

Ontwerp van een model voor didactische evaluatie

E. DE CORTE (Aangesteld navorser N.F.W.O.)

Centrum voor Psychopedagogisch en Didactisch Onderzoek, K. Univ. Leuven

Tijdens de laatste tien tot vijftien jaar hebben de opvattingen over didactische evaluatie een aanzienlijke ontwikkeling doorgemaakt. Zeer belangrijke aspecten in deze evolutie zijn de verbreding van de evaluatie naar inhoud en de differentiatie naar functie.

Sinds Tyler werd het onderwijs nagenoeg uitsluitend geëvalueerd via de resultaten ervan bij de leerlingen¹. Tegenwoordig duidt men dit aan als *produkt-evaluatie*; ze beoogt na te gaan, in hoever de doelstellingen van het onderwijs bereikt zijn bij de leerlingen die dit onderwijs hebben gevolgd. De verbreding van het begrip is tot stand gekomen als gevolg van de belangrijke rol die in de U.S.A. aan de evaluatie werd toebedeeld in het kader van de leerplanvernieuwing en -ontwikkeling. Dit heeft immers tot het inzicht geleid, dat rechtstreeks en systematische onderzoek van het onderwijsproces zelf, dat de bekomen resultaten heeft opgeleverd, eveneens een belangrijk aspect van de didactische evaluatie uitmaakt. Men spreekt in dit verband van *proces-evaluatie*.

Wat de functie van de evaluatie betreft, maakt men in de recente Amerikaanse literatuur onderscheid tussen formatieve en summatieve evaluatie; daaraan moet ons inziens een derde functie worden toegevoegd, nl. de continue evaluatie. *Formatieve evaluatie* komt neer op het verzamelen van evaluatieve informatie tijdens het proces van curriculumontwikkeling, met het oog op de voortdurende verbetering en aanpassing van dit curriculum. *Summatieve evaluatie* bestaat in het evalueren van een afgewerkt onderwijsleerplan; ze kan in feite beschouwd worden als het sluitstuk van de formatieve evaluatie. Onder *continue evaluatie* tenslotte verstaan we het permanent

evalueren van een curriculum, nadat tot het invoeren en gebruiken ervan in de praktijk werd beslist.

Als gevolg van de begripsverruiming heeft de theoretische studie nopens de evaluatie zich tijdens het voorbije decennium vooral gericht op het ontwerpen van z.g. modellen, die steunen op de nieuwe opvattingen over de inhoud van de evaluatie. Deze modellen trachten het proces van didactische evaluatie in kaart te brengen door het aanduiden van de verschillende didactische variabelen met hun relaties en interacties, die object moeten zijn van evaluatie. Een dergelijk model biedt aldus een referentiekader en een basis voor het ontwikkelen van een methodologie en een strategie voor didactische evaluatie. Op grond van de verdere studie kan het model vervolledigd en aangepast worden. Door de progressieve aangroei van de kennis over de inhoud van en over de methoden en criteria voor evaluatie kan het uitgebouwd worden tot een evaluatietheorie.

In onderhavige bijdrage wordt een model voor didactische evaluatie voorgesteld. Bij de opbouw ervan hebben wij ons geïnspireerd op het werk terzake van Cronbach, Taylor & Maguire, Stake, Scriven en Guba & Stufflebeam².

Aangezien de evaluatie volgens de verbrede visie dient gericht te zijn op het totale didactische proces, ligt het voor de hand om bij de opbouw van een evaluatiemodel uit te gaan van een analyse van dit proces. Meer bepaald hebben wij ons daarbij gebaseerd op de analyse van het didactisch proces, voorgesteld door Van Gelder³. Daar Van Gelders didactisch schema vrij algemeen bekend is, kan het volstaan hier de fasen

ervan bondig te omschrijven, vooraleer het evaluatiemodel te ontwikkelen.

Analyse van het didactisch proces volgens Van Gelder.

1. De eerste fase van het didactisch proces bestaat in het bepalen van de doelstellingen, die men door middel van het onderwijs bij de leerlingen wil nastreven en bereiken.
2. Naast het omschrijven van de doelen, stelt zich de vraag naar de didactische beginsituatie van de leerlingen, d.w.z. men moet nagaan op welk niveau de leerlingen staan bij de aanvang van het onderwijsleerproces.
3. Het ontwikkelen van onderwijsleersituaties vormt de volgende stap in het didactisch proces en heeft betrekking op de keuze en de programmering van de didactische middelen, waarmee men zal pogen de leerlingen te brengen van de beginsituatie naar de doelstellingen. Onderwijsleersituaties – door ons ook onderwijsstrategieën genoemd – worden geconstitueerd door volgende drie componenten: de leerstof, de leeractiviteit van de leerling en de onderwijsactiviteit van de leerkracht.
4. De uitvoeringsprocedure: de voorafgaande voorbereiding van het onderwijsproces is uiteindelijk gericht op het didactisch handelen van de leerkracht en het leren van het kind in de klas, m.a.w. op de onderwijsleerpraktijk.
5. Het bepalen van de resultaten van de didactische activiteit of, anders uitgedrukt, de fase van de evaluatie van het onderwijsrendement. Dit beantwoordt aan wat hierboven produkt-evaluatie werd genoemd.

Deze didactische analyse vormt de basis voor het evaluatiemodel dat we thans zullen beschrijven en waarin aan het gehele didactisch proces aandacht wordt geschonken. Op het einde van deze bijdrage wordt gepoogd het model in een schema weer te geven.

A. *Produkt-evaluatie*

Het meest voor de hand liggende uitgangspunt

voor de opbouw van een evaluatiemodel, is de vaststelling dat binnen het didactisch proces zelf een evaluatiemoment besloten ligt, m.n. de evaluatie van het rendement of produkt-evaluatie⁴.

Elke onderneming, zo ook het onderwijs, die bepaalde doelstellingen nastreeft, moet bereid zijn regelmatig na te gaan in hoever de geleverde inspanningen en de aangewende middelen daadwerkelijk tot de realisering van de doelstellingen hebben bijgedragen. M.a.w. het onderwijs moet periodiek zijn rendement evalueren. Het doel van deze evaluatie ligt vanzelfsprekend niet in zichzelf; ze is gericht op het verzamelen van informatie als basis voor het nemen van beslissingen over eventuele initiatieven tot verbetering van de middelen om de doelstellingen te bereiken, over eventuele adviezen voor een meer adequate toepassing van de gebruikte middelen door de practici, en zelfs over het al of niet verder nastreven van bepaalde doelstellingen. Bovendien kan het regelmatig vaststellen van de onderwijsresultaten op zich zelf reeds motiveren tot hernieuwde inspanningen om het onderwijsrendement te verhogen.

In de produkt-evaluatie kan men twee stappen onderscheiden, nl. globale en analytische evaluatie van het rendement enerzijds en diagnostische evaluatie anderzijds.

a. *Globale en analytische evaluatie van het rendement*

Met globale evaluatie wordt bedoeld, de beschrijving en de beoordeling van het rendement van het onderwijs *in een omschreven domein van doelstellingen als geheel* bij een bepaalde leerling of in een bepaalde klas, school of groep van scholen. Analytische evaluatie betekent de beschrijving en beoordeling van het rendement van het onderwijs in een bepaalde klas, school of groep scholen, met betrekking tot *elke doelstelling afzonderlijk*.

b. *Diagnostische evaluatie*

De globale en analytische evaluatie leidt tot een

beschrijving en een beoordeling van het rendement van het onderwijs in een bepaald domein van doelstellingen. Daarmee is men evenwel nog niet toe aan een *verklaring* van een eventueel tekort aan rendement, m.a.w. deze evaluatie stelt ons niet in kennis van de oorzaken van de tekorten, hoewel ze daaromtrent wel zekere aanwijzingen zal geven. Om echter op een gefundeerde wijze aan remediërend onderwijs te kunnen doen, moet men beschikken over meer nauwkeurige informatie nopens deze oorzaken. Daartoe moet de globale en analytische evaluatie gevolgd worden door diagnostisch onderzoek, dat men vooral kan richten op deze gebieden van doelstellingen waarbij een te zwak globaal en analytisch rendement werd geconstateerd.

In deze diagnostische evaluatie kunnen nog twee vormen onderscheiden worden:

1. De descriptieve diagnose, d.w.z. het precies *lokalisieren* van de lacunes in het kennen en kunnen, eventueel ook van de fase in het leerproces waar de ontsporing plaats gehad heeft. Hiertoe kan men de z.g. diagnostische schoolproeven gebruiken. In dergelijke proeven moet voor elke doelstelling een reeks opgaven voorkomen, die de opeenvolgende stappen of moeilijkheden in het bereiken van de doelstelling weerspiegelen.

2. Etiologische diagnose, d.w.z. het *opsporen van de oorzaken* van de vastgestelde lacune of ontsporing in het verloop van het leerproces. Hierbij dient men de volgende groepen van factoren in aanmerking te nemen.

- Factoren die betrekking hebben op het didactisch proces zelf, nl. de eventuele onaanpaktheid van de doelstelling voor de betrokken groep leerlingen of van de onderwijsstrategie die in een bepaalde school werd gevolgd om ze te bereiken. Het zal duidelijk zijn dat het onderkennen van deze factoren slechts mogelijk is door rechtstreeks onderzoek van het onderwijsproces zelf. M.a.w. dit aspect van de etiologische diagnose behoort in feite reeds tot de proces-evaluatie.

- Factoren die betrekking hebben op de leerlingen, m.n. de capaciteitsfactoren en de karakteriële correlaten, evenals het sociaal-cultureel milieu waaruit de leerlingen afkomstig zijn.

In het kader van het Centrum voor Psychopedagogisch & Didactisch Onderzoek (C.P.D.O.) te Leuven loopt sinds enkele jaren een onderzoeksproject over de globale en analytische evaluatie. Hierbij ligt het accent op het vaststellen van het onderwijsrendement bij groepen van leerlingen (een klas, een school of een groep van scholen). Het wordt nochtans niet uitgesloten, dat het globaal rendement zowel van individuele leerlingen als van een groep eventueel in het kader van eenzelfde onderzoek, m.a.w. door middel van eenzelfde evaluatie-instrument, kunnen vastgesteld worden⁵.

Wat de diagnostische evaluatie betreft, schrijft Tyler dat, hoewel in de literatuur een aantal basisprincipes voor diagnostisch onderzoek voorkomen, er slechts weinig tests beschikbaar zijn die eraan beantwoorden⁶.

In het kader van de ontwikkeling van programma's voor geïndividualiseerd onderwijs, werd vooral aan het 'Learning Research and Development Center' van de 'University of Pittsburgh', onder leiding van de gekende didacticus Robert Glaser, reeds interessant onderzoek verricht op het gebied van de descriptieve diagnostische evaluatie⁷. Ook het werk van de eveneens vooraanstaande onderwijskundige Benjamin S. Bloom over de ontwikkeling van 'diagnostic-progress tests' dient hier vermeld te worden⁸. In dezelfde lijn liggen de ideeën van Robert M. Gagné over de constructie van evaluatie-instrumenten, op basis van een regressieve taakanalyse uitgaande van de eind-doelstelling van een bepaald onderdeel van het onderwijs⁹.

Ook door het C.P.D.O. werd reeds onderzoek over descriptieve diagnostische evaluatie verricht, dat nauwe verwantschap vertoont met bovenvermelde Amerikaanse studies, nl. door Vandenberghe en Swinnen in het kader van het project 'Wetenschappelijke begeleiding van het scholenexperiment Polyvalente Beroepsschool'¹⁰.

Het is belangrijk voor ogen te houden, dat vanuit functioneel standpunt zowel formatieve, als summatieve en continue produkt-evaluatie noodzakelijk zijn. De formatieve produkt-evaluatie, d.i. evaluatie tijdens de curriculumconstructie zal gegevens opleveren, die oriënterend zijn voor de proces-evaluatie. De summatieve produkt-evaluatie is te beschouwen als een (voorlopige) eindtoets vooraleer het onderwijsleerplan in de praktijk verspreid wordt. Continue produkt-evaluatie tenslotte is onontbeerlijk als basis voor de proces-evaluatie van de onderwijsleerpraktijk. Deze continue evaluatie is niet alleen macro-didactisch, maar ook micro-didactisch van fundamenteel belang.

Uit het cyclisch en progressief karakter van het didactisch proces, waarvan de produkt-evaluatie de laatste fase is, volgt dat deze vorm van evaluatie in dubbel opzicht ook te beschouwen is als evaluatie van de beginsituatie. In de eerste plaats kan de rendementsevaluatie zekere aanwijzingen verschaffen nopens de juistheid van de informatie en onderstellingen over de beginsituatie, waarvan het onderwijsproces is uitgegaan. Uit de bevindingen kan bijvoorbeeld het vermoeden ontstaan, dat men de mogelijkheden en het niveau van de leerlingen bij de aanvang van het onderwijs niet juist heeft geschat. In de tweede plaats betekent de produkt-evaluatie aan het einde van een bepaalde onderwijsperiode meteen evaluatie van de nieuwe beginsituatie van de leerlingen in het licht van de volgende stap in hun schoolloopbaan.

Wat de instrumenten voor produkt-evaluatie betreft, zal het duidelijk zijn dat niet alleen objectief scorebare prestatieproeven van het papier-enpotlood type in aanmerking komen. Zoals reeds door Tyler werd gesteld, dienen alle technieken en methoden, die valide gegevens kunnen opleveren, aangewend¹¹. Bij wijze van voorbeeld vernoemen we: minder objectief scorebare prestatieproeven, observatie van de prestaties en de gedragspatronen van de leerlingen, vragenlijsten, beoordelingschalen, houdingsschalen, belangstellingsproeven, sociografische technieken, enz.

B. *Proces-evaluatie*

Zoals reeds gezegd, was de produkt-evaluatie tot voor een tiental jaren nagenoeg de enige bestudeerde en toegepaste vorm van systematische evaluatie en werd slechts in de recente literatuur een tweede vorm naar voren gebracht, m.n. de proces-evaluatie. Globaal omschreven bestaat deze vorm erin het didactisch proces zelf (zowel in zijn voorbereidende fasen als in de fase der praktische uitvoering), dat een bepaald produkt of rendement heeft opgeleverd, intentioneel en systematisch te onderzoeken. Zelfs de laatste fase van het didactisch proces, m.n. de produkt-evaluatie, kan object zijn van dit onderzoek.

Ten opzichte van de produkt-evaluatie is de proces-evaluatie complementair. Dit is reeds duidelijk gebleken bij de beschrijving van de inhoud van de diagnostische evaluatie. Langs de globale en analytische evaluatie en de daaropvolgende descriptieve diagnose kan men weliswaar de zwakke punten in het onderwijs lokaliseren, doch komt men niet tot een verklaring voor deze lacunes. Dit laatste is nochtans een noodzakelijke voorwaarde, zowel voor een constructieve herziening van het didactisch proces voor zover het bepaald wordt door de didactici en de leerkrachten, als voor een efficiënte remediëring voor zover het te zwakke rendement verband houdt met tekorten bij de leerlingen.

De uitbouw van het model, wat de proces-evaluatie betreft, berust op de basisgedachte dat alle momenten van het didactisch proces, evenals de interacties ertussen het voorwerp moeten zijn van evaluatie. We maken nochtans uitzondering voor de beginsituatie die, zoals daareven gezegd, grotendeels geëvalueerd wordt via het rendement van het voorafgaand onderwijs. Het ligt voor de hand dat men de bevindingen van de produkt-evaluatie verder kan aanvullen met informatie over de sociaal-culturele achtergrond van de leerlingen en met gegevens verzameld door de diensten voor psychopedagogische begeleiding.

Voor elk aspect van de proces-evaluatie zal hierna gepoogd worden het object van evaluatie te omschrijven en een eerste aanduiding te geven

betreffende de mogelijke methoden technieken en evaluatiecriteria. Inzake methoden, geldt ook hier, dat alle procédés die relevante informatie kunnen aanbrengen, geschikt zijn¹².

Wat de criteria betreft, kunnen op dit ogenblik slechts zeer tentatieve indicaties worden gegeven. Op dit punt staat de theorie en het onderzoek immers nog in het beginstadium. Men moet hier onderscheid maken tussen relatieve en meer absolute criteria. In verband met de studie van onderwijsmethoden worden in de didactiek reeds lang relatieve criteria gebruikt; m.a.w. een methode wordt beoordeeld door vergelijking, vooral wat de resultaten betreft, met een andere didactische werkvorm. Deze klassieke werkwijze wordt ook gevolgd voor curriculum-evaluatie. Cronbach heeft deze opzet, door Scriven 'comparative evaluation' genoemd, sterk bekritiseerd om volgende twee redenen. Allereerst omdat dergelijke vergelijkende studies gewoonlijk weinig positieve resultaten opleveren, in deze zin dat ze meestal geen noemenswaardige verschillen aan het licht brengen. Daarnaast oppert Cronbach methodologische bezwaren, gelegen in het feit dat beide curricula op zichzelf en de vergelijkingsgroepen (leerlingen én leerkrachten) die in het onderzoek betrokken worden, meestal zo zeer van elkaar afwijken dat eventuele verschillen moeilijk eenduidig te interpreteren zijn. Daarom pleit Cronbach voor niet-vergelijkende, intensieve studie van een curriculum, op kleine schaal en in goed gecontroleerde omstandigheden. Scriven gebruikt daarvoor de term 'non comparative evaluation'. Deze auteur neemt terzake zelf een ander standpunt in dan Cronbach. Hij meent dat naast niet-vergelijkend onderzoek de comparatieve evaluatie moet behouden blijven en haalt hiervoor vooral praktische redenen aan. De bevindingen van vergelijkende studies zijn belangrijk o.m. voor de gebruiker die een keuze moet doen tussen beschikbare curricula. Verder zijn dergelijke onderzoeken vaak gemakkelijker uitvoerbaar dan een niet-vergelijkende evaluatie, waarbij zich het moeilijke probleem van meer absolute criteria stelt¹³.

Het zijn vooral deze laatste soort criteria die wij hier op het oog hebben. Ze zijn vooralsnog

hoogstens impliciet aanwezig en de explicitering van deze absolute criteria is een moeilijke opgave. Men bevindt zich hier immers op het domein van de waarde-oordelen, dat op zichzelf reeds zeer complex is en waarvoor het exact onderzoek tot nog toe steeds allergisch is geweest. Stake & Denny hebben de huidige situatie op het stuk van de criteria kernachtig uitgedrukt in de volgende zin: 'At the present time, an evaluator is faced with a discouraging search for standards'¹⁴.

1. Doelstellingenevaluatie

Dat de onderwijsdoelstellingen een centrale plaats innemen in het didactisch proces hoeft nauwelijks gezegd te worden. Immers, alle andere momenten van het proces zijn er in mindere of meerdere mate van afhankelijk. Daarmee is meteen het grote belang van de doelstellingen-evaluatie aangetoond. Scriven stelt zelfs, ons inziens terecht, dat de waarde van de produkt-evaluatie staat of valt met het al of niet evalueren van de doelstellingen. Het heeft weinig zin, aldus deze auteur, te onderzoeken of de doelstellingen van het onderwijs bereikt zijn, indien deze doeleinden de moeite niet waard zijn om nagestreefd te worden¹⁵.

Ondanks het fundamenteel belang ervan, is op het stuk van de doelstellingenevaluatie nog weinig onderzoek verricht. Dit is niet zo verwonderlijk, wanneer men bedenkt dat zelfs de doelstellingsinventarisatie en -explicitering nog in een beginstadium verkeren¹⁶. Een belangrijke verklaring hiervoor is daareven gegeven, m.n. de waarde-gebondenheid van de doelstellingen, waardoor het gebied systematisch ontweken werd door het exact onderzoek; daardoor wordt de gehele problematiek beheerst door theoretische bespiegelingen en subjectieve meningen, die weliswaar interessant zijn, doch waaruit slechts algemene en vage doelomschrijvingen resulteren, die geen of althans geen eenduidige aanknopingspunten bieden voor de ontwikkeling van het didactisch proces.

Een tweede reden waarom de doelstellingen-evaluatie tot nog toe weinig aandacht heeft gekregen, ligt wellicht hierin, dat het evaluatie-

onderzoek in het algemeen de laatste jaren sterk geïntegreerd is geweest in het proces van curriculuminnovatie. Hierbij gaat men uiteraard uit van doelstellingen, waarvan men meent dat ze waardevol zijn. Dit kan tegenstrijdig lijken met de bewering, dat op het gebied van de doelstellingeninventarisatie nog slechts geringe voorde-lingen zijn gemaakt. Men moet echter voor ogen houden, dat de meeste Amerikaanse curriculumprojecten betrekking hebben op een bepaald leervak; de klemtoon valt daarbij vooral op de leerstofvernieuwing in het bijzonder en op de ontwikkeling van meer adequate onderwijsleer-situaties in het algemeen, mede door de pogin-gen tot integratie van onderwijstechnologische bevindingen en instrumenten. Het voorafgaande impliceert meteen een dubbele kritiek op de Amerikaanse curriculumbeweging; m.n. 1 het gebrek aan een integrale en dus niet vakgebonden aanpak van de didactische vernieuwing, 2 het feit dat de potentiële externe dynamiek van het didactisch proces, d.w.z. de betrokkenheid op de sociaal-culturele realiteit, onvoldoende geactua-liseerd wordt.

Het submodel over het evalueren van doel-stellingen zullen we hier niet beschrijven, daar dit reeds in een andere publikatie gebeurd is¹⁷. Gezien de weinige gegevens waarop kon voort-gebouwd worden, moet het betreffende concept van doelstellingenevaluatie als een eerste ver-kenning van het terrein worden beschouwd¹⁸.

2. *Evaluatie van de onderwijsstrategieën*

Volgens het voorgesteld didactisch model wordt de onderwijsleersituatie geconstitueerd door de volgende drie, nauw met elkaar samenhangende componenten: de leerstof, de leeractiviteit van het kind en de didactische werkvormen en hulp-middelen van de leerkracht.

Evaluatie van onderwijsleersituaties bestaat in het verzamelen van informatie over de waarde van de genoemde drie componenten, als basis voor het nemen van beslissingen tot herziening en vernieuwing van de onderwijstechniek. Vanuit functioneel standpunt is de formatieve evaluatie hier van uitzonderlijke betekenis; het is immers

tijdens het proces van curriculumontwikkeling, dat men de efficiëntie van de middelen maximaal moet waarborgen en hiertoe is de evaluatie als uitgangspunt voor het aanbrengen van verbeteringen een noodzakelijke voorwaarde. Dit neemt nochtans niet weg dat ook de summatieve en de continue evaluatie belangrijk zijn. De forma-tieve evaluatie zal trouwens uitlopen op summa-tieve evaluatie als (voorlopige) afsluitingsfase van de leerplanconstructie; verder moeten de onderwijsmiddelen vanzelfsprekend permanent het voorwerp zijn van evaluatie wanneer ze in de praktijk operationeel geworden zijn. Dit laatste behoort echter tot de evaluatie van de onderwijs-leerpraktijk.

De hiernavolgende beschouwingen nopens het evalueren van de onderwijsleersituatie hebben tot doel het verder denken over deze kwestie te stimuleren.

a. *Objecten van strategie-evaluatie*

Voor een verheldering van het evaluatiepro-bleem zijn het intentioneel karakter van de onderwijsstrategie en de interacties tussen de drie componenten richtinggevend.

De onderwijsleersituatie beoogt de leerlingen te brengen van de beginsituatie naar de doel-stellingen. Daaruit volgen twee principes die men bij de ontwikkeling van onderwijsstrate-gieën op het oog dient te houden en die meteen ook object moeten zijn van evaluatie: *de aan-sluiting tussen de beginsituatie en de strategieën en de overeenstemming tussen de doelstellingen en de strategieën*. Stake gebruikt in dit verband de term *contingentie*¹⁹. Wij verkiezen de term *continuïteit*, omdat deze beter het dynamisch karakter van de relaties uitdrukt.

Bij de continuïteit tussen de beginsituatie en de strategieën komt het erop aan, na te gaan of de aangewende middelen aansluiten bij het ontwikke-lingsniveau van de leerlingen. Is de leerstof in de voorgestelde aanbiedingsvorm bevattelijk voor de leerling? Wordt in de keuze van leer- en onderwijsactiviteiten rekening gehouden met de ontwikkelingsfase waarin de kinderen ge-komen zijn? Bijvoorbeeld komt het spel als

leeractiviteit voldoende aan bod in een curriculum voor de eerste klassen van het basisonderwijs?

De evaluatie van de continuïteit tussen de onderwijsleersituaties en de doelstellingen bestaat erin, te onderzoeken in hoever de ontwikkelde strategieën een adequate en volledige operationalisering van de doelstellingen belichamen. Wanneer het paraat kunnen reproduceren van een bepaalde informatie als doel gesteld is, kan deze informatie als dusdanig aangeboden worden in een aangepaste vorm en is het memoriseren en regelmatig herhalen de aangewezen leeractiviteit. Als men echter het functioneel en inzichtelijk verwerven van bepaalde principes op het oog heeft, zal de presentatie onder de vorm van problemen dienen te gebeuren en zal het zelfontdekkend verwerven en verwerken beklemtoond worden. Na het tot stand komen van het inzicht, zal men de leerlingen in de oefenfase gevarieerde opdrachten voorleggen om aldus het inzicht te laten functioneren en te verstevigen. In verband met het vormleeronderwijs in de lagere school hoort men vaak klagen, dat een belangrijk doel 'inzicht in de eigenschappen van de figuren' onvoldoende gerealiseerd wordt. De gebrekkige continuïteit tussen doelstelling en strategie biedt hier wellicht een verklaring. Het zelfontdekken van de eigenschappen van de figuren en het uitvoeren van gevarieerde constructie-opgaven, noodzakelijk om het inzicht functioneel te verwerven, komen niet voldoende aan bod en men verliest zich al te snel in het oplossen van vaak stereotype oefeningen over oppervlakteberekening.

Meteen is in het voorafgaande een belangrijk evaluatie-object aangeduid, dat echter moeilijk los te maken is van de continuïteitsevaluatie tussen doelstellingen en strategieën, m.n. *de interne coherentie tussen de drie componenten* van de strategie.

Op een facet van de continuïteitsevaluatie wensen we nog even de aandacht te vestigen, m.n. de volledigheid van de strategie ten opzichte van de doelstellingen. Het kan immers voorkomen dat bij de vertaling van de doelstellingen

in onderwijsleersituaties bepaalde aspecten verloren gaan of verwaarloosd worden.

Een component uit de onderwijsleersituatie die het voorwerp moet zijn van afzonderlijke evaluatie is *de leerstof*. Door de snelle vooruitgang van de wetenschappelijke kennis in de diverse vakdisciplines bestaat het gevaar, dat de inhoud van het onderwijs voorbijgestreefd geraakt en levensvreemd wordt. Daarom is een continue leerstofevaluatie noodzakelijk. Het is daarnaast ook wenselijk tijdens de curriculumontwikkeling, m.a.w. formatief, aan leerstofevaluatie te doen, meer bepaald met betrekking tot de aanbiedingsvorm, de aangebrachte structuur en de ordening.

Een laatste aspect van de strategie-evaluatie tenslotte is de *uitvoerbaarheid* in de realiteit. Onderwijsstrategieën moeten immers niet alleen didactisch verantwoord, maar ook in de school praktisch toepasbaar zijn. Men moet hier vooral rekening houden met de mogelijkheden van de leerkrachten, met de structuur van het onderwijs, maar ook met de beschikbare financiën.

b. *Methoden van strategie-evaluatie*

Inzake methoden van strategie-evaluatie dient men eveneens alle mogelijke, beloftevolle technieken aan te spreken. Dit braakliggend methodologisch veld biedt voorsnog zeer veel ruimte voor de inventiviteit van de evaluator. Vooral de volgende twee categorieën van onderzoekstechnieken moeten in het licht van de strategie-evaluatie uitgewerkt worden: beoordelings- en observatiemethoden. Het gebruik ervan werd in de recente literatuur door meerdere evaluatiedeskundigen, zoals Cronbach, Scriven, Stake, Taylor & Maguire en Grobman aanbevolen²⁰.

In verband met beoordelingsmethoden schrijven Stake & Denny:

As to judgments, we need (a) to be willing to admit them into our studies and (b) to devise instruments for collecting them²¹.

Daarmee is meteen gezegd, dat dergelijke onderzoeksinstrumenten nog onvoldoende aanvaard

zijn in het didactisch onderzoek en dit ondanks de ruime aanwending ervan in de andere sociale wetenschappen.

Beoordelingstechnieken kunnen adequaat toegepast worden in verband met diverse aspecten van strategie-evaluatie. Nopens de continuïteit tussen de beginsituatie en de strategieën kan men het oordeel inwinnen van ontwikkelingspsychologen, didactici en didactisch georiënteerde leerpsychologen, leerkrachten en, voor bepaalde onderwijsniveaus, van de leerlingen zelf. Wat de leerkrachten betreft, is vooral de mening interessant van hen die met de ontwikkelde onderwijstechnieken werken tijdens de experimentele toepassing. Dezelfde groepen van personen kunnen evaluerend optreden met betrekking tot de continuïteit tussen doelstellingen en strategieën. Voor dit aspect kan men bovendien leerstof-specialisten inschakelen, die meteen de aangewezen personen zijn voor de leerstofevaluatie. Het is duidelijk dat de verschillende categorieën van consultants niet alleen een waarde-oordeel kunnen vellen over het materiaal dat hen wordt voorgelegd, maar dat zij bovendien ook adviezen en suggesties voor herwerking kunnen formuleren.

Aansluitend bij Stake & Denny zouden we aldus de ontwikkeling van technieken voor de registratie en de herwerking van oordelen en opinies als een voorname taak voor de methodologie van de didactische research willen stellen.

Hetzelfde geldt trouwens voor de observatiemethoden²², die zeer veel dienst kunnen bewijzen bij de formatieve evaluatie van onderwijsleersituaties in het kader van de experimentele toepassing op beperkte schaal en in goed gecontroleerde omstandigheden. Zowel de niet-participerende observatie door onderzoekers, als de participerende observatie door de leerkrachten maar eventueel ook door leden van het onderzoeksteam zijn adequaat; aan de leerkrachten kan men tevens vragen hun ervaringen en bevindingen in een soort logboek neer te schrijven.

Een modelstudie waarin de observatiemethode werd toegepast is het onderzoek van Gallagher nopens een onderdeel uit één van de meest be-

kende Amerikaanse curriculumprojecten, m.n. de 'Biological Sciences Curriculum Study'. Gallagher registreerde in zes klassen, waarin BSCS-materiaal werd gebruikt door getrainde leerkrachten, de verbale interacties tijdens drie lessen. De onderzoeker ontwierp tevens een interessant systeem voor classificatie en analyse van deze interacties²³.

Vermeldenswaard is ook het observatieschema, ontwikkeld in het kader van het 'Individually Prescribed Instruction Project' van het 'Learning Research and Development Center' aan de University of Pittsburgh. Dit schema laat eveneens een nauwkeurige beschrijving toe van wat in de klas gebeurt. Het bestaat uit vijf rubrieken, waarin telkens een aantal items worden onderscheiden: 'independent work (by the pupils), teacher-pupil work, non instructional use of pupil time, pupil-pupil activity and group activity'²⁴.

Aanknopingspunten voor de elaboratie van observatietechnieken vindt men o.m. in de literatuuroverzichten van Boyd & Devault, van Medley & Mitzel en van Meux²⁵.

Bij de ontwikkeling van de onderzoeksinstrumenten zal de geleidelijke operationalisering van de te beoordelen en te observeren variabelen meteen leiden tot de omschrijving en de specificatie van de evaluatie-criteria. Bovendien zal de constructie van dergelijke methoden belangrijke resultaten van algemeen-didactische aard opleveren, m.n. een beter inzicht in de efficiëntie van bepaalde interactiepatronen in de klas en de verdere concretisering van een aantal fundamentele didactische principes, zoals het wekken van belangstelling, de aanschouwelijkheid, de zelf-activiteit.

3. *Evaluatie van de onderwijsleerpraktijk*

Onderwijsstrategieën worden ontworpen met het oog op het didactisch handelen en het leren in de concrete klassituatie. De noodzaak aan evaluatie van deze onderwijsleerpraktijk ligt voor de hand. Immers, de doelstellingen die men voorhoudt, kunnen pas bereikt worden, indien ze in de realiteit, m.a.w. feitelijk door de leerkrachten

worden nagestreefd. En hoe geperfectioneerd de ontwikkelde onderwijsstrategieën ook zijn, of ze het beoogde rendement zullen opleveren hangt in laatste instantie af van de toepassing in de praktijk. M.a.w. het curriculum moet het voorwerp zijn van voortdurende, periodieke evaluatie, wanneer het in de realiteit van het onderwijs verspreid en operationeel geworden is. Het gaat hier uiteraard om continue evaluatie.

In het voorafgaande is meteen een belangrijk doel van de praktijk-evaluatie vervat, m.n. onderzoeken of de vooropgezette doelstellingen feitelijk worden nagestreefd en of de ontworpen strategieën op adequate wijze worden toegepast. Dit is uiteindelijk gericht op het nemen van beslissingen nopens het geven van adviezen aan de leerkrachten en eventueel ook nopens het organiseren van recyclage-sessies. In verband met dit doel van de praktijk-evaluatie nemen we van Stake de term *congruentie-evaluatie* over. Deze auteur bedoelt daarmee de overeenkomst tussen de 'intended and observed antecedents', tussen de 'intended and observed transactions' en tussen de 'intended and observed outcomes'²⁶. Dit laatste aspect is duidelijk wat wij produkt-evaluatie genoemd hebben.

De evaluatie van de onderwijsleerpraktijk heeft evenwel nog een tweede doel, nl. het permanent verschaffen van feedback-informatie aan de curriculumontwerpers als basis voor verdere herziening en verbetering van de onderwijsstrategieën.

Uit deze doelomschrijving kunnen thans de objecten van de praktijk-evaluatie afgeleid worden.

Met het oog op het tweede doel – de voortdurende revisie en aanpassing van de strategieën – kan men hier de tweevoudige continuïteitsevaluatie, met inbegrip van het onderzoek naar de interne coherentie binnen de onderwijsleersituatie, verder doorzetten. Ook over de praktische uitvoerbaarheid van de strategieën kunnen bijkomende gegevens verzameld worden. In de fase van de praktijk-evaluatie zal vooral belang gehecht moeten worden aan het oordeel en de ervaringen van competente leerkrachten en zo mogelijk ook aan de mening van de leer-

lingen. Deze informatie kan aangevuld worden met de resultaten en bevindingen van niet-participerende observatie in een aantal klassen. Meteen is duidelijk geworden, dat de methoden en tevens de criteria voor evaluatie van de onderwijsleersituaties ook adequaat zijn voor de praktijk-evaluatie.

Wat de congruentie-evaluatie betreft, kunnen volgende objecten onderscheiden worden.

1. De congruentie tussen de doelstellingen die in de praktijk feitelijk worden nagestreefd door de leerkrachten en de doelstellingen die het curriculum voorhoudt. Het hoeft wellicht geen betoog dat hier discrepanties kunnen voorkomen.

In het evaluatiemodel van Stake ontbreekt dit object van de congruentie-evaluatie, dat tevens als een aspect van de doelstellingen-evaluatie te beschouwen is. Voor het onderzoek van deze doelstellingencongruentie is de niet-participerende observatie de aangewezen onderzoeksmethode.

2. Congruentie tussen de beginsituatie van de leerlingen, die het curriculum onderstelt en de werkelijke beginsituatie van de leerlingen die ermee onderwezen worden. Tijdens de curriculumontwikkeling moet er vanzelfsprekend naar gestreefd worden, zoveel mogelijk op basis van empirische gegevens, de aansluiting tussen het verstrekte onderwijs en het ontwikkelingsniveau te waarborgen. Bij de toepassing van het programma in de praktijk kan nochtans blijken, dat men zich hier, ondanks de evaluatie in de voorafgaande fase, op sommige punten vergist heeft. Ook hieromtrent kan zowel de observatie als het oordeel van de leerkrachten informatie aanbrenge. Deze gegevens kunnen verder aangevuld worden met eventuele resultaten, verzameld in het kader van de psychopedagogische begeleiding.

3. Congruentie tussen de ontwikkelde onderwijsleersituaties en de feitelijk toegepaste onderwijsstrategieën. Het gebrek aan onderwijsrendement, vastgesteld langs de produkt-

evaluatie kan te wijten zijn aan een verkeerde of inadequate toepassing van de strategieën in de praktijk. Daarom is dit aspect van de congruentie-evaluatie van uitzonderlijk belang. Nogmaals is de observatie de aangewezen onderzoeksmethode. De oordelen en ervaringen die de leerkrachten mededelen, kunnen eveneens relevante gegevens over de strategie-congruentie opleveren. Hieruit blijkt meteen de nauwe samenhang tussen het congruentie-onderzoek en de evaluatie van de praktijk in functie van de verbetering van het curriculum. Men zal trouwens opmerken dat deze beide aspecten slechts bij abstractie te onderscheiden zijn.

4. *Evaluatie van de didactische evaluatie*

Het bepalen van het onderwijsrendement is een evaluatie-procedure die, overeenkomstig het aangenomen didactisch model, tevens een inherente fase van het onderwijsproces is. Als dusdanig kan men de produkt-evaluatie zelf als object van evaluatie nemen. Dit komt neer op het verzamelen van informatie over de efficiëntie en de validiteit van de methoden, criteria en strategieën van produkt-evaluatie.

Het ligt nu echter voor de hand dat, zo de rendementsevaluatie op haar waarde wordt onderzocht, ook de proces-evaluatie aan deze toetsing moet onderworpen worden. Meteen is hiermee een verklaring gegeven van bovenstaande titel.

Een gedetailleerde uitwerking van een procedure voor dit soort van evaluatie zullen we hier niet geven. De elaboratie van een wetenschappelijk verantwoord en praktisch realiseerbaar systeem voor de didactische evaluatie zelf, is vooralsnog reeds een voldoende uitdaging voor de research en wanneer op dat stuk een zekere vooruitgang gemaakt is, zal men over een betere basis beschikken voor het ontwerpen van een concept voor toetsing van dit systeem.

Niettemin willen we er op wijzen, dat Guba & Stufflebeam aan deze problematiek reeds aandacht hebben geschonken. Deze auteurs geven een aantal criteria die de evaluator dient te

hanteren om zijn eigen activiteit te evalueren. Deze criteria zijn van tweeërlei aard, nl. wetenschappelijke normen en nuttigheidscriteria²⁷.

De wetenschappelijke criteria die Guba & Stufflebeam voorstellen zijn de volgende.

- De interne validiteit van de bekomen informatie ten aanzien van de variabelen, waarop de evaluatie betrekking heeft.
- De externe validiteit, d.w.z. de generaliseerbaarheid van de informatie tot andere situaties dan deze waarin ze verworven werd.
- De betrouwbaarheid, waarmee de repliceerbaarheid van de informatie bedoeld wordt. De nauwe samenhang tussen dit criterium en het voorafgaande valt onmiddellijk op. De repliceerbaarheid vormt immers de empirische basis voor de generaliseerbaarheid. Daarom kunnen beide o.i. samengenomen worden als één criterium, dat men zowel generaliseerbaarheid als betrouwbaarheid kan noemen.
- De objectiviteit van de informatie, waaronder de eenduidige interpreteerbaarheid verstaan wordt.

Daarnaast vermelden Guba & Stufflebeam de volgende nuttigheidscriteria voor het beoordelen van evaluatieve informatie.

- De relevantie van de informatie ten opzichte van de beslissingen nopens het didactisch proces, die moeten genomen worden.
- Het relatief belang van de gegevens in functie van deze beslissingen.
- De volledigheid van de informatie, waarmee bedoeld wordt dat de gegevens moeten betrekking hebben op alle factoren die in het beslissingsproces een rol spelen.
- De geloofwaardigheid ('credibility') van de informatie: 'the information must be trusted by the decision-maker and those he must serve (i.e., those who will be affected by the decisions)'.
- De 'timeliness' van de informatie, d.w.z. dat de evaluatieresultaten moeten voorhanden zijn op het ogenblik dat een bepaalde beslissing moet genomen worden. Hierbij is de vol-

gende bedenking van de auteurs vermeldenswaard: 'It is better in the evaluative situation to have reasonably good information on time than perfect information too late'.

- De doorstroming van de informatie naar allen die bij de te nemen beslissing betrokken zijn ('pervasiveness').
- De efficiëntie van het verzamelen van de evaluatieve informatie in termen van de geveerde tijd en inspanningen en van de kostprijs.

C. *Uitwerking en toepassing van het evaluatiemodel*

De uitwerking van het geschetste integraal evaluatiemodel tot een concreet evaluatiesysteem en de toepassing daarvan in het proces van curriculumontwikkeling, vertegenwoordigen een omvangrijke en gecompliceerde opdracht. Na de voorafgaande beschrijving hoeft dit ongetwijfeld nauwelijks gezegd te worden.

De omvang en de complexiteit van de opgave is evenwel geen excuus om de aanpak ervan op de lange baan te schuiven. Immers, voor het progressief doorbreken van het statisch karakter van het onderwijssysteem en voor de continue verbetering en vernieuwing ervan is de didactische evaluatie van fundamenteel belang²⁸.

Men moet zich trouwens niet laten afschrikken. Wat de uitwerking van het model tot een verantwoorde evaluatiestrategie betreft, is een benadering vanuit deelaspecten mogelijk. Deze partiële aanpak komt bijvoorbeeld tot uiting in het onderzoeksproject van het C.P.D.O., dat gericht is op de ontwikkeling van een methode voor de constructie van instrumenten voor globale en analytische produkt-evaluatie.

Inzake praktische toepassing in het proces van

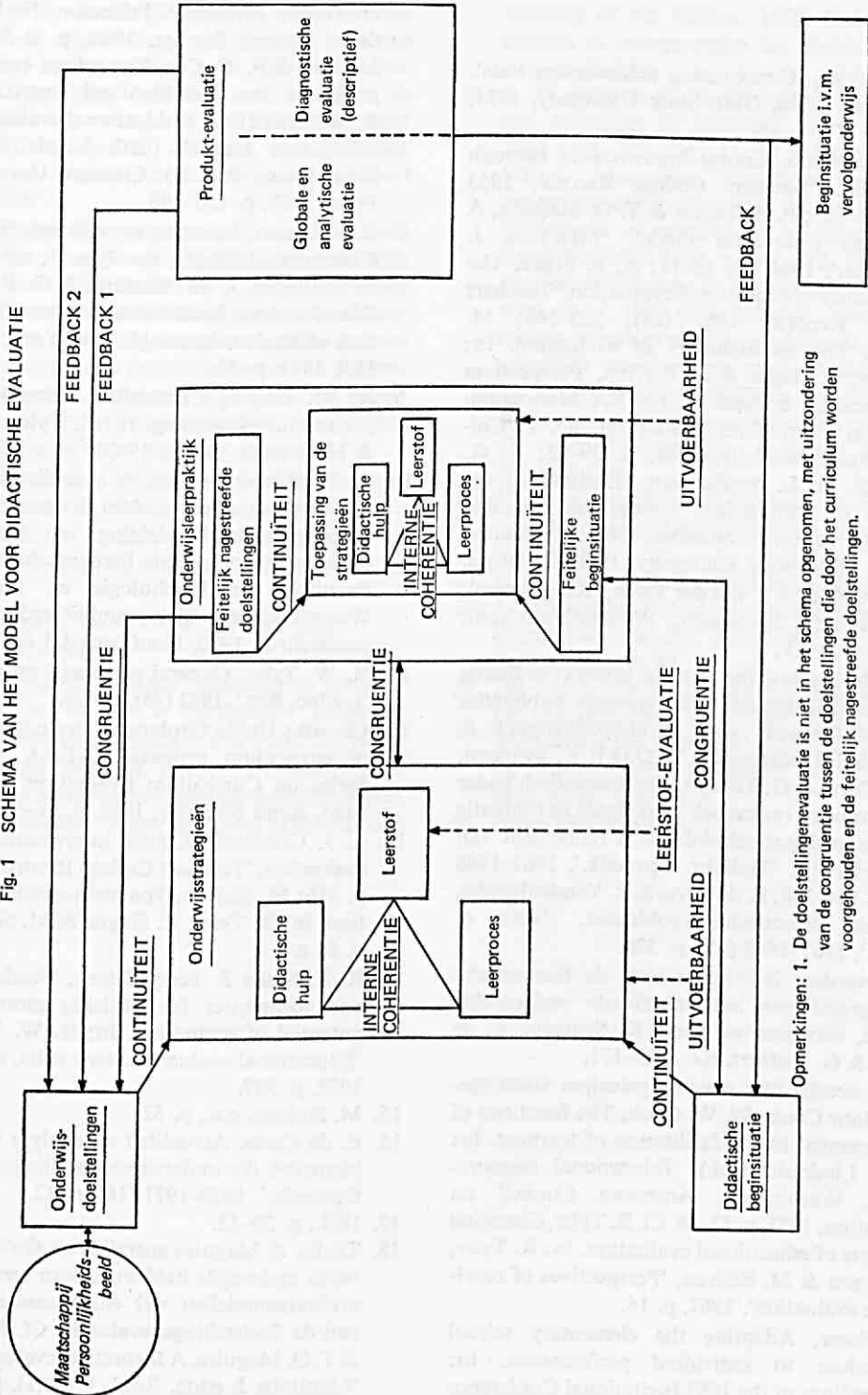
curriculumconstructie zal men realistisch tewerk moeten gaan, omdat een integrale uitvoering van de evaluatiemomenten uit het model wellicht niet mogelijk zal blijken, maar evenmin noodzakelijk zal zijn²⁹. De klemtoon zal nu eens hier en dan daar liggen. Prioriteiten zullen bepaald moeten worden, rekening houdend met de aard en de inhoud van het curriculum, met de beschikbare middelen en mogelijkheden en met de reeds bereikte fase in de ontwikkeling van het onderwijsleerplan. Ook zullen de resultaten van de produkt-evaluatie richtinggevend zijn voor de proces-evaluatie.

Zowel bij de partiële uitwerking van het model tot een evaluatiestrategie als bij de selectieve toepassing in de realiteit, is het evenwel belangrijk steeds het model in zijn geheel als achtergrond en referentiekader voor ogen te houden. Zo niet bestaat het gevaar dat men het deelaspect, waarmee men zich op een gegeven moment bezighoudt, gaat verabsoluteren ten nadele van andere belangrijke momenten. Dit leidt vanzelfsprekend tot verarming en verenging van de visie op de evaluatieproblematiek, een kwaal waarmee ook volgens Grobman nog steeds niet is afgerekend³⁰.

Tenslotte menen wij dat de uitbouw van een wetenschappelijk gefundeerd evaluatiesysteem, evenals het efficiënt functioneren ervan, slechts verwezenlijkt kan worden, indien daartoe speciale centra worden opgericht. Het spreekt vanzelf dat deze organen hun taak slechts naar behoren zullen kunnen vervullen indien zij over voldoende gespecialiseerd personeel kunnen beschikken. Daarom moet tevens voorzien worden in de opleiding van goed geschoolde evaluatoren.

Tot besluit van dit artikel volgt hierna nog een overzichtelijk schema van het ontworpen model voor didactische evaluatie (Fig. 1).

Fig. 1 SCHEMA VAN HET MODEL VOOR DIDACTISCHE EVALUATIE



Literatuur

1. R. W. Tyler, 'Constructing achievement tests'. Columbus, Ohio, Ohio State University, 1934, 102 pp.
2. L. J. Cronbach, Course improvement through evaluation. 'Teachers College Record', 1963 (64), 672-683; P. A. Taylor & T. O. Maguire, A theoretical evaluation model. 'Manitoba J. educ. Res.', 1966 (1), 12-17; R. E. Stake, The countenance of educational evaluation. 'Teachers College Record', 1967 (68), 523-540; M. Scriven, The methodology of evaluation. In: R. Tyler, R. Gagne & M. Scriven, 'Perspectives of curriculum evaluation'. (AERA Monograph Series on Curriculum Evaluation, no. 1) Chicago, Rand McNally, 1967, p. 39-83; E. G. Guba & D. L. Stufflebeam, 'Evaluation: the process of stimulating, aiding, and abetting insightful action'. Columbus, Ohio, Evaluation Center, Ohio State University, 1968, III-70 pp.
3. L. van Gelder & I. van der Velde, 'Kind, school, samenleving'. Groningen, Wolters-Noordhoff, 1968, p. 121 e.v.
4. Voor de beschrijving van de produkt-evaluatie werd gebruik gemaakt van vroegere publikaties van het Centrum voor Psychopedagogisch & Didactisch Onderzoek (C.P.D.O.): K. Swinnen, E. de Corte & G. Tistaert, Een theoretisch kader voor studie en onderzoek betreffende de evaluatie van het pedagogisch-didactisch rendement van het onderwijs. 'Tijdschr. Opvoedk.', 1967-1968 (13), p. 144-146; E. de Corte & R. Vandenberghe, Centrale didactische problemen. 'Nova et Vetera', 1967-1968 (45), p. 326.
5. Voor verdere informatie over de theoretische achtergrond van het betreffende onderzoeksproject, verwijzen we naar: K. Swinnen, E. de Corte & G. Tistaert, o.c., 129-171.
6. Tyler verwijst hier naar de principes vooropgesteld door Cook: W. W. Cook, The functions of measurement in the facilitation of learning. In: E. F. Lindquist (Ed.), 'Educational measurement'. Washington, American Council on Education, 1951, p. 37-38. Cf. R. Tyler, Changing concepts of educational evaluation. In: R. Tyler, R. Gagne & M. Scriven, 'Perspectives of curriculum evaluation'. 1967, p. 16.
7. R. Glaser, Adapting the elementary school curriculum to individual performance. In: 'Proceedings of the 1967 Invitational Conference on Testing Problems'. Princeton, N. J., Educational Testing Service, 1968, p. 3-36; C. M. Lindvall & R. C. Cox, The role of evaluation in programs for individualized instruction. In: R. W. Tyler (Ed.), 'Educational evaluation: new roles, new means'. (68th Yearb. Nat. Soc. Stud. Educ., Part II). Chicago, Univ. Chicago Press, 1969, p. 156-188.
8. B. S. Bloom, Learning for mastery. 'Evaluation Comment', 1968 (1), no. 2, p. 9; zie eveneens: B. S. Bloom, J. Th. Hastings & G. F. Madaus, 'Handbook on formative and summative evaluation of student learning'. New York, McGraw-Hill, 1971, p. 53.
9. R. M. Gagne, Curriculum research and the promotion of learning. In: R. Tyler, R. Gagne & M. Scriven, o.c., p. 19-38.
10. R. Vandenberghe, 'Studie over de strategie van onderwijsverandering. Met toepassing op een aspect-gerichte begeleiding van het scholenexperiment Polyvalente Beroepsschool'. Leuven, Faculteit der Psychologie en Pedagogische Wetenschappen, niet gepubliceerd doctoraatsproefschrift, 1970, Hoofdstuk III & IV.
11. R. W. Tyler, General statement on evaluation. 'J. educ. Res.', 1942 (35), p. 496.
12. Cf. ook: Hulda Grobman, 'Evaluation activities of curriculum projects.' (AERA Monograph Series on Curriculum Evaluation, no. 2) Chicago, Rand McNally, 1968, p. 51-52.
13. L. J. Cronbach, Course improvement through evaluation. 'Teachers College Record', 1963 (64), p. 676; M. Scriven, The methodology of evaluation. In: R. Tyler, R. Gagne & M. Scriven, o.c., p. 63 e.v.
14. R. E. Stake & Terry Denny, Needed concepts and techniques for utilizing more fully the potential of evaluation. In: R. W. Tyler (Ed.), 'Educational evaluation: new roles, new means.' 1969, p. 387.
15. M. Scriven, o.c., p. 52.
16. E. de Corte, Actualiteit en analyse van de problematiek der onderwijsdoelstellingen. 'Tijdschr. Opvoedk.', 1970-1971 (16), p. 12.
17. Ibid., p. 30-32.
18. Taylor & Maguire enerzijds en Guba & Stufflebeam anderzijds hebben in hun respectievelijke evaluatiemodellen wel enige aandacht besteed aan de doelstellingenevaluatie. Cf. P. A. Taylor & T. O. Maguire, A theoretical evaluation model. 'Manitoba J. educ. Res.', 1966 (1), p. 14; E. G.

- Guba & D. L. Stufflebeam, 'Evaluation: the process of stimulating, aiding, and abetting insightful action'. 1968, p. 51. (Het evaluatie-model van Guba & Stufflebeam wordt in de literatuur soms vermeld als het CIPP-model, naar de evaluatie-aspecten die erin onderscheiden worden, nl. 'context, input, process and product evaluation'.)
19. R. E. Stake, The countenance of educational evaluation. 'Teachers College Record', 1967 (68), p. 533. Stake spreekt over de contingentie tussen 'antecedents' en 'transactions' en tussen 'transactions' en 'intended outcomes'. Met antecedenten duidt Stake de beginsituatie van de leerlingen aan. Met transacties bedoelt hij het geheel van onderwijs- en leeractiviteiten; de leerstof als dusdanig valt niet onder dit begrip.
 20. L. J. Cronbach, o.c., p. 677; M. Scriven, o.c., p. 48-49; R. E. Stake & Terry Denny, o.c., p. 379-380 & p. 388; P. A. Taylor & T. O. Maguire, o.c., p. 16; Hulda Grobman, o.c., p. 53 e.v.
 21. R. E. Stake & Terry Denny, o.c., p. 388.
 22. Ibid., p. 379-380.
 23. J. J. Gallagher, Teacher variation in concept presentation in BSCS curriculum program. 'BSCS Newsletter', january 1967, no. 30, 8-18.
 24. J. L. Yeager & C. M. Lindvall, 'Evaluating an instructional innovation through the observation of pupil activities.' Paper presented at the Annual Meeting of the AERA, 1967. Het observatieschema is overgenomen in: Hulda Grobman, o.c., p. 58-59.
 25. R. D. Boyd & M. V. Devault, The observation and recording of behavior. 'Rev. educ. Res.', 1966 (36), 529-551; D. M. Medley & H. E. Mitzel, Measuring classroom behavior by systematic observation. In: N. L. Gage (Ed.), 'Handbook of research on teaching'. Chicago, Rand McNally, 1963, p. 247-328; M. O. Meux, Studies of learning in the school setting. 'Rev. educ. Res.', 1967 (37), 539-562. Voor een interessant overzicht van 79 verschillende observatieschema's verwijzen we naar: Anita Simon & G. E. Boyer (Eds.) 'Mirrors for behavior: an anthology of classroom observation instruments'. (2 vols). Philadelphia, Research for Better Schools, 1970.
 26. R. E. Stake, The countenance of educational evaluation. 'Teachers College Record', 1967 (68), p. 532-534.
 27. E. G. Guba & D. L. Stufflebeam, o.c., p. 67-69.
 28. J. S. Ahmann, Aspects of curriculum evaluation: a synopsis. In :R. Tyler, R. Gagne & M. Scriven, o.c., p. 89.
 29. Dit wordt ook door Hastings opgemerkt t.a.v. het evaluatiemodel van Stake. J. Th. Hastings, 'Evaluating the curricula'. Urbana, Ill., Center for Instructional Research and Curriculum Evaluation, 1967, p. 8.
 30. Hulda Grobman, o.c., p. 33.