

Wat leren kinderen op school? (II)*

Verslag van een literatuurstudie en een studiereis ter verkenning van de mogelijkheden van een nationaal evaluatieproject in het funderend onderwijs

J. M. WIJNSTRA

Centraal Instituut voor Toetsontwikkeling
(Cito), Arnhem

Samenvatting

De laatste jaren is in algemeen-maatschappelijke, politieke en onderwijskundige discussies herhaaldelijk de vraag gesteld: Wat leren kinderen nu eigenlijk op school? (vgl. bijv. de recente nota van de Bewindslieden van Onderwijs en Wetenschappen, getiteld *Kwaliteit van het onderwijs*). In zijn algemeenheid betreft deze vraag de kwaliteit en het niveau van het onderwijs, die echter op verschillende manieren kan worden uitgelegd. Enerzijds wordt gevraagd naar de geldende, respectievelijk de feitelijke nagestreefde doelstellingen en het daarop afgestemde vormingsaanbod, anderzijds wordt gevraagd wat de leerlingen van het genoten onderwijs hebben opgestoken.

Het Cito heeft aan dit onderwerp in 1979 en 1980 een voorstudie gewijd waarvan in dit artikel (in twee delen) verslag wordt gedaan. Naast literatuurstudie is ook een studiereis naar de Verenigde Staten gemaakt om nader kennis te maken met het project *National Assessment of Educational Progress (NAEP)* en enkele state assessment-programs. Deze projecten zijn bedoeld om informatie te leveren over wat kinderen kunnen en weten, en het vaststellen van veranderingen daarin, ten dienste van het beleid op verschillende niveaus.

In het eerste deel van dit artikel is, na een algemene inleiding, de achtergrond en opzet van NAEP beschreven: de doelstellingen, de instrumentontwikkeling, de steekproefrekening en afname, de verwerking en de rapportage van de gegevens. Uitvoerig is vervolgens ingegaan op een tweetal externe evaluaties van het project:

een onderzoek van de Carnegie Corporation, die in de beginjaren veel geld in het project heeft gestoken, en een onderzoek van de Amerikaanse rekenkamer. In beide onderzoeken wordt getwijfeld of de oorspronkelijke doelstellingen van het project, het vormen van een informatiebron voor het identificeren van problemen, het stellen van prioriteiten en het vaststellen van veranderingen ten behoeve van beleidsinstanties, curriculumontwikkelaars, onderzoekers en practici, voldoende gerealiseerd worden. Tenslotte is aandacht geschonken aan veranderingen die de laatste jaren hebben plaatsgevonden, zowel naar aanleiding van de kritiek als naar aanleiding van budgetbeperkingen.

In dit tweede deel van het artikel wordt aandacht geschonken aan de ontwikkelingen in de afzonderlijke staten. In 42 staten lopen op het ogenblik assessment-programs. Er is echter inmiddels een tendens om deze projecten, onder invloed van de 'minimum competency movement', om te buigen in de richting van beoordeling (en certificering) van individuele leerlingen, in plaats van evaluatie van het onderwijs als zodanig, zoals oorspronkelijk de opzet was. In par. 3 wordt verder het state assessment-program in Minnesota wat uitvoeriger beschreven. Dit project bestaat uit twee delen: het 'statewide assessment-program' en de 'local assessment option', waarop schooldistricten kunnen inschrijven om het onderwijs in hun district te evalueren. Deze 'local assessment option' lijkt een zeer effectief middel te zijn om de vorm en inhoud van het onderwijs ter discussie te stellen en om naar aanleiding daarvan tot veranderingen te besluiten.

De slotparagraaf bevat een aantal overwegingen m.b.t. een educational assessment-project in Nederland. Een dergelijk project wordt geplaatst in het kader van (de beleidsbepaling m.b.t.) de basisvorming in het funderend onderwijs. Er worden twee hoofdfuncties aangegeven: nagaan of het onderwijs aan haar doel beantwoordt en het leveren van informatie ten behoeve van de (planning van) curriculumontwikkeling. Voorts worden enkele aantekeningen gemaakt bij een mogelijke projectopzet: de inventarisatie van vormingsdoelen, het vormingsaanbod en de

* Het eerste deel van dit artikel is geplaatst in *Pedagogische Studiën*, 1982, 59, 223-232.

peiling van de beheersing bij de leerlingen aan het einde van het funderend onderwijs. Hierbij wordt ook aandacht geschonken aan de koppeling van dieptestudies en lokale of regionale vervolgmogelijkheden aan een dergelijk project.

3 State assessment-programs

3.1 Inleiding

Zoals in de vorige paragraaf werd beschreven past NAEP in de 'accountability movement' die in de loop van de zestiger jaren ontstond. Ook in de 50 staten met hun ruim 12000 schooldistricten heeft deze beweging zijn uitwerking niet gemist. Accountability is in zeker opzicht in de afzonderlijke staten en schooldistricten zelfs belangrijker dan op federaal niveau. De onderwijswetgeving is per staat geregeld – met grote verschillen wat de controle op het onderwijs betreft – en de kosten van het onderwijs worden grotendeels door de staat en/of de schooldistricten gedragen. Aan het eind van de zestiger en het begin van de zeventiger jaren heeft dit tot gevolg gehad dat een groot aantal staten assessment-programs heeft ingesteld. In staten met een sterke overheidscontrole werd dit meestal bij wet geregeld, in staten met een grotere mate van autonomie voor de schooldistricten als een service van het State Department of Education aan de districten.

Tot op zekere hoogte was dit geen nieuw verschijnsel, in die zin dat massale testprogramma's, georganiseerd door de staat, niet onbekend waren. Dyer & Rosenthal (1971) melden bijv. dat in 1967 in 42 staten 74 testprogramma's draaiden. In de meeste gevallen maakte men echter gebruik van gestandaardiseerde toetsen, ten behoeve van 'guidance', 'placement' of selectie van individuele leerlingen. Vanuit het vrij vage accountability-begrip ontstond echter behoefte aan evaluatie van het onderwijs als zodanig om de effectiviteit en efficiëntie vast te stellen. Hoewel het accountability-begrip veel ruimer is dan het op grote schaal afnemen van toetsen¹, is het afnemen van toetsen in veel gevallen toch wel de kern geworden van de accountability-activiteiten. Gestimuleerd o.a. door NAEP, zijn de toetsprogramma's echter wel anders van opzet geworden dan de oude programma's. Hoewel niet overal toegepast is matrix-steekproeftrekking een veel voorkomend kenmerk, naast het

gebruik van leerdoelgerichte toetsen, al of niet ontleend aan NAEP. Kauffman (1979) maakt melding van assessment-programs in 42 Amerikaanse staten.

In hoeverre in deze assessment-programs de 'minimum competency'-gedachte is doorgedrongen, valt uit zijn verslag niet op te maken. Geïnspireerd door de 'declining test scores' en de 'back to basics movement' is de laatste vier à vijf jaar de 'minimum competency testing movement' ontstaan, die in een aantal gevallen de oorspronkelijke accountability-gedachte verdringt. Minimum competency tests zijn leerdoel-, resp. domeingerichte toetsen die afgenomen worden om in de eerste plaats het niveau van de afzonderlijke leerlingen vast te stellen, om daar remediatie en eventueel sancties aan te verbinden als blijven zitten en het niet uitreiken van het high schooldiploma (zie bijv. Piphoo, 1978; Miller, 1978). Deze nieuwe movement, die zich aanvankelijk vooral op de senior high school richtte, heeft in enkele jaren een geweldige groei gehad. Piphoo (1978) meldt dat er al in 33 staten wetgevende maatregelen zijn genomen.

Een goed voorbeeld van deze ontwikkeling vinden we in de staat Florida. Op grond van een Accountability-act werden in deze staat tussen 1971 en 1974 assessment-programs uitgevoerd in de klassen 3 en 5 op steekproefbasis voor de vakken reading, writing en mathematics. In 1974 werd echter de steekproefbasis verlaten. In 1976 werd een nieuwe Educational Accountability Act aangenomen die belangrijke wijzigingen tot gevolg had. Vanaf 1977 werden ook alle klassen 8 en 11 in de 'basic skills assessment' betrokken waarvoor 'minimum performance standards' werden vastgesteld door het State Department of Education. Tevens werd voor klas 11 een 'functional literacy test' verplicht gesteld, bedoeld om de toepassingsvaardigheid met betrekking tot de basic skills te meten. Wie de standards van de functional literacy test, bestaande uit een communications en mathematics deel, ook na meerdere afnames, niet haalt, krijgt bij het verlaten van de high school geen diploma, maar een Certificate of Completion².

Dit minimum high school-examen heeft in Florida en de rest van de Verenigde Staten veel kritiek gekregen, en wel van twee kanten. Bedoeld om een minimaal niveau van het diploma te garanderen, lopen er nog een aantal gerechtelijke procedures waarin het onderwijs wordt

beschuldigd tekort te zijn geschoten in het onderwijzen van de basic skills. Van andere zijde is op meerdere plaatsen de arbitraire wijze van 'standard setting' door het State Department of Education bekritiseerd (vgl. o.a. Glass, 1978a; 1978b).

Op aanraden van NAEP hebben we de state assessment-programs van de staten Connecticut, Minnesota en Rhode Island bezocht. Alle drie de bezochte assessment-programs komen qua opzet ten dele overeen met NAEP en maken ten dele ook gebruik van NAEP-materiaal. Verder kennen alle drie een zogenaamde 'local assessment option'. In Connecticut kunnen schooldistricten in zijn geheel inschrijven voor de lopende assessment-cyclus met een rapportage op het niveau van de afzonderlijke school en het district als geheel. In Minnesota bestaat deze mogelijkheid ook, maar daarnaast kan ook materiaal, met rapportageservice, van vorige ronden worden besteld. In Rhode Island is naast het 'statewide' assessment-program, dat net als in de andere staten op steekproefbasis werkt, een itembank ontwikkeld waaruit scholen of schooldistricten toetsen op maat kunnen laten samenstellen. Om redenen van plaatsruimte wordt in dit artikel alleen het assessment-program van de staat Minnesota wat uitvoeriger besproken³.

3.2 Minnesota Educational Assessment-Program

Het assessment-program in de staat Minnesota bestaat uit twee componenten: de 'statewide assessment' (op steekproefbasis) en de 'local assessment option', ook wel aangeduid als 'Piggyback program'. De 'statewide assessment' wordt door Allen & Sushak (1979) omschreven als een middel om op basis van de prestaties van de leerlingen de stand van het onderwijs in de staat te analyseren en om aanbevelingen te doen over gewenste ontwikkelingen. De 'local assessment option' is een middel om dit proces op het niveau van het schooldistrict te herhalen. Nadrukkelijk worden de volgende functies uitgesloten:

- onderlinge vergelijking van schooldistricten;
- beoordeling van docenten;
- toewijzing van faciliteiten aan districten op basis van assessment-resultaten.

In tegenstelling tot vele andere state assessment-programs is het in Minnesota niet voortgekomen uit wettelijke maatregelen, maar geïnitieerd door het State Department of Educa-

tion als onderdeel van het beleidsinstrumentarium van de staat, mede ook ten behoeve van de schooldistricten. Het assessment-program wordt uitgevoerd door een sectie van de Division of special services⁴, in samenwerking met de Division of instruction.

Het assessment-program in Minnesota is in 1973/74 gestart met reading. Nadat vervolgens in 1974/75 mathematics aan de orde was geweest is de continuïteit een poos in gevaar geweest. De projectleider vertrok en in de schooljaren 1975/76 en 1976/77 zijn er geen veldactiviteiten geweest. Vanaf 1977/78 geldt het volgende programma:

1977/78: reading, social studies;

1978/79: mathematics, science;

1979/80: writing, mathematic numeracy;

1980/81: humanities, reading literacy;

1981/82: reading, social studies;

1982/83: mathematics, science.

Hieruit blijkt dat de kern bestaat uit reading, mathematics, social studies en science.

Inmiddels past het assessment-program, en met name de local option, goed in de onderwijswetgeving van de staat Minnesota. In 1976 werd namelijk een wet van kracht waarin de schooldistricten verplicht werden ieder jaar hun schoolwerkplan bij te stellen op basis van evaluatie en hierover te rapporteren. De local option van het assessment-program kan hiertoe goed gebruikt worden. Het State Department of Education en de Educational Cooperative Service Units (vergelijkbaar met onderwijsbegeleidingsdiensten) kunnen hierbij technische assistentie verlenen.

Instrumentontwikkeling

Aan de instrumentontwikkeling voor een assessment-cyclus ligt een doelstellingenlijst ten grondslag die door de Division of instruction in samenwerking met een Objectives task force en een Review committee wordt samengesteld. De leden van de Objectives task force en het Review committee kunnen zowel vakspecialisten als practici zijn. De doelstellingenlijsten worden onder de titel Some Essential Learner Outcomes in (SELO) gepubliceerd. De SELO's worden soms in een zodanige vorm gegoten dat ze ook zonder meer in een workshop op plaatselijk niveau gebruikt kunnen worden om het gehele proces te herhalen. De eigenlijke doelstellingenlijst wordt geordend naar domein, doelstelling en subdoelstelling. De doelstellingen en subdoelstellingen vormen

naderhand de primaire rapportage-eenheden.

De eigenlijke itemconstructie (meerkeuze-vorm) vindt op basis van de SELO's plaats door commissies bestaande uit vakspecialisten. Om de vergelijking met nationale en regionale gegevens van NAEP mogelijk te maken, wordt ongeveer het derde deel van de opgaven uit de NAEP-voorraad geselecteerd. Ook bij de itemconstructie en -selectie worden de Review committees weer ingeschakeld voor de screening van de opgaven. Het aantal betrokken personen kan vooral bij 'moeilijker' vakken zoals social studies zeer groot zijn. De projectleider noemde voor social studies een aantal van 500; gebruikelijk is echter een aantal van 50. Voordat de opgaven definitief worden opgenomen, vindt er een proeftoetsing plaats.

Steekproeftrekking en toetsafname

In het Minnesota Educational Assessment-Program wordt evenals bij NAEP, gebruik gemaakt van matrix sampling: een deel van de leerlingen maakt slechts een deel van de opgaven. De opgaven worden zodanig in afnamepakketten geordend dat ze in ongeveer 45 minuten gemaakt kunnen worden. Met uitzondering van lezen worden alle afnamepakketten voorzien van geluidsbanden, waardoor de instructies en de opgaven ook ten gehore kunnen worden gebracht om interferentie van de leesvaardigheid zo veel mogelijk uit te sluiten.

De primaire steekproefeenheid is de klas (4, 8 en 11). De steekproef wordt gestratificeerd naar de grootte van het schooldistrict getrokken. Eventuele parallelklassen binnen dezelfde school in de steekproef krijgen hetzelfde pakket opgaven om – als tegenprestatie – een rapportage op schoolniveau te kunnen leveren. In de toekomst wil men per district deze regel aanhouden, om ook op districtsniveau te kunnen rapporteren.

De toetsafnames vinden plaats onder leiding van toetsassistenten. In de toekomst zal dit waarschijnlijk door de leerkrachten of districtspersoneel gaan gebeuren, voornamelijk om financiële redenen. De leerlingen noteren hun antwoorden op optisch leesbare formulieren.

Rapportage

Een belangrijk element in de rapportage is de vaststelling van normen. Dit gebeurt vooraf door een commissie van vakspecialisten. Per

opgave geven ze een range aan van p-waarden die geldt als acceptabel. Vervolgens worden de resultaten gerapporteerd op een vijfpuntschaal: strength, potential strength, acceptable, potential need, need. Bij de interpretatie van de data, het opstellen van de rapporten en het formuleren van de aanbevelingen worden weer dezelfde Review committees ingeschakeld. Met name scores in de categorie 'need' krijgen bijzondere aandacht.

Zijn de rapporten gereed, dan volgt een fase van verspreiding van de gegevens in het onderwijs. De volgende activiteiten worden in dit kader ontplooid:

- aanbieding van de rapporten aan de Minnesota State Board of Education;
- presentatie aan de pers;
- organisatie van workshops, samen met andere organisaties, om bekendheid te geven aan de resultaten en om te discussiëren over de 'zwakke plekken' in het onderwijs met het oog op mogelijke verbeteringen.

De local assessment option (Piggyback program)

De local assessment option is een mogelijkheid voor een schooldistrict om een meting van de onderwijsopbrengst in alle scholen van het district uit te laten voeren. Hierbij wordt gebruik gemaakt van materialen en procedures die gebruikt zijn bij een eerdere of de lopende statewide assessment. Meer dan éénderde van alle schooldistricten maakt tegenwoordig gebruik van deze mogelijkheid.

De werkzaamheden worden in hoofdzaak uitgevoerd door medewerkers van het schooldistrict. Dit houdt in:

- het analyseren van de geldende leerdoelen en het beoordelen van het beschikbare toetsmateriaal;
- het vastleggen van grenzen voor acceptabele resultaten;
- het analyseren van de resultaten;
- het opstellen van rapporten met daarin aanbevelingen en plannen.

Het State Department verzorgt het toetsmateriaal, de verwerking en de rapportage tegen kostprijs: \$ 1.00 per leerling. Verder verzorgen zij een intensieve voorlichting aan de districten door middel van workshops. De rapportage geeft per school van het district de resultaten weer per opgave en subobjective, in combinatie met een foutenanalyse. Verder worden ook de corresponderende resultaten van de statewide

assessment vermeld.

In de zomer van 1978 werd begonnen met een evaluatie-onderzoek door de Evaluation Unit (1979) van het Office of Planning and Evaluation om na te gaan welke invloed de local assessment option heeft op het onderwijs in Minnesota en in welke richting veranderingen gewenst zijn. Het onderzoek bestond uit:

- een schriftelijke enquête onder de hoofden van scholen in niet deelnemende districten;
- een schriftelijke enquête onder de betrokken personeelsleden in de deelnemende districten;
- een telefonische en een schriftelijke enquête onder de centrale contactpersonen in de deelnemende districten.

De vragen aan de niet deelnemende districten betroffen het op de hoogte zijn met de local assessment option, de redenen van niet deelnemen en de behoefte aan evaluatie-instrumenten op bepaalde gebieden. In de meeste niet deelnemende districten bleek men wel globaal bekend te zijn met de mogelijkheden. Gebrek aan informatie, geen behoefte aan verdere toetsing en financiële redenen waren de belangrijkste argumenten om niet deel te nemen. In het algemeen blijkt men geen negatieve opvatting te hebben over het project. Gevraagd naar de onderdelen die eventueel het meest interessant zouden zijn, werden lezen en rekenen genoemd. In de rubriek (nog) niet beschikbare opties werden schrijfvaardigheid (stellen) en 'basic life skills' het meest genoemd. De scholen uit de deelnemende districten wensen een uitbreiding van het aanbod in dezelfde richting.

Uit de enquêtes onder de betrokkenen in de deelnemende districten blijkt verder een grote tevredenheid met de local assessment option waaraan men veel waarde toekent voor de besluitvorming op plaatselijk niveau. Het programma lijkt zeer effectief om veranderingen in het onderwijs te bewerkstelligen. Vele respondenten kunnen concrete veranderingen noemen die ze in hun klas, school of district hebben aangebracht naar aanleiding van de gevonden resultaten. Een ander kenmerk is verder een grote betrokkenheid van leerkrachten, ouders en bestuurders bij de verschillende aspecten van de local option. Kennelijk functioneert het als een goed middel om doel en functioneren van het onderwijs ter discussie te stellen en indien gewenst bij te stellen. In dit opzicht sluiten de ervaringen in Minne-

sota goed aan bij die in Connecticut (Report on the evaluation of the 1976-77 mathematics assessment, z.j.).

4 Overwegingen m.b.t. een educational assessment-project in Nederland

4.1 Inleiding

In de voorgaande paragrafen is een schets gegeven van wat er op het gebied van de educational assessment de laatste tien à twintig jaar in de Verenigde Staten is gebeurd. In deze slotparagraaf zal ik eerst nader ingaan op mogelijke doelen en functies van een assessment-project in Nederland. Vervolgens worden globaal enkele aspecten van een dergelijk project beschreven. De vraag naar de technische uitvoerbaarheid wordt niet aan de orde gesteld. In dit opzicht bieden de Amerikaanse ervaringen voldoende antwoord. Bovendien is in het begin van de vijftiger jaren in Noord-Brabant een onderzoek uitgevoerd dat in veel opzichten als een prototype kan worden beschouwd (Rapport over een onderzoek naar de stand van het gewoon lager onderwijs in Noord-Brabant, 1957)⁵.

Voor een goed begrip van het volgende is het nodig om vooraf het kader te schetsen waarin een assessment-project zou moeten passen. Hoewel om begrijpelijke, praktische redenen de start van een dergelijk project in het basisonderwijs zou kunnen liggen, biedt het begrip 'basisvorming' m.i. een beter en ruimer denk-kader. Onder basisvorming wordt verstaan: de kennis, het inzicht, de vaardigheden en houdingen die ieder nodig heeft om te kunnen functioneren in de maatschappij. Deze omschrijving is ontleend aan de Adviesaanvraag aan de Wetenschappelijke Raad voor het Regeringsbeleid met betrekking tot de basisvorming (1980). Het begrip basisvorming is geïntroduceerd in het Ontwikkelingsplan voor het voortgezet onderwijs (1979). Hoewel het hierin meer specifiek gekoppeld wordt aan het elementair onderwijs – het basisonderwijs en de voorgestelde algemene tweejarige brugperiode – wordt ook ruimte gelaten voor een ruimer interpretatie door koppeling aan de eerste fase van het voortgezet onderwijs. In deze paragraaf koppel ik het begrip basisvorming aan het (funderend) onderwijs tot aan het einde van de leerplichtige leeftijd, afgezien van de specifieke beroepsopleiding⁶.

4.2 Doelen en functies van een assessment-project

De vraag: Wat leren kinderen op school?, kan op verschillende manieren worden uitgelegd. De Innovatiecommissie Basisschool (1978) kiest voor een inventarisatie van het vormingsaanbod. In de Amerikaanse assessment-projecten ligt het accent op wat kinderen weten en kunnen. Hoe de vraag ook uitgelegd wordt, altijd gaat daaraan de vraag vooraf: Wat willen we dat kinderen leren, resp. onderwezen krijgen? Daarbij komt nog dat kinderen niet alleen op school leren. De meest fundamentele uitleg is m.i. de volgende:

- Wat vinden we de moeite waard om aan de opgroeiende generatie over te dragen?
- Wat hebben kinderen daarvan feitelijk geleerd?
- Welke bijdrage heeft de school daartoe geleverd, d.w.z. waaruit heeft het vormingsaanbod bestaan?

Het is er de laatste decennia niet duidelijker op geworden wat de taak en functie van de school is. De tijd is lang voorbij dat je kon zeggen: als mensen hebben leren lezen, schrijven en rekenen, hebben ze genoeg geleerd als algemene levensvoorbereiding, afgezien van de specifieke beroepsopleiding. Het onderwijs heeft er taken en functies bij gekregen, en over het gewicht van deze taken bestaat geen eenstemigheid. Enkele voorbeelden van mogelijke nieuwe taken zijn: Engels in het basisonderwijs, maatschappijleer, gezondheidsvoorlichting, milieu-educatie, etc. De beschikbare onderwijstijd kan evenwel slechts eenmaal gebruikt worden. Als je het een doet, heeft dat consequenties voor andere zaken. Kortom, welke bijdrage de school levert, resp. zou moeten leveren aan de basisvorming van een nieuwe generatie, is niet al te duidelijk, terwijl ook een discussie daarover niet gemakkelijk is te voeren, bij gebrek aan feitelijke gegevens. Tegen deze achtergrond moet ook de adviesaanvraag van de regering aan de Wetenschappelijke Raad voor het Regeringsbeleid gezien worden.

Een assessment-project zou m.b.t. deze problematiek een verhelderende rol kunnen spelen, als daarin antwoord zou worden gegeven op de drie bovengenoemde vragen. Met De Groot (1974a, p. 333) zou het doel van een assessment-project ook als volgt kunnen worden omschreven: 'Al degenen die onderwijs betalen (overheid, ouders, belastingbetalers),

ontvangen (leerlingen), geven (docenten), of trachten te verbeteren (pedagogen, onderzoekers, planners) hebben er belang bij dat zulk onderwijs aan zijn *doel* beantwoordt, en dat geregeld wordt nagegaan, of en in hoeverre dat zo is'. In dit opzicht is er een vergelijkende mogelijk met het Sociaal en Cultureel Planbureau, dat – o.a. gebruik makend van leefsituatie-surveys – om de twee jaar een Sociaal en Cultureel Rapport publiceert. Deze rapporten geven een beschrijving van de stand van zaken op maatschappelijk en cultureel gebied en de bewegingen die daarin gaande zijn, opdat 'alle betrokkenen bij dit beleid (uitvoerende werkers, politici, bestuurders, ambtenaren, studenten of zomaar geïnteresseerden), op gezette tijden, dwars door hun dagelijkse beslommeringen heen, worden herinnerd aan het lot van de mensen voor wie alle beleidsinspanningen bedoeld zijn'. (Sociaal en Cultureel Rapport 1974, 1975, voorwoord).

Naast deze belangrijke algemene doelstelling zou een assessment-project ook een handvat kunnen zijn voor de evaluatie en bijstelling van onderwijsleerplannen en schoolwerkplannen, resp. de ontwikkeling van eindtermen (vgl. de discussienota 'Naar het onderwijsleerplan basisschool' van de Projectgroep Leerplanning Basisschool, 1979). Bijl (1969) geeft voor deze functie van een assessment-project een uitstekende rationale avant la lettre in zijn brochure 'Over leerplanonderzoek'. Hij koppelt daarin de indeling van het vormingsaanbod aan een structurering van de levens-taak. 'Maar daarmee hebben wij nog geen programma met alle themata en stof-doelen waarover al is gesproken. Welke weg leidt naar dit voor het schoolbedrijf zo uiterst urgente overzicht van vereisten? Daarmee staan wij voor de *centrale vraag* van deze beschouwingen, die wij over leerplan-onderzoek wilden bieden.

Om tot zulke programma's te komen moet zeer veel werk worden verzet. Ze ontstaan niet achter een bureau. Ze kunnen in onze tijd ook niet worden geproduceerd door een commissie van deskundigen, die ze dan achter hun respectieve bureaus neerschrijven. Wat de levenstaak gaat inhouden moet in het samenleven zelf op wetenschappelijke wijze worden vastgesteld. Er is dus research nodig. (. . .)

Het onderzoek moet dus uitgaan naar wat er gekend en gekund wordt, maar er moet ook – democratisch! – gevraagd worden naar wat men wenselijk acht. Ten slotte beslist echter

over wat aan correctie op het aangetroffen beeld vanuit de wenselijkheid wordt aangebracht, het opvoedingsgezag, d.w.z. de organen die met dat gezag zijn bekleed'. (Bijl, 1969, p. 37, 38).

Vanuit deze gedachtengang zou een assessment-project van veel belang kunnen zijn voor landelijke instanties die zich bewegen op het terrein van de curriculumontwikkeling: de SLO, de landelijke pedagogische centra, educatieve uitgevers, etc. Een dergelijk project zou ook voor de rijksoverheid aanwijzingen kunnen opleveren voor het ontwikkelen van nieuwe beleidsinitiatieven. De overheid kan natuurlijk geen algemeen geldende minimumdoelstellingen m.b.t. de basisvorming voorschrijven. De bedoeling van een assessment-project is dan ook niet een beperking van de onderwijsvrijheid. Het kan wel bevorderen dat er van de onderwijsvrijheid een bewuster gebruik wordt gemaakt. In dit opzicht zijn de resultaten van een dergelijk project van veel belang voor bijv. vakbonden, besturenbonden en ouderorganisaties. Het tot stand brengen van een educational assessment-project zou daarom uitstekend passen in een constructief onderwijsbeleid. In een periode dat er belangrijke wijzigingen in het onderwijssysteem op stapel staan (bijv. de wet op het basisonderwijs, de discussie over geïntegreerd voortgezet onderwijs in de eerste fase), kan een assessment-project een brede maatschappelijke discussie over het onderwijs stimuleren.

Een landelijk assessment-project zou, met de nodige voorzieningen, ook van belang kunnen zijn voor regionale en plaatselijke begeleidings- en beleidsorganen. De Amerikaanse ervaringen in o.a. Minnesota en Connecticut wijzen erop dat het succes van een assessment-project voor een groot deel bepaald wordt door plaatselijke of regionale vervolgmogelijkheden. Naar mijn mening zou een dergelijke optie ook in Nederland essentieel zijn voor het welslagen van een landelijk project. Alleen dan kan een brede discussie over het onderwijs op gang komen.

Er is gesteld dat een assessment-project voor de rijksoverheid aanwijzingen kan opleveren voor het ontwikkelen van nieuwe beleidsinitiatieven. Om hiervan echter optimaal gebruik te kunnen maken, is het gewenst om aan het assessment-project de mogelijkheid van dieptestudies te verbinden. Dit geldt zowel voor beleidsvoorbereidende als -evaluerende on-

derzoeken. In het assessment-project zelf dienen geen onderzoeken ter evaluatie van beleidsmaatregelen te worden opgenomen. Die zouden de doelstelling en functie kunnen vertroebelen. Dergelijke onderzoeken zouden er wel als dieptestudies aan gekoppeld kunnen worden, waarbij bijv. gebruik wordt gemaakt van hetzelfde instrumentarium. Het assessment-project levert dan de basisgegevens waartegen andere gegevens afgezet kunnen worden.

4.3 Aantekeningen bij een mogelijke projectopzet

Een assessment-project zou in principe de gehele basisvorming moeten bestrijken. Dit schept niet geringe dekkingsproblemen. De vormingsdoelen omvatten meer dan alleen maar 'instructional objectives', die in termen van leerlinggedrag te specificeren zijn. Als tegenhanger van de 'instructional objectives' heeft Eisner (1969, p. 15, 16) het begrip 'expressive objective' geïntroduceerd: 'An expressive objective does not specify the behavior the student is to acquire after having engaged in one or more learning activities. An expressive objective describes an educational encounter: It identifies a situation in which children are to work, a problem with which they are to cope, a task in which they are to engage; but it does not specify what from that encounter, situation, problem or task they are to learn. An expressive objective provides both the teacher and the student with an invitation to explore, defer, or focus on issues that are of peculiar interest or import to the inquirer. An expressive objective is evocative, rather than prescriptive'.

Juist deze 'expressive objectives' zijn vanuit een 'educational measurement'-benadering de struikelblokken voor de dekking van het totale curriculum, resp. de basisvorming. In een serie lezingen en artikelen heeft De Groot (1974a, b, c, d; 1975; 1977; 1978) een bijzonder waardevolle bijdrage geleverd om tot een oplossing van het dekkingsprobleem te komen vanuit een classificatie van leereffecten (en daarmee samenhangende vormingsdoelen) op twee dimensies:

- kennis van de wereld versus zelfkennis;
- kennis van regels versus kennis van uitzonderingen.

Het is de taak van het onderwijs de leerlingen disposities of gedragsrepertoires te laten verwerven. Doelstellingen op het gebied van de

algemene en affectieve vorming zijn daarbij ook herleidbaar tot repertoires waarvan de leerling weet moet hebben: ze moeten rapporteerbaar zijn in termen van verworven (zelf)kennis, inzicht en vaardigheden. Toegespast op de herprogrammering van het wetenschappelijk onderwijs geeft De Groot (1974b) aan wat dit betekent voor de programma-ontwikkeling en toetsingsmogelijkheden. Hij onderscheidt daarbij vier soorten desiderata, concrete maar nog niet in detail geoperationaliseerde doelstellingen:

1. desiderata qua vakkennis en vakinzicht;
2. desiderata qua algemene vaardigheden;
3. desiderata qua verrijkend inzicht;
4. desiderata qua zelfinzicht.

De desiderata in de categorieën 1 en 2 zijn om te zetten in objectieve toetsingsprocedures, resp. intersubjectieve beoordelingsprocedures. De desiderata in de categorieën 3 en 4 kunnen eventueel via 'learner reports' (vgl. Van der Kamp, 1980) aan de orde worden gesteld. Daarnaast geldt met name voor deze categorieën dat de relatie met het vormingsaanbod aangegeven moet worden: 'Sommige typen doelstellingen horen wel thuis in de conceptdoelstellingslijst, maar moeten dan onmiddellijk in verband met het onderwijsprogramma worden gebracht: Hoe moet men dit opzetten om de gewenste demonstraties van vaardigheden (2), kennismakingen met andere ideeën en gebieden (3) en ervaringen qua zelfinzicht (4) te garanderen of althans te bevorderen?' (De Groot, 1974b, p. 224).

Wat betekent dit nu voor de opzet van een assessment-project? De eerste fase van een dergelijk project zou moeten bestaan uit een inventarisatie van wenselijke vormingsdoelen op een bepaald vormingsgebied. De indeling in vormingsgebieden behoeft daarbij niet noodzakelijk gekoppeld te zijn aan de gebruikelijke indeling in schoolvakken. De classificatie van De Groot biedt ook aanknopingspunten voor een indeling, terwijl ook de structuur van de levenstaak zoals door Bijl (1969) voorgesteld, een alternatief biedt. De tweede fase van een assessment-project is tweeledig en omvat:

- een inventarisatie van het vormingsaanbod in het betreffende onderwijs voor alle door De Groot genoemde soorten desiderata;
- een peiling van de beheersing van de vormingsdoelen in de categorieën 1 en 2 van De Groot.

Gezien het geschetste kader gaan de gedachten

naar het moment van de peiling van de beheersing, resp. de inventarisatie van het vormingsaanbod, in de eerste plaats uit naar het einde van het funderend onderwijs, resp. de leerplichtige leeftijd. Er zijn echter ook andere momenten denkbaar. Om redenen van praktische haalbaarheid komt het einde van het basisonderwijs – tenminste in de tegenwoordige situatie – in aanmerking. In dat geval zou een assessment-project een bijdrage kunnen leveren aan het formuleren van eindtermen voor het basisonderwijs. Vanuit relevantie-overwegingen is echter een peiling bij jonge volwassenen minstens zo belangrijk. Een dergelijke peiling heeft als voordeel dat nagegaan kan worden of wenselijk geachte kennis, vaardigheden en houdingen behouden blijven, of – in het geval dat deze aan het einde van het funderend onderwijs nog in onvoldoende mate aanwezig zijn – alsnog zijn verworven.

De instrumentontwikkeling ten behoeve van de peilingen en de inventarisatie van het vormingsaanbod, doorloopt in principe de normale stadia op basis van een lijst van vormingsdoelen. Eén aspect hiervan wijkt af van de gangbare procedures. Zowel bij het opstellen van de lijst van vormingsdoelen als bij de verdere instrumentontwikkeling dient een breed samengesteld forum te worden ingeschakeld. De Amerikaanse werkwijze, maar ook de indertijd in Noord-Brabant gevolgde procedure (zie noot 5) bieden hiervoor goede aanknopingspunten. Een dergelijk forum zou ook in de waarderings- en rapportagefase ingeschakeld moeten worden.

Wat de verzameling van de gegevens betreft, biedt de in Amerikaanse projecten gevolgde matrix-steekproeftrekking aanbeveling vanuit economische overwegingen. Het gaat niet om een niveaubepaling van individuele leerlingen, maar om een landelijke niveauschatting (met diverse uitsplitsingen) en dan is het niet nodig dat alle leerlingen in de steekproef aan dezelfde peiling deelnemen. Hoewel een periodieke herhaling niet noodzakelijk is, ligt dit wel voor de hand. Als dit het geval is, moeten er maatregelen worden getroffen om een vergelijking met een vorige ronde mogelijk te maken.

Als een assessment-project bij moet dragen om de betrokkenheid bij (de inhoud en vormgeving van) het onderwijs te vergroten, dienen in een dergelijk project lokale of regionale vervolgmogelijkheden te worden opgenomen, waarop ingeschreven kan worden. Hiertoe

dienen adequate servicemogelijkheden te worden ontwikkeld. Aanknopingspunten bieden de servicevormen in diverse Amerikaanse staten en ook het evaluatierapport van de Eindtoets Basisonderwijs van het Cito. Een dergelijke voorziening is waarschijnlijk essentieel voor het succes van een assessment-project en moet er daarom integraal deel van uitmaken.

Een mogelijkheid om dieptestudies aan een assessment-project te koppelen, kan in principe in een wat minder vaste relatie staan. Daarom ga ik daarop hier niet verder in. Wel dienen in het algemene landelijke assessment-project voldoende onafhankelijke variabelen te worden opgenomen om een goede koppeling mogelijk te maken. Hierbij valt o.a. te denken aan kenmerken van de leerling zelf, zowel als aan kenmerken van het gezin en de school.

4.4 Slot

Een assessment-project zoals in deze paragraaf geschetst, zou van veel belang kunnen zijn in de ontwikkeling van het Nederlandse onderwijs. Enerzijds kan het een verhelderende rol spelen, waardoor een betere discussie over de inhoud en vormgeving van het onderwijs mogelijk wordt, anderzijds kan het een belangrijk instrument zijn in de curriculumontwikkeling. Van directe invloed op de beleidsontwikkeling van de rijksoverheid moet men geen te hoog gespannen verwachtingen hebben. De Amerikaanse ervaringen wijzen uit dat dit soort invloeden via een omweg tot stand komen. Voor de overheid bestaan er echter wel mogelijkheden om een assessment-project op te nemen in een omvattend beleidsinstrumentarium.

De Amerikaanse ervaringen wijzen erop dat – tenminste in de beginfase – een assessment-project in het onderwijs als bedreigend ervaren kan worden. Ook in Noord-Brabant bestonden er volgens onze gegevens indertijd reserves m.b.t. het stand van zaken-onderzoek. Er bestaat angst dat de overheid door middel van een dergelijk project meer greep probeert te krijgen op de inhoud van het onderwijs en dat dit tot uniformering en aantasting van de onderwijsvrijheid zal leiden. Als in Nederland tot de instelling van een assessment-project zou worden besloten, dient aan dit aspect ruime aandacht te worden besteed.

Noten

1. Aan andere aspecten, bijv. de ontwikkeling van 'educational indicators', programmering en budgettering, accountability procedures op lokaal niveau, e.d. is veel aandacht besteed in het Cooperative Accountability Project, een studieproject dat door zeven staten van 1972-1975 werd uitgevoerd (zie bijv. Gunter, 1974; Wilsey & Schroeder, 1974; Grady, 1974; Browder, 1975).
2. Bij de eerste ronde, eind 1977, zakte 8% voor het communications deel en 36% voor het mathematics deel. Voor de leerlingen uit klas 11 die in 1978 voor de eerste keer aan de toets deelnamen, waren deze percentages resp. 3 en 26.
3. Gegevens over de andere bezochte staten zijn ter inzage bij de auteur (Cito, Postbus 1034, 6801 MG Arnhem, tel. 085-45 55 55).
4. Afgezien van de niet onaanzienlijke bijdrage uit andere afdelingen, bestaat de assessment staf in Minnesota uit vijf personen, de administratieve assistentie niet meegerekend. De kosten bedragen ruim \$ 300.000,- per jaar. In dit bedrag zijn de salariskosten van de functionaris voor de local assessment option niet inbegrepen; deze wordt uit andere bron betaald.
5. In het kader van de ontwikkeling van een Provinciaal Welvaartsplan stelde het Provinciaal Bestuur van Noord-Brabant in 1950 een commissie in onder voorzitterschap van Souren om een onderzoek te houden naar de stand van zaken in het lager onderwijs. Op basis van een leerstofprogramma dat vooraf aan een breed samengesteld forum voor commentaar werd voorgelegd, ontwikkelde de commissie een serie van ongeveer 25 toetsen en proeven voor nagenoeg alle lagere schoolvakken (inclusief de expressievakken!) om de verworven kennis, inzicht en vaardigheden aan het einde van de lagere school te onderzoeken. In het schooljaar 1951/52 namen bijna alle scholen in de provincie (\pm 800) deel aan het schoolvorderingenonderzoek met een kwart van de zesde klas leerlingen (totaal bijna 6000). Aan de proeven voor de expressievakken nam een steekproef van ongeveer 100 volledige zesde klassen deel (bijna 2000 leerlingen). Op verzoek van het Provinciaal Bestuur werd ook een vergelijking met enkele streken buiten Noord-Brabant in de onderzoek-sopzet opgenomen. Daartoe vond in 1955 een beperkte herhaling van het onderzoek plaats met de schoolvorderingstoetsen op een dertigtal scholen in Amsterdam, Enschede, Hengelo en Rotterdam. In 1957 bracht de commissie verslag uit in haar Rapport over een onderzoek naar de stand van het gewoon lager onderwijs in Noord-Brabant.
6. Het manuscript voor dit artikel werd afgesloten voordat de nota 'Verder na de basisschool' (Ministerie van Onderwijs en Wetenschappen, 1982) was verschenen. In deze nota wordt overigens in

het kader van het voortgezet basisonderwijs ook het begrip 'basisvorming' gehanteerd en de paragrafen 4.2-4.4 zijn evenzeer van toepassing op het voortgezet basisonderwijs als op het funderend onderwijs.

Literatuur

- Adviesaanvraag aan de Wetenschappelijke Raad voor het Regeringsbeleid met betrekking tot de basisvorming*, z.p., 1980.
- Allen, W. & E. Sushak, *Strategies and processes for utilizing statewide assessment data to facilitate educational changes at the local district level*. Paper prepared for presentation at the 9th Annual Conference on Large-Scale Assessment, Denver, 1979.
- Browder, L. H., *Who's afraid of educational accountability?* Denver: Cooperative Accountability Project, 1975.
- Bijl, J., *Over leerplanonderzoek*. Groningen: Wolters-Noordhoff, 1969 (Acta Paedagogica Ultrajectina XXIII).
- Dyer, H. S. & E. Rosenthal, *State educational assessment: programs*. Princeton N.J.: Educational Testing Service, 1971.
- Eisner, E. W., Instructional and expressive objectives: their formulation and use in curriculum. In: W. J. Popham, E. W. Eisner, H. J. Sullivan & L. L. Tyler, *Instructional objectives*. Chicago: Rand McNally, 1969 (AERA Monograph Series on Curriculum Evaluation no. 3).
- Evaluation Unit, *A study of the piggy-back option of the Minnesota State Assessment Program*. St. Paul Minn.: Minnesota Department of Education, 1979.
- Glass, G. V., Minimum competence and incompetence in Florida. *Phi Delta Kappan*, May, 1978, 602-605 (a).
- Glass, G. V., Standards and criteria. *Journal of Educational Measurement*, 1978, 15, 237-261 (b).
- Grady, M. J., *Using educational indicators for program accountability*. Denver: Cooperative Accountability Project, 1974.
- Groot, A. D. de, Over fundamentele ervaringen: prolegomena tot een analyse van gesprekken met schakers. *Pedagogische Studiën*, 1974, 51, 329-341 (a).
- Groot, A. D. de, Hoe stelt men eindtermen op? *Universiteit en Hogeschool*, 1974, 20, 213-232 (b).
- Groot, A. D. de, The problem of evaluating national educational systems. In: H. F. Crombag & D. N. de Gruijter (eds.), *Contemporary issues in educational testing*. The Hague: Mouton, 1974 (c).
- Groot, A. D. de, To what purpose, to what effect? Some problems of method and theory in the evaluation of higher education. In: W. A. Verreck (ed.), *Methodological problems in research and development*. Amsterdam: Swets & Zeitlinger, 1974 (d).
- Groot, A. D. de, Categories of educational objectives and effect measures: a new approach discussed in the context of second-language learning. In: A. J. van Essen & J. P. Menting (eds.), *The context of foreign-language learning*. Assen: Van Gorcum, 1975.
- Groot, A. D. de, Na de school op eigen kracht. *Pedagogische Studiën*, 1977, 54, 157-167.
- Groot, A. D. de, Wat neemt de leerling mee van onderwijs? Gedragsrepertoires, programma's, kennis-en-vaardigheden. In: *Handboek voor de onderwijspraktijk*. Deventer: Van Loghum Slaterus, 1978.
- Gunter, E., *Indicators and statewide assessment*. Denver: Cooperative Accountability Project, 1974.
- Innovatiecommissie Basisschool, *Innovatieplan basisonderwijs*. 's-Gravenhage: Staatsuitgeverij, 1978.
- Kamp, M. van der, *Wat neemt de leerling mee van kunstzinnige vorming?* 's-Gravenhage: Staatsuitgeverij, 1980 (SVO-reeks nr. 29).
- Kauffman, J. D., *State assessment programs: current status and a look ahead*. Bensenville Ill: Scholastic Testing Service, 1979.
- Miller, B. S., *Minimum competency testing: a report of four regional conferences*. St. Louis: CEMREL, 1978.
- Ontwikkelingsplan voor het voortgezet onderwijs*, 1979. (Bijlage X bij de Memorie van Toelichting bij hoofdstuk VIII van de Rijksbegroting voor het jaar 1980. Tweede Kamer, zitting 1979-1980, 15800).
- Pipho, C., Minimum competency testing in 1978: a look at state standards. *Phi Delta Kappan*, May, 1978, 585-588.
- Projectgroep Leerplanontwikkeling Basisschool, *Naar het onderwijsleerplan basisschool*. Enschede: Stichting voor de Leerplanontwikkeling, 1979 (2 delen).
- Rapport over een onderzoek naar de stand van het gewoon lager onderwijs in Noord-Brabant*. 's-Hertogenbosch: Provinciaal Bestuur van Noord-Brabant, 1957.
- Report on the evaluation of the 1976-77 mathematics assessment*, z.p., z.j. (intern memo Connecticut State Department of Education).
- Sociaal en cultureel rapport 1974*. 's-Gravenhage: Staatsuitgeverij, 1975.
- Wiley, C. E. & G. B. Schroeder, *Roles of participants in educational accountability*. Denver: Cooperative Accountability Project, 1974.

Curriculum vitae

J. M. Wijnstra (1946) studeerde na zijn opleiding tot onderwijzer tussen 1967 en 1973 pedagogiek aan de Rijksuniversiteit te Utrecht (specialisatie onderwijskunde). Vanaf 1971 was hij als projectleider verbonden

den aan het Project Friesland (SVO-project 0181). Het eindverslag van dit project vormde tevens zijn proefschrift 'Het onderwijs aan van huis uit friestalige kinderen' (promotie cum laude, 1976; het proefschrift verscheen als nr. 2 in de SVO-reeks). Van 1976 tot 1978 was hij onderwijskundig onderzoeker bij het Schoolpedagogisch Centrum Oostelijke Mijnstreek te Heerlen. Vanaf eind 1978 tot heden is hij medewerker van de afdeling Basisonderwijs van het Cito met als hoofdfunctie projectleider van de Eindtoets Basisonderwijs. Daarnaast werd hij onder meer belast met de Voorstudie National Assessment

waarvan in dit artikel verslag wordt gedaan. Publicaties in diverse binnenlandse en buitenlandse tijdschriften over toetsontwikkeling, met name m.b.t. taalbeheersing, evaluatie van onderwijs en onderwijs aan anderstalige leerlingen.

Werkadres: Cito, afdeling Basisonderwijs, Postbus 1034, 6801 MG Arnhem; tel. 085 - 45 55 55.

Manuscript aanvaard 24-11-'81