

# De invloed van een toetsprogramma\*

E. WARRIES,

*Stichting Research Instituut voor de Toegepaste Psychologie aan de Universiteit van Amsterdam, te Amsterdam.*

## 1. *Probleemstelling*

Schooltoetsen en studietoetsen hebben als evaluatiemiddel het bezwaar dat ze betrekkelijk eenzijdig de effecten van onderwijs meten. Er zijn veel doelen van onderwijs die niet met studietoetsen kunnen worden geëvalueerd. Het gaat in het onderwijs uiteindelijk om meer dan alleen het bijbrengen van kennis die in de vorm van meerkeuze-vragen afvraagbaar is. Dit bezwaar wordt door sommigen wel eens wat breder uitgemeten dan gerechtvaardigd is; met name wordt wel onderschat wat toetskonstruktors aan werkelijke inzichten van de leerlingen kunnen toetsen. Toch blijft het risico bestaan dat de toepassing van studietoetsen om de resultaten van het onderwijs te evalueren wel eens een negatieve invloed op het onderwijs-leerproces zou kunnen hebben. Dit vermeende risico vormde de vraagstelling waarvan wij in dit onderzoek zijn uitgegaan. Wij hebben ons bepaald tot de invloed die de jaarlijkse eindtoetsingen zouden kunnen hebben op het zesde klas programma in het basisonderwijs. Onze probleemstelling was daarbij: Welke meningen hebben onderwijzers, met betrekking tot het betreffende schooltoets-programma, over de invloed ervan op inhoud en didactiek van het onderwijs in de zesde klas.

## 2. *Situering van het onderzoek*

Sinds 1966 wordt door een groot aantal lagere

\* De uitvoering van het projekt waarbinnen dit onderzoek viel, werd mogelijk gemaakt door de Stichting voor Onderzoek van het Onderwijs SVO, projektnummer 0138.

scholen in ons land deelgenomen aan een centraal gemaakt toetsprogramma bij de overgang naar het voortgezet onderwijs. Deze jaarlijks vernieuwde reeks van opgaven over taal, rekenen en algemene kennis werd aanvankelijk door het toenmalige Nutsseminarium voor Pedagogiek en het R.I.T.P. te Amsterdam verzorgd. Sedert 1970 gebeurt dit door het Centraal Instituut voor Toetsontwikkeling, C.I.T.O., te Arnhem. Over de opzet van dit testprogramma, de uitvoering en de bedoeling ervan is ondermeer geschreven door Deen (1968), Warries (1969), de Samenwerkende Instituten (1967) en C.I.T.O. (1971).

In 1970 en 1971 werd door het R.I.T.P., met steun van S.V.O., een breed opgezet onderzoek uitgevoerd over de functie van dit toetsprogramma in het onderwijs. Daarbij werd gebruik gemaakt van verschillende onderzoeksmethoden waarmee de werking van de schooltoets en de meningen erover in het onderwijs werden gepeild. Dit artikel handelt over een stuk van een schriftelijke enquête die in het voorjaar van 1970 is ingesteld onder alle scholen voor basisonderwijs die in 1969 in 9 verschillende gebieden aan de Schooltoets hadden deelgenomen. Deze enquête had ten doel een inzicht te krijgen in de klachten en meningen over de Schooltoets zoals die door schoolhoofden en onderwijzers van zesde klassen werd ervaren. De enquête bevatte ook – en hier komen wij aan het stuk onderzoek waarover dit artikel gaat – 28 beweringen over de invloed van de Schooltoets op het schoolprogramma. De invullers werd gevraagd aan te geven of en in hoeverre zij het eens waren met het gestelde. De uitkomsten van de analyse van het antwoordmateriaal volgen hieronder, na een korte beschrijving van de opzet van de enquête.

### 3. Opzet en uitvoering van de enquête

De 28 uitspraken waarover gerapporteerd wordt vormden deel van een lijst van 96 beweringen over de Schooltoets. Deze waren merendeels afkomstig uit gesprekken met onderwijzers die kritisch tegenover de toets stonden. Het was de bedoeling van de leden van het onderzoekteam\* vooral de meningen over een veronderstelde negatieve uitwerking van de toets te verkennen. Bij een kategorisering achteraf bleken de negatieve meningen inderdaad te overheersen in de lijst van 28 uitspraken die in Tabel 2 is weergegeven. Er waren 3 positieve uitspraken, 16 negatieve en 9 uitspraken die moeilijk als positief of negatief te klassificeren waren. De 16 negatieve uitspraken zijn door ons later nog onderverdeeld in 6 *algemene* klachten, zoals b.v. 'De toets heeft een negatieve invloed op het onderwijs in de zesde klas' en 10 *specifieke* (maar niet op details van het onderwijs-leerproces ingaande) klachten, zoals b.v. 'Door de Schooltoets is er geen tijd voor de expressievakken'. In Tabel 2 zijn de uitspraken naar hoogte van het antwoordpercentage per groep gerangschikt van hoog naar laag.

De 770 scholen uit in totaal 9 gemeenten (soms met omstreken erbij), die in 1969 deelnamen aan de toets werden alle aangeschreven om medewerking. Van 610 scholen kwamen de enquêtes terug. Deze scholen weken in gemiddelde schooltoets-score niet af van de groep van 770 scholen (Chikwadrattoets 5%). De uitkomsten mogen dus beschouwd worden als geldend voor alle scholen die in 1969 aan de Schooltoets hebben meegedaan. Het respons-percentage van 79% is hoog in vergelijking met wat doorgaans gevonden wordt. Kennelijk had het onderwerp, dat zich als 'klachten en meningen over de Schooltoets' laat karakteriseren, de belangstelling van een zeer groot deel van de schoolhoofden en zesde-klasse-onderwijzers. In totaal kwamen van de

\* Drs. D. J. Bos, J. de Jong-Anema kandidate psychologie, A. K. M. C. Visser-Oerlemans project-assistente, Dr. E. Warries, allen in dienst van het R.I.T.P.

610 scholen 908 bruikbare ingevulde enquête-formulieren terug.

Omdat de lijst van 100 beweringen te lang had geleken om in te vullen, was deze in twee gelijke helften gesplitst en tot twee afzonderlijke versies, A en B, gemaakt. Daarbij is geprobeerd de twee versies inhoudelijk ongeveer parallel te maken. Elke versie bevatte pas aan het eind de (helft van de 28) uitspraken over de invloed van de toets op het programma. Versie A werd door 472 personen ingevuld en versie B door 436 personen. De verdeling over de verschillende gemeenten is in Tabel 1 aangegeven. De gemeente Rijswijk is in berekeningen per gemeente wegens het geringe aantal respondenten niet meegeteld. De volgorde waarin de vragen in de oorspronkelijke lijst voorkwamen was aldus: Versie A: de nummers 20, 5, 7, 14, 19, 8, 3, 10, 28, 11, 1, 21, 25, 22. Versie B: de nummers 24, 6, 4, 15, 18, 9, 2, 16, 27, 12, 13, 17, 26, 23. De veertien uitspraken werden op het enquêteformulier afzonderlijk ingeleid met de woorden: 'Deze laatste vragen gaan uitsluitend over de invloed van de Schooltoets op de inhoud en de didactiek van het basisonderwijs in de zesde klas'. In de hierna volgende beschrijving van uitkomsten worden de aantallen respondenten niet meer vermeld; wij volstaan met de constatering dat elke vraag in totaal door ongeveer 450 personen is beantwoord.

Tabel 1 Aantallen respondenten in de negen populaties

Gemeente:	Afkorting:	Versie A:	Versie B:
Amsterdam	ASD	202	176
Breda	BRE	37	23
Haarlem	HAA	79	75
Heerlen	HEE	66	74
Haarlemmermeer	HMR	31	32
Amstelveen	AVN	20	16
Rijswijk	RIJS	2	4
Enschede	ENS	15	15
Zutphen	ZUT	20	21
Totaal aantal respondenten		472	436

### 4. De analysering van het materiaal

Bij elk van de 28 uitspraken was door de circa

450 respondenten aangegeven of en in hoeverre zij het eens waren met het gestelde. Het antwoord werd gegeven door één van vier antwoordmogelijkheden aan te kruisen: helemaal mee eens, mee eens, oneens, helemaal oneens. Het is gebruikelijk deze vier alternatieven op te vatten als een schaal, bijvoorbeeld lopend van +1,  $+\frac{1}{2}$  naar  $-\frac{1}{2}$  en -1, waarop het antwoord gewogen wordt. Wij hebben zulke gewichtsfactoren niet gebruikt, maar ons in de analyse bepaald tot het tellen van alle personen die 'helemaal mee eens' of 'mee eens' aangekruisd hadden. Op deze manier gaat informatie verloren, maar dit verlies bleek zo gering te zijn\* dat wij ons veilig meenden te kunnen beperken tot deze eenvoudige telling van het aantal vóórstemmers. In Tabel 2 is het totale percentages vóórstemmers van de 9 gemeenten (inclusief Rijswijk) per item weergegeven. In Tabel 4 zijn voor de negatieve uitspraken nog de percentages per gemeente per item vermeld.

In onze analyse van het antwoordmateriaal wordt allereerst ingegaan op de totale antwoordpercentage zoals deze in Tabel 2 voorkomen. Daarna wordt een vergelijking getrokken tussen de gemeenten, op het punt van klachten over de Schooltoets. Verder wordt nog gekeken naar wat in elke gemeente als opvallende reacties op de klachten kan worden beschouwd. In de samenvatting en nabeschuiving wordt ingegaan op de betekenis van de gevonden feiten.

Tabel 2 De 28 uitspraken

(Achter het nummer staat het percentage bevestigende antwoorden.

De veertien nummers met een \* zijn afkomstig uit de enquête versie A; de andere veertien komen uit versie B).

### DRIE POSITIEVE UITSPRAKEN

- 1.\* (42%) Het stillezen komt nu meer aan bod dan vóór de toets

\* De correlatie tussen de aldus berekende itemscores en itemscores na weging van de antwoorden bleek nagenoeg perfect te zijn! Zie R.I.T.P.-memorandum nr. 011 febr. 1972 door E. Warries en Edith van Eck, kandidaat psychologie.

2. (39%) Je krijgt nieuwere opvattingen over het taalonderwijs
- 3.\* (38%) Door de toets krijg je nieuwere opvattingen over het rekenonderwijs

### TIEN SPECIFIEKE KLACHTEN

4. (40%) Er is geen gelegenheid meer om de kinderen te stimuleren iets onder woorden te brengen
- 5.\* (31%) Nieuwere opvattingen over het onderwijs in de zaakvakken worden door de toets tegengehouden
6. (28%) Door de toets weet je niet meer wat je van de zaakvakken moet behandelen
- 7.\* (18%) Een gezamenlijk werkstuk kan nu niet meer gemaakt worden
- 8.\* (14%) Aan het maken van opstellen kom je in de zesde klas niet meer toe
9. (12%) Er is geen tijd meer om de kinderen belangstelling voor kunst bij te brengen
- 10.\* (10%) Door de Schooltoets is er geen tijd voor de expressievakken
- 11.\* (9%) Door de toets hebben wij in onze zesde klas geen tijd voor radio- of televisielessen
12. (5%) Door de toets is er minder tijd voor zangonderwijs
13. (5%) Het zelf lezen van boeken kreeg vóór de Schooltoets meer nadruk in de zesde klas

### ZES ALGEMENE KLACHTEN

- 14.\* (42%) De toets versterkt het intellectualisme in de school
15. (27%) De toets werkt verbalisme in de hand
16. (23%) De toets werkt belemmerend op de vernieuwing in de didactiek
17. (22%) Je remt met deze toets de hoognodige vernieuwingen af
18. (13%) De toets heeft een negatieve invloed op het onderwijs in de zesde klas
- 19.\* (9%) De toets bewerkt een ongezonde wedijver bij de leerlingen

### NEGEN ANDERE UITSPRAKEN

- 20.\* (73%) Als je als onderwijzer gewoon je werk doet, hoef je je van de Schooltoets niets aan te trekken
- 21.\* (69%) Verbetering van het Onderwijs bereik je door de toets niet
- 22.\* (57%) De toets werkt training in de hand

23. (51%) De Schooltoets is een soort eindexamen lagere school
24. (46%) Je voelt je gedwongen om met de opgaven van de vorige jaren of ander toetsmateriaal te gaan oefenen
- 25.\* (42%) Je hebt toch een zekere angst om niet klaar te komen met wat in de toets gevraagd wordt.
26. (31%) De Schooltoets legt een dwang op het programma in het laatste jaar
27. (20%) De toets oefent in de vijfde klas invloed uit op het programma
- 28.\* (6%) Door de toets krijg je een toename van doubleren vóór de zesde klas.

5. *De totale antwoordpercentages: enkele uitkomsten*

De antwoordpercentages die in Tabel 2 tussen haakjes vermeld zijn, hebben een gemiddelde waarde van 26 en een standaardafwijking van 19. De mediane waarde is 27,5. Op grond van de statistische verdeling van de als itemscores opgevatte antwoorden van de tweemaal circa 450 onderwijzers valt het volgende op:

- De drie positieve uitspraken scoren alle drie hoger dan de mediaan.
- De 9 'andere' uitspraken scoren merendeels, n.l. 7 stuks, hoger dan de mediaan.
- De 16 klachten scoren merendeels, n.l. 12 stuks, lager dan de mediaan.
- Het verschil in beantwoording tussen de 16 negatieve en de 12 positieve + neutrale uitspraken is, gemeten met een chi-kwadraattoets significant ( $P < 0,0005$ ).
- De sterkst onderschreven uitspraken (twee standaardafwijkingen boven het gemiddelde) zijn de 'neutrale' nummers 20 en 21, waarin gezegd wordt dat 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 uitspraken (één standaardafwijking boven het gemiddelde) zijn de 'neutrale' nummers 22, 23 en 24 die alle drie betrekking hebben op hetzelfde algemene onderwerp: de toets opgevat als examen waarvoor gewerkt moet worden.
- Sterk afgewezen uitspraken (één standaardafwijking beneden het gemiddelde) zijn de negatieve nummers 21, 13 en de 'neutrale' 28: Een mogelijke negatieve invloed van de toets op het zangonderwijs en op het zelf lezen van boeken wordt afgewezen. Ook wordt ontkend dat de toets het doubleren in de vijfde klas zou bevorderen.
- Binnen de groep van zestien negatieve uitspraken zijn vijf klachten die door meer dan 25% worden onderschreven. De kritiek van meer dan 25% van de respondenten richt zich op de invloed op de zaakvakken (5 en 6: onzekerheid en tegengaan van vernieuwing), op 'intellectualisme' (14) en 'verbalisme' (15) en op de gestelde onmogelijkheid de kinderen zelf iets onder woorden te laten brengen (4).

6. *Globale verschillen tussen de gemeenten*

In Tabellen 3 en 4 zijn per gemeente de antwoordpercentages op de drie positieve en de zestien negatieve uitspraken gegeven. Nadat de antwoorden op de positieve uitspraken gespiegeld waren, hebben wij allereerst nagegaan of globaal gesproken er verschil bestaat in de mate van bevestiging of ontkenning van deze 19 uitspraken. Vervolgens hebben wij gekeken hoe die verschillen dan lagen: welke gemeente heeft meer klachten dan anderen\*.

\* Beide analyses werden uitgevoerd door Edith van Eck, candidate psychologie.

Tabel 3 Antwoordpercentages op drie positieve uitspraken

Nr.	ASD	BRE	HAA	HEE	HMR	AVN	ENS	ZUT
1.	45	32	35	39	65	40	40	45
2.	28	48	48	53	31	30	20	57
3.	36	38	41	44	49	20	40	35

Tabel 4 Antwoordpercentages op 16 negatieve uitspraken

Nr.	TIEN SPECIFIEKE KLACHTEN	ASD	BRE	HAA	HEE	HMR	AVN	ENS	ZUT
4	zelf iets onder woorden	44	35	36	43	44	44	<u>13</u>	38
5	zaakvakken	31	27	19	44	45	30	<u>47</u>	<u>10</u>
6	zaakvakken	30	35	19	39	34	19	<u>7</u>	<u>10</u>
7	gezamenlijk werkstuk	19	24	13	15	23	55	<u>40</u>	10
8	opstellen	<u>12</u>	11	11	20	19	15	<u>13</u>	10
9	belangstelling voor kunst	<u>16</u>	13	7	10	13	6	7	<u>14</u>
10	expressievakken	14	5	9	8	<u>16</u>	5	7	<u>0</u>
11	radio- of televisielessen	12	5	9	6	<u>3</u>	5	<u>27</u>	<u>0</u>
12	zangonderwijs	9	0	8	3	0	0	0	0
13	zelf lezen van boeken	5	4	5	5	6	0	0	0
	ZES ALGEMENE KLACHTEN								
14	intellectualisme	44	46	27	<u>38</u>	61	50	47	35
15	verbalisme	33	26	<u>24</u>	<u>24</u>	28	19	13	19
16	vernieuwing in de didactiek	<u>35</u>	9	16	11	25	25	7	10
17	hoognodige vernieuwingen	32	26	11	11	31	31	<u>0</u>	19
18	negatieve invloed	21	4	7	4	19	13	<u>7</u>	10
19	ongezonde wedijver	11	5	<u>5</u>	<u>3</u>	10	<u>20</u>	7	5
	10 × gemiddelde	230	172	141	177	236	211	151	119
	10 × standaard afwijking	126	141	88	151	166	172	161	114

Per uitspraak werden daartoe aan de acht gemeenten rangordegetallen van 1 tot 8, toegekend voor de mate waarin de uitspraak bevestigd werd. Zo had in Tabel 4 item 14 de volgende ranggetallen op zijn rij: 4, 5, 1, 3, 8, 7, 6, 2. Per gemeente werd het gemiddelde van deze 19 rangordegetallen berekend. Met behulp van de Friedman-toets werd nagegaan of deze gemiddelde rangnummers ongeveer gelijk waren. Dit was niet het geval (Chikwadraat .99 = 18,475), zodat wij aangenomen hebben dat er verschil tussen de gemeenten bestaat.

Om uit te vinden *hoe* dat verschil tussen de gemeenten lag, is voor elk tweetal van gemeenten het gemiddelde verschil tussen de antwoordpercentages op de items getoetst met een t-toets voor afhankelijke steekproeven\*. In tabel 5 is de uitkomst van deze 28 vergelijkingen in kaart gebracht. Als de tabel van links naar rechts wordt gelezen, dan betekent een minteken dat de gemeente van de horizontale rij negatiever tegen-

$$* t = \frac{\sum d \sqrt{n-1}}{S.A.}$$

Tabel 5 Paarsgewijze vergelijking van de gemeenten

	ASD	BRE	HAA	HEE	NMR	AVN	ENS	ZUT
ASD	0	—*	—**	—*	—	—	—	—**
BRE	+*	0	—	—	+	+**	—	—
HAA	+**	+	0	+*	+*	+*	+	—
HEE	+*	+	—	0	+	+	+	—
HMR	+	—	—*	—	0	+	+	—**
AVN	+	—**	—*	—	—	0	—	—**
ENS	+	+	—	—	—	+	0	—
ZUT	+**	+	+	+	+**	+**	+	0

over de toets staat (meer klachten sterker onderschrijft) dan de gemeente waarvan de naam boven de betreffende kolom staat. Een plusteken staat voor het tegendeel. Eén ster betekent significantie op 5% niveau, twee sterren op 1% niveau. Zo onderschrijft bijvoorbeeld Haarlem minder klachten dan alle andere gemeenten behalve Zutphen: zeer significant is dit in vergelijking met Amsterdam, en significant in vergelijking met Haarlemmermeer en Amstelveen.

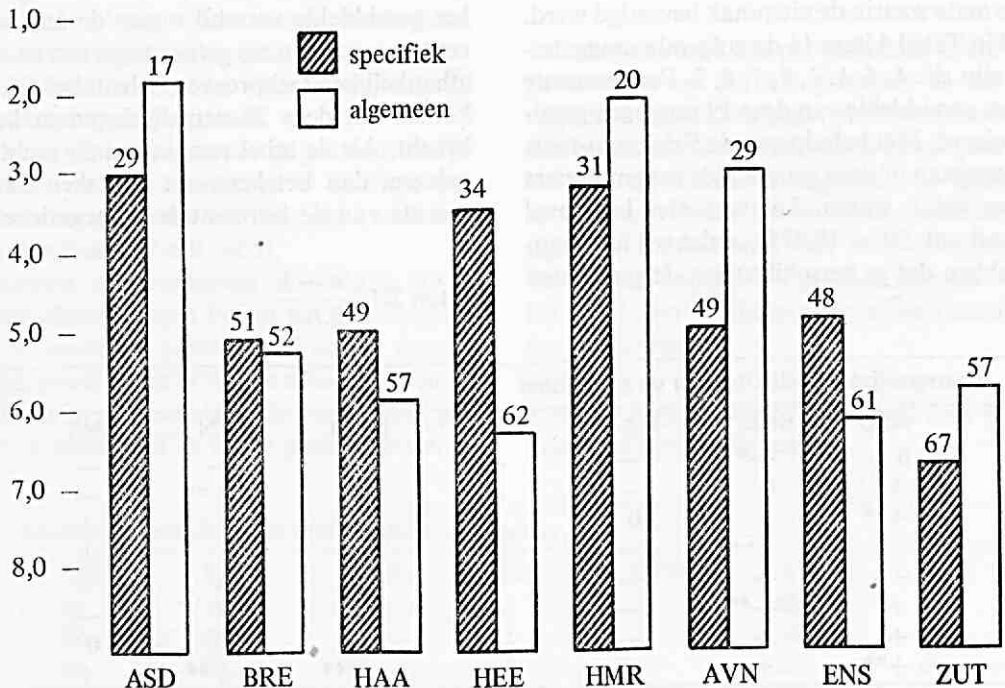
De resultaten van deze 28 vergelijkingen spreken voor zichzelf in de tabel. Het meest opvallend is de relatief negatieve opvatting in Amsterdam en Amstelveen en de relatief positieve houding in Zutphen.

7. Verschillen op de algemene en de specifieke klachten

Wij hebben 6 algemene negatieve uitspraken en 10 specifieke negatieve uitspraken onderscheiden. Dit leek zinvol omdat het basisonderwijs misschien in sommige gemeenten geen algemene

bezwaren had tegen de Schooltoets maar wél, door bijvoorbeeld gebrekkige organisatie of slechte aanpassing bij de lokale situatie, specifieke bezwaren tegen de toets onderschreef. Ook is het andersom denkbaar: geen specifieke klachten maar wél een algemene negatieve houding. Per gemeente is het gemiddelde berekend van de respectievelijk 6 en 10 rangordegetallen die waren ontstaan door de hierboven reeds genoemde toekenning van rangordegetallen van 1 tot 8 per uitspraak. Deze rangordegetallen gebaseerd op de in Tabel 4 vermelde percentages ja-zeggens op de 6 en de 10 uitspraken bepalen de hoogte van de kolommen in figuur A. Een laag rangordegetal, omgezet in een hoge kolom, wijst op een relatief sterkere onderschrijving van de klachten; men staat als het ware vooraan in de rij van klagers. In de figuur zijn de gemiddelde rangordegetallen, met 10 vermenigvuldigd, boven de kolommen gezet. De gearceerde kolom geeft de gemiddelde rangorde op de specifieke klachten aan. Het verschil in hoogte van de twee kolommen per gemeente zou men op kunnen vatten als een indicatie van de 'gerichtheid' of

Fig. A Het onderschrijven van specifieke en algemene klachten. Gemiddelde rangorde in 8 gemeenten.



'zakelijkheid' van de kritiek van de respondenten: hoe hoger de gearceerde kolom in vergelijking met de andere, des te zakelijker is de kritiek.

Uit de gegevens in Figuur A blijkt onder meer het volgende:

- a. Bij de algemene klachten zijn er drie geografisch dicht bij elkaar liggende gebieden, Amsterdam, Amstelveen en Haarlemmermeer, waar het rangordegetal ver 'boven' het gemiddelde (van 4,5 bij acht gemeenten) ligt.
- b. Bij de specifieke klachten is vooral Zutphen ruim onder het gemiddelde.
- c. Heerlen en Enschede hebben de meest 'zakelijke' kritiek, Amstelveen de minst 'zakelijke'.

8. *De negatieve uitspraken: enkele uitkomsten per gemeente*

In het voorgaande zijn uitkomsten besproken van het totale antwoordmateriaal en van enkele verschillen tussen de gemeenten. Wij kijken nu nog in elke gemeente naar de meest opvallende klachten, daarbij rekening houdend met het antwoordpatroon uit de andere gemeenten. Als bijvoorbeeld overal de negatieve uitspraak nummer 12 over de invloed van de Schooltoets op het zangonderwijs door de onderwijzers ontkend wordt, heeft het geen zin de aandacht te vestigen op het zeer lage percentage (4%) in de gemeente Breda. Wèl bijvoorbeeld is het de moeite waard op uitspraak 15 over verbalisme in Haarlem te wijzen want deze uitspraak blijkt, als we het uitrekenen, in Haarlem relatief sterker onderschreven te worden dan dit elders het geval is.

We hebben door een bepaalde rekenwijze de opvallende uitspraken per gemeente geïdentificeerd. De berekening die wij daartoe hebben uitgevoerd is als volgt. Voor elke gemeente is over de 16 klachten-percentages het gemiddelde en de standaardafwijking berekend. Elk percentage is vervolgens omgezet in een z-score (het aantal standaard-afwijkingen dat dit percentage van het gemiddelde verwijderd ligt) voor die gemeente. Als zeer opvallend (dubbele onderstreping in Tabel 4) hebben we beschouwd de z-score met de grootste afwijking van het gemiddelde op die rij, mits dit er niet meer dan één was. De z-score of

twee scores die daar op volgde(n) hebben wij ook nog als opvallend (één onderstreping in Tabel 4) beschouwd. De aldus gedefinieerde opvallende uitspraken zijn, per gemeente, de volgende(n):

In *Amsterdam* zijn de onderwijzers optimistischer dan elders over het maken van opstellen waaraan 'je niet meer toe komt', maar ze zijn somberder over de vernieuwing in de didactiek, waarop de toets belemmerend zou werken.

In *Breda* heeft men sterker dan elders de neiging krachtig te onderschrijven dat je door de toets niet meer weet wat je met de zaakvakken aanmoet.

In *Haarlem* zegt men meer dan elders dat de toets verbalisme in de hand werkt. De onderwijzers in Haarlem zijn verder niet zo bang voor ongezonde wedijver bij de leerlingen.

In *Heerlen* vindt men minder dat de toets het intellectualisme in de school versterkt. Ook hier is men niet zo bang voor ongezonde wedijver.

In *Haarlemmermeer* vinden meer respondenten dat er door de toets tijd van de expressievakken afgaat.

In *Amstelveen* vreest men meer dan elders de ongezonde wedijver door de toets en ook is men daar banger dat een gezamenlijk werkstuk nu niet meer gemaakt kan worden.

In *Enschede* ziet men nog wèl gelegenheid om de kinderen te stimuleren zelf iets onder woorden te brengen. De onderwijzers zijn hier ook minder bevreesd dan elders dat door de toets de hoognodige vernieuwingen worden afgeremd of dat nieuwere opvattingen in de zaakvakken worden tegengehouden. Wèl is men meer dan elders van mening dat je door de toets niet meer weet wat je van de zaakvakken moet behandelen.

In *Zutphen* wordt wat minder dan elders door de onderwijzers een slechte invloed onderschreven op de vernieuwing van de zaakvakken en de beschikbare tijd voor de expressievakken. Men lijkt iets pessimistischer dan elders over de tijd die nu overblijft om de kinderen in de zesde klas belangstelling voor kunst bij te brengen.

## 9. Samenvatting

Dit artikel bevat het verslag van een onderzoek naar de invloed van de Amsterdamse Schooltoets op de inhoud en didactiek van het onderwijs in de zesde klas van het basisonderwijs. Dit gebeurde door de meningen van onderwijzers te vragen. Er is daartoe een schriftelijke enquête ingesteld onder onderwijzers die 770 scholen representeerden in negen gemeenten. Zij gaven hun opinie over 28 uitspraken betreffende de invloed van de toets op het onderwijs.

Uit de uitgevoerde berekeningen met het antwoordmateriaal bleek het volgende.

De positief gekleurde en neutrale uitspraken werden door de onderwijzers sterker onderschreven dan de negatieve beweringen. Een positieve invloed, in die zin dat van verbetering van het onderwijs kan worden gesproken, onderkennen de meeste respondenten niet. Men heeft sterk de neiging het toetsprogramma als een externe zaak te beschouwen waarvoor je als onderwijzer traint – hoewel dat niet zou moeten – en waar je in je zesdeklas-programma naar toe werkt. De druk van de toets is overigens niet zo groot, dat er geen tijd meer zou zijn voor iets anders dan rekenen of taal.

Wat de zaakvakken betreft, is er onzekerheid door de toets, soms ook een verwijt van conservatisme. Sommige onderwijzers zijn bang voor 'intellectualisme' en 'verbalisme' in de school door het toetsprogramma.

Globaal gesproken bleek er verschil te bestaan tussen de negen gemeenten in de mate van bevestiging van de voorgelegde positieve en negatieve uitspraken. Het meest opvallend was de (relatief) negatieve opvatting in Amsterdam en Amstelveen en de (relatief) positieve houding in Zutphen. De laatste gemeente heeft vooral (relatief) weinig specifieke klachten, in Amsterdam, Amstelveen en ook Haarlemmermeer overheersen vooral de algemene klachten. Een enkele gemeente verschilde van anderen op het punt van de 'zakelijkheid' van de kritiek.

In elke gemeente werd nog, rekening houdend met antwoordpatronen elders in het land, nagegaan welke klachten over de invloed van de toets het sterkst onderschreven werden.

## 10. Discussie

De resultaten van dit kleine onderzoek onder een grote groep onderwijzers moeten geïnterpreteerd worden met het oog op twee dingen: (a) de beperktheid van de objectieve toetsmethode, die het voorwerp van de kritiek der ondervraagden was en (b) het groter kader van wat Tyler (1949) heeft betiteld als de fundamentele uitgangspunten van leerplan en onderwijs.

Het toetsprogramma waarover het hier gaat, tegenwoordig als 'Schooltoetsen Basisonderwijs' door het C.I.T.O. georganiseerd, heeft van nature de enorme beperking dat het uitsluitend datgene meet wat met vierkeuze-vragen te meten valt en slechts handelt over wat onder de gangbare wetgeving gevraagd kan worden bij de overgang naar V.W.O. en Havo. Het is goed bij de discussie hier rekening mee te houden. De makers van het toetsprogramma pretenderen ook niet dat ze een evaluatie instellen naar het bereiken van alle mogelijke doelen van de lagere school.

Het groter kader waarin het onderzoek geplaatst moet worden behoort eigenlijk door de lezer zelf te worden aangebracht. Vragen naar wat de doelen van het basisonderwijs zijn en hoe we die moeten bereiken, zijn hier niet te beantwoorden. Wel moeten de uitkomsten van de berekeningen gezien worden tegen de achtergrond van het didactisch proces, waarvan bijv. ten onzent Van Gelder (1967) een theoretische analyse heeft gegeven. Wij zullen op zijn minst er rekening mee dienen te houden dat de bepaling van de resultaten, of kortweg de *evaluatie*, niet alleen in wisselwerking staat met het onderwijsleerproces, of kortweg de *instructie*, maar ook met datgene wat het bevoegd gezag, de onderwijzers en de ouders met de school nastreven, of kortweg de *doelen*.

We hebben nu de meningen van onderwijzers gevraagd over de invloed van een bepaalde, steeds meer in zwang rakende evaluatievorm. We hebben die invloed geoperationaliseerd op een manier die een sterke aansluiting bij de bestaande opvattingen van onderwijzers garandeert. De meningen zijn verder representatief voor de scholen in negen grote gebieden. Wat hebben



onze inspanningen nu opgeleverd en wat doen we met het resultaat? Wij zijn gekomen tot de volgende vijf conclusies.

1. We concluderen dat de stemming onder het onderwijspersoneel en de hoofden niet zo somber is als wel eens verondersteld is in het verleden. De opvattingen over een slechte invloed nemen, blijktens onze berekeningen, niet de grootste plaats in. Maar . . . ze zijn er wel. Zolang meer dan één op de tien onderwijzers serieuze klachten onderschrijft, behoren zowel de toetsconstructeurs als het bevoegd gezag ongerust te blijven en over verbetering te blijven denken.  
Dat overigens de toetsconstructeurs zich aanhoudend zorgen maken over de inhoudelijke aspecten van dit programma moge blijken uit het onderzoek dat door mevrouw Peters-Sips (1970) werd ingesteld naar meningen over de opgaven en uit de jaarlijkse verantwoording die door het C.I.T.O. (1969) tevoren openbaar wordt gemaakt.
2. Hoewel het onderzoek research-technisch niet tekort schiet, zou men toch willen dat er wat meer naar de details was gevraagd. Wat denken de onderwijzers over de verschillende onderdelen van de toets, zoals bijv. geanalyseerd voor taal door Wesdorp (1969-70), voor rekenen door Stroomberg (1971) en voor algemene kennis door Sandbergen (zonder jaartal). Het blijft natuurlijk de vraag of eenieder in het onderwijs wel duidelijke meningen heeft over de invloed van *delen* van de evaluatie op *delen* van de instructie. In het vooronderzoek dat leidde tot het opstellen van de door ons gebruikte uitspraken, hebben wij niet zulke detailopvattingen gevonden. Misschien is het daarom beter om niet méér en verfijnder aandacht te schenken aan de invloed van de evaluatie op de instructie, maar de nadruk te leggen op de opvattingen over de doelstellingen, tot in details, van het onderwijs in de verschillende vakken.
3. Uit het onderzoek bleek ondermeer angst voor de vermeende invloed van de toets in de richting van 'verbalisme' en 'intellectualisme'.

Ook dit is een zaak die – als we het goed begrijpen – te maken heeft met de wisselwerking tussen doelstellingen en zowel instructie als evaluatie. Als inderdaad de respondenten met deze twee woorden bedoelen dat het verbaal en intellectueel functioneren van het kind in de cultuur in de toets teveel aandacht krijgt ten koste van andere vaardigheden, dan is dit een ernstige zaak. Het zou dan beslist noodzakelijke zijn om – gegeven nu eenmaal de invloed van een 'examen' (en de schooltoets wordt door zeker de helft van de onderwijzers als zodanig opgevat) – de eind-evaluatie van het basisonderwijs *breder* op te zetten en ook, zoals bijv. De Groot (1971) suggereert, ander eindgedrag te toetsen – zij het dan minder objectief.

4. Uit ons onderzoek is gebleken dat er nogal wat verschillen tussen de gemeenten onderling zijn in de mate waarin men algemene, nogal vage, klachten en specifieke klachten over de invloed van de toets onderschrijft. Voor een deel zullen de verschillen in reactie op deze klachten, zeker als ze gericht zijn op een bepaald specifiek aspect van het toetsprogramma, te maken hebben met verschillen in door-denking van de problemen en met plaatselijk op gang gekomen discussies in het onderwijs. Men zou kunnen zeggen dat veel specifieke klachten een teken zijn van de bewustwording van de onderwijzers. Aan de andere kant moet er rekening mee gehouden worden, en dit is dan een zaak voor het bevoegd gezag, dat sommige van de gevonden verschillen tussen gemeenten waarschijnlijk op een andere manier verklaard kunnen worden. Men zou bijvoorbeeld kunnen zeggen dat de relatief negatieve instelling in Amsterdam en omstreken te maken heeft met de omstandigheid dat dáár de toets het eerste geïntroduceerd is en gemaakt is: een profeet wordt in zijn eigen land nu eenmaal niet geëerd. Maar evenzeer kan men de globale verschillen tussen gemeenten op een andere manier verklaren: door de kwaliteit van de introductie van het toetsprogramma aan het onderwijzend personeel en de hoofden, door de gevoerde overlegproce-

dure en door de didactische begeleiding die men geeft aan de scholen. Het is altijd eenvoudiger een extern examen in de scholen te brengen dan een algemeen aanvaard en integreerbaar stuk evaluatie te introduceren. Hier ligt, de eer van de profeten terzijde gelaten, duidelijk een taak voor de lokale overheden.

5. Tenslotte komen we nog even terug op het vermeende verbalistische, intellectualistische karakter van de Schooltoets. (Wij gaan er even van uit dat met de eerste term *niet* bedoeld wordt dat de Schooltoets naar niet-gefundeerde en oppervlakkige kennis vraagt, naar namen en niet naar inhoud en inzichten). Het is van belang er rekening mee te houden dat dit Toetsprogramma niet alleen een *evaluerende* functie heeft maar ook wat ik eerder (Warries 1971) heb genoemd een *selecterende* taak. Het zal goed zijn als diegenen die de opgaven samenstellen, en ook diegenen die daar van gebruik maken, steeds bedenken dat de huidige verschijningsvorm van de toets op de historische realiteit van deze twee functies teruggaat. Het is niet onmogelijk dat schriftelijk rekenen en taal mede zoveel nadruk krijgen omdat ze goed te gebruiken zijn om succes in het algemene voortgezette onderwijs te voorspellen. Het is het overdenken waard of, door de instelling van het brugjaar in het voortgezet onderwijs, de selectieve functie van de Schooltoetsen Basisonderwijs langzamerhand op de achtergrond mag raken. Het zou dan verder een mooi resultaat van rationele doordenking van het onderwijsbeleid zijn als, op grond van een op gang gebrachte meningsvorming over de doelstellingen van het basisonderwijs, de huidige evaluatie-procedure sterk verbreed zou worden en ook de beoordeling van niet goed objectieveerbare vormingsdoelen zou omvatten.

#### Literatuur

- C.I.T.O., Schooltoetsen Basisonderwijs. Verantwoording inhoudstoets 1970. *CITO-uitgave* dec. 1969.
- C.I.T.O., Verslag schooltoetsen basisonderwijs 1970 *Cito-publicatie* nr. 10, februari 1971.
- Deen, N., Gebruiksmogelijkheden van Schooltoetsen. *Syllabus Nutsseminarium voor Pedagogiek*, 1968.
- Gelder, L. van, Leerplanontwikkeling en de analyse van het didactisch proces. *Pedagogische Studiën*, 44, 6, 1967.
- Groot, A. D. de, *Standpunt over onderwijs, democratie en wetenschap*, Mouton 1971, blz. 63.
- Peters-Sips, M. W. H., Verslag van de enquête 'Inhoud Schooltoets 1970', Arnhem, CITO, 1970. *CITO Memo*, nr. 5.
- Samenwerkende Instituten  
*Amsterdamse Schooltoetsen* no. 3 in de serie Empirische Studies over Onderwijs, Wolters, Groningen, 1967.
- Sandbergen, S., Het toetsen van 'algemene kennis' op de basisschool. *Intern Rapport R.I.T.P.* z.j.
- Stroomberg, H. P., De doelstellingen van het rekenonderwijs. *Discussienota R.I.T.P.*, augustus 1971.
- Tyler, Ralph W., *Basic Principles of Curriculum and Instruction*. Univ. of Chicago 1949, 30th Impression 1970.
- Warries, E., Proefwerken met vierkeuze-vragen voor het basisonderwijs. *Pedag. Stud.* 46, 1, 22-41 en 46, 3, 161-165, 1969.
- Warries, E., Drie redenen om te toetsen in het onderwijs. *Ped. Stud.* 48, 152-161, 1971. Ook verschenen in *Persoon en Gemeenschap* 23, 6, 272-282, 1971.
- Wesdorp, H., Het moedertaalgedeelte uit de Amsterdamse Schooltoets; een overzicht na vijf jaar. In: *Moer*; tijdschrift voor het onderwijs in het Nederlands. Jaargang 1969/70, nr. 2, blz. 49-59.