

Leergangevaluatie: de curvo-strategie in de praktijk*

J. TERWEL, W. SCHUIT en T. BEUMER

Vakgroep Onderwijskunde Rijksuniversiteit Utrecht

Samenvatting

Leergangevaluatie staat in de literatuur bekend als een moeilijke en tijdrovende aangelegenheid. Aan het nut hiervan wordt vaak getwijfeld. Hebben evaluatie-experts er al de grootste moeite mee, nog moeilijker wordt het wanneer de evaluatie wordt uitgevoerd door schoolteams binnen de context van bijvoorbeeld een scholengemeenschap. Deze 'tour de force' is met behulp van de curvo-strategie uitgevoerd in 4 projecten. Het ging hierbij om schoolteams die, met steun van interne en/of externe begeleiders, zelf hun leergangen evalueerden.

Deze projecten zijn uitgevoerd in een M.A.V.O./L.B.O.-scholengemeenschap, een kleuterschool en een Nieuwe Leraren Opleiding. In dit artikel presenteren we de opzet en resultaten van het empirisch onderzoek naar het functioneren van de curvo-strategie in deze projecten. Centraal staat hierbij de vraag: Is de curvo-strategie uitvoerbaar, effectief en efficiënt bij het structureren van het proces van leergangevaluatie? De resultaten van het onderzoek geven aan dat de curvo-strategie deze krachtproef redelijk heeft doorstaan in drie van de vier projecten. De strategie bleek een belangrijk hulpmiddel bij het evalueren in de praktijk. Een grondige bezinning op enkele uitgangspunten en vooronderstellingen achter de deelstrategie is echter gewenst. De deelstrategie kan op verschillende punten worden verbeterd waardoor de vele waardevolle momenten nog beter tot hun recht komen.

* Dit artikel is een vervolg op het artikel van J. Terwel, *Formatieve evaluatie in de curvo-strategie*, *Pedagogische Studiën*, 1982, 59, 175-193. In het eerste artikel ligt het accent op een beschrijving en theoretische analyse van de deelstrategie voor formatieve evaluatie. Dit (tweede) artikel gaat over het empirisch onderzoek.

1 Inleiding

De curvo-strategie geeft aanwijzingen voor het maken van leergangen door ontwikkelings-teams. Een (belangrijk) deel van deze strategie betreft de formatieve evaluatie. Dit artikel gaat over het functioneren van dit deel van de strategie bij het evalueren van leergangen.

De opzet van dit artikel is als volgt. Allereerst geven we een korte beschrijving van de praktijksituaties waarin de curvo-strategie is uitgevoerd. Dan volgt een paragraaf over het doel en de opzet van het onderzoek. Daarna volgt een overzicht van respectievelijk de vraagstellingen, hypothesen, methoden en instrumenten. Vervolgens geven we een resumé van de resultaten en conclusies. In een laatste discussie-paragraaf gaan we in op de betekenis en opbrengst van dit onderzoek. In dit artikel geven we alleen de grote lijn en de belangrijkste resultaten en conclusies. Voor details en achtergronden verwijzen we naar het onderzoeksrapport waarop dit artikel is gebaseerd (Beumer, Schuit, Terwel, 1982) en naar de vijf procesverslagen waarin uitgebreid verslag wordt gedaan van de ervaringen met de strategie in de verschillende projecten.

2 De praktijksituaties

Om na te gaan of de deelstrategie in de praktijk bruikbaar is, hebben we gezocht naar ontwikkelingsgroepen die bereid waren met de curvo-strategie te werken. Tenslotte zijn we in zee gegaan met vier ontwikkelingsgroepen die de strategie hebben gebruikt bij het ontwikkelen en evalueren van leergangen. Zo ontstonden vier samenwerkingsprojecten tussen curvo en ontwikkelteams die we met de volgende namen hebben aangeduid: *curnil*, *curlant*, *curtel* en *cursol*. Het woordje 'cur' slaat op curvo, terwijl met de tweede lettergreep een bepaald ontwikkelteam wordt aangeduid. Het gaat om vier projecten die in heel verschillende situaties hebben gewerkt aan de ontwikkeling van verschillende leergangen. Het ging echter steeds om schoolteams (d.w.z. groepen die verbonden

zijn aan een school of opleiding) aangevuld met enkele begeleiders vanuit curvo. In één geval (curtel) waren daarnaast nog andere externe begeleiders aanwezig. Per project geven we een overzicht van enkele belangrijke algemene gegevens. Daarnaast wordt, zeer in het kort, informatie gegeven over het evaluatieproces en de rol die de deelstrategie daarbij heeft gespeeld.

Het curnil-project

Het curnil-project, op de MAVO/LBO scholengemeenschap Nieuwelant te Schiedam, was gericht op het formatief evalueren van drie leerjaren voor gedifferentieerd onderwijs in het brugjaar voor de vakken Nederlands, Engels en Wiskunde.

Projecttitel	: Curvo - Nieuwelant (curnil)
Scholingsgebieden	: Nederlands, Engels, Wiskunde
Scholingsniveau	: Brugklas MAVO/LTO/LEAO
Uitvoerende instantie	: Medewerkers Nieuwelant i.s.m. curvo
Periode	: maart 1978 - december 1980

Figuur 1 Gegevens over het curnil-project

De curvo-strategie is in dit project alléén gebruikt bij de formatieve evaluatie. Er heeft dus geen volledig geïntegreerd ontwikkelings- en evaluatieproces plaatsgevonden. Het samenwerkingsproject tussen Nieuwelant en curvo is gestart op het moment dat de leerjaren in eerste versie (vrijwel) klaar waren, terwijl de concept-discussie min of meer was afgerond. Ten behoeve van het gebruik van de curvo-deelstrategie en om een verantwoorde evaluatie mogelijk te maken, heeft een reconstructie en explicitering plaatsgevonden van de concept-discussie. Dit is bij het begin van de samenwerking op schrift gesteld in enkele memo's. Vervolgens is in twee achtereenvolgende evaluatieronden een evaluatieplan opgesteld en uitgevoerd. Voor meer informatie verwijzen we naar twee procesverslagen (Terwel, 1980 en Beumer en Terwel, 1981). De inhoudelijke resultaten zijn neergelegd in een eindevaluatierapport (Van den Berg en Terwel, 1980). Dit rapport is in een artikel samengevat (Bjorkman, Beumer en Terwel, 1981).

Het curlant-project

Het curlant-project, eveneens op de MAVO/LBO scholengemeenschap Nieuwelant te Schiedam, was gericht op het ontwikkelen en formatief evalueren van een lessenserie Mechanische Techniek waarbij speciaal aandacht is geschonken aan de differentiatieproblematiek.

Projecttitel	: Curvo - Nieuwelant (curlant)
Scholingsgebied	: Mechanische Techniek
Scholingsniveau	: LTS 4
Uitvoerende instantie	: Leerkrachten en medewerkers van Nieuwelant i.s.m. curvo
Periode	: mei 1980 - mei 1981

Figuur 2 Gegevens over het curlant-project

In dit project is een lessenserie ontwikkeld en formatief geëvalueerd waarbij de curvo-strategie vanaf het begin is gevolgd tot en met de eerste evaluatieronde. Er is dus een geïntegreerd ontwikkelings- en evaluatieproces gerealiseerd. Voor meer informatie verwijzen we naar het procesverslag (Herfs, 1981) en het evaluatierapport (Reints, 1981).

Het curtel-project

Het curtel-project, op de basisschool in Almelo, was gericht op het ontwikkelen en formatief evalueren van een lessenserie voor Natuuronderwijs.

Projecttitel	: Curvo - Telgenkamp - NOB (curtel)
Scholingsgebied	: Natuuronderwijs
Scholingsniveau	: Kleuterschool
Uitvoerende instantie	: Kleuterleidsters en leerkrachten Telgenkamp i.s.m. curvo
Periode	: januari 1980 - oktober 1980

Figuur 3 Gegevens over het curtel-project

De lessenserie, die is opgezet rondom het thema water, was bestemd voor de jongste kleuters. Via prikacties² poogde men greep te krijgen op Natuuronderwijs. Eén van die prikacties zou gestructureerd worden aan de hand van de curvo-strategie. Na de voorlopige vaststelling van doel en inhoud van de leergang

besloot de curtel-groep de curvo-strategie niet meer als nauwgezet te volgen stappenschema te gebruiken. Men stapte over op de (incrementele) NOB-werkwijze, die men reeds eerder beproefd had. Om ervoor te zorgen dat curvo toch voldoende onderzoeksgegevens zou kunnen vergaren op de Telgenkamp, werd besloten om de deelstrategie voor formatieve evaluatie zo zorgvuldig mogelijk uit te voeren, om te komen tot een bijstelling van de ontwikkelde lessenserie. Niettemin heeft de uitvoering van de eerste evaluatieronde van de deelstrategie maar zeer ten dele plaatsgevonden. Voor meer gegevens verwijzen we naar het procesverslag (Herfs en Treffers, 1981).

Het cursol-project

Het cursol-project van de vakgroep Huishoudkunde aan de Stichting Opleiding Leraren te Utrecht, was gericht op het ontwikkelen en formatief evalueren van een cursus 'Besluitvorming' voor 4e jaars studenten 'Huishoudkunde' aan de Nieuwe Leraren Opleiding (N.L.O.).

Projecttitel	: Curvo - S.O.L. (cursol)
Scholingsgebied	: Huishoudkunde
Scholingsniveau	: Nieuwe Leraren Opleiding
Uitvoerende instantie	: Docenten van de Vakgroep Huishoudkunde i.s.m. curvo
Periode	: september 1979 - juni 1981

Figuur 4 Gegevens over het cursol-project

In dit project is de curvo-strategie vanaf het begin gevolgd tot en met de eerste evaluatieronde. Hierbinnen zijn enkele evaluatie-activiteiten vervallen of anders verlopen dan de deelstrategie voorschrijft. Voor meer informatie verwijzen we naar het procesverslag (Glas, e.a., 1981).

3 Doel en opzet van het onderzoek

Algemene typering: evaluatie-onderzoek

Zoals in de samenvatting reeds werd gesteld gaat het om een onderzoek naar het functioneren van de deelstrategie voor formatieve evaluatie. Dat wil zeggen: het gaat om de vraag naar de uitvoerbaarheid, effectiviteit en efficiëntie van de curvo-strategie bij het evalueren

van leergangen. Het gaat om een waardebeoordeling van de deelstrategie voor formatieve evaluatie. Daarom kan het onderzoek worden getypeerd als *evaluatie-onderzoek*. Aan ons onderzoek onderscheiden we 3 aspecten: verbetering van de strategie, informatieverstrekking aan toekomstige gebruikers en het leveren van een bijdrage aan theorievorming rondom formatieve evaluatie. Deze aspecten, die onderling in een zekere spanningsverhouding staan, vormen de gezichtspunten waaronder dit onderzoek kan worden gezien. Kort samengevat kan het onderzoek worden getypeerd als een evaluatie-onderzoek met praktijkgerichte en theoriegerichte aspecten. Omdat wordt gewerkt met 4 projecten, d.w.z. een zeer kleine populatie van onderzoeks-eenheden, zou men het onderzoeksdesign kunnen typeren met de term case-studies.

4 De vraagstellingen en hypothesen

Bij het ontwikkelen van de vraagstellingen en hypothesen is aangesloten bij de *basispretentie* van de deelstrategie, nl. dat de aanwijzingen praktisch uitvoerbaar, effectief en efficiënt zijn bij het opstellen en uitvoeren van een evaluatieplan. Hieruit zijn in totaal drie hoofdvragen afgeleid voor het onderzoek. In het vervolg van dit artikel geven we deze drie hoofdvragen aan met de Romeinse cijfers I, II en III (zie Figuur 5).

- | |
|---|
| <p>I Is de curvo-deelstrategie voor formatieve evaluatie uitvoerbaar en effectief bij het opstellen van het evaluatieplan?</p> <p>II Zijn het evaluatieplan en de daarbij behorende aanwijzingen uit de curvo-deelstrategie m.b.t. de uitvoering van dat plan uitvoerbaar en effectief?</p> <p>III Is het proces van opstellen en uitvoeren van het evaluatieplan efficiënt verlopen?</p> |
|---|

Figuur 5 Overzicht 3 hoofdvragen van het onderzoek

Elke hoofdvraag is uiteengelegd in een reeks deelvragen. Deze vragen hebben betrekking op de mate waarin de kenmerken van de deelstrategie feitelijk zijn gerealiseerd, de verwachte effecten en de kosten en baten van de evaluatie.

Hoofdvraag I (m.b.t. evaluatieplan) in deelvragen uiteengelegd:

- A. Zijn de kenmerken m.b.t. fasering en methodiek gerealiseerd?
- B. Zijn de kenmerken m.b.t. het produkt gerealiseerd? D.w.z. voldoet het evaluatieplan aan de eisen?
- C. Zijn de kenmerken m.b.t. het groepsfunctioneren gerealiseerd?
- D. Is het oordeel van de groepsleden positief over het proces?
- E. Is het proces rationeel verlopen?
- F. Is het oordeel van de groepsleden positief over het produkt (evaluatieplan)?
- G. Is het oordeel van *evaluatiedeskundigen* positief over het produkt (evaluatieplan)?
- H. Is het oordeel van *inhoudelijk deskundigen* positief over het produkt (evaluatieplan)?

Hoofdvraag II (uitvoeren van de evaluatie) in deelvragen uiteengelegd:

- I. Zijn de kenmerken m.b.t. fasering en methodiek gerealiseerd?
- J. Zijn de kenmerken m.b.t. het produkt gerealiseerd? D.w.z. voldoet het evaluatierapport aan de eisen?
- K. Is het oordeel van de groepsleden positief over het proces?
- L. Is het oordeel van de groepsleden positief over de resultaten?
- M. Is het oordeel van *evaluatiedeskundigen* positief over het evaluatierapport?
- N. Is er een beter produkt ontstaan?
- O. Hebben de groepsleden een positieve houding tegenover evalueren gekregen of behouden?

Hoofdvraag III in deelvragen uiteengelegd:

- P. Wat waren de kosten van het opstellen en uitvoeren van het evaluatieplan?
- Q. Wat waren de baten van het opstellen en uitvoeren van het evaluatieplan?
- R. Staan de kosten en baten in een redelijke verhouding?

In de reeks deelvragen A t/m R kan men, grofweg, 3 soorten onderscheiden:

Uitvoerbaarheidsvragen (kenmerken A, B, C, I en J)

Deze vragen hebben betrekking op de realisering van kenmerken uit de deelstrategie. De antwoorden op deze vragen geven aan of en in hoeverre de strategie volgens de bedoelingen is uitgevoerd³. Dit noemen we *discrepantie-analyse*.

Effectiviteitsvragen (effecten D, E, F, G, H, K, L, M, N, O)

Deze vragen hebben betrekking op de realisering van effecten. De antwoorden op deze vragen geven een indicatie voor de effectiviteit van de deelstrategie. Dit noemen we de *effect-analyse*.

Efficiëntievragen (P, Q, R)

Deze vragen naar de kosten, baten en de verhouding hiertussen. Dit betreft de *kosten-baten-analyse*.

De deelvragen worden afzonderlijk beantwoord, d.w.z. er worden geen relaties tussen deelvragen gelegd. Over de relatie tussen de realisering van kenmerken en het optreden van de effecten hebben we de volgende *basishypothese* opgesteld:

Als de handelingsaanwijzingen uit de curvo-deelstrategie worden uitgevoerd dan zullen de beoogde resultaten optreden.

Deze basishypothese is bewust op deze wijze geformuleerd, omdat dit in principe de mogelijkheid biedt om: a. de hypothese te weerleggen, b. per project de hypothese te toetsen, d.w.z. elke case als testcase voor de hypothese te gebruiken. De hypothese kan, in de woorden van De Groot (1972) worden gezien als een deterministische hypothese. Per case (project) kan worden nagegaan of de hypothese en de daaruit afgeleide voorspellingen kunnen worden weerlegd.

Een belangrijke vraag betreft de invulling van de 'ALS-' en de 'DAN-kant' van de hypothese. Bij de 'ALS-kant' gaat het om die kenmerken uit de deelstrategie die essentieel zijn voor het optreden van de verwachte effecten (zie uitvoerbaarheidsvragen). De 'DAN-kant' van de hypothese hebben we ingevuld met verwachte resultaten of effecten (zie effectiviteitsvragen).

Uit de genoemde basishypothese zijn twee hypothesen afgeleid voor de twee deelprocessen binnen formatieve evaluatie, namelijk één hypothese voor het opstellen en één voor het uitvoeren van het evaluatieplan. Deze twee hypothesen hebben we als volgt geformuleerd:

Hypothese I: *Als de handelingsaanwijzingen uit de curvo-deelstrategie m.b.t. het opstellen van het evaluatieplan worden uitgevoerd dan leidt dit tot de beoogde effecten (processen en resultaten of produkten).*

Hypothese II: *Als het evaluatieplan en de handelingsaanwijzingen uit de deelstrategie*

m.b.t. het uitvoeren van het evaluatieplan worden uitgevoerd dan leidt dit tot de beoogde effecten (processen en resultaten of produkten).

De twee hypothesen corresponderen met de hoofdvragen I en II van het onderzoek. Uit elke hypothese zijn 5 voorspellingen afgeleid d.w.z. voor elk verwacht effect één voorspelling.

Hypothese I is als volgt uiteengelegd. De 'ALS-kant' van elke voorspelling bestaat telkens uit de kenmerken A, B en C van de curvo-deelstrategie. De 'DAN-kant' van elke voorspelling bestaat telkens uit één effect: D, E, F, G of H. (Voor de betekenis der letters zie par. 4).

De 5 voorspellingen die uit hypothese I zijn afgeleid kunnen in formele taal als volgt worden samengevat:

- Voorspelling I,1 : (AABAC) → D
I,2 : (AABAC) → E
I,3 : (AABAC) → F
I,4 : (AABAC) → G
I,5 : (AABAC) → H

M.a.w. als de kenmerken (A) fasering en methodiek, (B) evaluatieplan en (C) groepsfunctioneren alle drie zijn gerealiseerd, dan voldoet het proces aan de criteria D en E en voldoet het evaluatieplan aan de criteria F, G en H.

Hypothese II is als volgt uiteengelegd. De 'ALS-kant' van elke voorspelling bestaat telkens uit de kenmerken I en J van de curvo-deelstrategie. De 'DAN-kant' van elke voorspelling bestaat telkens uit één effect: K, L, M, N of O.

De 5 voorspellingen die uit hypothese II zijn afgeleid kan men in formele taal als volgt samenvatten:

- Voorspelling II,1 : (IAJ) → K
II,2 : (IAJ) → L
II,3 : (IAJ) → M
II,4 : (IAJ) → N
II,5 : (IAJ) → O

M.a.w. als de kenmerken I (fasering en methodiek) en J (evaluatieplan) zijn gerealiseerd dan voldoet het proces aan het kenmerk K (oordeel groepsleden) en voldoen de resultaten of produkten aan de criteria L (oordeel groepsleden over het nut), M (oordeel evaluatiedeskundigen over rapport), N (beter produkt) en O (positieve houding tegenover evalueren).

We willen er op wijzen dat dus naast elkaar vragen en hypothesen (en voorspellingen) zijn geformuleerd. De vragen kunnen we wat genuanceerder beantwoorden, de hypothesen en voorspellingen vragen om een antwoord in termen van 'waar' of 'niet waar'.

5 Methoden, instrumenten en criteria

Voor het verzamelen van de gegevens hebben we gebruik gemaakt van een breed arsenaal van methoden, technieken en instrumenten. Hieronder geven we een overzicht waarbij per vraag wordt aangegeven welke methode van gegevensverzameling is gebruikt.

De criteria voor een positief antwoord hebben we via discussie in de onderzoeksgroep zoveel mogelijk vóór het proces van gegevensverzamelingen vastgesteld. Bij de hypothesen zijn we van één 'alles of niets criterium' uitgegaan. Een effect is positief of niet positief vanuit de voorspelling gezien. Bij de vraagstellingen daarentegen hebben we drie rubrieken gehanteerd: gerealiseerd of positief, niet gerealiseerd of negatief en gedeeltelijk gerealiseerd (positief noch negatief). Bij de discrepantie-analyse met betrekking tot de kenmerken A, B, I en J is gebruik gemaakt van 'checklists' waarmee de procesverslagen, evaluatieplannen en evaluatierapporten zijn geanalyseerd. Bij de vragenlijsten is vooraf als criterium vastgelegd dat de gemiddelde score significant in positieve richting moet afwijken van 3,5 (bij een vijf-puntschaal met 5 als positief schaalende). Bij de interviews met deskundigen is aan de deskundigen gevraagd om hun totaaloordeel in termen van voldoende/onvoldoende te geven. Hierbij legden we de criteriumbeslissing dus bij de deskundigen. Uiteraard werden wel uniforme dimensies aangereikt via de interviewvragen. Bij de interviews met groepsleden hebben we géén criteria vooraf geformuleerd. Evenmin hebben we gevraagd om een totaaloordeel te geven in termen van voldoende/onvoldoende (zoals bij de deskundigenbevraging). Bij de interviews met groepsleden hebben we zelf, achteraf, een voorzichtige conclusie getrokken over het eindoordeel van de groepsleden aan de hand van de uitgeschreven protocollen van de interviews. Bij de vragen rondom de efficiëntie hebben we vooraf géén criteria geformuleerd i.v.m. het exploratieve karakter van dit deel van het onderzoek. We

Vraag	Methoden/instrumenten
A, B, I en J getrouwe uitvoering strategie	Participerende observatie en analyse van proces- verslagen, evaluatieplannen en evaluatierapporten, d.w.z. discrepantie-analyse naar de realisering van kenmerken
C groepsfunctioneren conform strategie?	Analyse van vergaderprotocollen en vragenlijst over het groepsfunctioneren
D oordeel groepsleden over het proces	Interviews met groepsleden over het proces van uit- voeren van de strategie
E rationeel proces?	Interviews met groepsleden en analyse van vergader- protocollen met behulp van rationaliteitsschaal
F oordeel groepsleden over evaluatieplan	Vragenlijst voor groepsleden over het evaluatieplan
G en M oordeel deskundigen over produkten	Interviews met evaluatiedeskundigen over evaluatie- plannen en evaluatierapporten
H oordeel inhoudelijk deskundigen over evaluatieplan	Interviews met inhoudelijk deskundigen over evalua- tieplannen
K, L, N, en O oordeel groepsleden	Vragenlijst voor groepsleden over processen, resul- taten en produkten
P kosten van evalueren	Tijdschrijfformulier voor groepsleden
Q baten van evalueren	Diversen (zie o.a. effecten D t/m H en K t/m O)

Figuur 6 *Overzicht methoden en instrumenten*

volstaan hier met het weergeven van de kosten en baten. Tevens is geprobeerd een rangorde aan te brengen tussen de projecten m.b.t. de kosten en baten.

6 Resultaten, conclusies

Hieronder vatten we de resultaten samen voor de twee afzonderlijke deelprocessen: opstellen en uitvoeren van het evaluatieplan. Voor elk van deze deelprocessen hebben we één vraag

en één hypothese opgesteld m.b.t. de uitvoerbaarheid en effectiviteit.

Als antwoord op vraag I en II naar de uitvoerbaarheid en effectiviteit van de deelstrategie bij resp. het opstellen en het uitvoeren van het evaluatieplan trekken we de volgende conclusie: De deelstrategie is gedeeltelijk uitvoerbaar en gedeeltelijk effectief gebleken bij het opstellen en bij het uitvoeren van het evaluatieplan.

De hypothesen I en II zijn strikt genomen niet bevestigd. Met strikt bedoelen we dat de

Kenmerken en effecten	projecten				
	Curnil I	Curnil II	Curlant	Curtel	Cursol
VRAAG I EN HYPOTHESE I: OPSTELLEN EVALUATIEPLAN					
Kenmerk A: Fasering en methodiek	0	0	+	0	0
Kenmerk B: Evaluatieplan	-	+	+	-	-
Kenmerk C: Groepsfunctioneren	+	+	0	-	+
Effect D: Oordeel groepsleden over het proces	0	0	0	-	0
Effect E: Rationaliteit van het proces	0	0	0	-	0
Effect F: Oordeel groepsleden over het produkt (evaluatieplan)	+	+	0	0	0
Effect G: Oordeel evaluatiedeskundigen over het produkt (evaluatieplan)	0	+	0	0	-
Effect H: Oordeel inhoudelijk deskundigen over het produkt (evaluatieplan)	0	+	+	0	0
VRAAG II EN HYPOTHESE II: UITVOEREN EVALUATIEPLAN					
Kenmerk I: Fasering en methodiek	+	+	+	-	+
Kenmerk J: Evaluatierapport	+	+	+	0	0
Effect K: Oordeel groepsleden over het proces	0	0	0	0	0
Effect L: Oordeel groepsleden over de resultaten	0	0	0	0	0
Effect M: Oordeel evaluatiedeskundigen over het rapport	+	+	-	0	-
Effect N: Is er een beter produkt ontstaan?	0	0	0	0	0
Effect O: Houding groepsleden tegenover evalueren	+	+	0	0	+

- + = gerealiseerd (in enkele gevallen vrijwel geheel gerealiseerd)
 - = niet gerealiseerd of negatief volgens de gestelde criteria
 0 = gedeeltelijk gerealiseerd, positief noch negatief

Figuur 7 Samenvatting realisering kenmerken en effecten

hypothesen in *alle* effecten gerealiseerd moeten zijn. Wanneer we echter de afzonderlijke projecten en de afzonderlijke voorspelde effecten bekijken, kunnen we de volgende conclusies trekken.

Hypothese I kunnen we alleen toepassen op het curnil-II en het curlant-project, omdat alleen in deze twee projecten de handelingsaanwijzingen bij het opstellen (bijna) geheel zijn uitgevoerd. Van de 5 voorspellingen die we hebben afgeleid uit hypothese I gaan er 3 op in het curnil-II-project en 1 in het curlant-project (zie hiervoor de +-tekens in Figuur 7 bij de effecten).

Wanneer we *alle* 5 projecten bekijken dan blijkt de globale richting die in de hypothese wordt gesuggereerd overeen te komen met de bevindingen. Van een sterk rechtstreeks 'verband' tussen de mate van betrouwbaarheid van uit-

voering van de deelstrategie en de realisering van de verwachte effecten is echter geen sprake⁴.

Hypothese II kunnen we alleen toepassen op curnil-I en II, curlant en cursol, omdat in deze projecten het evaluatieplan en de daarbij behorende aanwijzingen uit de curvo-deelstrategie (bijna) geheel zijn uitgevoerd. Bij curnil-I en II gaan 2 van de 5 voorspellingen op. Bij curnil-II gaan bovendien 2 voorspellingen bijna op, d.w.z. de effecten L en N voldeden bijna aan het criterium. Bij cursol gaat één voorspelling op. Bij curlant gaat géén van de 5 voorspellingen op.

Uitvoerbaarheid

De mate waarin de deelstrategie *als geheel* volgens de bedoelingen is uitgevoerd verschilt per project (zie de kenmerken A, B, C, I en J). We

concluderen dat in het curlant-project en in de tweede ronde van het curnil-project de deelstrategie (vrijwel) geheel volgens de bedoelingen is uitgevoerd. Bij de projecten curnil-I en cursol is de deelstrategie gedeeltelijk uitgevoerd. Het curtel-project is voortijdig afgebroken. In dit project is de deelstrategie niet volgens de bedoelingen uitgevoerd. In *alle* projecten zijn aanpassingen, weglatingen en afwijkingen op meer of minder belangrijke onderdelen van de deelstrategie te constateren.

Wanneer we de projecten in een volgorde plaatsen naar de mate waarin de deelstrategie als geheel is uitgevoerd, krijgen we de volgende ordening:

1. curlant en curnil II; 2. curnil I; 3. cursol; 4. curtel.

Uit analyse van de procesverslagen, evaluatieplannen, evaluatierapporten en andere stukken uit de projecten blijkt dat de strategie een belangrijke steun was bij het moeilijke werk van evalueren. Er bestond veel waardering voor de consequentie en doordachte opzet van de deelstrategie. De indeling in A, B en C lijn bleek zeer zinvol. Allerlei deelprocedures (o.a. deskundigen raadplegen) bleken erg waardevol. Toch waren er ook veel problemen bij de uitvoering van de deelstrategie; we wijzen hiervoor naar ons eindrapport.

Effectiviteit

Ook de mate van effectrealisering voor het *proces als geheel* verschilt per project (zie effecten D t/m H en K t/m O). Van de in totaal 50 effecten, d.w.z. 5 evaluatieronden en per ronde 10 effecten, werden er 10 gerealiseerd, 5 effecten werden niet gerealiseerd, terwijl 35 effecten gedeeltelijk werden gerealiseerd.

De conclusie met betrekking tot het *proces* (effecten D, E en K) is in vrijwel alle projecten positief noch negatief. Alleen het curtel-project scoort negatief op twee van de drie proces-criteria. In de interviews (D) komen vrijwel dezelfde positieve en negatieve aspecten naar voren zoals die al bij de uitvoerbaarheidsvraag (discrepantie-analyse) werden vermeld. Het oordeel van de groepsleden over de *resultaten of het nut van evaluatie* (effecten L en N) is positief noch negatief te noemen. De *evaluatieplannen* (effecten F, G, H) voldeden meestal gedeeltelijk aan de criteria. D.w.z. naar het oordeel van groepsleden, evaluatiedeskundigen en inhoudelijk deskundigen. Eén project voldeed aan alle 3 criteria m.b.t. het evaluatie-

plan (curnil-II). In de overige 4 projecten zijn de verwachte effecten gedeeltelijk gerealiseerd. De evaluatieplannen vertoonden vaak tekorten op de punten: vraagstelling, instrumenten en het plan voor verwerking en analyse van de gegevens. De relatie tussen evaluatieplan en curriculuminhoud was verschillende keren problematisch, terwijl het proces van uitvoeren in de klas vaak te weinig aandacht kreeg in de evaluatieplannen. Soms was het evaluatieplan te pretentief (sociaal-wetenschappelijk) en te veel omvattend gelet op de omvang en de inhoud van de leergang (kanon op een mug). De deelstrategie blijkt te formeel, te weinig flexibel, te veelomvattend te zijn, sommige eisen zijn niet haalbaar of zinvol in situaties als deze.

De strategie legt grote nadruk op het anticiperen op discrepanties tussen vooraf geformuleerde eisen en de realisatie hiervan. In bijna alle projecten slaagde men er niet in om deze eisen en de drie typen vragen vooraf te formuleren.

Het gecombineerde oordeel van de evaluatie-deskundigen over de *evaluatierapporten* (effect M) leidt tot de volgende conclusie: van de 5 evaluatierapporten waren twee rapporten adequaat, twee rapporten waren niet adequaat en één rapport was gedeeltelijk adequaat. De rapporten voldeden vaak niet aan minimale eisen m.b.t. instrumenten, verwerking en analyse, conclusies en verbeteringssuggesties. Ook de vormgeving van de rapporten liet vaak te wensen over. In de rapporten wordt te weinig aandacht besteed aan het *proces* van uitvoeren in de klas en aan de *inhoud* van de leergang.

De houding van de groepsleden tegenover evalueren was in drie van de vijf gevallen positief. In de overige twee gevallen was de houding niet positief maar ook niet negatief.

Wanneer we de projecten in een rangorde plaatsen naar de mate van effectrealisering, krijgen we de volgende ordening: 6. curnil-II; 2. curnil-I; 3. curlant; 4. cursol; 5. curtel.

Efficiëntie

De efficiëntievraag hebben we uiteengelegd in drie deelaspecten: de kosten, de baten en de verhouding tussen kosten en baten.

De *kosten* van de evaluatie uitgedrukt in het totaal aantal werkuren verschilden sterk in de projecten. Het uitvoeren van één evaluatieronde kostte 140 uur in het minst kostbare project en 2600 uur in het meest kostbare pro-

ject. De spreiding van de kosten tussen de projecten wordt echter aanzienlijk teruggebracht wanneer men het omrekent naar onderwijsuren van de leergang. Per onderwijsuur kostte het evalueren gemiddeld ongeveer 10 uur. Ruw geschat komt het evalueren volgens de curvo-strategie, die immers minstens twee evaluatieronden voorschrijft, op 20 evaluatieuren per onderwijsuur. In Figuur 8 geven we een overzicht van de kosten per project.

projecten	totaal	per onderwijsuur
curnil-I	2600 u	7
curnil-II	1770	5
curlant	140	8
curtel	220	11
cursol	640	12

Figuur 8 *Kosten in uren besteed aan evaluatie*

De *baten* (voor een deel samenvallend met de effecten) zijn lang niet in alle gevallen opgetreden. Evaluatieplannen en evaluatierapporten voldeden niet altijd aan de eisen. Aan het nut van de evaluatiegegevens werd soms getwijfeld, terwijl de bijdrage van de evaluatiegegevens aan de verbetering van de leergang vrij gering is geweest. Uiteraard verschilden de projecten onderling ook wat betreft de realisering van de batens.

Wanneer we de projecten in een rangorde plaasten naar de verhouding tussen kosten per onderwijsuur en batens (*efficiëntie*), dan ontstaat het volgende rijtje: 1. curnil-II; 2. curnil-I; 3. curlant en curtel; 4. cursol. Het curnil-project is dan het meest efficiënt en het cursol-project het minst efficiënt.

Het evalueren volgens de curvo-strategie heeft veel tijd gekost. Deze tijd was vaak niet of nauwelijks op te brengen door de schoolteams waarmee is samengewerkt. De tijdsfactor was één van de oorzaken waarom curvo-medewerkers een relatief groot aandeel in de evaluatieactiviteiten hadden.

7 *Discussie en aanbevelingen*

Hieronder gaan we in op de betekenis van onze bevindingen. Eerst maken we enkele algemene opmerkingen over de betekenis met het oog op de aard van dit onderzoek. Vervolgens gaan we in op de vraag naar de opbrengst van dit onderzoek vanuit de drie gezichtspunten die we in par. 3 al noemden, nl.: a) verbetering, b) in-

formatie aan toekomstige ontwikkelingsgroepen en c) theorievorming.

Methodologische opmerkingen vooraf

De vraag naar de betekenis van de resultaten heeft o.a. te maken met aspecten als betrouwbaarheid, validiteit en generaliseerbaarheid. Wanneer we deze aspecten als normen opvatten in klassieke, methodologische zin dan voldoet ons onderzoek daaraan niet. We hebben vrijwel nergens nagegaan of onze instrumenten *betrouwbaar* waren. Strikt genomen is ook de *validiteit* discutabel: hebben we wel echt gemeten wat we wilden meten, nl. de invloed van de strategie in het evaluatieproces? Zijn we er wel in geslaagd om de 'strategie-factor' te isoleren uit het geheel van variabelen? Slaan onze bevindingen wellicht ook terug op de groepen, de bijdragen van curvo-mensen in het proces, de aard van de leergangen of de contexten waarbinnen is gewerkt? Waarschijnlijk wel! De 'strategie-factor' is slechts één van de vele variabelen in een zeer complex proces. De *generaliseerbaarheid* van onze bevindingen is strikt genomen ook problematisch. We hebben 'slechts' vier cases onderzocht. We kunnen niet 'hard maken' dat in een vijfde geval de resultaten hetzelfde zullen zijn. In een bepaalde fase van het curvo-onderzoek zijn twee hypothesen geformuleerd over de samenhang tussen de uitvoering van de deelstrategie en bepaalde effecten. Daarmee gaf de toen aanwezige curvo-groep aan dat men de curvo-deelstrategie als een *voldoende* voorwaarde opvatte voor het optreden van bepaalde effecten. Strikt genomen mag je dan achteraf, als de effecten niet in de verwachte mate zijn opgetreden, niet verwijzen naar factoren buiten de strategie.

Tegen de achtergrond van de onderzoeksresultaten en vooral de ervaringen in de projecten lijkt het echter niet zinvol deze strikte denkwijze te handhaven. Leergangevaluatie blijkt een zeer complex proces te zijn waarin strategiefactoren, personen (posities en rollen), situaties, waarde-oriëntaties in interactie optreden. Het denkkader achter de hypothesen is te simpel voor deze complexe problematiek.

De problemen van betrouwbaarheid, validiteit en generaliseerbaarheid hangen rechtstreeks samen met de aard van dit onderzoek. Hoe men het ook bekijkt, we hebben te maken met slechts vier cases met vaak heel kleine aantallen respondenten bij interviews en vragenlijsten. Ook het vergelijken tussen projec-

ten is niet zonder problemen. Bepaalde gebeurtenissen in een project, die geheel buiten de strategie kunnen liggen, kunnen verantwoordelijk zijn voor een lagere score, een lagere waardering op het moment van afname.

De bepaling van de *criteria* is ook voor discussie vatbaar. We hebben zoveel mogelijk geprobeerd om de criteria vóór het beschikbaar komen van de gegevens vast te leggen. Zijn we te streng geweest bij de bepaling van deze criteria? Als we bijvoorbeeld al die scores die wij op grond van onze criteria het predicaat 'gedeeltelijk gerealiseerd' meegaven als positieve scores opvatten, ontstaat er een veel gunstiger plaatje. M.a.w. de hoogte van het criterium is van grote invloed op de conclusies. Afdoende methoden voor het bepalen van criteria in situaties als deze zijn ons niet bekend. We hebben de meeste criteria vooraf vastgesteld via discussie totdat we consensus hadden bereikt. In een aantal gevallen (bijvoorbeeld bij de interviews met groepsleden, effect D) hebben we zelf *achteraf* vastgesteld of het proces van uitvoering van de strategie positief of negatief werd gewaardeerd. Terugblikkend zou het beter geweest zijn in dit bepaalde geval aan de groepsleden een totaal-oordeel te vragen in termen van positief of negatief. Bij de externe deskundigen (effect G) hebben we dat *wel* gedaan. D.w.z. de deskundigen hebben we heel nadrukkelijk gevraagd hun eendoordeel over de evaluatieplannen in de vorm van een cijfer of voldoende/onvoldoende kwalificatie te geven.

Samenvattend kunnen we stellen dat vanuit 'klassieke', strikte, methodologische criteria gezien ons onderzoek enkele discutabele punten bevat. Vanuit dit standpunt bezien kan de betekenis van ons onderzoek worden gerelativeerd.

Bij deze constatering hoeven we het echter niet te laten. We hebben namelijk langs andere dan statistische weg getracht om betrouwbaarheid, validiteit en tot op zekere hoogte ook generaliseerbaarheid *op te bouwen*. We hebben dit trachten te realiseren door de deelstrategie in sterk verschillende situaties toe te passen en door een breed arsenaal van methoden en technieken te gebruiken. De betrouwbaarheid en geldigheid van onze conclusies is gaandeweg 'opgebouwd' in telkens nieuwe situaties, via andere mensen en verschillende methoden en technieken. De procesverslagen, verkregen via intensieve participatie in de projecten, vormen

a.h.w. het landschap waartegen deze conclusies reliëf krijgen. Uiteraard moeten de conclusies steeds in relatie met de vier cases worden bezien (zie par. 2).

De opbrengst van het onderzoek

De gegevens uit ons onderzoek bieden tal van aanknopingspunten voor verbetering van de deelstrategie en bevatten een schat aan gegevens voor toekomstige gebruikers. We verwijzen hiervoor naar het eindrapport, Beumer, Schuit, Terwel, 1982, naar het artikel van J. Terwel, 1982 en naar de S.V.O.-brochure 'Een leergang maken', 1982⁵.

Naast deze praktische opbrengsten geeft ons onderzoek aanleiding tot enkele theoretische reflecties. Twee hoofdaspecten willen we kort aanstippen. Allereerst het aspect van de evaluatietheorie. Het gaat hierbij om twee dilemma's of dimensies in de literatuur rondom curriculumevaluatie en -implementatie:

- a) algemene, formele versus inhoudsspecifieke, contextgebonden strategie,
- b) de perspectieven van getrouwheid en wederzijdse aanpassing.

De curvo-(deel)strategie kan enigszins schematiserend worden getypeerd als een algemene, formele strategie waarbij het accent ligt op een implementatie-opvatting die als 'getrouwheids-perspectief' kan worden getypeerd.

Bij het ontwikkelen van een 'evaluatie-theorie' voor situaties waarbij leerkrachten een centrale rol spelen, zal een optimale positie moeten worden gekozen op de twee dimensies a en b. Ons onderzoek kan een bijdrage leveren aan de (impliciete) evaluatie-theorie die aan de curvo-deelstrategie ten grondslag ligt. Deze bijdrage zou hierin kunnen bestaan dat met behoud van de waardevolle elementen uit de curvo-opvatting een accentverlegging plaatsvindt die rekening houdt met recente ontwikkelingen in de literatuur over curriculum-evaluatie, -implementatie en -innovatie. We denken hierbij aan opvattingen van o.a. Wolf (1979), House (1977), Moser (1980), Fullan (1979) en Morocco (1979 en 1980). Vooral in situaties waarbij 'Teachers as Evaluators' optreden zou meer aandacht moeten komen voor besluitvorming, groepsfunctioneren, feedback procedures en discussiemethoden.

Een ander aspect van de 'theoretische meeropbrengst' van ons onderzoek is een methodologische reflectie. In ons onderzoek is gewerkt met een combinatie van participerende

en meer afstandelijke methoden. In het bijzonder het participerende aspect, met name de rol van de onderzoeker in de projecten, biedt aanknopingspunten voor discussie op methodologisch gebied. Enerzijds kan worden gezegd dat dit onderzoek niet mogelijk zou zijn geweest zonder directe, intensieve participatie van onderzoekers in de ontwikkelingsgroepen. De 'eerste hands' ervaringen en onderzoeksgegevens die langs deze weg zijn verkregen zijn van grote betekenis geweest bij de beantwoording van de onderzoeksvragen. Tegelijk vormt de bijdrage van de curvo-medewerkers in de projecten een factor die vanuit een bepaalde opvatting over onderzoek problematisch kan zijn. Deze bijdrage betekent de inbreng van een 'actie-component' in het proces⁶. De curvo-medewerkers probeerden (uiteraard) mee te werken aan het goed verlopen van het proces en aan het ontwikkelen van goede produkten (evaluatieplannen en evaluatierapporten). Mede door hun inbreng zijn effecten gerealiseerd die anders misschien niet zouden zijn opgetreden. De toeschrijving van de effecten aan de 'onafhankelijke variabele' (de strategie) wordt hierdoor problematisch. Het isoleren van de strategiefactor in dit complexe interactieproces van personen, situaties en rollen blijkt een hachelijke onderneming. Toch vormt dit denken in termen van onafhankelijke en afhankelijke variabelen de achtergrond van onze hypothesen. Dit denkkader eist de getrouwe uitvoering van de strategie als onafhankelijke variabele. Wanneer deze getrouwe uitvoering niet of niet geheel kan worden gerealiseerd valt over toeschrijving van de effecten strikt genomen niets meer te zeggen.

Deze getrouwe uitvoering werd, mede dankzij de sturing door de curvo-medewerkers, slechts in twee projecten (vrijwel geheel) gerealiseerd. Hoewel het denken in termen van hypothesen wel verhelderend is geweest moeten we constateren dat het hypothesen toetsen in situaties als deze op grote problemen stuit.

In het kader van dit artikel kunnen we hierop niet nader ingaan. We verwijzen naar het verslag van de curvo-studiedag op 23 februari 1982. Voorts verwijzen we naar de dissertatie van Hoeben 1981, waarin de methodologische problematiek centraal staat.

Slotopmerkingen

In de twee artikelen over formatieve evaluatie in Pedagogische Studiën hebben we veel aan-

dacht besteed aan punten die verbetering behoeven in de deelstrategie. Toch is onze totaalindruk allerm minst negatief. De deelstrategie blijkt een heldere, consequente opbouw te bezitten. De strategie bevat vele waardevolle deelprocedures voor leergangevaluatie. Er is ons geen strategie bekend die zulke gedetailleerde richtlijnen geeft voor de planning van een leergangevaluatie. Voor zover wij kunnen nagaan bestaat er geen strategie voor leergangevaluatie die zó uitgebreid is beproefd in de praktijk. De door ons gesignaleerde problemen zijn voor een niet onbelangrijk deel praktijkproblemen, die niet alléén via een verbetering van de strategie kunnen worden opgelost. Wanneer men leerkrachten een wezenlijke bijdrage wil laten leveren aan de vormgeving van het onderwijs door middel van leergangontwikkeling en evaluatie dan zal ook aan verbetering van de condities moeten worden gewerkt.

Noten

1. De curvo-strategie geeft aanwijzingen voor het ontwikkelen van leergangen. Een deel van deze aanwijzingen betreffen de formatieve evaluatie als onderdeel van het ontwikkelingswerk. In dit artikel beperken we ons tot dit deel van de curvo-strategie.
2. Onder een prikactie verstaat de projectgroep NOB, het opzetten van een plan voor een aantal lessen in samenwerking met leerkrachten, het uitvoeren van die lessen, het observeren en tot slot het evalueren ervan.
3. Overigens valt de uitvoerbaarheidsvraag niet helemaal samen met de vraag of de strategie volgens de bedoelingen is uitgevoerd. Het 'waarom' van het al of niet uitgevoerd zijn en de condities waaronder dienen hierbij te worden betrokken. In de procesverslagen en de discrepantie-analyse is hiermee rekening gehouden.
4. Overigens mogen we strikt genomen niet over een verband spreken. Onze deterministische, criteriumgerichte aanpak bij het formuleren van hypothesen laat geen uitspraken toe over alle 5 projecten.
5. Onze praktijkervaringen met de curvo-strategie als geheel zijn in een vlot leesbare S.V.O.-brochure samengevat. De brochure heet: *Een leergang maken*. De curvo-strategie in de praktijk. Den Haag, S.V.O. februari 1982.
6. Hoe complex de zaak ligt blijkt als we naar de verschillende rollen van de curvo-participanten kijken. Ze waren:
 - a. participierend observerator ten behoeve van het curvo-onderzoek;

- b. bewaker van de curvo-strategie – bij afwijken van de strategie moest er gereageerd worden;
- c. uitlegger van de curvo-strategie, als de strategie niet duidelijk was;
- d. leverancier van deskundigheid, die niet altijd in de projectteams aanwezig was;
- e. leverancier van arbeidstijd, door taken te verrichten waar de projectteams geen tijd voor hadden.

Literatuur

- Berg, G. v.d., J. Terwel (red.), *Eindevaluatierapport. Formatieve evaluatie van interne differentiatie*. Utrecht, Vakgroep Onderwijskunde, IPAW, 1980.
- Beumer, T. en J. Terwel, *Formatieve evaluatie van interne differentiatie volgens het BHE-model, curnil-project, tweede evaluatieronde*. Utrecht, Vakgroep Onderwijskunde, IPAW, maart 1981. (Curvo-procesverslag)
- Beumer, T., W. Schuut, J. Terwel, *Evalueren met de curvo-strategie*. Deel-eindrapport curvo-2 onderzoek. Utrecht, Vakgroep Onderwijskunde, IPAW, 1982.
- Bjorkman, C., T. Beumer en J. Terwel, *Interne differentiatie in heterogene brugklassen*. *School*, 1981, 9, nr. 6.
- Fullan, M., *The relationship between evaluation and implementation in curriculum*. Ontario Institute for Studies in Education, 1979.
- Glas, M., M. Y. de Kok-Damave, K. Treffers, H. v.d. Wielen, *De ontwikkeling van een cursus 'Besluitvorming' voor 4e jaars studenten van een SOL*. Utrecht, Vakgroep Onderwijskunde, IPAW, juli 1981. (Curvo-procesverslag)
- Groot, A. D. de, *Methodologie*, 's-Gravenhage: Mouton & Co., 1972.
- Herfs, P., *Ontwikkeling van een lessenserie op het gebied van de Mechanische Techniek*. Utrecht, Vakgroep Onderwijskunde, IPAW, augustus 1981. (Curvo-procesverslag)
- Herfs, P. en K. Treffers, *Ontwikkeling van een leergang voor natuuronderwijs*. Utrecht, Vakgroep Onderwijskunde, IPAW, oktober 1981. (Curvo-procesverslag)
- Verslag curvo-studiedag van 23 februari 1982*, Utrecht, Vakgroep Onderwijskunde, IPAW, 1982.
- Hoeben, W., e.a., *Onderzoek naar curriculumontwikkeling*. Onderzoeksplan van het curvo-2-project. Utrecht, Vakgroep Onderwijskunde, IPAW, 1978.
- Hoeben, W., *Praktijkgericht onderzoek en de groei van kennis*. 's-Gravenhage: SVÖ-reeks 46, 1981. (Dissertatie)
- House, E. R., *The logic of evaluative arguments*. Center for the Study of Evaluation, University of California. C.S.E. monograph series in evaluation, Los Angeles: 1977.
- Morocco, Catherine Cobb, *The Role of Formative Evaluation in Developing and Assessing Educational Program*. *Curriculum Inquiry*, 1979, 9:2, 137-148.
- Morocco, Catherine Cobb, *A 'Working Assumptions' Approach to Early Program Evaluation*. Paper presented at AERA Annual Meeting Boston, april 1980.
- Moser, H., *Evaluation als Handlungsforschung*. In: W. v.d. Grift & W. Stelling, *Verslag van het symposium 'Evaluatie van onderwijsvernieuwingen*. Amersfoort, 3-4 juni 1980.
- Reints, A. (red.), *Naar een interne differentiatie voor het vak mechanische techniek, Eindrapport van het curlant-project*. Utrecht, Vakgroep Onderwijskunde, IPAW, 1981.
- Terwel, J., *Formatieve evaluatie van interne differentiatie volgens het BHE-model, curnil-project, eerste evaluatieronde*. Utrecht, Vakgroep Onderwijskunde, IPAW, maart 1980. (Curvo-procesverslag)
- Terwel, J., *Formatieve evaluatie in de curvo-strategie*. *Pedagogische Studiën*, 1982, 59, 175-193.
- Wolf, R. L., *The use of Judicial Evaluation: methods in the formulation of Educational policy*. *Educational Evaluation and policy analysis*, mei/juni, vol. no. 3, 1979.

Curricula vitae

J. Terwel is vanaf 1963 werkzaam geweest in verschillende vormen van onderwijs; o.a. LBO en leraaropleidingen. Studeerde onderwijskunde in Groningen. Hij werkt sinds 1976 als wetenschappelijk medewerker bij de vakgroep onderwijskunde van de R.U. te Utrecht; is vooral werkzaam op het gebied van curriculumontwikkeling en evaluatie en de opleiding van leraren, participeerde van 1976-1982 in het curvo-project. Was hoofdverantwoordelijke voor dit deel van het curvo-onderzoek.

W. Schuut studeerde economie en kwam in het onderwijs terecht. De praktijk werd aangevuld met een studie onderwijskunde. Na zijn leraarschap in het VWO was hij als onderwijskundige werkzaam bij een organisatieadviesbureau, bij de Technische Hogeschool Twente en bij de Rijksuniversiteit Utrecht. Participeerde van 1981-1982 in het curvo-project.

T. Beumer, doctoraal student in de westerse sociologie. Part-time werkzaam in de onderwijsresearch (Rijksuniversiteit Utrecht en CITO Arnhem). Participeerde van 1978-1982 in het curvo-project.

Adres: Vakgroep Onderwijskunde Rijksuniversiteit Utrecht, Heidelberglaan 1, 3584 CS-Utrecht

Manuscript aanvaard 15-6-'82