

Periodiek peilingsonderzoek: de opzet van het project PPON geïllustreerd aan de rekenpeiling einde basisonderwijs

J. M. WIJNSTRA

*Instituut voor Toetsontwikkeling (Cito),
Arnhem*

Samenvatting

In 1986 is de voorbereiding van het project Periodieke Peiling van het Onderwijsniveau (PPON) gestart om, voorlopig beperkt tot het basisonderwijs, periodiek inhoudelijke gegevens over de stand van zaken te verzamelen. Deze gegevens moeten een basis verschaffen voor een discussie over de inhoud en het niveau van het onderwijs.

Na de inleiding worden in deze bijdrage eerst het doel en de achtergrond van het project toegelicht in relatie tot het recente beleid met betrekking tot de eindtermen. Daarna komen de organisatie, planning en werkwijze van het project in het basisonderwijs aan de orde. De werkwijze wordt in par. 4 geïllustreerd aan de eerste, in 1987 uitgevoerde peiling, de rekenpeiling einde basisonderwijs. Hierin wordt de nadruk gelegd op de bevordering van de mogelijkheden om op basis van de rapportage een discussie te voeren over de inhoud en het niveau van het onderwijs.

1 Inleiding

Wat moeten kinderen leren op school? En leren ze dat ook werkelijk? Over deze en soortgelijke vragen is een eindeloze discussie mogelijk, juist door het gebrek aan inhoudelijke gegevens over het onderwijs. Om aan de discussie over de inhoud en het niveau van het onderwijs een meer rationele basis te verschaffen is in 1986 in opdracht van de Minister van Onderwijs en Wetenschappen de voorbereiding van het project Periodieke Peiling van het Onderwijsniveau (PPON) van start gegaan om – voorlopig beperkt tot het basisonderwijs – periodiek inhoudelijke gegevens

over de stand van zaken te verzamelen.

Hoewel er ook relaties zijn met de voorziene eindtermen basisonderwijs, is het project daarvan geen uitvloeisel. De initiatieven tot het opzetten van periodiek peilingsonderzoek dateren al van het eind van de jaren zeventig. Onder verwijzing naar een advies uit de SVO-nota Opstaan tegen het Zittenblijven (Doornbos, 1969) heeft de Innovatiecommissie Basisonderwijs (ICB) reeds in 1978 gepleit voor een periodiek onderzoek naar de kwalitatieve ontwikkeling van het basisonderwijs, waarbij de nadruk zou moeten liggen op de inventarisatie van het leer- en vormingsaanbod (ICB, 1978). In de beleidsnotitie bij dit innovatieplan heeft de toenmalige minister positief gereageerd op deze aanbeveling. Hoewel er reeds eind 1979 een technisch vooroverleg over dit onderwerp werd gestart met de SLO, de SVO en het Cito om de mogelijkheden te verkennen, heeft het tot juni 1985 geduurd voor de beleidsnotitie verscheen die ten grondslag ligt aan de opdracht om peilingsonderzoek in het onderwijs te gaan uitvoeren. In die tussentijd verscheen de nota Kwaliteit van het Onderwijs (1981), terwijl in verschillende memories van toelichting bij het onderwijshoofdstuk uit de rijksbegroting aandacht werd besteed aan de kwaliteitsproblematiek. Aan de verstrekking van de opdracht ging verder een door de Stichting Centrum voor Onderwijsonderzoek (SCO) te Amsterdam uitgevoerde voorstudie op het gebied van het moedertaalonderwijs vooraf waarin de haalbaarheid van periodiek peilingsonderzoek werd verkend (zie Wesdorp, 1985; Wesdorp, Van den Bergh, Bos, Hoeksma, Oostdam, Scheerens & Triescheijn, 1986).

In deze bijdrage wordt de opzet van het project PPON in het basisonderwijs geschetst¹. Na een beschrijving van het doel en de achtergrond van het project worden de organisatie, planning en werkwijze toegelicht. Bij wijze van illustratie wordt de bijdrage afgerond met een beschrijving van de eerste uitgevoerde peiling, de rekenpeiling einde basisonderwijs. Daarbij wordt de nadruk gelegd op de bevordering van de mogelijkheden om op basis van de rapportage een discussie te

voeren over de inhoud en het niveau van het onderwijs.

2 Doel en achtergrond van PPON

In de beleidsnotitie van juni 1985 die aan de opdracht ten grondslag ligt, wordt gesignaleerd dat de aandacht voor het vraagstuk van de kwaliteit van het onderwijs de laatste jaren sterk is toegenomen. Tegelijkertijd wordt vastgesteld 'dat de onderwijskundige en maatschappelijke discussies over het niveau en de kwaliteit van het onderwijs worden bemoeilijkt door het veelal ontbreken van empirische gegevens over de resultaten van het onderwijs'. Om in deze leemte te voorzien is daarom aan de Stuurgroep PPON, waarin aanvankelijk het Instituut voor Onderzoek van het Onderwijs (SVO) en het Instituut voor Toetsontwikkeling (Cito) samenwerkten, opdracht gegeven peilingsonderzoek te gaan uitvoeren, te beginnen in het basisonderwijs.

In de aangehaalde beleidsnotitie wordt het algemene doel van peilingsonderzoek omschreven als 'systematisch bij te dragen tot het verkrijgen van een beeld van het leeraanbod en de effecten van het onderwijs' met als voornaamste functies:

- het bieden van een referentiekader voor de school voor de beoordeling van de eigen doelstellingen en resultaten;
- het leveren van indicaties voor het verrichten van nadere (diepte-)studies en evaluatie-onderzoek die van belang kunnen zijn voor de kwaliteitsbeoordeling en kwaliteitsbevordering van het onderwijs;
- het bijdragen aan de ontwikkeling van toetsen voor het gebruik door scholen, die een vergelijking met landelijke trends mogelijk maken;
- het bieden van een (minimale) empirische basis voor de meer algemene maatschappelijke discussie over kwaliteit en niveau van het onderwijs.

PPON moet in principe het gehele curriculum van het basisonderwijs bestrijken. De peilingsonderzoeken richten zich voorlopig primair op de stand van zaken aan het einde van het basisonderwijs, met uitzondering van Nederlandse taal en rekenen waarvoor ook peilingsonderzoeken halverwege het basisonderwijs worden uitgevoerd. Peilingsonderzoek is in principe steekproefonderzoek, dit in tegen-

stelling tot bijvoorbeeld de bestandstellingen van het Centraal Bureau voor de Statistiek en het Ministerie van Onderwijs en Wetenschappen. Het is de taak van het project PPON op basis van gegevensverzameling bij steekproeven van scholen en leerlingen per leer- en vormingsgebied, periodiek herhaald, een algemeen inzicht te verschaffen in het actuele leer- en vormingsaanbod, de onderwijsleerresultaten en de relaties daartussen. Daarbij dient ook aandacht te worden geschonken aan relaties van deze variabelen met andere relevante kenmerken van scholen en leerlingen, zodat in de rapportage relevante subgroepen afzonderlijk kunnen worden beschreven. Het is echter uitdrukkelijk niet de bedoeling - en in de gekozen opzet ook niet goed mogelijk - uitspraken over individuele scholen of leerlingen te doen. Wel levert PPON een referentiekader voor de zelfevaluatie van scholen en wordt nagegaan in hoeverre de peilingsinstrumenten voor zelfevaluatie kunnen worden ingezet.

Gezien het tijdstip waarop de beleidsnotitie tot stand kwam, kon in de weergegeven opsomming van functies nog niet worden verwezen naar de (toekomstige) eindtermen basisonderwijs, waarmee uiteraard ook relaties zullen gaan bestaan als het wetsvoorstel dienaangaande wordt aangenomen. Volgens de Memorie van Toelichting bij dit wetsvoorstel zijn eindtermen 'beschrijvingen van kwaliteiten van leerlingen op het gebied van kennis, inzicht en vaardigheden die de school bij haar onderwijsactiviteiten in elk geval als doelstelling moet hanteren. Eindtermen maken deel uit van de kwaliteitseisen die aan het basisonderwijs (...) worden gesteld. In die zin zijn eindtermen richtlijnen voor scholen voor de na te streven resultaten van leerlingen'. Gezien deze formulering zijn eindtermen op te vatten als een normstelling in het (cyclische) proces van kwaliteitsbeheersing (vgl. de nota Kwaliteit van het Onderwijs, 1981) en worden deze ook van belang voor de inhoud van de peilingsonderzoeken.

Door dit gewijzigde kader, dat enkele jaren geleden - althans in het basisonderwijs - nog moeilijk voorstelbaar was (vgl. Wijnstra, 1982) en zijn oorsprong vindt in het WRR-rapport over de basisvorming (WRR, 1986), wordt (impliciet) de functie van PPON ook verruimd, in die zin dat het project (mede) een functie krijgt in de landelijke kwaliteitsbe-

heersing van het onderwijs. In de terminologie van de aangehaalde nota vormen de peilingsonderzoeken een instrument om op het niveau van het onderwijssysteem de mate van normrealisering waar te nemen (de tweede fase in het proces van kwaliteitsbeheersing).

Dit betekent dat in de toekomst de eindtermen in elk geval in de peilingsinstrumenten vertegenwoordigd moeten zijn om een afweging en waardering van de bevindingen in relatie tot de normstelling (de derde fase in het proces van kwaliteitsbeheersing) mogelijk te maken. Hierdoor vindt overigens tegelijkertijd een concretisering van de eindtermen plaats, zodat ook een oordeel over de haalbaarheid daarvan mogelijk is. Men kan zich in dit verband zelfs afvragen of het formuleren van eindtermen zinvol is als deze niet operationeel gedefinieerd worden. Peilingsonderzoek behoeft zich echter niet te beperken tot de eindtermen. Als dat het geval was, zou een discussie over ontwikkelingen in het onderwijs nodeloos worden gefrustreerd. Tegelijkertijd wordt ook het gevaar van een conserverend effect voorkomen. Juist met het oog op een discussie over nieuwe ontwikkelingen is het van belang ook zaken die nog geen gemeengoed zijn, maar wel al een behoorlijke ingang hebben gevonden, in de peilingen op te nemen. Dit principe is ook het uitgangspunt bij de voorbereiding van de peilingsonderzoeken in de periode dat er nog geen officiële eindtermen zijn.

PPON levert inhoudelijke beschrijvingen van de stand van zaken op basis waarvan een gefundeerde discussie over de inhoud en het niveau van het onderwijs gevoerd kan worden en eventueel maatregelen kunnen worden getroffen ter handhaving, respectievelijk bijstelling van de normen, zoals de vierde fase in het proces van kwaliteitsbeheersing wordt genoemd in de nota *Kwaliteit van het Onderwijs* (1981). Het waarden van de bevindingen en het doen van aanbevelingen maken echter geen deel uit van de opdracht. PPON rekent het echter wel tot zijn taak de discussie over de peilingsuitkomsten te bevorderen.

3 *Organisatie, planning en werkwijze*

3.1 *Organisatie*

Oorspronkelijk is de opdracht tot het uitvoeren van het peilingsonderzoek verleend aan een samenwerkingsverband van SVO en Cito, die daartoe de Stuurgroep PPON hadden gevormd. Voorzien was dat het Cito de periodiek herhaalde peilingen zou uitvoeren en dat via SVO dieptestudies naar aanleiding van bevindingen uit de peilingen zouden worden uitgevoerd. Voor het geheel zou één budget beschikbaar zijn dat door SVO zou worden beheerd. Deze taakverdeling bestaat nog steeds, de vorm van het samenwerkingsverband en het financieel beheer zijn echter herzien. Enerzijds gaf de invoering van de Wet op de Onderwijsverzorging (WOV) hiertoe aanleiding, anderzijds bleek de gezamenlijke verantwoordelijkheid van SVO en Cito voor zowel de programmering als de uitvoering problemen op te leveren. De Cito-activiteiten worden nu op basis van artikel 73 van de WOVS gefinancierd (1.6 miljoen op jaarbasis), terwijl SVO en Cito een coördinatiegroep hebben gevormd waarin afstemmingsoverleg plaatsvindt. SVO en Cito hebben daarnaast de gezamenlijke adviesgroep met vertegenwoordigers van de Vereniging Samenwerken de Landelijke Pedagogische Centra (VSLPC), de Vereniging Werkverband van Plaatselijke en Regionale Onderwijsbegeleidingsdiensten (WPRO), het Instituut voor Leerplanontwikkeling (SLO) en waarnemers van het Ministerie van Onderwijs en Wetenschappen en de Inspectie Basis- en Speciaal Onderwijs met een nagenoeg ongewijzigde taakomschrijving gehandhaafd. De taken van deze adviesgroep betreffen met name advisering over de gebruikswaarde van de projectopbrengst en de doorwerking daarvan in het onderwijs.

3.2 *Planning*

In een periode van ongeveer vijf jaar moeten alle leer- en vormingsgebieden uit de Wet op het Basisonderwijs (WOB) in peilingsonderzoeken aan de orde komen. Zeker ook in het licht van de eindtermen moet het als een gelukkige beslissing worden gezien dat het gehele curriculum moet worden bestreken. Daarmee wordt voorkomen dat eenzijdige accenten worden gelegd die ongewenste neveneffecten kunnen hebben. Na een evaluatie van de

eerste cyclus, waarover gerapporteerd wordt aan de Centrale Commissie voor Onderwijs-overleg (CCOO) en de Tweede Kamer, begint daarna in principe de volgende cyclus van peilingsonderzoeken, zodat ook vergelijkingen over de tijd mogelijk worden. Voor de eerste cyclus geldt op dit moment het tijdschema dat in Tabel 1 is weergegeven.

Tabel 1 *Tijdschema voor de eerste cyclus van peilingsonderzoeken in het basisonderwijs*

Rekenen	1987
Nederlandse taal	1988/1989
Sociale wereldoriëntatie	1990
Natuuronderwijs en gezondheids- educatie	1990/1991
Verkeerseducatie	1991
Engelse taal	1991
Bewegingsonderwijs	1987/?
Kunstzinnige vorming	1987/1988/?

De rekenpeilingen medio en einde basisonderwijs zijn in 1987 uitgevoerd en de rapportage daarover is inmiddels beschikbaar (Wijnstra, 1988, 1990; Verhelst & Eggen, 1989). In 1988 heeft de gegevensverzameling voor de taalpeiling einde basisonderwijs plaatsgevonden (rapportage voorjaar 1990), terwijl de mediopeiling op dit gebied in het najaar van 1989 is uitgevoerd. Bij de voorbereiding en verwerking van de taalpeilingen wordt samengewerkt met de SCO te Amsterdam.

Zoals eerder aangegeven worden op de overige leer- en vormingsgebieden alleen peilingsonderzoeken aan het einde van het basisonderwijs uitgevoerd. Wereldoriëntatie is in de planning in twee grote onderdelen gesplitst: sociale wereldoriëntatie (alle kennisgebieden uit de WOB met uitzondering van natuur) en natuuronderwijs, met inbegrip van gezondheidseducatie. De voorbereidingen voor wereldoriëntatie zijn al in de loop van 1986 gestart, in samenwerking met het Cito-project Itembank Wereldoriëntatie. Voor wereldoriëntatie is relatief veel voorbereidingstijd uitgetrokken, omdat de leerplanontwikkeling en toetsconstructie op dit gebied, in vergelijking met andere leer- en vormingsgebieden, nog niet erg ver ontwikkeld zijn.

Ter voorbereiding van peilingen op het gebied van de kunstzinnige vorming (de expressie-activiteiten uit de WOB) en het bewegingsonderwijs zijn in de jaren 1987 en

1988 in samenwerking met de SLO inventarisaties van het leer- en vormingsaanbod uitgevoerd met behulp van vragenlijsten (Brink, Wijnstra & Janssens, 1989a, 1989b; Zwarts & Wijnstra, 1989). In 1989 is de voorstudie naar de vorm en inhoud van peilingsonderzoeken op deze leer- en vormingsgebieden afgerond, wederom in samenwerking met de SLO. Gelet op het karakter van het onderwijs op deze gebieden diende nader bepaald te worden welke doelstellingen voor meting van kennis en vaardigheden op leerlingniveau in aanmerking komen en welke doelstellingen via een meting van het onderwijsaanbod moeten worden gepeild. Waarschijnlijk zal de eerste peiling in 1992 plaatsvinden. Welk onderdeel dit betreft, is nog niet vastgesteld. Dit verklaart de vraagtekens in Tabel 1.

3.3 *Werkwijze*

Naast instrumenten om achtergrondkenmerken van scholen en leerlingen te inventariseren, zijn – gelet op de doelstellingen van PPO – twee soorten, gedeeltelijk complementaire instrumenten nodig: instrumenten om het leer- en vormingsaanbod te inventariseren en instrumenten om vaardigheden (met inbegrip van kennis en inzicht) bij leerlingen te meten. In deze paragraaf wordt een aantal uitgangspunten voor de instrumentontwikkeling geformuleerd.

De instrumentontwikkeling voor een leer- en vormingsgebied start met het opstellen van een structurele beschrijving van het betreffende gebied in de vorm van een geordende lijst van leer- en vormingsdoelen op basis van aanwezige expertise, beschikbare leergangen, handboeken, enzovoort. Na een commentaarfase wordt de domeinbeschrijving vastgesteld, waarna deze dienst gaat doen als uitgangspunt voor de feitelijke instrumentontwikkeling. Zoals ook aangegeven in de vorige paragraaf is het mogelijk om in de peilingsonderzoeken elementen op te nemen die in het onderwijs (nog) geen gemeengoed zijn. Als op voorhand de inhoud van een leer- en vormingsgebied zou worden beperkt tot wat algemeen gangbaar is, of – in de toekomst – de wettelijke eindtermen basisonderwijs, snijdt dat de mogelijkheid af wellicht toch interessante of belangrijke aspecten voor de ontwikkeling van het onderwijs in de discussie over de bevindingen te betrekken. Dit impliceert ook dat de inhoud van het peilingsinstru-

mentarium niet gelegitimeerd behoeft te worden in de zin van de verantwoording van beperkende keuzen, zoals dat voor leerplan- en leergangontwikkeling geldt.

Dezelfde domeinstructurering leent zich voor de ontwikkeling van geheel verschillende soorten toetsen, afhankelijk van het gebruiksdoel. Als het om een algemene niveauindicatie gaat, kan met een betrekkelijk korte, maar inhoudelijk gevarieerde toets worden volstaan. De resultaten op een dergelijke toets lenen zich echter niet goed voor verdere inhoudelijke uitsplitsing. In het ontwerp van PPON heeft voorop gestaan dat de leer- en vormingsgebieden ruim gedekt moeten worden en dat de resultaten een inhoudelijke discussie over het onderwijs mogelijk moeten maken, bijvoorbeeld in het kader van de ontwikkeling van eindtermen. Dit heeft tot het uitgangspunt geleid dat de instrumenten een beschrijving van de vaardigheden van de leerlingen mogelijk moeten maken op nauwkeurig omschreven, didactisch betekenisvolle eenheden van beperkte omvang.

Dit niveau van detaillering komt ongeveer overeen met wat Bock, Mislevy en Woodson (1982) 'indivisible curricular elements' noemen en in dit opzicht komt de opzet van de PPON overeen met die van het California Assessment Program (CAP; vgl. Pandey, 1986). Dezelfde principes zijn ook terug te vinden in het proefonderzoek van Wesdorp et al. (1986). Een dergelijke opzet leidt tot een vrij groot aantal rapportage-eenheden per leer- en vormingsgebied. Dat is voor een discussie over de inhoud en het niveau van het onderwijs of voor de planning van ontwikkelingswerk eerder een voor- dan een nadeel. Uiteraard zullen de vaardigheden van de leerlingen op verwante rapportage-eenheden gecorrigeerd zijn. Gelet op het doel van het onderzoek is dit echter van secundair belang.

Een dergelijke modulaire opbouw heeft daarnaast diverse andere voordelen, vooral met het oog op de vergelijking over de tijd. Bij kleine eenheden is de kans het grootst dat veranderingen in het curriculum een uniform effect hebben op alle opgaven van een module, zodat veranderingen in de vaardigheid van de leerlingen goed te interpreteren zijn. Als uitgegaan zou worden van meer omvattende curriculum-eenheden, zouden accentwijzigingen in het curriculum nagenoeg onoverkomelijke problemen opleveren bij de vergelijking over

de tijd. Een verder voordeel is dat nieuwe aspecten op een gebied gemakkelijk als modules aan het instrumentarium kunnen worden toegevoegd zonder dat de vergelijkbaarheid van scores op andere modules met eerdere peilingen wordt aangetast (verouderde schalen kunnen ook zonder bezwaar worden weggelaten). Betrekkelijk kleine, inhoudelijk homogene eenheden bieden ten slotte – voorzover dat van toepassing is – een betere mogelijkheid itemresponsmodellen toe te passen dan meer omvattende eenheden. In tegenstelling tot de klassieke testtheorie is het met behulp van itemresponsmodellen mogelijk constante schalen te definiëren, onafhankelijk van de vaardigheid van de leerlingen. Daardoor is het ook mogelijk in een later stadium nieuwe opgaven aan de schaal toe te voegen, zodat in een volgende peilingsronde met gedeeltelijk nieuwe opgaven de verdeling van de vaardigheid in de populatie op een constante schaal kan worden weergegeven. Vervanging van opgaven is noodzakelijk, omdat een deel van de opgaven als illustratie in de rapportage moet worden gebruikt.

De benodigde opgaven voor de peilingsinstrumenten zullen in het algemeen worden geconstrueerd door commissies van ervaren leraren uit het basisonderwijs onder leiding van een projectmedewerker. De vorm van de opgaven zal afhankelijk zijn van de te meten vaardigheid, waarbij het uitgangspunt is dat het domein zo goed mogelijk gedekt moet worden. In dit opzicht worden op voorhand geen beperkingen aangebracht. Als een vaardigheid beter met open vragen kan worden gemeten dan met meerkeuzevragen, wordt de voorkeur gegeven aan open vragen. Als groepsgewijze, schriftelijke afname van de opgaven geen adequate mogelijkheid is, zal gebruik worden gemaakt van individuele afnameprocedures.

Complementair aan de instrumenten voor het meten van de onderwijsleerresultaten dienen instrumenten voor de inventarisatie van het leer- en vormingsaanbod te worden samengesteld. Deze zullen over het algemeen bestaan uit schriftelijke vragenlijsten voor de groeps- en vakleraren en/of de directeur van de school, maar ook andere instrumentaties zijn denkbaar. De inhoud en vorm van deze instrumenten zal sterk afhankelijk zijn van het leer- en vormingsgebied. Waar dat van toepassing is, zal ook aandacht worden

besteed aan aspecten die niet direct de bedoe-
ling hebben tot bepaalde leerresultaten te lei-
den ('expressive objectives' in de terminologie
van Eisner, 1969) en die niet in de toetsen ver-
tegenwoordigd zijn.

4 *Rekenpeiling einde basisonderwijs: een illustratie*

4.1 *Inleiding*

In 1987 zijn de hiervoor beschreven uit-
gangspunten en werkwijzen voor de eerste
keer toegepast in de rekenpeilingen. In het
voorjaar is de peiling einde basisonderwijs
uitgevoerd en in het najaar de peiling medio
basisonderwijs. De peilingen werden op af-
zonderlijke steekproeven van scholen uitge-
voerd. In deze paragraaf worden als illustratie
van het voorgaande enkele aspecten nader be-
licht op basis van de rekenpeiling einde basis-
onderwijs. Daarbij wordt de nadruk gelegd
op de bevordering van de mogelijkheden om
naar aanleiding van de bevindingen een dis-
cussie te voeren over de inhoud en het niveau
van het onderwijs. Voor (methodologische)
details zij verwezen naar Bokhove, Brink en
Janssen (1988a), Eggen (1989), Verhelst
(1989), Verhelst en Eggen (1989) en Wijnstra
(1988, 1989).

4.2 *Dataverzameling en -verwerking*

Afgezien van de instrumenten om gegevens
over achtergrondkenmerken van scholen en
leerlingen te verzamelen, werden de instru-
menten gebaseerd op een indeling van het re-
kendomein in 27 onderwerpen, verdeeld over
tien hoofdonderwerpen (zie Tabel 2). Aan de-
ze indeling ligt een kruising van inhouds- en
gedragscategorieën ten grondslag.

In samenwerking met een tweetal commis-
sies van ervaren leraren uit het basisonderwijs
zijn bijna 500 opgaven geconstrueerd: 32 voor
het onderwerp basisoperaties en zestien à
twintig voor elk van de overige onderwerpen,
voor het merendeel in open vraagvorm. Aan
deze vorm werd de voorkeur gegeven om tot
een zo volledig mogelijke dekking te komen
en om zo dicht mogelijk aan te sluiten bij de
praktijk van het onderwijs. Alleen wanneer de
meerkeuzevorm duidelijke voordelen had,
bijvoorbeeld bij schattingsopgaven, werd de-
ze vorm gekozen. In psychometrisch opzicht
heeft de open vraagvorm als voordeel dat geen

rekening behoeft te worden gehouden met ef-
fecten van raden.

De toetsen werden onder leiding van toets-
assistenten op ruim 300 scholen gemaakt. Om
tijdsredenen en om het clustereffect te beper-
ken maakte iedere leerling slechts een deel van
de toetsen. Feitelijk werden op deze wijze zes
(equivalente) substeekproeven gecreëerd met
ieder 1200 à 1300 leerlingen. In aanvulling op
de klassikale toetsafnames werden per klas,
afhankelijk van het aantal leerlingen, twee of
drie leerlingen systematisch getrokken die in
een individuele sessie drie tot vijf van de eer-
der gemaakte opgaven over één onderwerp
nog eens opnieuw kregen voorgelegd met de
vraag aan te geven welke procedure bij de
oplossing was gevolgd. Dit betrof in totaal
dertig opgaven over acht onderwerpen, met
name hoofdreken- en toepassingsonderwer-
pen. De procedures werden door de toetsas-
sistenten volgens een geprecodeerd catego-
rieënsysteem ingedeeld of uitgeschreven, als
de werkwijze niet direct was onder te brengen.

Alle 27 onderwerpen konden met toepas-
sing van een itemresponsmodel succesvol
worden geschaald waarbij minder dan 5%
van de opgaven buiten beschouwing moest
blijven, omdat deze onvoldoende in het model
pasten. Onder aanname van een bij benade-
ring standaardnormale verdeling van de vaar-
digheden van de leerlingen (door de toege-
paste transformaties) is bij alle schalen het
landelijk gemiddelde van de leerlingen in jaar-
groep 8 op 250 gesteld en de standaarddeviatie
op 50. Door de toepassing van een item-
responsmodel kunnen de moeilijkheidsgraad
van de opgaven en de vaardigheid van de leer-
lingen op dezelfde schaal worden afgebeeld
(zie voor een voorbeeld Figuur 1). In de geko-
zen opzet voor de rapportage geeft de schaal-
waarde van een opgave aan dat een leerling
met een overeenkomstige vaardigheidsscore
een kans van 80% heeft een dergelijke opgave
goed te maken. Bij lager op de schaal ge-
plaatste opgaven heeft dezelfde leerling een
grotere succeskans; bij moeilijker, hoger op
de schaal geplaatste opgaven is die kans uiter-
aard kleiner. Naast de schaal zijn de percen-
tielen vermeld (P.), dat wil zeggen de percen-
tages leerlingen die een score onder of gelijk
aan de betreffende schaalwaarde hebben be-
haald. In aanvulling op deze gegevens kan aan
de hand van de gegevens uit de individuele af-
names inzicht worden gegeven in de gehan-

Tabel 2 *Rekenonderwerpen peiling einde basisonderwijs*

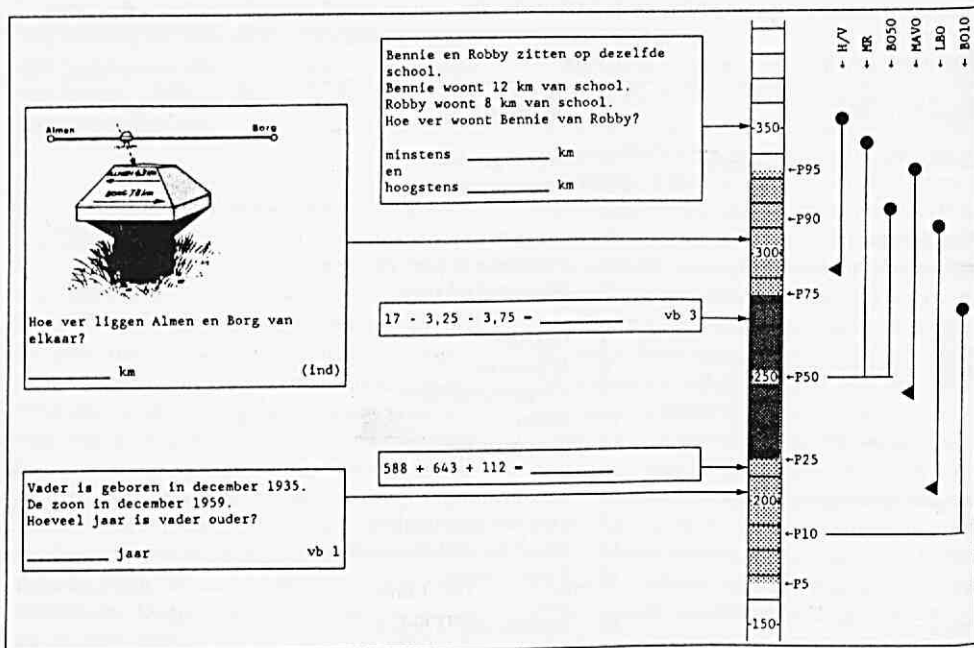
Hoofdonderwerp	Onderwerp
Basisoperaties	1 Basisoperaties
Getallen: basiskennis en begrip	2 Gehele getallen
	3 Kommagetallen
Hoofdrekenen (gehele en kommagetallen)	4 Optellen en aftrekken
	5 Vermenigvuldigen en delen
	6 Schattend rekenen
Cijferen (gehele en kommagetallen)	7 Optellen
	8 Aftrekken
	9 Vermenigvuldigen
	10 Delen
	11 Toepassingen
Breuken/gemengde getallen	12 Basiskennis en begrip
	13 Optellen en aftrekken
	14 Vermenigvuldigen en delen
	15 Toepassingen
Procenten	16 Basiskennis en begrip
	17 Handig procentrekenen
	18 Toepassingen
Verhoudingen	19 Basiskennis en begrip
	20 Toepassingen
Meten (lengte, omtrek, oppervlakte, inhoud, gewicht)	21 Basiskennis en begrip
	22 Tellen/aflezen van het aantal maateenheden
	23 Berekenen van het aantal maateenheden
	24 Maatsystemen
	25 Toepassingen
Kalender en klok	26 Toepassingen
Geld	27 Toepassingen

teerde oplossingswijzen. De in Figuur 1 met (ind) gemerkte opgave is een voorbeeld van een opgave die ook individueel werd afgenomen.

Om zicht te krijgen op de omvang en de inhoud van het rekenonderwijs vulden de groepsleraren van de jaargroepen 6, 7 en 8 een vragenlijst in. Naast enkele vragen over de methode, differentiatie en de beschikbare tijd voor het rekenonderwijs, stond hierin de (relatieve) verdeling van de tijd over de in de domeinbeschrijving onderscheiden (hoofd)onderwerpen centraal. De groepsleraar van jaargroep 8 kreeg verder per onderwerp drie a priori naar moeilijkheidsgraad geordende voorbeeldopgaven voorgelegd met het ver-

zoek aan te geven tot hoe ver de behandeling van de rekenstof was gevorderd voor de gemiddelde leerling (in Figuur 1 zijn twee van dergelijke met (vb) gemerkte opgaven afgebeeld). Voor de deelnemende leerlingen moest op de leerlingenlijst tevens worden vermeld welke leerlingen eventueel de stof die nodig is om de gemakkelijkste voorbeeldopgave op te lossen, nog niet hadden gehad.

In de rapportage worden de aanbod- en vaardigheidsgegevens met elkaar in verband gebracht, terwijl ook relaties met achtergrondkenmerken van scholen en leerlingen worden geëxploreerd.



Figuur 1 Schaal 4: Optellen en aftrekken (hoofdrekenen)

4.3 Gerealiseerd en gewenst niveau

In de rapportage zoals weergegeven in Figuur 1, wordt het gerealiseerde niveau van leerlingen aan het einde van het basisonderwijs beschreven. Of dit ook een acceptabel niveau is, is een geheel andere vraag, die niet door PPON kan worden beantwoord. Omdat dit de eerste peiling is, zijn ook nog geen vergelijkingen met eerdere peilingen mogelijk. Ter bevordering van de discussie is besloten enige indicaties te verzamelen over het gewenste niveau van enkele groepen leerlingen. Daartoe zijn in het voorjaar van 1988, toen de schaalconstructie voorlopig was afgerond, enquêtes gehouden onder groepsleraren van jaargroep 8, ouderleden van medezeggenschapsraden en docenten uit het LBO, MAVO en HAVO/VWO.

De enquêtes bestonden uit afbeeldingen van de schalen over de onderwerpen 2-27 met de illustratieve opgaven, maar zonder gegevens over de peilingsresultaten. De ouderleden van de medezeggenschapsraden en de leraren van de basisscholen werd een oordeel gevraagd over het gewenste niveau van de gemiddelde leerling op ieder van de schalen. De groepsleraren van de basisscholen werd tevens

gevraagd aan te geven welk niveau tenminste door 90% van de leerlingen zou moeten worden gehaald. De docenten uit (de brugklas van) het voortgezet onderwijs werd gevraagd aan te geven welk niveau voor hun schooltype aan het begin van het voortgezet onderwijs gewenst werd geacht.

Parallel aan de vaardigheidsschaal in Figuur 1 wordt het door de respondenten gewenste niveau en het verschil met het (geschatte) feitelijk gerealiseerde niveau aangegeven. Over het algemeen blijken aanzienlijke discrepanties. Bij de afgebeelde schaal leggen de groepsleraren van jaargroep 8 bijvoorbeeld gemiddeld het minimaal gewenste niveau voor 90% van de leerlingen (in Figuur 1 aangeduid als BO10) op 278. Dit wordt aangegeven door het bolletje. Het verschil met het niveau dat 90% van de leerlingen werkelijk bereikt (percentiel 10) wordt door de verticale lijn aangegeven en bedraagt bijna 90 punten op de vaardigheidsschaal, bijna twee standaarddeviaties. Het gewenste niveau voor deze groep leerlingen ligt hoger dan het feitelijke landelijke gemiddelde. De groepsleraren plaatsen het gewenste niveau van de gemiddelde leerling (BO50) op 317, de ouderleden van de mede-

zeggenschapsraden (MR) komen uit op 344, een niveau dat door nog geen 5% van de leerlingen wordt gehaald. Het gewenste niveau van de gemiddelde leerling wordt in de figuur door een lijn verbonden met percentiel 50. Gelet op het gerealiseerde niveau plaatsen ook de docenten uit het voortgezet onderwijs de grenzen hoog. De LBO-docenten leggen (gemiddeld) het gewenste niveau op 310. De MAVO-docenten komen uit op 332 en de HAVO/VWO-docenten (H/V) op 353. Deze niveaus worden door een lijn verbonden met de geschatte gemiddelden van deze groepen leerlingen op basis van gegevens uit de Eindtoets Basisonderwijs. Het vlaggetje aan de lijn wijst het betreffende schaalpunt ongeveer aan.

Om in de rapportage een tweede voorzet te geven voor een discussie over de inhoud en het niveau van het rekenonderwijs zijn een drietal deskundigen op dit gebied uitgenodigd een commentaar te schrijven op de uitkomsten van de peiling einde basisonderwijs. In de commentaren zou het vooral moeten gaan om de vraag wat het rekenonderwijs moet inhouden en welke resultaten dat moet opleveren, waarbij ook aanbevelingen werden gevraagd met betrekking tot de didactiek, ontwikkelingswerk, opleiding, begeleiding en onderzoek. Zonder verschuivingen in de tijdsverdeling binnen het rekenonderwijs uit te sluiten, werd de commentatoren verzocht uit te gaan van de randvoorwaarde dat voor het totaal van het rekenonderwijs in de toekomst (gemiddeld) niet meer tijd beschikbaar is dan in de tegenwoordige situatie (bijna vijf uur per week in de bovenbouw).

De commentaren zijn integraal opgenomen in een afzonderlijk deel van het rapport (zie Wijnstra, 1988). In het eerste hoofdstuk komt Huitema, medewerker van de Onderwijsbegeleidingsdienst in het Gewest Arnhem aan het woord. Hij is voorzitter van de Centrale Werkgroep Rekenen/Wiskunde-Onderwijs en tevens auteur van de rekenmethode *De Wereld in getallen* en het remediërend rekenpakket *Remelka*. Teunissen, universitair docent aan de Vakgroep Onderwijskunde van de Rijksuniversiteit in Utrecht, neemt het tweede hoofdstuk voor zijn rekening. Hij is op het gebied van het rekenonderwijs vooral bekend als auteur van zowel de eerste als de vernieuwde versie van de methode *Operator Rekenen*. In het laatste hoofdstuk geeft Treffers, hoogleraar aan de

Vakgroep Onderzoek Wiskundeonderwijs en Onderwijs Computercentrum (OW & OC) van de Rijksuniversiteit in Utrecht zijn visie. Hij werkte eerder mee aan het Wiskobas-project en promoveerde op het proefschrift *Wiskobas Doelgericht* (1978).

4.4 *Presentatie en verspreiding van de bevindingen*

Mede voorbereid door een serie artikelen van Bokhove en Janssen (1987), respectievelijk Bokhove, Brink en Janssen (1988a, 1988b) zijn de bevindingen uit de rekenpeiling einde basisonderwijs en de commentaren op 2, 3 en 4 november 1988 uitvoerig gepresenteerd op de jaarlijkse najaarsconferentie van de Nederlandse Vereniging tot Ontwikkeling van het Reken-Wiskundeonderwijs (NVORWO), waarbij in het programma ook nadrukkelijk de relatie werd gelegd met de eindtermen en met name de zogenaamde beredeneerde eindtermen voor rekenen/wiskunde (Treffers, 1988). Als vervolg op de conferentie zijn verdere artikelen gepubliceerd in het Tijdschrift voor Nascholing en Onderzoek van het Reken-Wiskundeonderwijs en Willem Bartjens (Bokhove & Janssen, 1988c, 1989a-e) en er zullen nog enkele volgen. Op deze manier kunnen ontwikkelaars, begeleiders en opleiders op het gebied van het rekenonderwijs op een succesvolle manier worden bereikt. De betreffende instituten hebben verder het volledige rapport over de rekenpeilingen ontvangen.

Alle scholen voor basis- en speciaal onderwijs hebben een brochure ontvangen met een beknopte schets van de rekenpeilingen. Met name voor een discussie over het schoolwerkplan van een individuele school blijft het echter een probleem dat de peilingsonderzoeken zich op macro-niveau bewegen. Hierdoor is het moeilijk de verbinding naar de eigen situatie van een individuele school te leggen. De mogelijkheid om de situatie van de eigen school te kunnen vergelijken met landelijke of regionale gegevens is in verschillende 'state assessment programs' in de Verenigde Staten zeer effectief gebleken om op lokaal niveau het doel en het functioneren van het onderwijs ter discussie te stellen en – indien gewenst – bij te stellen (met name de staat Minnesota is een goed voorbeeld; vgl. Wijnstra, 1982 en Richards, 1988).

Gezien de opzet van PPON is het in het kader van de gewone gegevensverzameling niet

mogelijk om op schoolniveau te rapporteren. Nog afgezien van de zinvolheid van een dergelijke rapportage in steekproefonderzoek, verhindert de toegepaste vorm van steekproef-trekking binnen de scholen dit, omdat niet alle leerlingen dezelfde opgaven hebben gemaakt. In de politieke besluitvorming over het peilingsonderzoek is deze mogelijkheid ook uitgesloten. Scholen dienen op vrijwillige basis deel te nemen, zonder in de rapportage herkenbaar te zijn om ongewenste sancties te voorkomen. Wel dienen de peilingsonderzoeken een kader te scheppen voor zelfevaluatie en dient te worden bijgedragen aan de ontwikkeling van toetsen die scholen in dit verband kunnen toepassen (vgl. par. 2).

In verband met een eventueel toekomstig gebruik van de PPON-instrumenten in het kader van de zelfevaluatie is in 1988 een proef genomen met een experimentele schoolrapportageservice. De scholen die in 1987 aan de rekenpeilingen hadden deelgenomen, werden in de gelegenheid gesteld het instrumentarium voor de rekenpeiling einde basisonderwijs in gecomprimeerde vorm af te nemen bij de leerlingen die in 1988 de basisschool zouden verlaten. Ruim een derde van de ongeveer 600 scholen heeft hieraan deelgenomen. Begin 1989 hebben zij een zodanige rapportage over combinaties van de oorspronkelijke schalen ontvangen dat een vergelijking mogelijk is met landelijke gegevens en met de categorie van scholen met een vergelijkbare samenstelling van de schoolbevolking. Uit de evaluatie is gebleken dat deze service door de scholen positief wordt gewaardeerd. Gezien het succesvolle verloop van deze proef heeft het Cito het voornemen deze service verder te ontwikkelen en algemeen beschikbaar te stellen, ook met uitbreiding naar andere leer- en vormingsgebieden. Deze activiteiten zullen in een afzonderlijk serviceproject worden ondergebracht.

5 Besluit

In de brochure *Over Leerplanonderzoek* behandelt Bijl (1966) als centrale vraag hoe het programma tot stand moet komen voor wat we tegenwoordig de basisvorming noemen. Zijn standpunt is dat de opstelling van een dergelijk programma geen schrijftafelwerk meer kan zijn, ook niet van een commissie van deskundigen, maar dat er onderzoek nodig

is. '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 wordt aangebracht, het opvoedingsgezag, d.w.z. de organen die met dat gezag zijn bekleed.' Hiermee geeft hij een uitstekende rationale avant la lettre voor het peilingsonderzoek in Nederland zoals dat in het voorgaande is beschreven.

PPON is te beschouwen als een diepte-investering om politiek en maatschappelijk gewenste processen mogelijk te maken. Uiteraard moet peilingsonderzoek op een technisch adequate manier worden uitgevoerd en in dit opzicht zullen ongetwijfeld verbeteringen mogelijk blijken te zijn. Het succes van PPON moet echter vooral blijken uit de manier waarop de gegevens gebruikt worden ten behoeve van de ontwikkeling van het onderwijs.

Noten

1. Inmiddels zijn ook proefonderzoeken in het speciaal en het voortgezet onderwijs uitgevoerd (zie Vergeer, Oostdam & Otter, 1988; Kuhlemeier & Van den Bergh, 1989) en de minister heeft het Cito medio 1989 opdracht verleend de peilingen naar het speciaal onderwijs uit te breiden. De consequenties van dit besluit, dat na de afsluiting van het manuscript bekend werd, blijven in dit artikel nog buiten beschouwing. In principe zal de eerste peilingsronde in het speciaal onderwijs – vooralsnog beperkt tot LOM- en MLK-scholen – een geheel gaan vormen met de tweede ronde in het basisonderwijs.

Literatuur

- Bock, R. D., R. Mislevy & C. Woodson, The next stage in educational assessment. *Educational Researcher*, 1982, 11, no. 3, 411, 16.
- Bokhove, J., H. Brink & J. Janssen, Periodiek peilingsonderzoek in het basisonderwijs (2). *Tijdschrift voor Nascholing en Onderzoek van het Reken-Wiskundeonderwijs*, 1988a, 6, nr. 3, 3-22.
- Bokhove, J., H. Brink & J. Janssen, Periodiek peilingsonderzoek in het basisonderwijs (3). *Tijdschrift voor Nascholing en Onderzoek van het Reken-Wiskundeonderwijs*, 1988b, 6, nr. 4, 47-57.

- Bokhove, J. & J. Janssen, Periodiek peilingsonderzoek in het basisonderwijs. *Tijdschrift voor Nascholing en Onderzoek van het Reken-Wiskundeonderwijs*, 1987, 6, nr. 1, 3-6.
- Bokhove, J. & J. Janssen, Periodiek peilingsonderzoek in het basisonderwijs (4). *Tijdschrift voor Nascholing en Onderzoek van het Reken-Wiskundeonderwijs*, 1988c, 7, nr. 2, 16-34.
- Bokhove, J. & J. Janssen, Periodiek peilingsonderzoek in het basisonderwijs (5). *Tijdschrift voor Nascholing en Onderzoek van het Reken-Wiskundeonderwijs*, 1989a, 7, nr. 3/4, 3-14.
- Bokhove, J. & J. Janssen, Peiling van het rekenonderwijs. *Willem Bartjens; Tijdschrift voor Reken-Wiskundeonderwijs in de basisschool*, 1989b, 8, 66-70.
- Bokhove, J. & J. Janssen, PPON-onderzoek naar oplossingswijzen. *Willem Bartjens; Tijdschrift voor Reken-Wiskundeonderwijs in de basisschool*, 1989c, 8, 125-129.
- Bokhove, J. & J. Janssen, PPON-onderzoek naar oplossingswijzen (2). *Willem Bartjens; Tijdschrift voor Reken-Wiskundeonderwijs in de basisschool*, 1989d, 8, 184-189.
- Bokhove, J. & J. Janssen, Periodiek peilingsonderzoek in het basisonderwijs (6). *Tijdschrift voor Nascholing en Onderzoek van het Reken-Wiskundeonderwijs*, 1989e, 8, nr. 2, 3-32.
- Brink, H., J. M. Wijnstra & F. J. G. Janssens, *Inventarisatie van het leer- en vormingsaanbod bij het bewegingsonderwijs in de basisschool*. Arnhem: Instituut voor Toetsontwikkeling, 1989 (PPON-rapport nr. 1).
- Brink, H., J. M. Wijnstra & F. J. G. Janssens, *Inventarisatie van het leer- en vormingsaanbod bij de beeldende vorming in de basisschool*. Arnhem: Instituut voor Toetsontwikkeling, 1989 (PPON-rapport nr. 2).
- Bijl, J., *Over leerplanonderzoek*. Groningen: Wolters, 1966 (Acta Paedagogica Ultrajectina XXIII).
- Doornbos, K., *Opstaan tegen het zittenblijven*. 's-Gravenhage: Staatsuitgeverij, 1969.
- Eggen, T. J. H. M., Itemcalibratie bij vertakt toetsen. In: W. J. van der Linden & L. J. T. van der Kamp (red.), *Meetmethoden en data-analyse* (pp. 75-88). Amsterdam: Swets & Zeitlinger, 1989.
- 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* (pp. 1-31). Chicago: Rand McNally, 1969 (AERA Monograph Series on Curriculum Evaluation no. 3).
- ICB, *Innovatieplan basisonderwijs*. 's-Gravenhage: Staatsuitgeverij, 1978.
- Kuhlemeier, H. & H. van den Bergh, *De proefpeiling Nederlands: een onderzoek naar de haalbaarheid van peilingsonderzoek in het voortgezet onderwijs*. Arnhem: Instituut voor Toetsontwikkeling, 1989.
- Kwaliteit van het Onderwijs. 's-Gravenhage: Ministerie van Onderwijs en Wetenschappen, 1981.
- Pandey, T. N., State of the art of large-scale assessment in the United States. In: W. J. van der Linden & J. M. Wijnstra (red.), *Ontwikkelingen in de methodologie van het onderwijsonderzoek* (pp. 5-25). Lisse: Swets & Zeitlinger, 1986.
- Richards, C. E., A typology of educational monitoring systems. *Educational Evaluation and Policy Analysis*, 1988, 10, 106-116.
- Treffers, A., Beredeneerde eindtermen rekenwiskunde voor het basisonderwijs. *Willem Bartjens; Tijdschrift voor Reken-Wiskundeonderwijs in de basisschool*, 1988, 8, 9-52.
- Verhelst, N. D., Informatiewinst bij vertakt toetsen. In: W. J. van der Linden & L. J. T. van der Kamp (red.), *Meetmethoden en data-analyse* (pp. 89-96). Amsterdam: Swets & Zeitlinger, 1989.
- Verhelst, N. D. & T. J. H. M. Eggen, *Psychometrische en statistische aspecten van peilingsonderzoek*. Arnhem: Instituut voor Toetsontwikkeling, 1989 (PPON-rapport nr. 4).
- Vergeer, M. M., R. J. Oostdam & M. E. Otter, *Taalonderwijs op LOM- en MLK-scholen; voorstudie periodieke peiling van het onderwijsniveau*. 's-Gravenhage: Het Instituut voor Onderzoek van het Onderwijs (SVO), 1988 (Selecta-reeks).
- Wesdorp, H., *Goed onderwijs, wat is dat?* 's-Gravenhage: Stichting voor Onderzoek van het Onderwijs, 1985 (Selecta-reeks).
- Wesdorp, H., H. van den Bergh, D. J. Bos, J. B. Hoeksma, R. J. Oostdam, J. Scheerens & B. Triesscheijn, *De haalbaarheid van periodiek peilingsonderzoek; een voorstudie op het gebied van het taalonderwijs in de lagere school*. Lisse: Swets & Zeitlinger, 1986 (SVO-reeks nr. 85).
- WRR, *Basisvorming in het onderwijs*. 's-Gravenhage: Staatsuitgeverij, 1986.
- Wijnstra, J. M., Wat leren kinderen op school? *Pedagogische Studiën*, 1982, 59, 223-232 en 277-287.
- Wijnstra, J. M. (red.), *Balans van het rekenonderwijs in de basisschool*. Arnhem: Instituut voor Toetsontwikkeling, 1988.
- Wijnstra, J. M. (red.), *Verantwoording van de rekenpeiling medio en einde basisonderwijs 1987*. Arnhem: Instituut voor Toetsontwikkeling, 1990 (PPON-rapport nr. 5).
- Zwarts, M. A. & J. M. Wijnstra, *Inventarisatie van het leer- en vormingsaanbod bij het muziekonderwijs in de basisschool*. Arnhem: Instituut voor Toetsontwikkeling, 1989 (PPON-rapport nr. 3).

Curriculum vitae

J. M. Wijnstra (1946) studeerde na zijn opleiding tot onderwijzer pedagogiek aan de Rijksuniversiteit te Utrecht (specialisatie onderwijskunde) en promoveerde in 1976 op het proefschrift 'Het onderwijs aan van huis uit friestalige kinderen'. Sinds 1978 is hij als medewerker verbonden aan het Insti-

tuut voor Toetsontwikkeling (Cito) te Arnhem, tegenwoordig als projectleider van het project Periodeke Peiling van het Onderwijsniveau (PPON).

Adres: Instituut voor Toetsontwikkeling (Cito)
Postbus 1034, 6801 MG Arnhem

Manuscript aanvaard 6-12-'89

Summary

Wijnstra, J. M. 'The design of the Dutch national assessment program in education illustrated with examples from the first mathematics assessment.' *Pedagogische Studiën*, 1990, 67, 139-150.

The operational preparations for the Dutch national assessment program in education started in 1986. The aim of the program is to periodically describe the contents of the curriculum in use and the achievement of students in primary education, so as to provide a basis for a more rational discussion about the contents and the level of education. In this paper the aim and background of the program are explained in relation to recent policy on a national core curriculum. Subsequently, the organization, planning, and procedures of the project are described. The procedures are illustrated with examples from the first mathematics assessment at the end of primary education, conducted in 1987. In this section special attention is paid to the way the findings are reported, in order to foster the public discussion of these issues.