernieuwing neig ik ertoe, voor de keuze geplaatst: interne of externe evaluator, de evaluatie van meet af aan een duidelijk onderscheidende functie te doen zijn en daarmee buiten het innovatieproces, buiten het onmiddellijk feitelijke ontwikkelingswerk te plaatsen. Wanneer zij wel verbonden blijft met het ontwikkelingswerk, dan in een duidelijk van de ontwikkeling onderscheiden functie. De ervaringen uit het verleden maken me enigszins sceptisch of dat in een ontwikkelingssteam of innovatieteam gerealiseerd kan worden. Een mogelijkheid om dat te bereiken lijkt overigens nog steeds de duidelijke definiëring van de taken die de evaluator dient uit te voeren. Wanneer overleg met de betrokkenen over het definitieve evaluatieplan heeft plaats-

gevonden, kan de zelfstandige positie van de evaluator behouden blijven.

Psychologisch gezien zal een onafhankelijke positie van de evaluator een pluspunt zijn. Juist een wat meer afstandelijke positie maakt het voor een evaluator wat eenvoudiger tot een – soms negatief – oordeel te komen over de resultaten van het ontwikkelingswerk.

Het garanderen, behouden en bewaken van die eigen plaats kan bovendien vooral geschied door de eigen inbreng, de professionele deskundigheid van de evaluator zowel in methodologisch-technische zin (evaluatievraagstelling, methoden en technieken en dergelijke) als in inhoudelijke zin: de mogelijkheid om bij de analyse van vraagstellingen gebruik te maken van onderwijskundige theorievorming en later de theoretische reconstructie van de resultaten van formatief evaluatieonderzoek.

Literatuur

Berg, G. van den & W. Hoeben, De evaluatiepraktijk en het gebruik van evaluatiegegevens. *Pedagogische Studiën*, 1984, 61, 139-151.

Berg, G. van den & W. Th. J. G. Hoeben, *Curriculumevaluatie en beslissen*. Voortgangsverslag van SVO-project 0532 over de periode 1 maart 1981 - 1 oktober 1981, Haren, RION, 1981, Eindrapport. Haren: RION, 1983.

Beumer, T., W. Schuit & J. Terwel, *Evalueren met de CURVO-strategie*; eindrapport CURVO-2 project, deel II, 1982.

Creemers, B. P. M. & P. Weeda, *Amerikaanse reisereningen*. Utrecht: Vakgroep Onderwijskunde, 1974.

Creemers, B. P. M., K. Meijer, F. Teunissen & E. Wardenaar, *Curriculumontwikkeling: probleem-inventarisatie en trends*. Utrecht: Vakgroep Onderwijskunde, 1975.

Creemers, B. P. M. & K. Meijer, *Formatieve curriculumevaluatie*. *INFO 10*, 1978, nr. 1, 3-28.

Creemers, B. P. M. & C. Terlouw, *Evaluation in the Netherlands*. *Studies in Educational Evaluation*, 1984, 10, 215-234.

Creemers, B. P. M., *De verdere ontwikkeling van de onderwijskunde*. *Nederlands Tijdschrift voor Opvoeding, Vorming en Onderwijs*, 1985, 1, 30-40.

Herb, P. & K. Treffers, *Ontwikkeling van een leer-gang voor natuuronderwijs*. Utrecht: Vakgroep Onderwijskunde, 1981.

Hoeben, W. Th. J. G., *Praktijkgericht onderzoek en de groei van kennis*. 's-Gravenhage: SVO-reeks 46, 1981.

Kok-Damave, M. de, *Ontwikkeling van een verkeerscurriculum*. Utrecht: Vakgroep Onderwijskunde, 1977.

Meijer, K., *Ervaringen met de CURVO-strategie* (twee procesverslagen en een enquête). Utrecht: Vakgroep Onderwijskunde, 1977a.

Meijer, K., *Ontwikkeling van een leergang 'De schoolomgeving'*. Utrecht: Vakgroep Onderwijskunde, 1977b.

Oyen, P. van, I. Th. M. Snellen & J. M. van Westerlaak, *Ambtenaren en Onderzoekers*; een vergelijkende studie naar de samenhang tussen onderzoek en beleid bij 45 onderzoeksprojecten in opdracht van 10 departementen. 's-Gravenhage: Staatsuitgeverij, 1984.

Scheerens, J., *Evaluatie onderzoek en beleid* (methodologische en organisatorische aspecten). 's-Gravenhage: SVO-reeks 68, 1983.

Terwel, J., *Formatieve evaluatie van interne differentiatie volgens het BHE-model, Curnil-project, eerste evaluatieronde*. Utrecht: Vakgroep Onderwijskunde, 1980.

Terwel, J., *Formatieve evaluatie in de CURVO-strategie*, *Pedagogische Studiën* 1982, 59, 175-193.

Terwel, J., *Onderwijs maken; naar een ander onderwijs voor 12-16 jarigen*. Harlingen: Flevodruk, SVO-reeks 77, 1984.

Vall, M. van de, *Sociaal beleidsonderzoek; een professioneel paradigma*. Alphen aan den Rijn: Samsom, 1980.

Curriculum vitae

B. P. M. Creemers (1942), hoogleraar onderwijskunde Rijksuniversiteit Groningen en directeur van het RION, Instituut voor Onderwijsonderzoek.

Adres: RION, Postbus 1286, 9701 BG Groningen

Manuscript aanvaard 7-11-'86

Summary

Creemers, B. P. M., 'Formative curriculum evaluation, a re-evaluation of the CURVO-strategy.' *Pedagogische Studien*, 1987, 64, 59-66.

A re-evaluation is presented of a general strategy for curriculum evaluation. This is based on experiences in curriculum projects using this strategy and on the evaluation of educational innovation in the Netherlands. The main conclusions are:

- Only a relative distinction exists between formative and summative evaluation. In both kinds of evaluation judgement about curriculum material is based on the effectiveness of the curriculum, and evaluation is directed to investigate the relationship between means and ends (causes and effects).
- For good evaluation it is necessary to formulate prior criteria and describe carefully the features of a curriculum. A distinction can be made between standard and specific features of a curriculum. There are characteristics in which different curricula are comparable (standard) and there are specific characteristics of a particular curriculum.
- Also development and evaluation can take place within one project. External evaluation is preferred for the sake of independence and to guarantee the quality of evaluation.