

Rekenmethoden-onderzoek op de proef gesteld

M. VAN DEN HEUVEL-PANHUIZEN
*Vakgroep Onderzoek Wiskundeonderwijs en
Onderwijscomputercentrum (OW & OC),
Rijksuniversiteit Utrecht*

Samenvatting

Een de laatste tijd veel gehoorde vraag is, of de gebruikte methode er wel toe doet. Harskamp c.s. beantwoorden naar aanleiding van hun implementatie- en effectonderzoek van reken-wiskundemethoden op de basisschool deze vraag ontkennend. Qua leerstof en qua leerresultaten blijken de methoden slechts marginaal te verschillen. In deze bijdrage wordt nagegaan in hoeverre deze conclusies en de bijbehorende aanbevelingen wel met recht worden verkondigd. Uit de analyse komt naar voren dat de basis ervoor maar zwak is.

1 Inleiding

Onderzoek naar de implementatie van nieuwe reken-wiskundemethoden en naar het effect van deze methoden op de leerresultaten van leerlingen is nodig, maar moeilijk. Nodig is het, omdat om het onderwijs en het rendement van het onderwijs te kunnen verbeteren, eerst bekend moet zijn hoe er in de praktijk met een bepaalde methode wordt gewerkt en wat dit oplevert aan leerresultaten. Moeilijk is het, omdat bij dit soort onderzoek zoveel variabelen in het spel zijn die bijna niet of zelfs helemaal niet te controleren zijn. Dit maakt het onderzoek niet onmogelijk, maar noopt wel tot de nodige voorzichtigheid met betrekking tot de conclusies en de aanbevelingen. Er dient dan ook kritisch te worden nagegaan of de conclusies en aanbevelingen wel met recht worden verkondigd. In deze bijdrage worden de conclusies die Harskamp en Suhre trekken naar aanleiding van hun implementatie- en effectonderzoek van reken-wiskundemethoden in het basisonderwijs aan zo'n kritische analyse onderworpen.

Via allerlei publikaties gericht op verschillende doelgroepen hebben Harskamp en Suhre laten weten:

- dat de methoden qua leerstof slechts heel weinig van elkaar verschillen (Harskamp & Suhre, 1987, 21 en 23; Suhre, 1988, 5),
- dat er een indicatie is, dat de (moderne) methoden worden gebruikt zoals bedoeld (Harskamp, 1988, 85)¹,
- dat er verschillen bestaan in de condities waaronder de methoden worden gebruikt (Harskamp & Suhre, 1988, 213-214), m.n. dat de moderne methoden veel op scholen worden gebruikt met leerlingen van een meer dan gemiddeld intelligentieniveau (Suhre, 1988, 6 en 5), dat de moderne methoden een uitdaging vormen voor intelligente kinderen en dat in heterogene klassen de stof voor zwakke leerlingen problemen op kan leveren (Harskamp & Suhre, 1987, 23),
- dat de verschillen in toetsprestaties en in rekenattitude tussen de leerlingen marginaal zijn (Harskamp & Suhre, 1987, 22 en 23; Harskamp & Suhre, 1988, 216; Harskamp, 1988, 86, 87 en 89),
- dat niet de methode, maar een aantal implementatiekenmerken – waaronder het percentage leerlingen waarmee de basisstof is behandeld – van invloed is op de leerprestaties (Harskamp & Suhre, 1987, 23; Harskamp & Suhre, 1988, 216; Harskamp, 1988, 88).

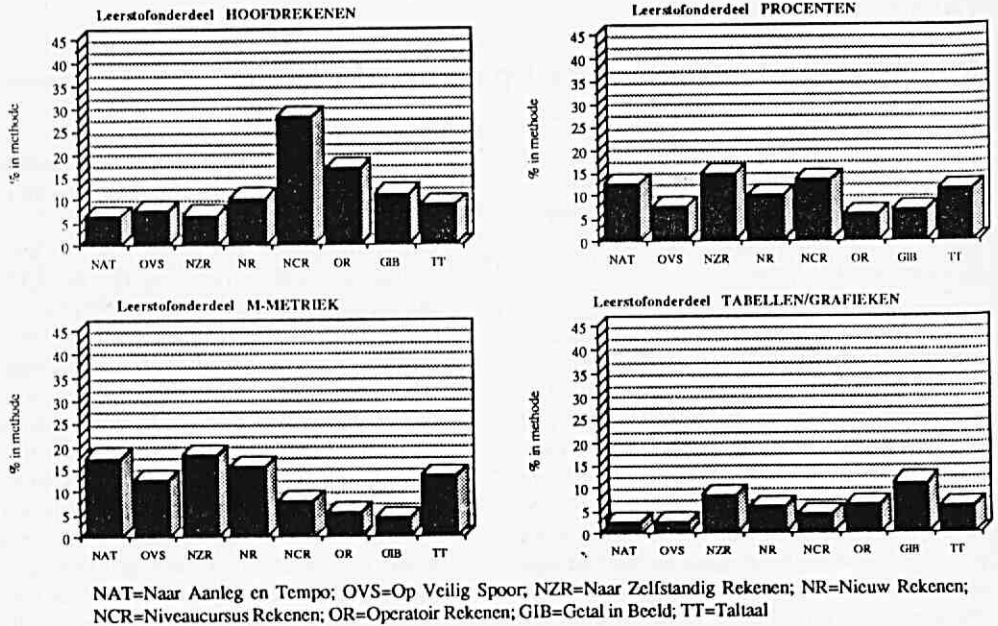
Zijn al deze conclusies wel houdbaar?

2 De inhoud van de reken-wiskundemethoden

De inhoudsanalyse van de methoden, die in het kader van het onderzoek van Harskamp en Suhre is uitgevoerd, omvat een leerstofanalyse en een didactische analyse.

2.1 Leerstofanalyse

Als eerste wordt ingegaan op de leerstofanalyse van de methoden en de conclusie dat de methoden wat betreft de relatieve aandacht voor de verschillende leerstofonderdelen weinig



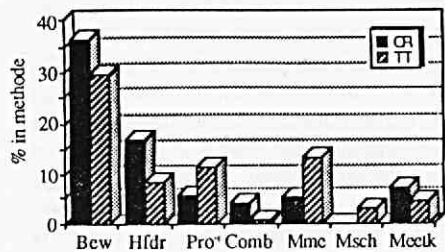
Figuur 1 Percentage aandacht dat de verschillende methoden besteden aan bepaalde leerstofonderdelen (percentages overgenomen uit Harskamp, 1988, 39)

van elkaar verschillen. De toegepaste analyseprocedure – per taak zijn drie punten te verdelen over een aantal vooraf onderscheiden leerstofcategorieën (Harskamp, 1988, 36-37) – is nogal zwak om als basis te dienen voor deze conclusie.

Een bezwaar is, dat er sprake is van een tamelijk grove indeling waardoor het niet altijd gemakkelijk is om uit te maken bij welke categorie een bepaalde opgave hoort. Maar ook zonder dit probleem is het nog de vraag of op grond van de gevonden percentages (o.c., 39) wel geconcludeerd kan worden dat “er (...) over het geheel genomen geen grote verschillen (zijn, mvdh) in het percentage leerstof dat per methode wordt besteed aan de verschillende leerstofonderdelen” (Harskamp, 1989, 3). Aan de hand van de door Harskamp gevonden percentages kan ook een andere conclusie worden getrokken. Namelijk, dat er met betrekking tot de aandacht die methoden besteden aan bepaalde leerstofonderdelen soms grote verschillen bestaan. Zie bijvoorbeeld de verschillen in relatieve aandacht die er tussen de methoden bestaan met betrekking tot de leerstofonderdelen hoofdrekenen, procenten, metrie en tabellen/grafieken (zie Figuur 1).

Dit zijn allemaal leerstofonderdelen waarop volgens Harskamp vrijwel alle methoden veel nadruk leggen (o.c., 3).

Bij de conclusie, dat er tussen de methoden over het geheel genomen geen grote verschil-



Bew = Bewerkingen
 Hfdr = Hoofdrekenen
 Pro = Procenten
 Comb = Combinatoriek
 Mmc = Meten – Metrie
 Msch = Meten – Schatten
 Meetek = Meetekunde

Figuur 2 Percentage aandacht dat de methoden OR en TT besteden aan de bepaalde leerstofonderdelen (percentages overgenomen uit Harskamp, 1988, 39).

len zijn in aandacht voor de leerstofonderdelen, moet bovendien worden betrokken, dat in de onderwijspraktijk bijna nooit sprake is van een keuze uit alle methoden. Er is meestal sprake van een vergelijking van twee of drie methoden. En dan kunnen de verschillen in relatieve aandacht voor de verschillende leerstofonderdelen ook groot zijn. Zoals bijvoorbeeld het geval is bij de methoden *Operatoir Rekenen* en *Taltaal* (zie Figuur 2).

Iets dat de vergelijking ook moeilijk maakt, is de breedte van het leerstofonderdeel bewerkingen. Dit leerstofonderdeel bevat zowel opgaven die betrekking hebben op de hoofdbewerkingen, als opgaven met breuken en opgaven met kommagetallen. Onder deze overkoepelende categorie kunnen veel verschillen schuilgaan. Het belangrijkste bezwaar tegen de uitgevoerde leerstofanalyse is dan ook, dat er geen inhoudelijke analyse van de leerstof heeft plaatsgevonden. Pas dan kan worden gezegd, of er wat de inhoud betreft tussen methoden verschillen bestaan.

Hoe verschillend de leerstofinhoud kan zijn, zelfs als het gaat om een zeer duidelijk omschreven leerstofonderdeel als het leren optellen en aftrekken, moge blijken uit het volgende voorbeeld dat betrekking heeft op de methoden *Naar Zelfstandig Rekenen* (NZR) en *De Wereld in Getallen* (WIG) (zie Van den Heuvel-Panhuizen, 1989). Bij NZR leren de kinderen in groep 3 optellen en aftrekken tot 20 en hieronder vallen ook de sommen waarbij het tiental wordt overschreden. Bij WIG daarentegen leren de kinderen in groep 3 al sommen maken tot 100, maar er worden nog geen sommen gemaakt waarbij sprake is van een tientaloverschrijding. Deze leerstofinhoudelijke verschillen zouden bij een louter kwantitatieve analyse, op de manier zoals Harskamp (1988, 36-37) die heeft toegepast, niet naar voren zijn gekomen. Alles zou dan onder de noemer van bewerkingen zijn gevallen².

2.2 *Didactische analyse*

Naast de leerstofanalyse heeft er in het onderzoek van Harskamp c.s. ook een didactische analyse van de methoden plaatsgevonden. Ondanks dat de methoden qua relatieve aandacht voor de onderscheiden leerstofonderdelen in grote mate met elkaar overeenstemmen, verschillen de methoden volgens Harskamp wél in de vakinhoudelijke vormgeving van de

leerstof (o.c., 1988, 83 en 15). Harskamp beschrijft uitgebreid de bevindingen van De Jong (1986) aangaande de aanwezigheid en de getrouwheid van Wiskobaskenmerken in methoden en gebruikt deze methodebeoordelingen om de bij zijn onderzoek betrokken methoden in te delen in modern en traditioneel.

Het gegeven dat De Jong bij de vakinhoudelijke beoordeling van de leerstof in tegenstelling tot Harskamp wel tot grote verschillen komt tussen de methoden, wordt door Harskamp geenszins als problematisch beschouwd. De Jongs typering van methoden is immers uitsluitend op de aanwezigheid en getrouwheid van Wiskobaskenmerken gebaseerd (Harskamp, 1988, 83). Bovendien hoeven volgens Harskamp vakinhoudelijke verschillen niet in te houden, dat de moderne methoden op andere leerstofonderdelen nadruk leggen dan de traditionele methoden (o.c., 9). Uit de door Harskamp (1988, 7-8) vermelde lijst van Wiskobaskenmerken valt echter op te maken dat deze opvatting niet houdbaar is. De Wiskobaskenmerken hebben niet alleen betrekking op de vormgeving van de leerstof, maar ook op de keuze voor een bepaalde leerstofinhoud. Vandaar ook dat het werk van Wiskobas is terug te vinden in de eindtermen van de basisschool (vgl. Treffers, 1989).

De didactische analyse die Harskamp c.s. zelf hebben uitgevoerd, betreft de vakdidactische kenmerken leerstofdifferentiatie en variatie in leerstof (Harskamp, 1988, 37).

Voor de *variatie in leerstof* is gebruik gemaakt van de eerder uitgevoerde leerstofanalyse. Aan de hand van de geïnventariseerde percentages per leerstofonderdeel en een aanvullende inventarisatie voor het leerstofonderdeel redactieopgaven, is de mate van variatie in de leerstof vastgesteld. Deze mate van variatie wordt berekend door "de percentages van redactieopgaven en hoofdrekenen en eigenschapsrekenen bij elkaar op te tellen. Dit percentage geeft aan welk deel van de totale leerstof besteed wordt aan andere activiteiten dan het inoefenen van algemene rekenvaardigheden" en "het percentage algemene rekenvaardigheden in 'kale' vorm is nu een restcategorie, te bepalen door de percentages voor redactieopgaven en hoofdrekenen en eigenschapsrekenen af te trekken van 100%" (o.c., 41)³. Dit is een operationalisatie die

veeleer betrekking heeft op de mate waarin een methode aandacht besteedt aan bepaalde leerstof (redactieopgaven, hoofdrekenen en eigenschapsrekenen) dan op de variatie in leerstof. Doordat er vanuit wordt gegaan dat het percentage redactieopgaven, hoofdrekenen, eigenschapsrekenen en de algemene rekenvaardigheden in kale vorm samen 100% vormen, is het bovendien een operationalisatie die uitgaat van een tamelijk beperkte kijk op de inhoud van het vak rekenen-wiskunde. Leerstofonderdelen als meten, tabellen/grafieken en meetkunde tellen hierbij blijkbaar niet mee. Ter zijde zij nog opgemerkt, dat bij de meting van de door de leerkrachten uitgevoerde leerstofvariatie een enigszins andere operationalisatie van de leerstofvariatie is gehanteerd. Er is gevraagd naar het percentage van de lestijd dat besteed wordt aan hoofdrekenen, redactieopgaven (of toepassingsproblemen uit reële situaties) en algemene rekenvaardigheden (o.c., 127)⁴. Het eigenschapsrekenen is hierbij als aparte categorie komen te vervallen en de algemene rekenvaardigheden (oefeningen op het gebied van cijferen, metriek stelsel, meten, breuken, procenten, enz.) bestaan volgens deze omschrijving blijkbaar niet meer noodzakelijkerwijs alleen uit kale opgaven.

Terug naar de leerstofvariatie in de methoden. De door Harskamp gekozen operationalisatie levert een beeld op van de mate van

leerstofvariatie in de verschillende methoden, dat met name door de plaats die de methode Niveaucursus Rekenen (NCR) in de rij inneemt, niet zo erg herkenbaar is. De hoge score van NCR is bijna volledig in tegenspraak met bijvoorbeeld de beoordeling die De Jong (1986, 183) geeft van NCR: "in alle opzichten een mechanistische methode, die de niveaucursus-organisatie centraal stelt ten koste van een actief, gedifferentieerd (...) en wiskundig rijk onderwijs". Harskamp zelf noemt de score van NCR ook opvallend (o.c., 41), maar ziet hierin klaarblijkelijk geen aanleiding om de gekozen operationalisatie te heroverwegen.

Een alternatief is bijvoorbeeld: per methode de leerstofonderdelen in volgorde van grootte zetten en dan vervolgens kijken naar het aantal verschillende leerstofonderdelen waaruit de eerste 60% of 80% (te beginnen bij het leerstofonderdeel met het hoogste percentage) van het totale leerstofaanbod bestaat. Als dit wordt toegepast op de door Harskamp bij de leerstofanalyse gevonden percentages (met alle gebreken vanden), zou dit leiden tot een ander beeld van de mate van leerstofvariatie in de methoden (zie Figuur 3). Een beeld dat waarschijnlijk meer overeenkomt met de werkelijke variatie in leerstof(onderdelen) in de methoden. NCR bijvoorbeeld blijkt hierin juist laag te scoren.

methode	aantal onderdelen in 60% van de leerstof	aantal onderdelen in 80% van de leerstof	'variatie in leerstof Harskamp (1988)
NCR	Bew Hfd	Bew Hfd Pro	45%
NZR	Bew Mm	Bew Mm Pro Tab	30%
OVS	Bew Mm	Bew Mm MI Pro - Hfd	32%
NAT	Bew Mm	Bew Mm Pro Ver Get	32%
NR	Bew Mm	Bew Mm Hfd Pro Get	38%
OR	Bew Hfd	Bew Hfd Mtk Tab Pro - Get	49%
GIB	Bew Get Hfd	Bew Get Hfd Tab Pro Eig	45%
TT	Bew Mm Pro	Bew Mm Pro Hfd Get Tab	46%

Bew=Bewerkingen; Hfd=Hoofdrekenen; Pro=Procenten; Mm=Meten-Metriek; MI=Meten-Aflezzen; Tab=Tabellen/Grafieken; Get=Getallen; Eig=Eigenschapsrekenen

Figuur 3 Variatie in leerstof(onderdelen) in methoden.

3 De implementatie van rekenwiskundemethoden

Via een schriftelijke bevraging van de leerkrachten is nagegaan in welke mate zij een aantal implementatiekenmerken toepassen. Geconcludeerd wordt, dat het feit dat de methoden, zoals verwacht, verschillen op de twee vakdidactische implementatiekenmerken een indicatie geeft dat de (moderne) methoden worden gebruikt zoals bedoeld (Harskamp, 1988, 85). Gemiddeld worden de moderne methoden zelfs iets beter geïmplementeerd dan de traditionele (Suhre, 1987, 353). Betekent dit nu, dat uit het onderzoek van Harskamp en Suhre blijkt dat de vernieuwende kenmerken – want daar gaat het bij implementatie toch om (vgl. Fullan & Pomfret, 1977) – van de realistische methoden inderdaad gerealiseerd worden in de praktijk? Nee. Immers, geen enkele van de vernieuwende kenmerken van de realistische methoden, zoals het gebruik van contexten, de aandacht voor modellen, de grote mate van eigen inbreng van de leerlingen, het interactieve karakter en de onderlinge samenhang van de verschillende leerstofonderdelen (zie Treffers & Goffree, 1985; Treffers & Feijs, 1987) is door Harskamp en Suhre onderzocht. In plaats daarvan is gekozen voor implementatiekenmerken “die door de leerkracht op betrouwbare wijze kunnen worden beantwoord” en “die vermoedelijk voorspellende waarde hebben ten opzichte van leerprestaties” (Harskamp, 1988, 27). Het interactieve karakter van het onderwijs, dat als eerste bij de *vakdidactische implementatiekenmerken* wordt genoemd (o.c., 15-16) en waarin alle kenmerken van de realistische didactiek worden ondergebracht (Harskamp, 1989, 4), is bij nader inzien dan ook niet onderzocht. Het zou te moeilijk zijn te operationaliseren door de te globale aanwijzingen in de methoden, het zou alleen op de introductie van nieuwe leerstof slaan en het zou geen voorspellende waarde hebben ten aanzien van de leerresultaten (Harskamp, 1988, 27).

Wat overblijft zijn twee andere vakdidactische implementatiekenmerken: variatie in de leerstof en leerstofdifferentiatie. Voor beide kenmerken is een significant verschil gevonden tussen de acht methoden (o.c., 56)⁵. Bovendien is er een samenhang gevonden tussen datgene wat de didactische analyse van de me-

thoden hierover heeft opgeleverd en de bevraging van de leerkrachten (o.c., 56). Het zijn deze beide gegevens die een indicatie geven dat de (moderne) methoden worden gebruikt zoals bedoeld (o.c., 85). Gezien de gekozen operationalisaties is de basis voor deze conclusie zwak. Voor de bezwaren tegen de operationalisatie van de variatie in leerstof kan verwezen worden naar de opmerkingen die zijn gemaakt bij de didactische analyse van de methoden. Ten aanzien van de operationalisatie van de leerstofdifferentiatie moet worden opgemerkt, dat de bij de theoretische achtergronden wel genoemde (o.c., 17), voor de realistische didactiek kenmerkende, procesdifferentiatie later niet terug komt bij de bevraging van de leerkrachten. Omdat de aanwijzingen omtrent procesdifferentiatie in de methoden te summier zijn, worden geen grote verschillen in instructiegedrag verwacht en wordt alleen gevraagd naar de mate waarin verschillende leerlingen in een klas verschillende leerstof krijgen tijdens de lessen (o.c., 28). Een bijkomend probleem is dat de categorieën die gebruikt zijn bij de methodenanalyse niet dezelfde zijn als bij de bevraging van de leerkrachten (vgl. o.c., 37 en 127).

Naast de twee vakdidactische implementatiekenmerken omvat het onderzoek ook nog drie *algemene implementatiekenmerken*: het percentage leerlingen waarmee de basisstof is behandeld, het aantal lessen per week en het aantal minuten dat men per les bezig is met rekenen. Ten aanzien van de laatste twee kenmerken is geen significant verschil tussen de acht methoden gevonden (o.c., 55-56). Bij het percentage leerlingen dat de basisstof heeft gedaan, is wel een significant verschil gevonden tussen de acht methoden, maar niet tussen de traditionele en de moderne methoden (o.c., 54). Ook deze kenmerken zeggen in feite niets over de mate waarin een bepaalde vernieuwing wordt gerealiseerd. Een leerkracht kan op deze manier gemakkelijk een hoge implementatiescore halen – door bijvoorbeeld veel lessen per week te geven of door lange lessen te geven – zonder dat iets van de vernieuwing in praktijk wordt gebracht. De kenmerken zijn veeleer algemene onderwijscondities dan implementatiekenmerken. Er moet een bepaalde tijd aan onderwijs worden besteed, wil het effect sorteren.

Met betrekking tot het percentage leerlin-

gen waarmee de basisstof is behandeld, ligt het wat moeilijker. Het is de vraag of dit ook als een algemene onderwijsconditie kan worden beschouwd of dat het eerder een gevolg is van de gebruikte methode. In § 5.2 wordt hier nog op teruggekomen.

4 *De condities waaronder de methoden worden gebruikt*

Hoe zit het met de condities waaronder de methoden worden gebruikt? Voor de beantwoording van deze vraag zijn gegevens verzameld over drie achtergrondkenmerken van de leerkrachten en twee achtergrondkenmerken van de leerlingen.

Met betrekking tot de *leerkrachten* is nagegaan hoeveel jaren ervaring zij hebben met de methode, of er sprake is van een leerling- of van een leerstofgerichte attitude en of ze wel of niet een verzaamd wiskunde-examen in hun vooropleiding hebben gehad. Alleen bij het aantal jaren ervaring is een significant verschil gevonden tussen de acht methoden (o.c., 58). Harskamp voegt hier aan toe, dat leerkrachten van moderne methoden gemiddeld genomen minder jaren ervaring hebben met hun methode dan leerkrachten van traditionele methoden (o.c., 58). Uit niets blijkt echter dat dit ook is getoetst. Het lijkt erop, dat aan het gegeven dat er een significant verschil bestaat tussen de acht methoden tevens de conclusie is verbonden dat er een verschil bestaat tussen de moderne en de traditionele methoden. Dit is een onjuiste handelwijze die meer dan eens voorkomt⁶.

Wat betreft de *leerlingen* heeft er aan het eind van de basisschool een meting van de intelligentie plaatsgevonden en is er een milieuscore vastgesteld. Voor de intelligentie is gebruik gemaakt van de ISI-test en de milieuscore is gebaseerd op het beroep van de vader (o.c., 48). Analyse van deze gegevens leidt tot de conclusie dat "voor de verwachting dat moderne methoden vaker in klassen worden gebruikt met een wat hoger intelligentieniveau (...) ondersteuning (is, mvdh) gevonden" en dat er "significante verschillen tussen methoden (zijn, mvdh) ten aanzien van de gemiddelde intelligentiescores van klassen" (o.c., 59). Er worden echter geen toetsingsgegevens vermeld waaruit kan worden afgeleid dat er ook een toetsing heeft plaatsgevonden die een sig-

nificant verschil heeft opgeleverd tussen de twee soorten methoden (zie o.c., 57). Pas als dit wel is gebeurd, kan er een conclusie worden getrokken over de relatie tussen intelligentie en type methode. En dan niet de conclusie die Harskamp heeft geformuleerd, maar de conclusie dat aan het eind van de basisschool de gemiddelde intelligentiescore hoger is in klassen met een moderne methode dan in klassen met een traditionele methode (zie ook de reactie van Freudenthal, 1989)⁷.

5 *Verschillen in leerprestaties*

De kern van het onderzoek wordt gevormd door de vraag of er verschillen zijn in leerprestaties tussen de methoden als rekening wordt gehouden met de omstandigheden waaronder de methoden worden gebruikt. Om de effecten van de methoden te meten is gebruik gemaakt van de CITO-eindtoets rekenen en de CITO-eindtoets tabellen & grafieken. Verder zijn aanvullende RION-toetsen ontwikkeld voor de leerstofonderdelen schatten van standaardmaten, meten, verhoudingen en combinatoriek. Ook is nog een vragenlijst gebruikt waarmee wordt gemeten welk belang kinderen hechten aan rekenen en of ze er plezier in hebben (Harskamp, 1988, 44). Er hebben twee analyses plaatsgevonden, een op leerling- en een op klassenniveau.

5.1 *Analyse op leerlingniveau*

Begonnen is met de analyse op leerlingniveau. Via covariantie-analyses is nagegaan of er – als er wordt gecorrigeerd voor intelligentie – verschillen zijn tussen de rekenmethoden ten aanzien van de rekenprestaties van de leerlingen.

Wat betreft het corrigeren voor intelligentie kunnen zich volgens Harskamp twee belemmeringen voordoen (o.c., 59). De eerste belemmering heeft te maken met de beïnvloedbaarheid van intelligentie door onderwijs. Omdat het hier gaat om de achteraf gemeten intelligentie, is het corrigeren voor intelligentie alleen toelaatbaar als het onderwijs dat de leerlingen hebben gehad geen invloed heeft op de prestaties op de intelligentietest op het eind van de basisschool. Op grond van de geraadpleegde literatuur (Detterman & Sternberg, 1982; Slavenburg, 1986) is Harskamp (1988, 20 en 59) inderdaad die mening

toegedaan. Daarmee is de eerste belemmering van de baan. Maar hoe gegrond is die mening?

Op de eerste plaats is het belangrijk, dat een onderscheid wordt gemaakt tussen de mogelijkheid om via speciale trainingen een duurzame verbetering van het intelligentieniveau te bewerkstelligen en de mogelijkheid dat een aantal jaren onderwijs leidt tot een verhoging van het intelligentieniveau (zie ook Freudenthal, 1989). Het is echter niet duidelijk waar de aangehaalde literatuur betrekking op heeft. Wat betreft de invloed van het onderwijs op het intelligentieniveau, kan tegenover de mening van Harskamp in ieder geval de mening van Lohman worden geplaatst. In een recente publikatie stelt Lohman (1989), dat nog te vaak wordt gedacht dat intelligentietesten de aangeboren bekwaamheid van de leerling meten. Volgens hem meten deze testen juist een van de belangrijkste uitkomsten van het onderwijs, namelijk de vaardigheid om datgene wat op school is geleerd over te dragen naar nieuwe situaties. Naast allerlei vaardigheden die dicht liggen bij wat de kinderen onderwezen is, de zogenoemde 'crystallized abilities', is ook de 'fluid ability' een produkt van onderwijs en ervaring. Een onderzoek dat betrekking had op verschillende taal/leesprogramma's toonde zelfs aan dat verschillende programma's verschillende scores bij deze twee soorten 'abilities' kunnen opleveren. Bij de erg gestructureerde programma's met veel 'drill and practice' waren de vorderingen bij de 'fluid ability' het laagst en die bij de vakspecifieke vaardigheden het hoogst, terwijl bij de minder gestructureerde en meer op probleem oplossen gerichte programma's het omgekeerde patroon was te zien. Of met de woorden van Lohman: "You get out a program what you put into it" (Lohman, 1989, 9).

De tweede belemmering heeft betrekking op een mogelijk verschillende uitwerking van methoden op de rekenprestaties van kinderen met een verschillende intelligentie. Daarom wordt door middel van multivariate variantie-analyse (met methode en intelligentie als onafhankelijke variabelen en de verschillende toetscores als afhankelijke variabelen) eerst gekeken of er een interactie bestaat tussen methode en intelligentie op de toetsprestaties. Dit blijkt niet het geval te zijn (Harskamp, 1988, 60).

Na de eliminatie van de twee mogelijke be-

lemmeringen is overgegaan tot de covariantie-analyse met intelligentie als covariaat. Eerst is er multivariaat getoetst met alle afhankelijke variabelen. Deze multivariate toetsing leverde een methode-effect op (o.c., 60). Onduidelijk is of het hier behalve de factor 'methode', ook de factor 'type methode' betreft. Toetsingsgegevens ontbreken.

De vervolgens uitgevoerde univariate covariantie-analyses leverden geen significant verschil op tussen de methoden ten aanzien van de totaalscores bij de CITO-toets rekenen en de CITO-toets tabellen & grafieken (o.c., 62) en leverden ook geen significant verschil op bij de RION-toets als geheel (o.c., 60). Alleen voor het onderdeel combinatoriek is een significant verschil gevonden tussen de acht methoden én tussen de twee typen methoden (o.c., 60).

Behalve het significante verschil bij het toetsonderdeel combinatoriek, meldt Harskamp ook dat er een opvallend hoge samenhang is ($r = .9$) tussen de relatieve aandacht van de acht methoden voor het leerstofonderdeel combinatoriek en de gemiddelde leerlingsscore per methode op dat onderdeel (o.c., 60). Ook voor twee andere onderdelen van de RION-toets, de onderdelen verhoudingen en meten zijn tamelijk sterke positieve correlaties (resp. $r = .9$ en $r = .7$) gevonden tussen de relatieve aandacht voor het betreffende leerstofgebied in de acht methoden en de gemiddelde toetsresultaten van de leerlingen per methode (o.c., 86)⁸. Harskamp hecht aan deze uitkomst kennelijk niet zoveel belang. Zijn conclusie is: "Sommige methoden leiden tot iets betere leerlingresultaten op onderdelen van de toetsen dan andere methoden, maar *over het geheel genomen zijn er dus geen verschillen tussen de methoden* (curs., mvdh) op zowel algemene rekenvaardigheden (CITO-toets) als op toepassingen (RION-toets)" (o.c., 86).

In de gevonden samenhang tussen de mate van aandacht voor een leerstofonderdeel en de prestaties op dat onderdeel ziet Harskamp een bevestiging van de 'conclusie' van Walker en Schaffarzick (1974), dat methoden vooral zullen verschillen in leerresultaten bij leerstofonderdelen die verschillende nadruk in de methoden krijgen (Harskamp, 1988, 86). De redenering van Harskamp c.s. is ongeveer als volgt: er zijn weliswaar lichte verschillen tussen methoden in nadruk op bepaalde leerstof-

onderdelen, maar over het geheel genomen maakt het voor de rekenprestaties niet uit met welke methode de kinderen les krijgen (Harskamp en Suhre, 1987, 23).

Dat is nu precies waar Walker en Schaffarzick voor waarschuwen! Zij begonnen – zoals zij zelf zeggen – hun studie naar onderzoeken waarbij curricula worden vergeleken zoals er altijd wordt begonnen, naïef, met het zoeken naar gegevens die zouden kunnen wijzen op de superioriteit van moderne curricula ten opzichte van traditionele curricula. Wat ze vonden was geen superioriteit, maar analogie. De curricula deden het beter met betrekking tot die leerstofonderdelen in de toetsen die het meest overeenkomen met wat het curriculum biedt. Dit blijkt echer zo'n vanzelfsprekende uitkomst te zijn, dat er door niemand aandacht aan wordt besteed. Walker en Schaffarzick willen daarentegen juist, dat er meer aandacht wordt besteed aan de voor de hand liggende uitkomst, dat verschillende curricula leiden tot verschillen in prestaties (vgl. Lohman, 1989). Aan het eind vatten ze het nog eens kernachtig samen: curricula moeten gezien worden als middel – en naar het schijnt als een machtig middel – om de leerprestaties van leerlingen te sturen. Kennelijk hebben Harskamp c.s. de essentie van wat Walker en Schaffarzick te zeggen hebben over het hoofd gezien. Ook zij gaan nogal achteloos voorbij aan de vanzelfsprekende uitkomst dat de mate van aandacht voor bepaalde leerstofonderdelen samenhang vertoont met de prestaties op die leerstofonderdelen.

Behalve naar de leerprestaties is door Harskamp c.s. ook gekeken naar de invloed van de methode op het belang dat kinderen hechten aan rekenen en het plezier dat ze er in hebben. Ook nu weer was er geen sprake van een significant verschil (Harskamp, 1988, 62). Voor-

standers van de “moderne (zogenoemde ‘realistische’) rekendidactiek” verwachten volgens Harskamp, dat moderne methoden behalve tot betere leerprestaties, ook tot een positievere houding van leerlingen ten opzichte van rekenen leiden (o.c., 11). Dit is een misverstand. Bij de aangehaalde auteurs (Treffers, 1982; Van den Heuvel-Panhuizen & Goffree, 1986) kan dit ook niet worden gevonden. Als voorstanders van de realistische didactiek het hebben over een betere attitude van leerlingen, dan hebben ze het op de eerste plaats over een *wiskundige attitude*, een *onderzoeksgerichte instelling* en *wiskundige aanpak* van problemen (vgl. Freudenthal, 1978; Streefland, 1980). Plezier eraan beleven kan er bij komen, maar daarin hoeven de kinderen niet te verschillen van de kinderen die mechanistisch rekenonderwijs krijgen. De resultaten van een hypothesen-onderzoek bij het MORE-project (Gravemeijer, Van den Heuvel-Panhuizen & Van der Ploeg, t.p.) laten duidelijk zien dat dit door deskundigen op het gebied van de realistische didactiek ook niet zozeer wordt verwacht.

Ook de uitspraak dat de uitgevers van de drie onderzochte moderne methoden (OR, TT, en GIB) het ontwikkelen van een positieve rekenattitude als doel stellen (Harskamp, 1989, 6), is niet in overeenstemming met hetgeen eerder is gemeld. Alleen GIB en TT hebben gevraagd om een instrument voor de rekenattitude en OR niet (Harskamp, 1988, 43-44). OR blijkt ook maar in beperkte mate belang te hechten aan het betreffende instrument. TT vindt het daarentegen zeer belangrijk en GIB, NAT en NCR vinden het belangrijk (zie Tabel 1). Deze scores hebben echter niet betrekking op het belang van de doelstelling zelf, maar op het belang dat wordt gehecht aan het instrument. Het zegt volgens Harskamp iets over de inhoudsvaliditeit ervan (o.c., 50)⁹.

Tabel 1 *Belang toetsen beoordeeld door uitgevers (Harskamp, 1988, 50, tabel 5.4).*

methode	OVS	NAT	NZR	NR	NCR	OR	GIB	TT
CITO-toets Rekenen	-	5	6	3	4	5	4	5
RION-toets Rekenen	-	5	?	3	5	4	5	6

1 = totaal onbelangrijk

2 = niet belangrijk

3 = nauwelijks belangrijk

3 = enigszins belangrijk

5 = belangrijk

6 = zeer belangrijk

	sign. verschil tussen methoden	voor IQ en milieu gecorrigeerde sign. correlaties tussen impl.en lkr.kenn. en toetsprestaties	relevante controle- variabelen
implementatie kenmerken			
% basisstof	+	+	•
lessen per week	-	+	•
minuten per les	-	-	
variatie in leerstof	+	+	
leerstofdifferentiatie	+	-	
leerkracht kenmerken			
ervaring methode	+	-	
leerl.-leerstof attitude	-	+	•
wiskunde in vooropl.	-	-	
leerling kenmerken			
intelligentie	+		•
milieu	-		•

Figuur 4 Overzicht implementatie-, leerkracht- en leerlingkenmerken en gebruikte controlevariabelen.

5.2 Analyse op klasseniveau

Bij de analyse op klasseniveau is nagegaan "of door middel van controle voor verschillen in implementatie-, leerkracht- en leerlingkenmerken, die niet het gevolg zijn van de gehanteerde methode, andere resultaten kunnen worden bereikt" dan bij de analyse op leerlingniveau (o.c., 87). Eerst is via de berekening van de partiële correlaties nagegaan welke kenmerken van belang zijn in verband met hun invloed op de klasseprestaties. Een keer is hierbij gecontroleerd voor de gemiddelde intelligentiescore van een klas en een keer voor de gemiddelde milieuscore (o.c., 64). Dit leverde vijf relevante controlevariabelen op: intelligentiescore, milieuscore, aantal rekenlessen, percentage leerlingen met basisstof en leerkrachtattitude (o.c., 65) (zie Figuur 4).

De twee vakdidactische implementatiekenmerken variatie in leerstof en leerstofdifferentiatie zijn niet als controlevariabele opgenomen, omdat deze variabelen mede beïnvloed worden door de factor methode (o.c., 65). Het algemene implementatiekenmerk percentage basisstof, waarvoor ook een significant verschil is gevonden tussen de methoden (o.c., 54), wordt daarentegen wel als controlevariabele meegenomen, terwijl dit kenmerk ook zou kunnen worden gezien als een effect van de methode. In de trant van: een goede methode leidt tot goede leerresultaten bij meer leerlingen, ofwel door een goede methode wordt met meer leerlingen de basisstof behandeld.

Harskamp zelf sluit ook niet uit dat er sprake kan zijn van een effect van de methoden, maar beschouwt dit zonder verdere toelichting als een ongewenst, niet door de methoden bedoeld effect (o.c., 85).

De uitgevoerde multivariate covariantie-analyse, waarbij gecontroleerd is voor de genoemde variabelen leverde geen significant methode-effect op ten aanzien van de CITO- en RION-toetsscores van de klassen (o.c., 66).

Hoewel Harskamp het niet erg plausibel acht, zou het volgens hem toch kunnen dat bepaalde methoden enige invloed hebben op het intelligentieniveau van klassen (o.c., 87). Vandaar dat ook nog een vergelijking is uitgevoerd zonder statistische controle voor intelligentie en de andere controlevariabelen. Voor de variabele percentage basisstof is in zekere zin wel gecontroleerd. De analyse is alleen gebaseerd op de klassen waarin met 75% of meer van de leerlingen de basisstof is behandeld. Ook nu weer leverde een multivariate variantie-analyse van de klasseprestaties geen significant verschil op tussen de acht methoden (o.c., 68).

Een secundaire analyse van de gegevens van Harskamp (o.c., 67) wijst echter uit, dat die minimum-eis van 75% bij de traditionele methoden leidt tot gemiddeld 41% teruggang in klassen, terwijl bij de moderne methoden sprake is van een gemiddelde teruggang van 25% (zie Tabel 2).

Tabel 2 *Teruggang aantal klassen per methode en per type methode bij de 75% eis.*

methode	OVS	NAT	NZR	NR	NCR	OR	GIB	TT
totaal aantal klassen	7	12	13	23	19	19	10	12
aantal klassen bij 75% eis	2	8	11	18	7	19	9	3
% vermindering per meth.	71	33	15	22	63	0	10	66
% vermindering per type meth.			41				25	

6 *Invloed implementatiekenmerken op leerresultaten*

Om de laatste onderzoeksvraag – welke verschillen in implementatie van rekenmethoden leiden tot verschillen in leerresultaten? (o.c., 5) – te beantwoorden, zijn alle factoren die samen bleken te hangen met de toetsprestaties van klassen, nadat er voor verschillen in intelligentiescore tussen klassen is gecorrigeerd, in een model gestopt. Dit zijn de implementatiefactoren percentage basisstof, aantal rekenlessen per week, variatie in leerstof en leerstofdifferentiatie (o.c., 71)¹⁰. Na toetsing van het model met behulp van het programma LISREL IV is de conclusie, dat de rol die de implementatiekenmerken spelen ongeacht de methode, nu duidelijker is dan bij de multivariate covariantie-analyse op klassenniveau. De belangrijkste factor voor de toetsprestaties blijkt het percentage basisstof te zijn. De factoren variatie in leerstof en aantal rekenlessen per week hebben een iets kleiner totaaleffect (o.c., 88).

7 *Aanbevelingen*

Na het onderzoek komt Harskamp met een aantal aanbevelingen. Verder onderzoek is volgens hem van belang om aan te tonen welke concrete lesgedragingen samenhangen met de hoeveelheid behandelde stof en de actieve leertijd van de leerlingen. Deze factoren zouden te zamen een algemeen model voor doelmatig rekenonderwijs vormen, dat een belangrijk gedeelte van de rekenprestaties kan verklaren (o.c., 91). Harskamp verwijst daarbij naar Brophy (1986), terwijl Brophy in plaats van te pleiten voor een algemeen model, juist vindt dat het onderwijsonderzoek minder 'content free' moet worden. Alleen op die manier kan volgens Brophy – die zelf

geen reken/wiskunde-didacticus is – het onderzoek het stadium van de kinderschoenen ontgroeien en een relevante basis verschaffen voor de onderwijspraktijk. Wil men verder komen dan moeten vakdidactici en algemeen onderwijskundigen gaan samenwerken, aldus Brophy (1986).

Bij de aanbevelingen richt Harskamp zich ook op de begeleiders en nascholers en op de makers van methoden. "Leerkrachten dienen doelmatig met een methode te kunnen werken in de zin van behandeling van leerstof en besteding van lestijd. Wellicht dat daarna vakdidactische kennis en vaardigheden van belang kunnen zijn om leerresultaten van leerlingen te verbeteren" (Harskamp, 1988, 93). Voor de methoden geldt, dat ze strikter dan nu het geval is, dienen aan te geven welke minimale mate van implementatie van algemene en vakdidactische kenmerken wordt verondersteld voor het bereiken van goede rekenresultaten (Harskamp, 1989, 12)¹¹. Harskamp besluit met: "Wanneer methodemakers er in slagen de methode op deze punten te verbeteren, zullen leerkrachten de methode effectiever kunnen gebruiken" (o.c., 12).

Met andere woorden: verbeterde methoden leiden tot een effectiever gebruik van die methoden; een effectiever gebruik van de methoden betekent dat met meer kinderen de basisstof wordt behandeld; vervolgens wordt er weer gecontroleerd voor deze variabele en... worden er weer geen verschillen tussen de methoden gevonden; wel blijkt het percentage basisstof een belangrijke factor te zijn voor de leerresultaten; op grond hiervan kunnen een aantal aanbevelingen worden gedaan gericht op de verbetering van de methoden... etc.

Samenvattend kan niet anders worden gezegd, dan dat de in de inleiding vermelde conclusies en aanbevelingen op een zeer zwakke basis zijn gebouwd¹². Ze wekken ten onrechte de indruk, dat de methoden qua leerstof en qua leerresultaten zo goed als niet verschillen. Over hoe het staat met de implementatie van de rekenmethoden zegt het onderzoek eigenlijk niets. Het heeft eerder betrekking op een aantal algemene onderwijscondities dan op het in praktijk brengen van gewenst onderwijs. De aanbevelingen die worden gedaan zijn deels niet onderzocht en voor zover ze wel voortvloeien uit de onderzoeksresultaten niet erg sterk te noemen.

Opgemerkt moet worden dat bij deze reactie op het onderzoek van Harskamp c.s. de door hen gekozen onderzoekspopzet als uitgangspunt is genomen. De kritiek heeft alleen betrekking op het onderzoek zoals het is uitgevoerd. Zaken als de keuze voor een eenmalige produkt-evaluatie en het ontbreken van gegevens over het feitelijke onderwijsgedrag zijn niet aan de orde gesteld.

Evenzo is buiten beschouwing gelaten, dat ook andere onderzoeken aantonen dat de conclusie dat de methode er niet toe doet niet houdbaar is. Zo blijkt uit de PPON-gegevens dat aan het eind van de basisschool de moderne methode Operator Rekenen op 10 van de 27 onderdelen significant hoger scoort dan de traditionele methode Naar Zelfstandig Rekenen (Treffers, 1990; zie ook Treffers, 1988). Ook de uitkomsten van het MORE-onderzoek, die in groep 3 al een methode-effect laten zien (Van den Heuvel-Panhuizen, 1989; Gravemeijer, Van den Heuvel-Panhuizen & Van der Ploeg, t.p.), zijn bij de kritiek niet in het geding gebracht.

Hoe nu verder? Hopelijk hebben de verschillende benaderingen van onderzoek van reken/wiskundeonderwijs in de toekomst elkaar nog wat te bieden en is de hier geuite kritiek op het onderzoek van Harskamp c.s. niet het begin van het einde. Want – zoals Brophy al eerder zei – om verder te komen in het onderzoek van reken/wiskundeonderwijs, is het noodzakelijk dat het tot een samenwerking komt van vakdidactici en algemeen onderwijskundigen. Brophy besluit het in 1986 gepresenteerde AERA-paper dan ook

met de woorden: "It is in that spirit that the comments and suggestions made in the present paper have been offered" (Brophy, 1986, 344). Hierbij wil ik me graag aansluiten.

Noten

1. Met de 'moderne methoden' worden hier bedoeld de methoden Taltaal, Getal in Beeld en Operator Rekenen. De methoden De Wereld in Getallen, Rekenwerk en Rekenen & Wiskunde die én van recentere datum zijn én meer kenmerken bevatten van de realistische didactiek vallen er dus buiten.
2. Van den Oever (1987, 241) omschrijft bij de beoordeling van het 'Eindrapport Vergelijking van rekenmethoden in het basisonderwijs' de door Harskamp c.s. gevolgde procedure ten aanzien van de methode-inhoud als een Procrustes-situatie.
3. Op een andere plaats wordt een andere operationalisering van de leerstofvariatie in methoden gegeven, nl. de verhouding tussen opgaven over hoofdrekken en redactierekenen enerzijds en kale sommen voor het inoefenen van algemene rekenvaardigheden anderzijds (Harskamp, 1988, 85).
4. Ook met betrekking tot de door leerkrachten uitgevoerde leerstofvariatie wordt op twee andere plaatsen weer een andere operationalisatie gehanteerd, nl. het percentage van de lestijd dat besteed wordt aan hoofdrekken en redactieopgaven *versus* (curs., mvdh) het percentage algemene rekenvaardigheden (Harskamp, 1988, 49 en 56).
5. Zonder dat uit toetsingsgegevens blijkt dat dit ook is getoetst, wordt hier later aan toegevoegd dat er bij deze kenmerken ook een verschil bestaat tussen de traditionele en de moderne methoden (Harskamp, 1988, 58-59, 84).
6. Bij de implementatiekenmerken variatie in leerstof en leerstofdifferentiatie (zie noot 5) en bij het leerlingkenmerk intelligentie (volgt nog) gebeurt dit ook.
7. Freudenthal richt zich in deze reactie vooral op de manier waarop in het onderzoek met het kenmerk intelligentie wordt omgesprongen. Twee voorname punten van kritiek zijn: a) de kronkel-formulering "Moderne rekenmethoden zullen vaker in klassen worden gebruikt met een wat hoger intelligentieniveau"; Freudenthal vraagt zich af waarom Harskamp voor deze formulering heeft gekozen, terwijl in feite wordt aange-toond dat in klassen waar gewerkt wordt met een moderne methode het intelligentieniveau hoger is; en b) de veronderstelling van Harskamp dat (reken)onderwijs geen invloed heeft op het intelligentieniveau van kinderen en de onvoldoende

- onderbouwing van deze veronderstelling. Een meer algemeen punt van kritiek is, dat de impact van moderne rekenmethoden een te ingewikkelde kwestie is om eenvoudigweg met statistiekjes over te beslissen.
8. Op pag. 60 staan andere correlaties: verhoudingen $r = .6$; meten $r = .9$.
 9. Op pag. 62 lijkt het of Harskamp dit zelf is vergeten. Verwijzende naar tabel 5.4 (o.c., 50) zegt hij, dat het ontwikkelen van een positieve rekenattitude een belangrijke doelstelling is van de meeste methoden.
 10. Er wordt hierbij verwezen naar tabel 5.10 (Harskamp, 1988, 64), terwijl in die tabel staat aangegeven dat leerstofdifferentiatie bij correctie voor intelligentie niet significant samenhangt met de toetsprestaties.
 11. Uit niets blijkt, dat Harskamp heeft onderzocht of het helpt als dit in de methode wordt aangegeven. Ook is onduidelijk waarop de uitspraak is gebaseerd, dat in vrijwel alle methoden een duidelijke planning ontbreekt van de te behandelen leerstof en de te besteden lessen in het cursusjaar (Harskamp, 1989, 12). Het voorbeeld dat Harskamp (1988, 92) hierbij noemt, dat bij De Wereld in Getallen (WIG) gras in 1986 in een gebruikersbulletin uiteen is gezet wat in deze methode als basisstof en als verrijgingsstof kan worden aangemerkt, klopt in ieder geval niet. Al in de eerste druk (van de deeltjes 1a en 1b) van WIG wordt aangegeven wat tot het minimumprogramma behoort, wat tot het basisprogramma en wat tot het uitbreidingsprogramma.
 12. Zie in dit verband ook de SVO-beoordeling van het betreffende RION-onderzoek (Van den Oever, 1987) en de kritische kanttekeningen die twee uitgevers bij dit onderzoek hebben gemaakt (Van Velthoven & Janssen, 1986).

Literatuur

- Brophy, J., Teaching and learning mathematics: Where research should be going. *Journal for Research in Mathematics Education*, 1986, 17, 5, 323-346.
- Detterman, D. K. & R. J. Sternberg, *How and how much can intelligence be increased*, Norwood: Ablex, 1982 (geciteerd door Harskamp, 1988).
- Gravemeijer, K.P.E., M. van den Heuvel-Panhuizen & D. A. van der Ploeg, *Ander rekenboek, andere rekenprestaties. Een vergelijking tussen twee reken-wiskundemethoden in groep 3*. Utrecht: OW & OC, ter perse.
- Harskamp, E. G., *Rekenmethoden op de proef gesteld* (Dissertatie). Groningen: RION, 1988.
- Harskamp, E., *Implementatie- en effectonderzoek van nieuwe reken-wiskundemethoden in het basisonderwijs*. Paper t.b.v. Symposium 'Evaluatie en de ontwikkeling van wiskundeonderwijs', Groningen, 23 juni 1989.
- Harskamp, E. & C. Suhre, Rekenmethoden op de proef gesteld. *School*, 1987, nov., 20-23.
- Harskamp, E. & C. Suhre, Rekenmethoden vergeleken: een effectonderzoek aan het einde van de basisschool. *Pedagogische Studiën*, 1988, 65, 5, 208-218.
- Heuvel-Panhuizen, M. van den & F. Goffree, *Zo rekent Nederland*. Enschede: SLO, 1986.
- Heuvel-Panhuizen, M. van den, De eerste uitkomsten. In: E. de Moor (Ed.), *Panama Cursusboek 7*. Utrecht: HMN(SOL)/OW&OC, 1989.
- Jong, R. de, *Wiskobas in methoden* (Dissertatie). Utrecht: OW&OC, 1986.
- Fullan, M. & A. Pomfret, Research on curriculum and instruction implementation. *Review of Educational Research*, 1977, 47, 335-397.
- Freudenthal, H., *Weeding and Sowing. Preface to a science of Mathematical Education*. Dordrecht/Boston: Reidel, 1978.
- Freudenthal, H., *Reactie naar aanleiding van het symposium 'Evaluatie en ontwikkeling van wiskundeonderwijs' - Groningen 23 juni 1989*. Brief, 17 juni 1989.
- Lohman, D. F., *Two implications of cognitive psychology for educational measurement*. Adress to the Division of Learning and Instruction Institute for Educational Research in the Netherlands, Utrecht, January 27, 1989.
- Oever, A. C. C. van den, Beoordeling van het eindrapport 'Vergelijking van rekenmethoden in het basisonderwijs' uitgevoerd door het RION. *Tijdschrift voor Onderwijsresearch*, 1987, 12, 4, 239-242.
- Slavenburg, J. H., *Onderwijsstimulering en Gezinsactivering*. 's-Gravenhage: SVO, 1986 (geciteerd door Harskamp, 1988).
- Streefland, L., *Makro-structurele verkenningen voor het wiskundeonderwijs*. Utrecht: IOWO, 1980.
- Suhre, C., Effectief gebruik van Rekenmethoden: Evaluatieonderzoek in het laatste jaar van de basisschool. *Tijdschrift voor Onderwijsresearch*, 1987, 12, 6, 344-354.
- Suhre, C., Rekenmethoden onderzocht. *Basis-schoolmanagement*, 1988, sept., 4-6.
- Treffers, A., Cijferen in het rekenonderwijs, toen en nu. *Pedagogische Studiën*, 1982, 59, 3, 97-115.
- Treffers, A. & F. Goffree, Rational analysis of realistic mathematics education - the Wiskobas Program. In: L. Streefland (Ed.), *Proceedings of the ninth international conference for the psychology of mathematics education. Volume II*. Utrecht: OW&OC, 1985, 97-121.

- Treffers, A. & E. Feijs, Realistisch reken-wiskundeonderwijs. *Willem Bartjens*, 1987, 6, 4 (juli), 195-193.
- Treffers, A., Over de merkbare invloed van onderwijsmethoden op leerprestaties. In: J. Wijnstra, *Balans van het rekenonderwijs in de basisschool*. Arnhem: CITO, 1988.
- Treffers, A., Over het nieuwe programma voor rekenen-wiskunde op de basisschool. *Nieuwe Wiskrant*, 1989, 8, 3, 10-18.
- Treffers, A., *Het voorkomen van ongecijferdheid op de basisschool* (Oratie). Utrecht: VOU/OW&OC, 1990.
- Velthoven, J.A.M. van & G.M. Janssen, Het RION-onderzoek: Een valse start of een goed begin? In: E.G.rHarskamp & C.J.M. Suhre, *Vergelijking van rekenmethoden in het basisonderwijs. Eindrapport*. Groningen: RION, 1986.
- Walker, D.F. & J. Schaffarzick, Comparing curricula. *Review of Educational Research*, 1974, 44, 1, 83-111.

Curriculum vitae

M. van den Heuvel-Panhuizen (1950) was werkzaam in het basisonderwijs en in het speciaal onderwijs. Studeerde onderwijskunde aan de Rijksuniversiteit te Utrecht en specialiseerde zich op het gebied van de reken-wiskundendidactiek van de basisschool. Heeft meegewerkt aan het NOT-programma 'Pluspunt', de methode 'Zo reken ik ook!' en SLO-publikaties over rekenen-wiskunde. Thans is zij werkzaam bij de vakgroep OW & OC van de RUU en houdt zich daar onder meer bezig met onderzoek naar gebruik en effecten van reken-wiskundemethoden (MORE-project).

Adres: Rijksuniversiteit Utrecht, Vakgroep OW & OC, Tiberdreef 4, 3561 GG Utrecht.

Manuscript aanvaard 25-1-'90

Summary

Heuvel-Panhuizen, M. van den. 'Research on mathematics textbooks tested.' *Pedagogische Studiën*, 1990, 67, 271-283

Do the learning results depend on the textbooks as used? This is an actual and also very important question. Harskamp and Suhre's answer to this question is not an affirmative one. They compared eight series of Dutch mathematics textbooks. The result of their research was that there is hardly any difference neither between the content of the textbooks nor between the test scores of the children taught with different textbooks. In the present article a critical reflection has been carried out with respect to the conclusions and recommendations of this research. The executed analysis shows that the conclusions and recommendations of Harskamp and Suhre have been founded too weakly to stand the criticism as given.