

De schooltoets na twaalf jaar

Deel II: Een evaluatie van de op het onderwijs gerichte functies: analyse van de stand van het onderwijs, concretisering van doelstellingen en beïnvloeding van het lager onderwijs

H. BLOK EN H. WESDORP
R.I.T.P., Amsterdam

Samenvatting

De CITO-Eindtoets Basisonderwijs bestaat 12 jaar. In 1966 werd de toets (in Amsterdam) geïntroduceerd; de constructeurs kenden 6 functies aan de toets toe. Deze waren a) Selectie voor het voortgezet onderwijs; b) Advies voor schoolkeuze aan de ouders; c) Het opsporen van reserve aan intellect; d) Analyse van de stand van het onderwijs; e) Concretisering van doelstellingen van het onderwijs en f) Beïnvloeding van het lager onderwijs. In het eerste deel van dit artikel, dat in het vorige nummer van Pedagogische Studiën verscheen, werden de functies a, b en c nader beschouwd. Aan de hand van empirische gegevens werd nagegaan hoe de selectie- en advies-functie in de huidige praktijk tot hun recht komen, en of er enige aanleiding is de ontdekking van 'reserve aan intellect' te verwachten. In het hiervolgende tweede deel van dit artikel worden de functies d, e en f nader bekeken. Aan de hand van enquête-resultaten wordt nagegaan of het 'bevoegd gezag' van de toetsresultaten gebruik maakt bij de beleidsvoering (functie d); op grond van o.a. een nauwkeurige analyse van het taalgedeelte van de schooltoets wordt nagegaan in hoeverre de functies e (concretisering van doelstellingen) en f (beïnvloeding van het lager onderwijs) worden vervuld.

4. Analyse van de stand van het onderwijs

Uit het in paragraaf 1. aangehaalde artikel van De Groot (1966) blijkt dat men zekere verwachtingen koesterde betreffende eventuele beleidsbeïnvloeding door toetsresultaten. Die verwachting was wel gerechtvaardigd. Als men de literatuur op dit punt naleest, blijken grootschalige evaluatieprogramma's wel vaker van belang voor het onderwijsbeleid (Wesdorp, 1979, par. 3.1.). In 1977 is door het RITP nagegaan in hoeverre van beleidsbeïnvloeding door schooltoetsresultaten in Nederland sprake kan zijn.

Het onderzoek werd verricht door Wolowitsj (1978). Nagegaan werd of toetsprogramma's aan het eind van het basisonderwijs gebruikt worden om het locale onderwijsbeleid mede vorm te geven. Meer specifiek was de onderzoeksvraag: is het beleid van bevoegde gezagen die over toetsinformatie beschikken actiever dan dat van gezagen die niet over dergelijke informatie beschikken? Met deze formulering wordt ook de opzet van het onderzoek in principe geschetst. Deze was als volgt: met behulp van een enquête zouden twee groepen van bevoegde gezagen worden benaderd. Enerzijds zou een groep worden benaderd die geen toetsinformatie ter beschikking kon hebben, omdat geen gebruik van toetsen werd gemaakt aan het eind van het basisonderwijs. Anderzijds zou een groep worden benaderd, die wel kon beschikken over toetsinformatie. Aan beide groepen zouden vragen omtrent de activiteit van het onderwijsbeleid worden gesteld. Nagegaan zou worden of het beleid in de wél-toetsende gebieden zich onderscheidde van dat in de niet-toetsende gebieden.

De belangrijkste doelgroep werd gevormd door het 'direct bevoegd gezag', voor het openbaar onderwijs het gemeentebestuur, voor het bijzonder onderwijs de schoolvereniging. Nog een andere doelgroep werd benaderd; ter controle op de resultaten van de enquêtes onder het bevoegd gezag werd tevens de 'uitvoerende macht', d.w.z. de hoofden van scholen en onderwijzers van zesde klassen basisonderwijs geënquêteerd.

De enquête bevatte vragen naar het soort toetsingsmiddel dat in gebruik was bij de overgang BO-VO; naar de toegankelijkheid van de aldus verkregen (toets-)informatie voor het bevoegd gezag; naar de mate waarin een aantal 'beleidsmaatregelen' werden getroffen. Deze laatste betroffen de mate waarin het bevoegd gezag uit eigen middelen bijdragen leverde aan, of pogingen deed extra Rijkssubsidie te verkrijgen voor de volgende kostenposten:

1. de aanschaf van onderwijsleermaterialen;
2. de aanschaf van hulpmiddelen in de sociaal-cultuur

- rele en emotionele sector;
- de salariëring van boventallige leerkrachten en extra vakleerkrachten;
 - het doen uitvoeren van experimenten in de school;
 - het treffen van voorzieningen betreffende het schoolgebouw en het meubilair.

Tenslotte werden ook vragen gesteld naar de overwegingen en informatiebronnen die bij de beleidsvoorbereiding een rol spelen.

Het bleek zeer moeilijk om voldoende 'niet-toetsgemeenten' bij het onderzoek te betrekken, omdat die er in Nederland zeer weinig zijn. Door het werk van Van Putten en Rijnbeek (1978) werden wij attent gemaakt op een aantal 'globale niet-toetsgebieden', te weten de provincies Friesland, Zeeland, Zuid-Zuid-Limburg, en Midden-Limburg. Besloten werd de provincies Zeeland en Friesland als niet-toetsgebieden in de steekproef te betrekken; als wél-toetsgebieden werden de provincies Overijssel en Gelderland (CITO-toets), en de provincie Zuid-Holland (Rotterdamse en Leidse Schooltoets) in de steekproef betrokken. In totaal werden in deze provincies 195 gemeentebesturen benaderd (responspercentage: 55%). Schoolverenigingen werden ook benaderd, 125 in totaal, enigszins anders verspreid (responspercentage: 27%). In de drie provincies – in dezelfde gemeenten – werden 454 scholen benaderd (respons: 53%).

Van belang voor de uitvoerbaarheid van het onderzoek was de vraag in hoeverre wij er in geslaagd

waren in de steekproeven voldoende toetsende en niet-toetsende gemeentebesturen en schoolverenigingen te representeren. Uit de antwoorden op vragen dienaangaande bleek het percentage niet-toetsende scholen zeer gering. Daarmee werd een vergelijking tussen wel- en niet-toetsende overheden onmogelijk.¹

Uit de enquête-resultaten bleek voorts, dat nog geen kwart – nl. 23% – van de gemeentebesturen de beschikking heeft over de resultaten van het onderzoek bij de overgang BO-VO, ongeacht de aard van het onderzoek (vooral toets of anderszins). Bij de schoolverenigingen was dat anders: daar zegt 82% de beschikking te hebben over de resultaten. Uit het eerste percentage blijkt al, dat voor een eventuele directe beïnvloeding van het beleid van het bevoegd gezag der openbare scholen de kansen niet optimaal zijn: men beschikt doorgaans eenvoudig niet over de informatie.

Tenslotte werden vragen gesteld omtrent de overwegingen die men laat gelden, en de informatiebronnen die men zegt te gebruiken bij de besluitvorming omtrent de eerder genoemde maatregelen. Voor de gemeentebesturen en besturen van schoolverenigingen ging het hier om de aanleidingen, overwegingen, en informatiebronnen die zij zelf in hun besluitvorming een rol laten spelen. Bij de scholen betreft het de mening van hoofden van scholen en/of leerkrachten der zesde klas over de volgens hen voor hun bevoegd gezag belangrijke aanleidingen etc. In Ta-

Tabel 6

Frequentietellingen over aanleidingen, overwegingen en informatiebronnen; de relevantie voor de besluitvorming

Informatiebron	gemeenten N = 107	schoolverenigingen N = 34	scholen N = 241
1. Politieke overtuiging en maatschappelijke inzichten	79	4	28
2. Resultaten van het onderzoek bij de overgang BO-VO	14	5	7
3. Verzoeken van leerkrachten van het BO	440	25	154
4. Adviezen van schooladviesdiensten	75	12	60
5. Adviezen van Rijks- en/of Gemeentelijke inspecteurs	204	10	95
6. Publicaties in dag- en weekbladen	1	4	7
7. Publicaties in de onderwijspers	3	8	9
8. Verslagen van wetenschappelijk onderzoek	3	4	4
9. Publicaties van het Ministerie van Onderwijs en Wetenschappen	15	2	24
10. Verzoeken van ouders, actiegroepen	29	7	29
11. Andere aanleidingen	70	12	27
Totaal aantal genoemde informatiebronnen	933	93	444
Gemiddeld per respondent	8,7	2,7	1,8

bel 6 worden de frequenties waarmee aanleidingen etc. als relevant voor de besluitvorming worden beschouwd, samengevat.

De belangrijkste aanleiding tot het nemen van de binnen dit onderzoek aan de orde gestelde beleidsmaatregelen vormen volgens de drie respondentengroepen de verzoeken van leerkrachten van het lager onderwijs (informatiebron 3.). Informatiebron 2. (de resultaten van het onderzoek bij de overgang BO-VO, waar het in dit onderzoek speciaal om ging) wordt zeer weinig genoemd als overweging, óók door de schoolverenigingen. De conclusie is daarom duidelijk: de informatie uit toetsprogramma's heeft geen invloed op het beleid. Enerzijds omdat de informatie bij de gemeentebesturen nauwelijks bekend is; anderzijds omdat deze informatie nòch door gemeentebesturen, nòch door schoolverenigingen – die er wèl over beschikken – als overweging voor beleid wordt genoemd. Misschien heeft de toetsinformatie enig effect via de leerkracht, die er uiteraard wèl over beschikt en die wèl een serieuze informatiebron voor het beleid vormt, getuigt Tabel 6.

5. Concretisering van doelstellingen

Hoewel de deskundigen het in het algemeen eens zijn over de concretiserings- of operationaliseringsfunctie van schooltoetsen, is niemand daar zo expliciet over als De Groot (1966), die met name ter operationalisering van communale doelstellingen de objectieve studietoets aanbeveelt. Ons staat de mogelijkheid open de schooltoets te analyseren en na te gaan welke doelstellingen de toets concretiseert. Vervolgens kunnen wij daarover een oordeel uitspreken. Ook kunnen wij de enquêtes die (o.a.) door het CITO zijn gehouden onder vnl. toetsgebruikers raadplegen. Wellicht geven zij informatie over de mate waarin de gebruikers de belangrijkste doelstellingen geoperationaliseerd achten. Eerst analyseren wij in het volgende als voorbeeld het taalgedeelte van de schooltoets (par. 5.1.); vervolgens gaan wij na welke meningen de schooltoetsgebruikers over de toets uitspreken (par. 5.2.).

5.1. Een analyse van het taalgedeelte 1966-1978

Wat bevat de schooltoets aan 'taal'-items? Ten eerste is er de zgn. stillestoets, die wij in het volgende niet verder beschouwen: Classificatie van de in deze toets gestelde vragen is een hachelijke zaak, wat o.a. door Van Deursen (1971b) is aangetoond. Ten tweede is er het deel dat nu 'Taalgebruik' heet, en

vroeger betiteld werd als 'Gemengde Taalopgaven'. Ten derde is er de spellingtoets en tenslotte is er een grammatica-onderdeel, dat niet in deze paragraaf doch wel in de volgende aan bod komt.

De toetsconstructeurs hebben van meet af aan getracht d.m.v. een zo gevarieerd mogelijk aanbod een diversiteit aan vaardigheden te operationaliseren, omdat slechts diversiteit het totaal aan taaldoelen enigszins zou representeren. De items van de taalopgaven zijn te classificeren in drie hoofdgroepen: de op 'semantische' vaardigheden gerichte opgaven, de 'syntactische' opgaven en de 'morfologische' opgaven. Er is een kleine restcategorie, die rand-vaardigheden bevat, zoals alfabetiseren en het kennen van afkortingen. Onder de semantische opgaven vallen items die gericht zijn op het definiëren van een (moeilijke) woordbetekenis, op het classificeren van woorden naar de (intensiteit van de) betekenis; op het kennen en toepassen van spreekwoorden of vaste uitdrukkingen; op het kiezen van een passend woord in een bepaalde context. De syntactische items betreffen het kiezen van de juiste werkwoordsvorm in een bepaalde zinsstructuur, het (re-)construeren van zinnen volgens bepaalde voorschriften, het bepalen van het aantal zinnen in een bepaalde alinea. De morfologische items houden zich bezig met woordvormingsproblemen van zelfstandige en bijvoeglijke naamwoorden en werkwoorden.

In de volgende Tabel (Tabel 7) hebben wij de verdeling van de taalitems over de hoofdcategorieën samengevat; het totaal-aantal taalitems heeft vooral in de beginjaren nogal geschommeld; sinds 1970 is een tweetal subtoetsen (= 60 items) aan 'taalgebruik' gewijd.

Het overzicht in Tabel 7 vergt enige toelichting. Aan de indeling in 'semantiek – syntaxis – morfologie' kan slechts worden vastgehouden als men ook de sinds 1973 in het taalgedeelte functionerende toets 'Taalgebruik 2' op dezelfde wijze indeelt. In dat jaar werd namelijk een subtoets ingevoerd, die zich van de tot dan gehanteerde taalopgaven onderscheidde doordat er een poging in gedaan werd aan de hand van een complete tekst (met fouten erin) bepaalde formuleer-technische zaken aan bod te doen komen. De hoofdcategorieën van deze toets zijn uiteraard, zo men wil, ook semantisch en syntactisch van aard. Wij hebben echter de oorspronkelijke betiteling van die categorieën gehandhaafd: 'passend en effectief taalgebruik' (te vergelijken met de semantische opgave 'in een bepaalde context het meest gepaste woord te kiezen'), en 'grammaticale opbouw van de

Tabel 7

Overzicht van de hoofdcategorieën van items in het taalgedeelte van de Schooltoets 1966-1978, in percentages van het totaal aantal taal-items

Jaar:	1966	67	68	69	70	71	72	73	74	75	76	77	78
Semantische opgaven + passend en effectief taalgebruik, vanaf 1973	63	54	77	60	73	72	35	35	40	42	47	33	33
Syntactische opgaven + grammaticale bouw van de zin, vanaf 1973	37	46	23	26	7	13	32	27	27	23	27	30	25
Organisatorisch-inhoudelijke zaken	-	-	-	14	7	7	7	18	18	18	13	18	25
Morfologie-opgaven	-	-	-	-	-	-	22	10	12	10	12	18	15
Rest-opgaven					13	8	5	10	3	7	2	-	2
Totaal aantal taal-items	30	48	35	35	60	60	60	60	60	60	60	60	55

zin' (te vergelijken met syntactische opgaven die structuren van zinnen aan de orde stellen). Daarnaast onderscheiden wij in Tabel 7 apart de organisatie- en inhoudelijke zaken, die sinds 1973 expliciet in de toets Taalgebruik 2 aan bod kwamen, maar al enkele jaren daarvoor ook in de losse items van Taalgebruik 1.

Als wij de inhouds-ontwikkeling willen kenschetsen, dan kunnen wij zeggen dat na een aanvankelijk wispelturig en nogal ondoorzichtig 'beleid' de inhoud vanaf 1972 vrij constant is gebleven. Kennelijk hebben in de beginjaren onzekerheid omtrent te toetsen doelstellingen en ook onvermogen op het gebied van de toetstechniek geleid tot een nogal zware druk op de semantische 'weetjes'. Vanaf 1972 is de verdeling gelijkmatig: ongeveer éénderde semantische zaken, éénderde syntactische, en de rest verdeeld over de andere categorieën. De toets is vanaf 1972 te kenschetsen als een 'gearriveerd instrument', dat niet of nauwelijks inhoudelijke schommelingen vertoont. Dit is dus de operationalisatie van een belangrijk deel der taal-doelstellingen. Op de inhoud van de 'constante' toets der jaren 1972-1978 gaan we daarom iets nader in.

Als we kijken naar de toetsinhoud van deze toetsen die recht moeten doen aan doelstellingen op het gebied van het taalgebruik, dan moeten we het volgende constateren: ondanks de invoering van de geheel anders georiënteerde toets Taalgebruik 2, die zoals boven werd aangeduid bepaalde taalvaardigheden 'geïntegreerd' aan de orde stelt, is natuurlijk het totaal-aanbod schamel te noemen. Dit impliceert

geen diskwalificatie van de CITO-toets, of van objectieve toetsen in het algemeen. Het is slechts de nuchtere constatering van een onvoldoende dekking der taal-doelstellingen (door dit instrument; maar ook andere instrumenten laten verstek gaan). Want zeer belangrijke aspecten van het taal(vaardigheids)-onderwijs ontbreken in de toets: het werkelijk (gericht) schrijven, maar ook het discussiëren, spreken, luisteren. Dat impliceert, dat deze toets - die 'constant' is en echt geen grote sprong voorwaarts meer kan maken - de grenzen van zijn concretiserings-mogelijkheden heeft bereikt. Als verduidelijking van doelstellingen door deze toets wordt bewerkstelligd, is het hoogstens de verduidelijking van een beperkt gedeelte van de doelen. De toetstechniek beperkt de operationalisatie-mogelijkheden, dat weet iedereen. Vooral op het gebied van het taalonderwijs mag daarom de schooltoets geen hoge verwachtingen wekken met betrekking tot de concretiseringsfunctie.

5.2. Enquêtes over de toets en toets-inhoud

Enkele enquêtes werden gehouden door het CITO; ze waren alle gericht op de inventarisatie van meningen over de concrete opgaven. Een eerste doelstelling van deze enquêtes was dan ook de toets te verbeteren op grond van gesignaleerde tekorten. Peters-Sips (1970) constateerde n.a.v. haar enquête-onderzoek dat de meningen van wèl- en niet-deelnemers aan de schooltoets omtrent het belang van bepaalde vak-onderdelen verschilden: gebruikers van de toets achten rekenen, spelling en begrijpend

lezen belangrijker dan niet-gebruikers. Algemene kennis daarentegen werd door niet-deelnemers belangrijker gevonden. Toch bleken in het algemeen de verschillen in opvatting tussen deelnemers en niet-deelnemers niet groot. Een vergelijkbare enquête werd gehouden door Van Deursen (1971a). Deze constateerde een redelijke mate van tevredenheid met de toets-items, hoewel de subtoetsen Rekenen en Woordbenoeming/Zinsontleding ernstige kritiek ondervonden m.b.t. het minimum karakter van de toetsen. Van Dam (1972) enquêteerde de schooltoetsgebruikers om uitsluitel te verkrijgen omtrent hun meningen t.a.v. het in opspraak geraakte onderdeel Spelling; hij constateerde dat de meerderheid der toetsgebruikers voor een verplichte spellings-toets was; ongeveer 1 op de 3 gebruikers koos voor een facultatieve toets. M.i.v. 1973 is daarop de mogelijkheid gecreëerd om niet aan de subtoets Spelling deel te nemen. Vanaf 1973 leverde het CITO bij de rapportage een totaalscore inclusief Spelling en een totaalscore exclusief Spelling; deze maatregel werd opgeheven in 1977. Rommes en Knops (1977) inventariseerden nogmaals de meningen van toetsgebruikers over de belangrijkheid van de diverse subtoetsen. Zij constateerden geringe verschillen in ervaren belangrijkheid tussen de peiling uit 1979 (Peters-Sips) en hun eigen peiling; alleen de belangrijkheid van de toets Algemene Kennis was aanzienlijk gedaald, maar dat neemt niet weg dat in totaal nog 85% der respondenten de toets als (zeer) belangrijk betitelde. Woordbenoeming/Zinsontleding werd door slechts 57% belangrijk gevonden.

De enquête-resultaten lijken erop te wijzen, dat met behulp van de schooltoets inderdaad door velen als 'communaal' ervaren onderwijsdoelen worden geoperationaliseerd. Maar het blijft de vraag of de geringe kritiek op de meeste subtoetsen en items inderdaad te interpreteren is als volledige tevredenheid. Tenslotte zijn belangrijke doelen duidelijk niet in de toets gerepresenteerd. Over de mate waarin zulk een ontbreken als acceptabel of gevaarlijk wordt ervaren, doen de enquête-resultaten uiteraard geen uitspraak. Evenmin is duidelijk in hoeverre de gebruikers van heden de toets als bedreigend en als negatieve beïnvloeding ervaren, of, omgekeerd, als een positieve steun.

6. Beïnvloeding van het lager onderwijs

In onze inleiding stelden wij al, dat een toets die inhoudsvalide is t.o.v. de belangrijke doelstellingen een positief effect kan hebben. De vraag is dus, of de

toetsconstructeurs een dergelijke valide instrumentatie hebben bewerkstelligd; vervolgens is het nog de vraag, of het onderwijs zich inderdaad, zoals voorspeld, richt naar de toets. Over de inhoudsvaliditeit willen wij – na onze analyse van het taalgedeelte in par. 5 – niet in details treden. Wij willen slechts het sub-toetsenaanbod en de scorings-service nader beschouwen. Alvorens dat te doen geven wij nog enkele enquête-resultaten m.b.t. het beïnvloedingsaspect.

6.1. Enquête-resultaten

In 1970 en 1971 werd door het RITP een onderzoek verricht dat onder andere in een enquête (voorjaar 1970) onder alle in 1969 deelnemende scholen de invloed van het toetsprogramma aan de orde stelde (Warries, 1972). De belangrijkste resultaten van deze enquête komen – sterk samengevat – op het volgende neer: het sterkst werden uitspraken onderschreven waarin gezegd wordt dat je je van de schooltoets niets hoeft aan te trekken als je als onderwijzer gewoon je werk doet en dat je geen verbetering van het onderwijs bereikt door de toets. Sterk onderschreven worden ook uitspraken over het examen-karakter van de toets, waardoor training met toetsmateriaal in de hand wordt gewerkt. Sterk afgewezen worden uitspraken die verbetering van het onderwijs door de toets postuleren. Een negatieve invloed van de toets op zangonderwijs en zelf lezen van boeken wordt afgewezen. De kritiek van de respondenten richt zich vooral op de invloed op de zaakvakken, op 'intellectualisme' en 'verbalisme' die in de hand zouden worden gewerkt, en op de gestelde onmogelijkheid de leerlingen zelf iets onder woorden te laten brengen. In het algemeen trekt Warries de conclusie dat de stemming onder het onderwijspersoneel en de hoofden niet zo somber is als wel eens verondersteld is. De opvattingen over een slechte invloed nemen niet de belangrijkste plaats in, maar ze zijn er wel.

In 1977 werd eveneens door het RITP een onderzoek verricht onder CITO-toets-gebruikende onderwijzers (Wesdorp, 1979). Het had ten doel na te gaan wat de attitude van de leerkrachten jegens de toets was. Een landelijke steekproef van onderwijzers was erbij betrokken. Enkele items in de enquête betroffen de invloed op didactiek en curriculum: belangrijke minderheden steunden uitspraken als 'Nieuwere onderwijsopvattingen in de zaakvakken worden tegengehouden' (20%); 'Allerlei waardevolle didactische werkvormen raken in het gedrang' (19%); 'Door de eindtoets komt er helaas teveel nadruk op prestaties te liggen' (50%). De resul-

taten uit 1977 vertoonden vrij veel overeenstemming met die uit 1970/71: ook nu weer klachten over de (negatieve) invloed, zij het minder sterk.

6.2. *Het subtoetsen-aanbod van de schooltoets 1966-1978*

In par. 5.1. analyseerden wij het taalgedeelte; daarbij classificeerden wij de items en trokken wij conclusies naar aanleiding van de inhoud der items. Niet alleen de inhoud der items, maar ook de hoeveelheden items per subtoets en de hoeveelheden subtoetsen per 'vaardigheid' zijn een indicatie van de mogelijke beïnvloeding van het onderwijs. Als we even afzien van de samenstelling van de toets in de beginjaren – we zagen hierboven al dat toetstechnische en onderwijskundige onzekerheden misschien tot een wisselvallig beleid leidden – moeten we het volgende constateren: In het algemeen wordt het *Rekenen* door ± 90 items in 3 subtoetsen vertegenwoordigd. Alleen in 1969 is een duidelijk experimentele toets met extra vraagstukken toegevoegd. Niet duidelijk is in hoeverre deze toets als beïnvloeder in bepaalde richting was bedoeld. In 1970 was hij weer verdwenen. De toets *Stillezen* omvat constant 2 subtoetsen van ± 30 items. Het *Taalgebruik* werd eerst wat dun vertegenwoordigd (men zie de tabel in par. 5.1.) met ± 30 items; vanaf 1970 zijn constant twee subtoetsen van elk 30 items onderdeel van het pakket. Tot zover de harde kern van het subtoetsen-aanbod: het 'verplichte pakket', bestaande uit *Rekenen* + *Lezen* + *Taal*. De andere subtoetsen zijn illustratiever voor de eventuele positieve onderwijsbeïnvloeding.

Wat betreft de toets *Algemene kennis*: oorspronkelijk – tot 1970 – was deze vertegenwoordigd met 2 subtoetsen, sindsdien met één; echter altijd 'facultatief' en wel zodanig dat in totaalscoreberekeningen géén rekening werd gehouden met de Algemene kennis-toetsprestaties. Vanaf 1977 is de toets echter niet meer facultatief; op grond van de uit enquêtes blijkende waardering voor de subtoets en het hoge deelnamepercentage (99%) werd de toets verplicht onderdeel. Omdat echter de meningen over de doelstellingen op het door de subtoets bestreken gebied niet eensluidend waren, werden de AK-scores niet opgenomen in de totaalscores, waarmee de toets toch iets van zijn facultatieve karakter bleef behouden. De scoreverwerking bleef hetzelfde; de toets kreeg echter het predicaat 'verplicht onderdeel'.

De *Spellingtoets* was oorspronkelijk een verplicht onderdeel; in 1973 werd – in zekere zin onder druk van leerkrachten, en in elk geval nadat gebruikers gehoord waren (Van Dam, 1972) – een andere scoreverwerking van kracht (vgl. 5.2.); daardoor werd

het mogelijk de Spellingtoets niet af te nemen, omdat ook een Taal-percentielscore werd geleverd zonder dat de Spellingtoets daarin werd betrokken. Het onderwijs kreeg zo de mogelijkheid het verminderde belang, dat men aan de spelvaardigheid hechtte, in het toetsgebruik te realiseren. Echter, sinds 1977 is deze maatregel teruggedraaid. Nadat vier jaren lang het spellinggedeelte kon worden weggelaten, werd dit weer onmogelijk in 1977. In de 'Verantwoording Inhoud 1977' wordt deze stap toegelicht met het argument, dat de totaalpercentielscore exclusief en die inclusief spelling voor vrijwel 100% aan elkaar gelijk waren.

Tenslotte de *Grammaticatoets*: deze is in elk geval sinds 1973 facultatief en sinds 1978 uit het reguliere toetsboekje verdwenen. Met deze maatregel volgt de toets belangrijke ontwikkelingen in het Taalonderwijs, die o.a. de explicietere aandacht aan zinsontleden/woordbenoemen steeds minder belangrijk maken, als men de ACLO-Moedertaal mag geloven (Advies Commissie voor de Leerplan Ontwikkeling, 1978).

Vatten wij ons overzicht van het subtoetsen-aanbod samen, dan constateren wij dat de Schooltoets in zijn harde kern betrekkelijk traditioneel is: *Rekenen*, *Lezen* en 'Taal' vormen de verplichte kern van de toets. Sinds 1977 is daartoe ook de Algemene Kennis-toets te rekenen, die echter door zijn aparte scoring de werkelijke verplichte status nog niet helemaal heeft. Maar men zou de statuswijziging toch als een positief punt kunnen beschouwen, zeker in het kader van deze paragraaf, die de beïnvloeding van het onderwijs tot onderwerp heeft. Immers, door de verplichtstelling is de reikwijdte van de toets verbreed; blijkens enquêteresultaten acht men de subtoets AK een aanwinst. Men mag dus enige positieve invloed verwachten.

Het verdwijnen van de toets *Woordbenoeming/Zinsontleding* is voor sommigen die het goed menen met het onderwijs, een negatief punt. Anderen – en de ACLO-M voorop – achten modern taalonderwijs zonder traditionele grammatica zeer goed mogelijk. Daarom is het waarschijnlijk een juiste beslissing geweest, de subtoets uit het pakket te halen en als geheel aparte toets ter beschikking te stellen. Hoe het onderwijs daar zelf over denkt, wordt momenteel door het RITP onderzocht.²

Tenslotte is de argumentatie voor het 'herstel' van de Spellingtoets als verplicht onderdeel (zonder de mogelijkheid spelling niet in de totaalscore op te nemen) bepaald zwak. Wij achten een hoog verband tussen scores met en scores zonder spelling geen onderwijskundig argument. Van belang is, hoeveel gewicht 'men' aan het onderdeel hecht. Dat bleek in

1972, toen eenderde der gebruikers weinig belang aan de subtoets hechtte, en tweederde veel. Een dergelijk enquêteresultaat is onderwijskundig relevanter dan de genoemde correlatiecoëfficiënt. Op grond van de CITO-enquêteresultaten 1972 achten wij het herstel van de Spellingtoets een maatregel, die althans in sommige scholen niet als positief beïnvloedend zal worden ervaren.

6.3. *Het service-aanbod: de evaluatieve mogelijkheden van de toets*

Tenslotte willen wij nog even stilstaan bij de sinds 1973 geïntroduceerde evaluatieve gebruiksmogelijkheid van de toets. Sinds dat jaar geeft het CITO naast de individuele leerlingcores ook een schoolrapport en een evaluatierapport. Het schoolrapport maakt het mogelijk de schoolprestaties op de diverse onderdelen te vergelijken met de landelijke gemiddelde prestatie op die subtoetsen. Tevens wordt een 'oordeel' gegeven, dat normaal verdeeld is: $\pm 14\%$ der scholen krijgt het oordeel 'laag'; $\pm 72\%$ het oordeel 'gewoon', en $\pm 14\%$ het oordeel 'hoog'. Het evaluatierapport maakt het de school mogelijk de schoolprestaties op de diverse toetsonderdelen (dus bv. voor de in par. 5.1. behandelde taalgebruiktoets 1: Morfologie, Syntaxis, Semantiek) te vergelijken met de landelijke prestaties. Dit evaluatierapport is dus een gedetailleerd schoolrapport. De 'evaluatie-service' moet apart worden besteld bij het CITO; hij maakt geen deel uit van het normale service-pakket. Uit het aantal scholen dat de evaluatieservice bestelde kan men opmaken in hoeverre dit evaluatief gebruik inderdaad plaatsvindt. Zo kan men een indruk krijgen van de mate waarin positieve beïnvloeding (door evaluatie) van het onderwijs een kans krijgt. In 1976, '77 en '78 waren de percentages deelnemers aan de evaluatieservice resp. 28,7%, 29,5% en 32,5%: de service wint kennelijk veld.³

Een andere indruk van de mate van evaluatief gebruik biedt ons de in 1978 gehouden enquête naar gebruik en waardering van eindprocedures en methoden bij de overgang van het basisonderwijs naar het voortgezet onderwijs. Deze enquête werd in opdracht van het CITO uitgevoerd door de Stichting voor Research en Marketing te Heerlen (CITO, 1978). Van de ondervraagde CITO-toetsgebruikers (185) bleken 24% de toets graag wel als evaluatie-instrument te hanteren (28% liever niet; 48% onbekend). Ook werd onderzocht in hoeverre de toets daadwerkelijk op evaluatieve wijze werd toegepast, d.w.z. in hoeverre de toets (en ook: andere procedures) ooit was/waarden gehanteerd om veranderingen in lesprogramma of -methodiek door te voeren. Van de

185 ondervraagde CITO-toetsgebruikers bleek 34% dit ooit gedaan te hebben (60% nooit; 6% onbekend). De CITO-toets blijkt meer dan de andere procedures in evaluatieve zin te worden gebruikt: van de 85 niet-CITO-toetsgebruikers bleken slechts 25% ooit evaluatieve effecten aan hun procedures te hebben gegeven.

Zowel de CITO-gegevens over de jaren 1976-'78 als de R + M-gegevens uit de enquête van 1978 duiden erop dat de toets in een minderheid van de gevallen als evaluatief instrument wordt gehanteerd. De positieve beïnvloeding van het onderwijs - onderwerp van deze paragraaf - kan ook via de evaluatieve functie derhalve geen bijzonder grote omvang aannemen; de stijgende belangstelling voor de evaluatie-service is echter een positief punt.

7. *Samenvatting en discussie*

Voor wij hierna de voornaamste resultaten samenvatten, willen wij een relativerende opmerking maken. Een belangrijk deel der hiervóór gepresenteerde onderzoeksresultaten is verkregen met behulp van enquêtes. De respons-percentages waren meestal laag. Dit betekent, dat de onderzoeksresultaten een beperkte geldigheid hebben. Dit dient in overweging genomen te worden, ook als wij in het volgende op grond van de gerapporteerde resultaten enige suggesties doen met de bedoeling de discussie over de functie van de schooltoets te heropenen. Feitelijk zijn wij niet de eersten die deze hernieuwde discussie noodzakelijk achten. Het CITO zelf verleende reeds in 1977 de Stichting voor Research en Marketing opdracht tot het invoeren van een onderzoek naar het gebruik en de waardering voor eindprocedures en -methoden bij de overgang (CITO, 1978).

In de inleiding onderscheidde wij zes functies van de schooltoets.

Een selecterende en adviserende functie blijkt door de toets nauwelijks vervuld te worden. Toch mag een geringe invloed niet ontkend worden. Twee positief te waarderen verschijnselen treden bij deze invloed op: de toets wordt vaker ten voordele dan ten nadele van de leerling gebruikt; in het licht van discriminatie naar milieu of sexe is de toets wellicht zelfs meer aanvaardbaar dan het advies van de onderwijzer of de wens van de ouders.

Wat betreft het opsporen van een reserve aan intellect vervult de toets inderdaad een functie. Er is een groep leerlingen met een duidelijk hogere score dan het eerste advies zou doen vermoeden. De

grootte van deze groep is niet gering en bedraagt naar schatting circa 6%. Er zijn echter twijfels of de toets meespeelt bij de effectuering van de ontdekking in een verhoogd tweede onderwijzersadvies.

De schooltoets blijkt de functie 'analyse van de stand van het onderwijs' niet te vervullen. Beleidsvoerders beschikken over het algemeen niet over de resultaten op de toets. Als ze er wel over beschikken, wordt de informatie uit de toets niet als overweging gehanteerd bij de besluitvorming.

Aangaande de vijfde functie, concretisering van doelstellingen, vervult de toets inderdaad een functie. Onze onderzoeken op dit punt hadden alleen betrekking op het taalgedeelte van de toets. Door velen als communiaal ervaren doelen van het taalonderwijs worden in de toets geoperationaliseerd. Tevens moet geconstateerd worden dat belangrijke doelen van het taalonderwijs niet in de toets geconcretiseerd worden, bijv. doelstellingen op het gebied van stel-, spreek- en luistervaardigheid. De inhoud van de taaltoetsen blijkt sinds 1972 nauwelijks meer veranderd te zijn. Vooruitgang op het gebied van concretisering van doelstellingen is niet meer te verwachten. De toets blijft een instrument dat bepaalde belangrijke doelen wel en andere, ook belangrijke doelen, niet concretiseert.

Als laatste functie behandelden wij de beïnvloeding van het lager onderwijs. Vooropgesteld dient te worden dat nog zelden effecten van de toets op het onderwijs, d.w.z. op leerplan, didactiek of leerresultaten, empirisch aangetoond zijn. Verwacht kan worden dat van de toets negatieve invloed uit kan gaan (training op getoetste onderdelen, verwaarlozing van niet-getoetste onderdelen), omdat leerkrachten aan de toets veel vaker een selectieve dan een evaluatieve functie toekennen (CITO, 1978). Maar evenzeer kunnen bepaalde positieve invloeden worden verwacht; met name is het niet uitgesloten dat de aandacht van de toets voor bepaalde 'moderne' vaardigheden in de Reken- en Algemene Kennis-toets het onderwijs heeft 'bevrijd' van ballast en suggesties voor zinvoller vulling heeft gedaan. Helaas konden wij in het bestek van dit artikel de genoemde toetsen niet nauwkeurig analyseren.

Van de zes functies vervult de toets er drie in geringe mate. Deze functies hebben alle te maken met de overgang naar het voortgezet onderwijs: selectie, advies en opsporen van reserve aan intellect. De functie 'analyse stand van het onderwijs' wordt in het geheel niet vervuld. De concretiseringsfunctie wordt in beperkte mate uitgeoefend. Over de invloed die de toets op het onderwijs uitoefent kunnen de verwachtingen derhalve niet al te hoog gestemd zijn.

Deze stand van zaken mag weinig bevredigend genoemd worden.

Deze conclusie sluit echter niet uit dat in vergelijking met de situatie van vóór 1966 (het toetsloze tijdperk) flinke verbeteringen tot stand zijn gekomen. De toets vervult immers enkele functies op het gebied van de BO-VO-overgang, al werden deze functies door de oorspronkelijke makers als minder belangrijk beschouwd. De toets heeft onmiskenbaar enige adviserende waarde voor ouders, leerkrachten en toelatingscommissies. Ook verschaft de toets een kleine groep leerlingen meer mogelijkheden op toelating tot een hoger schooltype. Allerwege wordt echter de situatie bij de BO-VO-overgang als onbevredigend aangeduid. De initiatieven om tot een ongedeelde doorstroming te komen nemen in kracht en aantal toe en velen zijn doordrongen van de nadelen van een discontinue schoolloopbaan. Het zou wellicht nu al aanbeveling verdienen daar waar een ongedeelde doorstroming gerealiseerd is, ontheffing te verlenen van het Koninklijk Besluit betreffende het tweede toetsingsgegeven.

De functie die met 'analyse van de stand van het onderwijs' werd aangeduid, zou zonder de massaaf te nemen schooltoets op een andere en betere wijze tot zijn recht kunnen komen. In een terugblik op 'vijftig jaar testen' doet Buros (1977) de volgende aanbeveling: ten behoeve van de bepaling van de stand van het onderwijs bij grote groepen leerlingen kan het gebruik van korte toetsen worden overwogen, die elk door slechts een fractie van de leerlingen hoeven te worden gemaakt. Eén voordeel is direct al evident. De benodigde toetstijd (± 7 uur per leerling bij de CITO-toets) wordt teruggebracht tot $1/2$ à 1 uur per leerling. Lang niet alle leerlingen behoeven meer lastig gevallen te worden voor het verzamelen van informatie, die vooral dienstig is aan beleidsvoerders. Het verzamelen van informatie kan beperkt blijven tot die leergebieden, waarvoor beleidsvoerders werkelijk over informatie dienen te beschikken. Dit is efficiënter dan de routinematige inzameling van zoveel mogelijk toetsgegevens bij alle leerlingen.

Ook met betrekking tot de vijfde functie van de schooltoets, concretisering van doelstellingen lijkt deze procedure voordelen te bieden. Veel belangrijke doelstellingen kunnen niet in massaal afneembare geprecodeerde instrumenten geoperationaliseerd worden. Wanneer 'massaal afneembaar' een minder stringente eis wordt, ontstaan er mogelijkheden om te experimenteren met meer arbeidsintensieve toetstechnieken ten dienste van moeilijk operationaliseerbare maar belangrijke doelstellingen.

Welke voordelen de procedure biedt met het oog op de zesde functie van de schooltoets, beïnvloeding van het lager onderwijs, is onder meer afhankelijk van de status die aan de procedure toegekend wordt. Richtende invloed van een toetsprogramma op het onderwijs hangt af van de status die het programma in de ogen van zijn gebruikers heeft (Wesdorp, 1979). Duidelijk is in ieder geval dat wanneer een toets geen selectieve functie meer heeft, de noodzaak tot excessieve training op getoetste onderdelen vervalt. Het lager onderwijs wordt bevrijd uit het centraliserende keurslijf waarin leerkrachten soms menen dat het zich bevindt. Daarmee verdwijnt de angst voor een belangrijk negatief effect op het onderwijs. Om een positieve invloed op het onderwijs veilig te stellen kan overwogen worden jaarlijkse toetsen ter beschikking te stellen voor leerkrachten die hun onderwijsresultaten willen evalueren. Gedacht kan worden aan handhaving van de schooltoets met alleen facultatieve onderdelen.

Noten

1. Vergelijking tussen wel- en niet-toetsende overheden zou ook onder gunstiger omstandigheden uiteraard een hachelijke zaak blijven. Immers, de twee te vergelijken groepen zouden op een aantal belangrijke variabelen kunnen verschillen. Vóór het onderzoek startte werd echter verwacht dat voldoende niet-toetsende gemeenten zouden worden gevonden, en dat een enigszins redelijke vergelijking mogelijk zou zijn met zorgvuldig gekozen toetsgemeenten.
2. In het kader van SVO-project 0412 is in het voorjaar van 1979 onderzocht welke motivering leerkrachten (van LO, LBO en AVO/VWO) hanteren voor hun grammatica-onderwijs, hoeveel tijd zij er aan besteden, en welke inhoud zij er aan geven. Een interimrapport dat onderzoeksliteratuur over de effectiviteit van grammatica-onderwijs samenvat benevens bovenstaande enquête-resultaten bevat is in september verschenen (Tordoir en Wesdorp, 1979).
3. Deze gegevens zijn ons beschikbaar gesteld door dr. J. Wijnstra, projectleider Eindtoets Basisonderwijs van het CITO.

Literatuur

- Advies Commissie voor de Leerplanontwikkeling Moedertaal, *Advies over het grammatica-onderwijs in de lagere school met het oog op de aansluiting van het voortgezet onderwijs op het lager onderwijs*. Enschede, 1978.
- Buros, O. K., Fifty years in testing: some reminiscences, criticisms, and suggestions. In: *Educational Researcher*, vol. 6, no. 7, 1977, 9-15.

- CITO, *Meningen over de eindtoets en plannen voor de toekomst*. CITO, Arnhem, 1978.
- Dam, P. R. L. van, *Spelling toetsen? Verslag van een enquête*. CITO, Arnhem, 1972.
- Deursen, J. I. van, *Enquêteverslag Inhoud Schooltoetsen Basisonderwijs*. CITO, Arnhem, 1971a.
- Deursen, J. I. van, *Verslag van een onderzoek naar de praktische bruikbaarheid van een classificatiesysteem van stillenzen bij de schooltoetsen basisonderwijs*. CITO, Arnhem, 1971b.
- Groot, A. D. de, Het nut van een schooltoets in de zesde klas L.O., *Het Schoolblad*, 1, 1966, 288-292.
- Peters-Sips, M. W. H., *Verslag van de enquête 'Inhoud Schooltoets 1970'*. CITO, Arnhem, 1970.
- Putten, K. van en K. Rijnbeek, *Klassikale tests en toetsen bij de overgang van het basisonderwijs naar het voortgezet onderwijs 1975-1976 in Nederland*. Doctoraalscriptie, Vakgroep onderwijskunde, Rijksuniversiteit, Leiden, 1978.
- Rommes, A. E. N. en W. J. E. Knops, *Verslag over de enquête met betrekking tot de inhoud van de Eindtoets Basisonderwijs 1977*. CITO, Arnhem, 1977.
- Tordoir, A. en H. Wesdorp, *Het grammatica-onderwijs in Nederland; een researchoverzicht betreffende de effecten van grammatica-onderwijs en een verslag van een onderzoek naar de praktijk van dit onderwijs in Nederland*. Staatsuitgeverij, 's-Gravenhage, 1979.
- Warries, E., De invloed van een toetsprogramma, *Pedagogische Studiën*, 49, 1972, 73-82.
- Wesdorp, H. (red.), i.s.m. H. Blok, E. de Graaff, A. Wolowitsj en S. Zijlmans, *Studiëtoetsen en hun effecten op het onderwijs*. RITP, Amsterdam 1979.
- Wolowitsj-Schelvis, J. J., *De invloed van toetsinformatie op het beleid*. RITP, Amsterdam, 1978.

Curricula vitae

H. Blok (1954) studeerde psychologie aan de Vrije Universiteit en later aan de Universiteit van Amsterdam (doctoraal 1979). Sinds 1976 werkzaam in de onderwijsresearch aan het R.I.T.P. Heeft zich tot nu toe bezig gehouden met onderzoek naar effecten van studietoetsen.

Adres: R.I.T.P., Prinsengracht 303, Amsterdam.

H. Wesdorp (1935) studeerde Nederlands (doctoraal 1965), was leraar Nederlands (1965-1969) en medewerker van het CITO (1969-1973); is sinds 1966 werkzaam in de onderwijsresearch aan het R.I.T.P. Deed onderzoek op het gebied van de evaluatie van het onderwijs in de moderne talen en de moedertaal, op het gebied van de doelstellingen van het taalonderwijs, over de beïnvloeding van het onderwijs door toetsgebruik. Promoveerde (1974) bij A. D. de Groot op 'Het meten van de productief-schriftelijke taalvaardigheid'.

Adres: R.I.T.P., Prinsengracht 303, Amsterdam.