

Huub van den Bergh, Jan Rijkers & Michel Zwarts

## Samenvatting

In deze studie zijn begripvend leesprestaties van achtstegroepers in zes opeenvolgende jaren ( $N = \pm 6767$  van basisscholen  $\pm 310$  per jaar) gerelateerd aan het type leesmethode dat op de school gebruikt is. Hierbij is een onderscheid gemaakt in drie typen leesmethoden: nieuwe methoden, die dekkend zijn voor de kerndoelen van het basisonderwijs (type 1), oudere methoden waarin kerndoelen noch leesstrategieën aan de orde komen (type 3), en methoden uit de tussenliggende periode, waarin een beperkte aandacht is voor de kerndoelen en/of leesstrategieën (type 2).

De gemiddelde leesvaardigheid van de leerlingen blijkt niet te variëren van type methode tot type methode: gemiddeld genomen kan géén onderscheid tussen de oudere en nieuwere methoden aangetoond worden. Wel blijken verschillen in varianties tussen de drie typen methoden aantoonbaar; zowel de verschillen tussen leerlingen, als de verschillen tussen leestaken zijn groter bij de methoden uit de derde categorie, dan bij de methoden uit de eerste of de tweede categorie.

Het schoolgewicht, als indicator voor de sociaal economische samenstelling van een school bleek, een (relatief) goede predictor van de verschillen in begripvend leesprestaties tussen scholen. Rekening houdend met het schoolgewicht en nodaalgebied – als indicator voor urbanisatiegraad – zijn de verschillen tussen scholen bijzonder klein ten opzichte van de verschillen tussen jaren. Daarom kan geconcludeerd worden dat schoolkarakteristieken die over het algemeen beschouwd worden als indicatoren van effectief onderwijs alleen een ondergeschikte rol kunnen spelen.

doelstellingen van het basisonderwijs. In 1995 is de kwaliteit van het onderwijs in begripvend en studerend lezen door de Inspectie van het Onderwijs geëvalueerd (Inspectie van het Onderwijs, 1996). In deze evaluatie is, naast de inhoud van dat onderwijs (het aanbod), met name aandacht besteed aan het instructieproces. Uitgangspunt bij de evaluatie was, dat goed onderwijs in begripvend en studerend lezen wordt gekenmerkt door een leerstofaanbod dat de kerndoelen bevat, inclusief gestructureerde en systematische aandacht voor strategische vaardigheden en door een methodiek waarbij de leraren tijdens het instructieproces de procedures en technieken van procesgericht, strategisch leesonderwijs hanteren. Op ongeveer één zesde deel van de basisscholen bleek de onderwijspraktijk op een goede manier deze kenmerken te vertonen. Op ongeveer de helft van de scholen ontbraken in het leerstofaanbod en/of in de methodiek essentiële onderdelen waardoor het leesonderwijs onvoldoende kwaliteit had. Het belangrijkste knelpunt op deze scholen was, dat het strategische karakter in het leesonderwijs ontbrak.

Daarnaast bleek, dat op slechts een klein deel van de scholen (5 - 10%) de leesvorderingen van de leerlingen systematisch werden gevolgd en het leesonderwijs werd aangepast aan de geconstateerde niveaoverschillen.

Eén van de belangrijkste determinanten van de kwaliteit van het leerstofaanbod en het instructieproces bleek de gebruikte methode (Inspectie van het Onderwijs, o.c.). Een noodzakelijke voorwaarde voor goed of voldoende strategisch leesonderwijs was de aanwezigheid van een recente methode; op scholen zonder dergelijke methode werd geen strategisch leesonderwijs aangetroffen. Tweederde deel van de scholen met een recente methode bleek zich de praktijk van procesgericht leesonderwijs eigen te hebben gemaakt.

## 1 Inleiding

De ontwikkeling van een goede vaardigheid in begripvend lezen is één van de belangrijkste

Eén van de redenen om nieuwe leesmethoden te ontwikkelen en op de markt te brengen is de

verbetering van de leesvaardigheid van de leerlingen. Een 'betere leesvaardigheid' is in dit opzicht een meer-dimensionaal begrip, dat gewenste uitkomsten aangeeft. Immers, met een 'betere leesvaardigheid' wordt niet uitsluitend bedoeld op decodeer- en begripvaardigheden, maar ook op (de verwerving van) leesstrategieën. Leesstrategieën verwijzen hierbij naar processen als: planning, evaluatie en regulatie van het eigen leesproces (Baker & Brown, 1984), het stellen van vragen aan de tekst, achterhalen van het hoofdthema, voorspellen van de clou van een tekst etc. (Paris, Wasik & Turner 1991; Aarnoutse 1991; van Hout-Wolters, 1992). In de leesmethoden die vanaf de negentiger jaren op de markt zijn gekomen, wordt ook aan deze strategieën aandacht besteed. Dit in tegenstelling tot de oudere leesmethoden.

In onderzoek zijn effecten van deze 'nieuwe' leesmethoden aangetoond. Dat wil zeggen: effecten op instrumenten die specifiek ontwikkeld zijn om na te gaan of de 'nieuwe' doelen bereikt zijn (Palincsar & Brown, 1984; Brand-Gruwel, 1995); leerlingen kennen na een training de geoefende leesstrategieën en kunnen die ook daadwerkelijk toepassen.

Hoewel in de 'nieuwe' methoden meer aandacht besteed wordt aan leesstrategieën en aan de functie van lezen, rijst de vraag naar de mate waarin deze methoden ook effecten hebben op de leesvaardigheid als geheel. Heeft het aanleren van een diversiteit aan leesstrategieën ook effect op het tekstbegrip van leerlingen, of is dit effect alleen beperkt tot zeer specifieke onderdelen van de leesvaardigheid?

In deze studie gaan wij in op de effecten van verschillende leesmethoden op de leesprestaties van leerlingen aan het einde van het basisonderwijs. Hiertoe worden de prestaties van leerlingen die met verschillende begrijpend-leesmethoden les kregen vergeleken. Het resultaat dat een leraar met zijn leerlingen bereikt is natuurlijk niet alleen afhankelijk van het gebruik van een bepaalde methode. Ook kenmerken van de leerlingen zijn van belang. Een school met een minder taalvaardige leerlingpopulatie kan niet als minder effectief gekarakteriseerd worden, enkel en alleen vanwege die leerlingpopulatie zonder rekening te houden met het instructieproces (vergelijk o.a. Bosker, 1991). Ook de ervaring die een school heeft

met een leesmethode moet wat dit betreft genoemd worden. Wanneer leraren al jaren met dezelfde methode werken en de methode door en door kennen, kan dat zijn weerslag hebben op het instructieproces, en zodoende op de bereikte prestaties (vergelijk: Chall, 1983). Bij een vergelijking zouden dan de 'oude' methoden bevoorrecht worden boven de 'nieuwe' methoden.

De resultaten uit deze studie vergelijken wij tot slot met de resultaten van de periodieke peiling van het onderwijsniveau (PPON), waarin geconcludeerd is dat leesonderwijs waarin geen aandacht is voor leesstrategieën tot lagere leesprestaties van leerlingen leidt (Sijtstra, 1997).

## 2 Gegevens

### Leesmethoden

Op grond van de inhoud en de gehanteerde didactische principes zijn de 24 meest gebruikte methoden voor begrijpend lezen in drie categorieën ingedeeld. In de eerste categorie vallen alleen de methoden die de kerndoelen van het basisonderwijs dekken en daarnaast in voldoende mate aandacht besteden aan leesstrategieën. Het betreft hier methoden die in de negentiger jaren op de markt zijn verschenen. In de derde categorie zijn voornamelijk de oudere leesmethoden ondergebracht; methoden waarin veel van de kerndoelen niet aan de orde komt en waarin een productgerichte benadering van het leesonderwijs gehanteerd wordt. De tweede categorie bestaat uit methoden die het midden houden tussen het eerste en het derde type; de inhoud dekt in meer of mindere mate de kerndoelen en slechts enkele van de aspecten van strategisch leesonderwijs komen aan de orde. De eerste categorie bestaat uit drie, de tweede uit zes en de derde uit vijftien verschillende methoden voor begrijpend lezen (zie: Tabel 1).

In een landelijke steekproef van 3453 basisscholen is nagegaan welke leesmethode gebruikt werd. Daarnaast hebben de scholen aangegeven hoe lang zij de huidige methoden gebruikten en spraken zij zich uit over hun plannen voor de toekomst. In totaal heeft 97% van de scholen de beide eerste vragen beant-

woord en 10% de derde vraag. In Tabel 1 wordt een overzicht gegeven van de resultaten.

**Tabel 1**  
*Mate van gebruik van de methoden voor begrijpend lezen (percentages)*

| Methoden                        | Gebruik in 1995 | Toekomstig gebruik |
|---------------------------------|-----------------|--------------------|
| <i>A-categorie</i>              | 20.2            | 59.6               |
| Lees je wijzer                  | 12.3            | 21.4               |
| Wie dit leest                   | 4.8             | 23.7               |
| Ik weet wat ik lees             | 3.1             | 14.5               |
| <i>B-categorie</i>              | 49.7            | 25.3               |
| Leeswerk                        | 25.2            | 15.0               |
| Begrijpend lezen/Taal totaal    | 15.0            | 4.0                |
| Taalkabaal                      | 7.6             | 1.4                |
| Lezen in Balans                 | 1.1             | 4.0                |
| Taaltoren                       | 0.8             | 0.9                |
| <i>C-categorie</i>              | 30.2            | 15.0               |
| Met inzicht leren lezen         | 5.5             | 1.5                |
| Stillezen                       | 5.1             | 0.0                |
| Eerste stap op het studiepad    | 3.6             | 0.3                |
| De leesmarkt                    | 2.1             | 0.0                |
| Moderne leesvormen              | 2.1             | 0.0                |
| Actief lezen                    | 0.8             | 0.0                |
| Tong en teken                   | 0.8             | 7.2                |
| Nieuwe leesvormen               | 0.6             | 0.0                |
| Jong studeren                   | 0.3             | 0.6                |
| Allemaal taal                   | 0.2             | 0.6                |
| Leeskisten                      | 0.2             | 0.0                |
| Taal actief                     | 0.1             | 0.0                |
| Bij jou en bij mij              | 0.1             | 0.0                |
| Taal voor het leven             | 0.1             | 0.0                |
| Geen methode                    | 8.6             | 4.9                |
| Aantal scholen                  | 3256            | 3463               |
| Ontbrekende antwoorden (aantal) | 197             | 107                |

Uit Tabel 1 blijkt dat in 1995 iets meer dan 20% van de basisscholen een methode uit de eerste categorie gebruikt. Bijna de helft gebruikt een methode uit de tweede categorie en bijna één derde een methode uit de derde categorie. Van de scholen die overwegen een nieuwe leesmethode in te voeren kiest meer dan zestig procent voor een van de methoden uit de eerste categorie. Het marktaandeel van methoden uit de tweede en derde categorie neemt navenant af.

### Tekstbegrip

De scholen die de vragen over het methodengebruik hebben beantwoord, zijn benaderd met de vraag of zij gebruik maken van de Cito-toets. Indien dat het geval was, is toestemming

gevraagd om de resultaten van de eindtoets van 1990 tot en met 1995 te mogen analyseren. Meer dan 98% van de gevraagde scholen verleende hiervoor toestemming. Van deze scholen zijn bij het Cito de resultaten opgevraagd. Voor de analyse is uit dit bestand een steekproef getrokken van ruim 300 scholen (zie Tabel 2 voor een overzicht).

De Cito-eindtoets bestaat uit de onderdelen 'informatieverwerking', 'rekenen' en 'taal'. In dit onderzoek is alleen gebruik gemaakt van (de scores op) dit laatste onderdeel. De taaltoets omvat vijf onderdelen waarin het tekstbegrip van de leerlingen gemeten wordt: (1) het gebruik van informatiebronnen, (2) het lezen van kaarten, (3) het lezen van tabellen en grafieken, (4) het lezen van teksten: identificatie van informatie, en (5) het lezen van teksten: deductie van informatie.

In dit onderzoek zijn de scores van de leerlingen gestandaardiseerd per taak, zodat voor elke taak het gemiddelde gelijk is aan nul en de standaarddeviatie gelijk is aan één. Vervolgens zijn de scores getransformeerd naar een (arbitraire) schaal met een minimum van nul, een maximum van tien en een gemiddelde van 7.2 ( $SD = 2.2$ )<sup>1</sup>.

Uit Tabel 2 blijkt dat voor alle drie de methoden het aantal scholen enigszins varieert van jaar tot jaar; niet elke school heeft elk jaar aan de Cito-eindtoets deelgenomen.

### Schoolgewicht, ervaring, klassengrootte en nodaalgebied

De toekenning van (personele) middelen aan scholen is (mede) afhankelijk van de samenstelling van de schoolbevolking, uitgedrukt in de zogenoemde leerlinggewichten. Allochtone leerlingen krijgen hierbij een gewicht toegewezen van 1.90, kinderen in een achterstandssituatie (bijvoorbeeld beide ouders hebben een relatief laag opleidingsniveau) een gewicht van 1.25 en de andere kinderen een gewicht van 1.00. Het gewogen gemiddelde van de leerlinggewichten is dan het schoolgewicht. Het schoolgewicht is derhalve een indicator voor de schoolpopulatie; op scholen met een hoog schoolgewicht zitten meer allochtone leerlingen en/of leerlingen uit achterstandssituaties dan op scholen met een lager schoolgewicht. Daar in scholen met verhoudingsgewijs veel

Tabel 2

Het aantal scholen, aantal leerlingen, de gemiddelde klassengrootte, het aantal scores, het gemiddelde aantal ervaringsjaren, en het schoolgewicht voor drie categorieën van leesmethoden per jaar (waar van toepassing: standaarddeviaties tussen haakjes)

|                                   | Jaar  |             |             |             |             |             |
|-----------------------------------|-------|-------------|-------------|-------------|-------------|-------------|
|                                   | 1990  | 1991        | 1992        | 1993        | 1994        | 1995        |
| <b>Methode: categorie 1</b>       |       |             |             |             |             |             |
| Aantal scholen                    | 106   | 108         | 111         | 115         | 123         | 123         |
| Aantal leerlingen                 | 2312  | 2414        | 2430        | 2460        | 2681        | 2685        |
| Klassengrootte (gem.)             | 21.8  | 22.3 (11.5) | 21.9 (11.3) | 21.4 (10.4) | 21.8 (10.6) | 21.8 (11.1) |
| Aantal scores                     | 11560 | 12070       | 12150       | 12300       | 13011       | 13425       |
| Ervaringsjaren                    | 2.94  | 2.93 (1.55) | 2.87 (1.53) | 2.91 (1.54) | 2.87 (1.56) | 2.89 (1.59) |
| Schoolgewicht                     | 1.14  | 1.14 (1.16) | 1.14 (1.16) | 1.13 (1.16) | 1.13 (1.16) | 1.13 (0.16) |
| Nodaalgebied (*10 <sup>-3</sup> ) | 71.3  | 71.9 (90.8) | 75.1 (90.3) | 70.7 (102)  | 70.0 (101)  | 70.0 (98.1) |
| <b>Methode: Categorie 2</b>       |       |             |             |             |             |             |
| Aantal scholen                    | 105   | 109         | 111         | 117         | 128         | 128         |
| Aantal leerlingen                 | 2371  | 2363        | 2474        | 2631        | 2890        | 2743        |
| Klassengrootte (gem.)             | 22.6  | 21.7 (14.2) | 22.3 (13.7) | 22.5 (14.4) | 22.6 (14.0) | 21.4 (12.3) |
| Aantal scores                     | 11855 | 11855       | 12370       | 13155       | 13345       | 13715       |
| Ervaringsjaren                    | 4.53  | 4.54 (1.08) | 4.52 (1.07) | 4.88 (1.08) | 4.47 (1.01) | 4.45 (1.12) |
| Schoolgewicht                     | 1.16  | 1.16 (1.21) | 1.15 (1.22) | 1.15 (1.22) | 1.15 (1.21) | 1.15 (1.21) |
| Nodaalgebied (*10 <sup>-3</sup> ) | 84.1  | 79.3 (126)  | 83.4 (115)  | 79.2 (124)  | 76.5 (122)  | 76.5 (117)  |
| <b>Methode: Categorie 3</b>       |       |             |             |             |             |             |
| Aantal scholen                    | 71    | 71          | 76          | 86          | 88          | 89          |
| Aantal leerlingen                 | 1541  | 1479        | 1630        | 1709        | 1894        | 1897        |
| Klassengrootte (gem.)             | 21.7  | 20.8 (12.8) | 21.4 (12.2) | 19.8 (10.5) | 19.4 (12.5) | 21.3 (11.6) |
| Aantal scores                     | 7705  | 7395        | 8150        | 8545        | 8525        | 9485        |
| Ervaringsjaren                    | 4.94  | 4.96 (1.40) | 4.96 (1.31) | 4.95 (1.33) | 4.97 (1.34) | 4.96 (1.29) |
| Schoolgewicht                     | 1.13  | 1.13 (1.15) | 1.13 (1.15) | 1.12 (1.15) | 1.11 (1.14) | 1.11 (1.14) |
| Nodaalgebied (*10 <sup>-3</sup> ) | 56.1  | 56.1 (55.8) | 53.4 (55.8) | 50.9 (54.9) | 50.8 (52.4) | 50.3 (51.5) |

allochtone leerlingen in het algemeen een lagere score op tekstbegrip behaald wordt, ligt het voor de hand om bij een methodevergelijking rekening te houden met het schoolgewicht. Het schoolgewicht lijkt te variëren tussen de drie methodecategorieën; het schoolgewicht is groter voor scholen uit de eerste twee methodecategorieën dan voor de scholen die een methode uit de derde categorie gebruiken. Op scholen waar een leesmethode uit de derde categorie gebruikt wordt, is de proportie sociaal zwakke leerlingen geringer dan op scholen waar een leesmethode uit de beide eerste categorieën in gebruik is.

In recente discussies (en beleid) omtrent de *effectiviteit van het basisonderwijs wordt gewezen op de invloed van klassengrootte*. Vandaar dat in Tabel 2, naast een overzicht van het gemiddelde schoolgewicht, ook de gemiddelde klassengrootte is weergegeven. Het blijkt dat de klassengrootte slechts in geringe mate varieert van jaar tot jaar, maar wel zijn de klassen op scholen waar een methode uit de derde categorie gebruikt wordt kleiner dan de klassen op scholen waarin een methode uit één van beide andere categorieën gehanteerd wordt.

Het ligt in de rede dat de ervaring met een methode van invloed is op de gerealiseerde prestaties. Bij een nieuwe methode zijn de docenten wellicht nog wat onwennig ten aanzien van de leerstof en de in de handleiding voorgestelde werkwijze en werkvormen. Vandaar dat de aan deze studie deelnemende scholen gevraagd is naar het aantal jaren dat de huidige methode in gebruik is (zie: Tabel 2). Uit deze tabel wordt duidelijk dat (zoals te verwachten was) de gemiddelde ervaring met de 'nieuwe' methoden duidelijk geringer is dan die met methoden van de tweede categorie, en met methoden van de derde categorie is de meeste ervaring opgedaan<sup>2</sup>.

In Tabel 2 zijn ook gegevens omtrent het nodaalgebied van de scholen opgenomen. Dit is een maat voor de urbanisatiegraad van – in dit geval – het voedingsgebied van een school. Heel duidelijk blijkt dat scholen die een methode uit de tweede categorie gebruiken zich in (gemiddeld) dichter bevolkte gebieden bevinden, terwijl met name de scholen die een methode uit de derde categorie gebruiken zich in de dunner bevolkte gebieden bevinden.

Enigszins gechargeerd, want de verschillen tussen scholen in één categorie zijn groot, kan gesteld worden dat de methoden uit de tweede categorie (verhoudingsgewijs) veel gebruikt worden in de grote steden, terwijl methoden van de derde categorie veel gebruikt worden op scholen in kleinere gemeenschappen.

### 3 Analyses

Bij de analyse van de gegevens moet rekening gehouden worden met de hiërarchische structuur van de gegevens. Immers, de observaties (scores) zijn genest binnen leerlingen, leerlingen zijn genest binnen klassen, en klassen zijn genest binnen scholen. Een model voor de gemiddelde score, en de verschillende variantiecomponenten (de variantie tussen scholen, de variantie tussen jaren (klassen) binnen scholen, de variantie tussen leerlingen binnen jaren en scholen, en de variantie tussen taken binnen leerlingen, jaren en scholen) kan geschreven worden als:

$$Y_{hijk} = \beta_0 + [e_{hijk} + u_{0ijk} + v_{00jk} + w_{000k}]. \quad (1)$$

Hierbij is  $Y_{hijk}$  de score op taak  $h$  ( $h = 1, 2, \dots, 5$ ) van leerling  $i$ , ( $i = 1, 2, \dots, I_j$ ), in jaar  $j$  ( $j = 1, 2, \dots, J_k$ ), van school  $k$  ( $k = 1, 2, \dots, K$ ).

Het bovenstaande model bestaat verder uit twee delen: een fixed deel en een random deel (tussen vierkante haken). In het fixed deel van het model is slechts één parameter gespecificeerd ( $\beta_0$ ); een schatting van de gemiddelde score. In het random deel van het model worden vier residuele scores onderscheiden. Deze residuen worden uitgedrukt als afwijkingsscores. Dat wil zeggen: op schoolniveau representeert  $w_{000k}$  de afwijking van school  $k$  van het geschatte populatiegemiddelde ( $\beta_0$ ). Het geschatte gemiddelde van school  $k$  is derhalve  $\beta_0 + w_{000k}$ . De tweede residuele score,  $v_{00jk}$ , representeert de deviatie van jaar  $j$  van het gemiddelde van school  $k$  (over alle jaren). De gemiddelde score van school  $k$  in jaar  $j$  is dus:  $\beta_0 + w_{000k} + v_{00jk}$ . Het derde en vierde residu,  $u_{0ijk}$  en  $e_{hijk}$ , representeren de afwijking van leerling  $i$  in jaar  $j$  van school  $k$ , en de afwijking van de score op taak  $h$  van de gemiddelde score van een leerling. We nemen, zoals gebruikelijk, aan dat de residuele scores normaal verdeeld zijn



met een verwachte waarde van nul, en een variantie van respectievelijk  $S^2_e$ ,  $S^2_u$ ,  $S^2_v$ , en  $S^2_w$  (de variantiecomponenten). Daarnaast wordt aangenomen dat de residuele scores van verschillende niveaus (i.e. taak binnen leerling, leerling binnen jaar, jaar binnen school, en school) niet gecorreleerd zijn.

Het basismodel in Vergelijking (1) kan uitgebreid worden met de variabelen waarin men geïnteresseerd is. We beperken ons tot een vijftal variabelen: het type leesmethode, het aantal jaren ervaring met deze methode, het schoolgewicht, de klassengroote en het nodaalgebied.

In het tweede model worden alleen het type leesmethode ingevoerd met behulp van drie dummy-variabelen,  $M1_{hijk}$ ,  $M2_{hijk}$  en  $M3_{hijk}$ . Afhankelijk van de gebruikte methode staat één van dummy-variabelen 'aan'. Dat wil zeggen dat deze dummy de waarde 1 heeft, en de twee andere dummy-variabelen de waarde 0. Voor elk type methode wordt het gemiddelde geschat (respectievelijk  $\beta_1$ ,  $\beta_2$  en  $\beta_3$ ) alsmede vier variantiecomponenten (de variantie tussen taken, de variantie tussen leerlingen, de variantie tussen jaren en de variantie tussen scholen)<sup>3</sup>.

$$Y_{hijk} = M1_{hijk} (\beta_1 + e_{1hijk} + u_{10ijk} + v_{100ijk} + w_{1000k}) + M2_{hijk} (\beta_2 + e_{2hijk} + u_{20ijk} + v_{200ijk} + w_{2000k}) + M3_{hijk} (\beta_3 + e_{3hijk} + u_{30ijk} + v_{300ijk} + w_{3000k}) \quad (2)$$

In het derde model worden drie variabelen toegevoegd: het aantal jaren ervaring met een methode, het schoolgewicht, en de klassengroote. Omdat het effect van deze variabelen afhankelijk kan zijn van de methode, wordt per variabele per methodecategorie het effect geschat; in feite worden dus de interacties tussen methode en de drie verklarende variabelen geschat. Dit model kan geschreven worden als:

$$Y_{hijk} = M1_{hijk} (\beta_1 + \beta_4 * ERV_{00jk} + \beta_7 * SCHGW_{00jk} + \beta_{10} * KLGR_{00jk}) + M2_{hijk} (\beta_2 + \beta_5 * ERV_{00jk} + \beta_8 * SCHGW_{00jk} + \beta_{11} * KLGR_{00jk}) + M3_{hijk} (\beta_3 + \beta_6 * ERV_{00jk} + \beta_9 * SCHGW_{00jk} + \beta_{12} * KLGR_{00jk}) + [M1_{hijk} (e_{1hijk} + u_{10ijk} + v_{100ijk} + w_{1000k}) + M2_{hijk} (e_{2hijk} + u_{20ijk} + v_{200ijk} + w_{2000k}) + M3_{hijk} (e_{3hijk} + u_{30ijk} + v_{300ijk} + w_{3000k})] \quad (3)$$

In dit model zijn  $M1_{hijk}$ ,  $M2_{hijk}$  en  $M3_{hijk}$  dezelfde dummy-variabelen als in Model II. Daarnaast wordt de invloed van de ervaring

( $ERV_{00jk}$ :  $\beta_4$ ,  $\beta_5$  en  $\beta_6$ ) geschat en het effect van het schoolgewicht ( $SCHGW_{00jk}$ :  $\beta_7$ ,  $\beta_8$  en  $\beta_9$ ), alsmede het effect van klassengroote ( $KLGR_{00jk}$ :  $\beta_{10}$ ,  $\beta_{11}$  en  $\beta_{12}$ ). Alle verklarende variabelen (ervaring, schoolgewicht en klassengroote) zijn gecentreerd rond het algemeen gemiddelde. Daarom geven de regressiegewichten  $\beta_1$ ,  $\beta_2$  en  $\beta_3$  de gemiddelde score voor een school met een gemiddeld aantal ervaringsjaren, een gemiddeld schoolgewicht, en een gemiddeld aantal leerlingen in een klas.

De laatste verklarende variabele die we introduceren is het nodaalgebied of de urbanisatiegraad. We kunnen urbanisatiegraad niet zomaar als een verklarende variabele (zoals bij voorbeeld ervaring) beschouwen. In feite zijn leerlingen genest binnen combinaties van school en urbanisatiegraad (vergelijk: Goldstein, 1995, hoofdstuk 8). Daarom moet in het model op schoolniveau een extra variantiecomponent geschat worden, die aangeeft hoe de verschillen tussen scholen variëren met urbanisatiegraad.

Tabel 3

Parameterschattingen voor de fixed en random effecten voor Model I (standaardfouten tussen haakjes)

| Fixed parameters                      | gemiddelde (se)    |        |
|---------------------------------------|--------------------|--------|
| $\beta_0$                             | 7.08               | (.04)  |
| Random parameters                     | variantiecomponent | $\rho$ |
| $S^2_w$ (variantie tussen scholen)    | .45 (.04)          | .094   |
| $S^2_v$ (variantie tussen jaren)      | .21 (.01)          | .048   |
| $S^2_u$ (variantie tussen leerlingen) | 1.84 (.01)         | .443   |
| $S^2_e$ (variantie tussen taken)      | 2.31 (.01)         |        |

## 4 Resultaten

De parameterschattingen behorende bij Model I staan in Tabel 3. Tussen haakjes zijn de standaardfouten weergegeven<sup>4</sup>.

Uit de tabel kan afgelezen worden dat het populatiegemiddelde 7.08 is. Op alle vier de niveaus zijn de variantiecomponenten significant; de parameterwaarden zijn veel groter dan 1.96 \* standaardfout. Om deze variantieschattingen te interpreteren is de intraklassecorrelatie ( $\rho$ ) een

handige maat. De intraklassecorrelatie wordt gevormd door de ratio van de variantie op één niveau en de som van de varianties op dat niveau en die op lagere niveaus (bij voorbeeld  $\rho_{\text{jaar}} = S^2_v / [S^2_v + S^2_u + S^2_e]$ ). We kunnen dus voor elk niveau, behalve het laagste, een intraklassecorrelatie berakenen. De intraklassecorrelatie is .044 voor de verschillen tussen jaren ( $\rho_{\text{jaar}}$ ), en .094 voor de verschillen tussen scholen ( $\rho_{\text{school}}$ ). Hieruit kan geconcludeerd worden dat, hoewel de verschillen tussen jaren in begrijpend lezen significant en betekenisvol zijn, deze duidelijk kleiner zijn dan de verschillen tussen scholen. Het komt er dus feitelijk op neer dat de verschillen tussen scholen relatief stabiel lijken te zijn over jaren.

Dit neemt niet weg dat, zoals altijd, de verschillen tussen leerlingen veel groter zijn dan de verschillen tussen scholen (en/of jaren). Opvallend is wel dat de verschillen tussen taken weer veel groter zijn dan de verschillen tussen leerlingen; de specifieke operationalisering d.m.v. een (sub)toets legt veel meer gewicht in de schaal dan veelal aangenomen wordt.

De parameterschattingen voor Model II, waarin een onderscheid tussen de drie typen methoden geïntroduceerd wordt, staan in Tabel 4. In dit model is het effect van de methoden opgenomen.

De (geschatte) gemiddelden voor de drie methodecategorieën liggen heel dicht bij elkaar. De verschillen in gemiddelden zijn dan ook niet significant (de toetsingsgroottheid  $\chi^2$  voor bijvoorbeeld het verschil tussen het eerste en derde type methode is 2.01;  $df = 1$ ;  $p = .16$ ).

Voor alle drie de methodecategorieën zijn er

duidelijke en significante verschillen tussen scholen. Daarbij blijkt de tussen-scholen-variantie te verschillen voor de drie methodecategorieën; de tussen-scholen-variantie voor de derde methodecategorie is significant kleiner dan die voor de tweede categorie ( $\chi^2 = 8.83$ ;  $df = 1$ ;  $p = .003$ ), maar noch tussen de eerste en de tweede methodecategorie, noch tussen de eerste en de derde methodecategorie zijn de verschillen significant ( $\chi^2 < 2.54$ ;  $df = 1$ ;  $p > .11$ ). Ook relatief zijn er duidelijke verschillen in varianties; op schoolniveau varieert de intraklassecorrelatie ( $\rho_{\text{school}}$ ) van .060 tot .120 voor respectievelijk de derde en tweede methodecategorie, waarmee aangegeven wordt dat ook verhoudingsgewijs de verschillen tussen scholen die methoden uit de tweede en derde categorie gebruiken groot zijn.

Voor alle drie de methodecategorieën zijn er significante verschillen tussen de jaren: in het ene jaar zijn de gemiddelde schoolprestaties hoger of lager dan in het volgende jaar. En de verschillen in prestaties tussen jaren zijn niet onaanzienlijk;  $\rho_{\text{jaar}}$  varieert van .044 tot .054 voor respectievelijk de eerste en tweede methodecategorie. Echter, de verschillen tussen methodecategorieën in dit opzicht zijn niet significant; de fluctuatie in schoolscores over jaren verschilt niet van methodecategorie tot methodecategorie.

De verschillen tussen leerlingen ( $S^2_u$ ) zijn voor de derde leesmethodecategorie duidelijk groter dan voor de beide andere methodecategorieën ( $\chi^2 > 16.69$ ;  $df = 1$ ;  $p < .001$ ). Gemiddeld over taken verschillen de leesprestaties van leerlingen die onderwijs krijgen met methoden uit de derde categorie meer van elkaar dan bij de andere twee methodecategorieën het geval is. Hetzelfde geldt voor de ver-

Tabel 4

Parameterschattingen voor de fixed en random effecten voor Model II ( $S^2$ : variantie; standaardfouten tussen haakjes;  $\rho$ : intraklassecorrelatie)

|                                      | Type 1     |        | Methode Type 2 |        | Type 3     |        |
|--------------------------------------|------------|--------|----------------|--------|------------|--------|
| Gemiddelde ( $\beta_1$ - $\beta_3$ ) | 7.14 (.06) |        | 7.05 (.07)     |        | 7.02 (.06) |        |
| Variantiecomponent                   | $S^2$ (se) | $\rho$ | $S^2$ (se)     | $\rho$ | $S^2$ (se) | $\rho$ |
| Tussen scholen                       | .42 (.06)  | .09    | .58 (.08)      | .12    | .29 (.05)  | .06    |
| Tussen jaren                         | .19 (.01)  | .04    | .23 (.02)      | .05    | .21 (.02)  | .05    |
| Tussen leerlingen                    | 1.82 (.03) | .44    | 1.74 (.02)     | .43    | 1.99 (.03) | .46    |
| Tussen taken                         | 2.29 (.02) |        | 2.28 (.02)     |        | 2.37 (.02) |        |

Tabel 5  
Parameterschattingen voor de fixed en random effecten voor Model III en IV (standaardfouten tussen haakjes;  $\rho$ : intra-klassecorrelatie)

|  | Type I                     |                            | Methode Type II            |                            |                            |                            | Type III                   |                            |
|--|----------------------------|----------------------------|----------------------------|----------------------------|----------------------------|----------------------------|----------------------------|----------------------------|
|  | Model III                  | Model IV                   | Model III                  | Model IV                   | Model III                  | Model IV                   | Model III                  | Model IV                   |
| Fixed parameters                             |                            |                            |                            |                            |                            |                            |                            |                            |
| Intercept ( $\beta_1 - \beta_3$ )            | 7.10 (.06)                 | 7.09 (.05)                 | 7.12 (.05)                 | 7.15 (.04)                 | 7.07 (.14)                 | 7.04 (.11)                 |                            |                            |
| Ervaring ( $\beta_4 - \beta_6$ )             | -0.03 (.03)                | -0.04 (.03)                | 0.03 (.04)                 | 0.00 (.03)                 | 0.09 (.15)                 | .07 (.12)                  |                            |                            |
| Schoolgewicht ( $\beta_7 - \beta_9$ )        | -2.58 (.32)                | -2.65 (.25)                | -2.97 (.22)                | -3.15 (.18)                | -2.33 (.38)                | -2.31 (.27)                |                            |                            |
| Klassengrootte ( $\beta_{10} - \beta_{12}$ ) | .002 (.003)                | .002 (.003)                | .005 (.003)                | .005 (.003)                | .003 (.003)                | .003 (.003)                |                            |                            |
| Random parameters                            |                            |                            |                            |                            |                            |                            |                            |                            |
| Variantiecomponent                           | S <sup>2</sup> (se) $\rho$ | S <sup>2</sup> (se) $\rho$ | S <sup>2</sup> (se) $\rho$ | S <sup>2</sup> (se) $\rho$ | S <sup>2</sup> (se) $\rho$ | S <sup>2</sup> (se) $\rho$ | S <sup>2</sup> (se) $\rho$ | S <sup>2</sup> (se) $\rho$ |
| Scholen                                      | .25 (.04) .05              | .21 (.05) .04              | .19 (.03) .04              | .18 (.05) .04              | .19 (.03) .04              | .04 (.04) .01              |                            |                            |
| Nodaalgebied                                 |                            | .03 (.04) .01              |                            | .00 (.04) .00              |                            | .14 (.04) .03              |                            |                            |
| Jaren  | .19 (.03) .04              | .19 (.02) .04              | .23 (.02) .05              | .24 (.02) .05              | .21 (.02) .05              | .21 (.02) .05              |                            |                            |
| Leerlingen                                   | 1.82 (.03) .44             | 1.82 (.03) .44             | 1.74 (.02) .43             | 1.74 (.02) .43             | 1.99 (.03) .46             | 1.99 (.03) .46             |                            |                            |
| Taken  | 2.29 (.02)                 | 2.29 (.02)                 | 2.28 (.02)                 | 2.28 (.02)                 | 2.37 (.02)                 | 2.37 (.02)                 |                            |                            |

schillen tussen taken ( $S^2_e$ ); ook hier zijn de verschillen duidelijk groter voor de derde methodecategorïe ( $\chi^2 > 7.15$ ;  $df = 1$ ;  $p < .007$ ) dan voor de beide andere categorïeën.

In het derde model zijn naast het gemiddelde per categorie van leesmethode ook de effecten van de ervaring opgedaan met de methoden, het schoolgewicht, en de klassengrootte geschat. Uit de fixed effecten blijkt dat de methodegemiddelden variëren van 7.07 (voor de derde methodecategorïe) tot 7.12 (voor de tweede methodecategorïe). Evenmin als bij Model II zijn deze verschillen significant ( $\chi^2$  voor het verschil tussen de tweede en derde methodecategorïe is slechts .10;  $df=1$ ;  $p=.75$ ).

Het aantal jaren ervaring met een methode lijkt geen effect te hebben op de leesprestaties van de leerlingen, de ratio van de parameterschattingen en hun standaardfouten is in alle gevallen kleiner dan ( $\pm$ ) 1.96 en de overschrijdingskans dus groter dan .05.

De derde variabele, het schoolgewicht heeft echter een duidelijk effect. Voor elke eenheid verandering in deze variabele verandert de gemiddelde score met -2.33 (voor methodecategorïe 3) tot -2.97 (voor methodecategorïe 2). Echter, aan deze verschillen tussen methodecategorïeën kan geen betekenis toegekend worden, aangezien zij niet significant zijn (Bij voorbeeld de toetsingsgrootte  $\chi^2$  voor het verschil in regressiegewichten tussen methode 2 en 3 is slechts .30;  $df=1$ ;  $p=.58$ ).

Het effect van klassengrootte blijkt voor

geen van de drie methodecategorïeën significant (en ook niet te verschillen van methodecategorïe tot methodecategorïe).

Uit de variantiecomponenten blijkt dat de verschillen tussen scholen conditioneel op beide verklarende variabelen nog steeds significant zijn. Dit neemt niet weg dat de afname in tussen-scholen-variantie door opnemen van deze variabelen substantieel is;  $\rho_{\text{school}}$  varieert van .040 voor de derde methodecategorïe tot .055 voor de eerste methodecategorïe (een verschil dat overigens niet significant is:  $\chi^2 < 1.49$ ;  $df=1$ ;  $p > .22$ ). De verschillen in tussen-scholen-variantie tussen methodecategorïeën, die in Model II nog duidelijk waren, kunnen derhalve toegeschreven worden aan verschillen in schoolgewicht. De varianties tussen jaren, leerlingen en taken zijn in Model III nauwelijks veranderd ten opzichte van de schattingen in Model II.

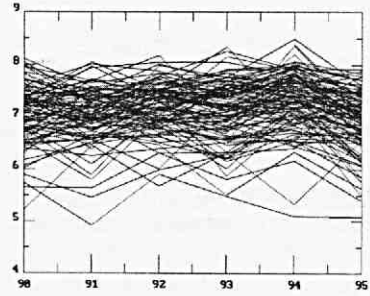
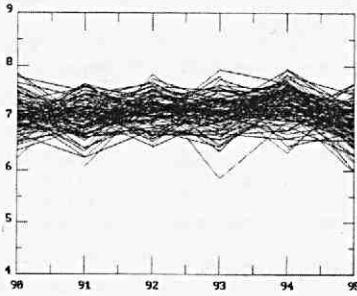
In het vierde en laatste model (zie Tabel 5) wordt naast de al genoemde variabelen ook het effect van urbanisatiegraad geschat. De schattingen van de fixed parameters zijn door het opnemen van de variantie tussen scholen ten gevolge van urbanisatiegraad nauwelijks beïnvloed. We beperken de bespreking daarom tot de variantiecomponenten. Opnemen van urbanisatiegraad leidt tot vermindering van de variantie tussen scholen voor elk van de methodecategorïeën. De variantie tussen scholen is voor de eerste twee methodecategorïeën nog steeds significant, maar is duidelijk kleiner dan



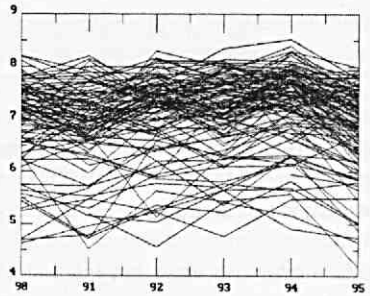
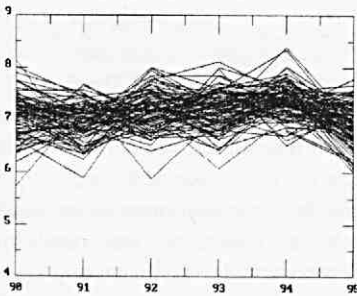
Conditioneel

Niet conditioneel

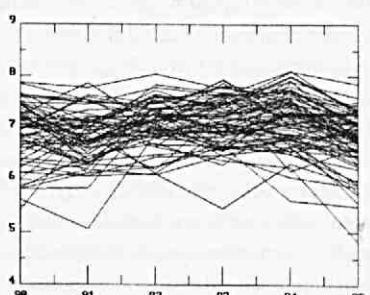
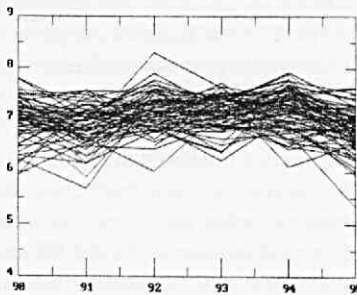
Methode Type 1



Methode Type 2



Methode Type 3



*Figuur 1.* Voorspelde scores (y-as) per jaar (x-as) per methodecategorie, conditioneel op ervaring, schoolgewicht en klassengrootte (links) en niet conditioneel (rechts). N.b. met de niet-significante effecten van 'ervaring' en 'klassengrootte' is geen rekening gehouden bij de constructie van de conditionele figuren).

die in voorgaande modellen. Voor de derde categorie van leesmethoden is de daling van de variantie tussen scholen zeer groot: deze daalt van .19 naar .04; de variantie zelf is niet meer significant. Voor deze methodecategorie is een groot deel van de verschillen tussen scholen derhalve niet toe te schrijven aan de school, of aan verschillen in gegeven onderwijs, maar aan de urbanisatiegraad van het voedingsgebied van de school. De instroom lijkt bepalend voor

de kwaliteit van een school. Dientengevolge neemt het relatieve belang van 'jaar' toe: als bekend is wat de grootte van de plaats is, dan kunnen er wel verschillen tussen jaren aangetoond worden, maar géén verschillen tussen scholen.

Om het belang van de effecten van 'jaar' te duiden is Figuur 1 geconstrueerd. In deze figuur representeert elke lijn één school. Voor de overzichtelijkheid zijn per methodecatego-

rie twee figuren geconstrueerd; één figuur waarin wel rekening gehouden wordt met het schoolgewicht en één figuur waarin hier géén rekening mee gehouden wordt.

Het eerste dat opvalt in Figuur 1 is dat de lijnen niet alle even lang zijn, een gevolg van het gegeven dat niet elke school alle jaren deelnam aan de Cito-eindtoets. Vergelijken we de figuren aan de linkerzijde van Figuur 1 met die aan de rechterzijde dan valt de invloed van het schoolgewicht op. De variantie tussen scholen kan voor een groot deel toegeschreven worden aan kenmerken van de schoolbevolking. Uit de figuren aan de rechterkant kan opgemaakt worden dat de verschillen tussen scholen relatief stabiel zijn over jaren; in sommige scholen hebben de leerlingen elk jaar relatief hoge scores behaald, terwijl op andere scholen de leerlingen elk jaar relatief lage scores behaald hebben. Echter, zodra rekening gehouden wordt met de verschillen in schoolgewicht (als indicatie voor het type leerlingen) zijn de verschillen tussen scholen over jaren lang niet zo duidelijk. Scholen kunnen het in het ene jaar goed doen en slecht in het andere jaar. Een groot deel van de verschillen tussen scholen kan derhalve toegeschreven worden aan instroomkarakteristieken van de leerlingen, in plaats van aan stabiele schoolkenmerken. Dit lijkt met name op te gaan voor scholen die een leesmethode uit de derde categorie hanteren; voor deze scholen kan men zich zelfs afvragen of enig verschil gerelateerd is aan de school, aangezien geen van de scholen zich gedurende alle onderzochte jaren aan de bovenkant van het spectrum of aan de onderkant van het spectrum bevindt.

## 5 Conclusies en discussie

De verschillen tussen scholen in leerlingprestaties ten gevolge van de gebruikte methode voor begrijpend lezen zijn onderzocht in zes opeenvolgende jaren. Hiervoor zijn de scores van de leerlingen op de Cito-eindtoets van deze scholen nader geanalyseerd. Om rekening te houden met de afhankelijkheden die zich in dergelijke datasets bevinden (leerlingen zijn genest in klassen, klassen in scholen etc.) zijn vier multiniveau-modellen gespecificeerd. Hierbij is rekening gehouden met de ervaring met de

methode, de klassengrootte, het schoolgewicht en de urbanisatiegraad van het voedingsgebied van de school.

De gemiddelde leesvaardigheid van leerlingen aan het einde van de basisschool bleek niet afhankelijk van de gehanteerde methodecategorie; alle drie de methodecategorieën geven in dit opzicht een vergelijkbaar resultaat. Het aantal jaren dat een methode in gebruik is, blijkt evenmin gerelateerd aan de behaalde leesprestaties; op scholen die pas kort met een methode werken zijn leesprestaties even goed als op scholen die al jaren met een methode vertrouwd zijn. Ook de klassengrootte blijkt niet gerelateerd aan de leesprestaties in de achtste groep. De vierde verklarende variabele, het schoolgewicht als indicator voor het type leerlingen op een school, is sterk gerelateerd aan de leesprestaties. Uit het gegeven dat deze variabele slechts een marginale invloed heeft op de verschillen tussen jaren kan opgemaakt worden dat het schoolgewicht een relatief stabiel kenmerk is van een school; de samenstelling van de schoolbevolking is relatief constant over jaren. Het schoolgewicht heeft daarentegen een grote invloed op de tussen-scholen-variantie. Met name de verschillen tussen scholen die een leesmethode uit de eerste en tweede methodecategorie gebruiken nemen enorm af wanneer rekening gehouden wordt met het schoolgewicht. Hieruit kan geconcludeerd worden dat scholen met een (relatief) hoog schoolgewicht oververtegenwoordigd zijn in de groepen die methoden uit de eerste of tweede categorie gebruiken en ondervertegenwoordigd zijn in de groep scholen die een leesmethode uit de derde categorie hanteert.

We merken hierbij wel op dat de beschikbare gegevens het niet toestonden om individuele leerlinggewichten in de analyse te betrekken. Tevens is geconstateerd dat leerlingen met een gewicht groter dan 1.00 ondervertegenwoordigd zijn op scholen die begrijpend leesonderwijs geven aan de hand van een methode uit de derde categorie. Op grond van deze gegevens kan opgemaakt worden dat de besproken verschillen tussen leerlingen tussen de verschillende categorieën van leesmethoden een onderschatting vormen van de werkelijke verschillen; in werkelijkheid zullen de verschillen, conditioneel op leerlinggewicht, voor de beide eerste categorieën leesmethoden kleiner zijn

dan gerapporteerd. Ter illustratie: de oververtegenwoordiging van leerlingen met een gewicht van 1.00 op scholen met een methode uit de derde categorie rechtvaardigt de verwachting dat op deze scholen de gemiddelde scores hoger zouden moeten kunnen liggen; de effecten van de achtergrondkenmerken van leerlingen op de leesprestaties zijn immers groot (vergelijk: Sijstra, 1997).

Naast de gemiddelde score per type leesmethode zijn ook verschillende variantiecomponenten geschat. Gebleken is dat de geobserveerde verschillen tussen leerlingen en tussen taken bij leesmethoden uit de derde categorie relatief groot zijn. Dat wil zeggen: op scholen waar methoden met een productgerichte benadering, zonder aandacht voor leesstrategieën gebruikt worden, zijn de prestaties van leerlingen minder homogeen dan op scholen waar een methode uit de eerste of tweede categorie gebruikt wordt. Daarnaast zijn bij scholen met een methode uit de derde categorie de verschillen tussen taken groter dan bij scholen die een methode uit de eerste of tweede categorie hanteren; bij de derde methodecategorie is de leesvaardigheid van de leerlingen minder homogeen dan bij gebruik van de eerste twee categorieën.

In het algemeen wijst de grootte van de verschillen tussen taken op het belang dat gehecht moet worden aan de specifieke operationalisatie van de leesvaardigheid. Dit betekent dat de generalisatiemogelijkheden van studies waarin leesvaardigheid op slechts één manier geoperationaliseerd is uiterst beperkt is; men kan niet naar het theoretische construct 'leesvaardigheid' generaliseren (zie ook Kuhlemeier & Van den Bergh (1998) voor een vergelijkbare conclusie). Hiermee wordt het belang van verschillende operationalisaties voor één en hetzelfde construct (leesvaardigheid) onderschreven (vergelijk: Clark, 1973).

De geobserveerde verschillen tussen scholen zijn weliswaar groot, maar zodra rekening gehouden wordt met het schoolgewicht, en in mindere mate met de urbanisatiegraad, blijken deze verschillen duidelijk kleiner dan een eerste indruk suggereert. Wanneer geen rekening gehouden wordt met verschillen in instroom (schoolgewicht) zijn de verschillen tussen

scholen ongeveer twee keer zo groot als de verschillen tussen jaren, hetgeen relatief stabiele schooleffecten suggereert. Houden we wel rekening met instroomverschillen, dan zijn de verschillen tussen scholen en de verschillen tussen jaren in dezelfde orde van grootte. Houden we daarnaast ook rekening met verschillen ten gevolge van het voedingsgebied van scholen, dan blijken de verschillen tussen jaren groter dan de verschillen tussen scholen. Hoewel dit onafhankelijk van de categorie van leesmethode geldt, is voor scholen die een methode uit de derde categorie gebruiken de afname in tussen-scholen-variantie enorm. We kunnen, als we rekening houden met schoolgewicht en urbanisatiegraad geen verschillen tussen scholen meer aantonen. Anders gezegd: alle verschillen tussen scholen kunnen toegeschreven worden aan de instroomkenmerken en aan kenmerken van het voedingsgebied; stabiliteit van schoolprestaties kan derhalve opgevat worden als stabiliteit van de leerlinginstroom.

Gezien de kleine verschillen tussen scholen, ongeacht de (categorie van gehanteerde) leesmethode, kunnen we ons afvragen in hoeverre verschillen tussen scholen verantwoordelijk zijn voor verschillen in leesprestaties. Gegeven de resultaten van deze studie lijken verschillen tussen scholen in leesprestaties veel meer afhankelijk van kenmerken van de schoolbevolking dan van de gebruikte methode. En gezien de grote verschillen tussen jaren lijkt er nauwelijks ruimte voor duidelijk traceerbare effecten van effectief, strategisch leesonderwijs. Of, algemener doen deze resultaten twifelen aan de legitimiteit van concepten als effectieve instructie (voor dit leergebied). Meer nadruk op zaken als het compenserend effect van scholen (Goldstein, 1987, p. 59 e.v.) lijken voor de hand te liggen.

Er zijn duidelijke verschillen tussen de drie methodecategorieën op zowel de variantie tussen leerlingen, als de variantie tussen taken. In beide gevallen springen de methoden van de derde categorie eruit; zowel de tussen-leerlingen- als de tussen-taken-variantie is bij de derde methode groter. Omdat de gemiddelden niet verschillen, kan hieruit afgeleid worden dat bij methoden van de derde categorie er (relatief) zowel meer achterblijvers zijn, als meer leerlingen met een hoge begrijpend leescore. Uit de grotere variantie tussen taken

blijkt dat bij de derde methodecategorie de leesvaardigheid minder homogeen, minder breed toepasbaar, ontwikkeld is.

Deze conclusies komen niet overeen met de uitkomsten van de tweede taalpeiling einde basisonderwijs (Sijstra, 1997). In die PPOON-studie wordt geconcludeerd dat de leesprestaties van leerlingen die leesonderwijs hebben gehad met een methode (in onze termen) uit de derde categorie significant minder zijn dan bij leerlingen die met een methode uit de eerste twee categorieën onderwijs hebben gehad.

De verschillen in conclusies kunnen (naar alle waarschijnlijkheid) niet worden toegeschreven aan de indeling van de methoden in categorieën; zowel de uitgangspunten voor de categorie-indeling, als de verdeling van de methoden over de categorieën zijn in hoge mate vergelijkbaar. Een verschil is wel dat de PPOON-studie is gebaseerd op één afnamejaar (1993) en onze studie op een periode van zes jaren. Daarnaast wordt in de PPOON-studie geen onderscheid gemaakt tussen school- en leerling-niveau, wat ongetwijfeld een te optimistisch beeld van de standaardfouten geeft, en dus leidt tot een te snelle verwerping van de hypothese dat er geen verschillen tussen de methoden zouden zijn; in deze studie is wél rekening gehouden met de hiërarchische structuur van de gegevens, zonder dat daarbij restricties aan verschillen in varianties opgelegd zijn. Een laatste kenmerkend verschil is dat in de PPOON-studie de leerlingprestaties op afzonderlijke vaardigheden binnen begrijpend lezen zijn geanalyseerd; in deze studie zijn de prestaties op de vijf toetsonderdelen voor het tekstbegrip van de Cito-toets geanalyseerd. En, dat deze geen (onafhankelijke) vaardigheden impliceren behoeft nauwelijks betoog.

Ondanks de methodische verschillen zijn de verschillen in de conclusies opmerkelijk. Het lijkt daarom gewenst om in een nader onderzoek naar de relatie tussen leesprestaties en (het leesonderwijs vanuit een bepaalde) leesmethode de beide onderzoeksopties te combineren. Het object zou dan kunnen zijn: de invloed van de gebruikte leesmethode op de (ontwikkeling van de) leesprestaties van leerlingen bij de diverse aspecten van begrijpend lezen, op taak-, leerling- en schoolniveau, aan het einde van het basisonderwijs, in relatie tot

leerling- en schoolkenmerken, over een periode van enkele jaren. Vanzelfsprekend wordt daarbij niet alleen aandacht besteed aan de methode, zoals in deze studie, maar aan de wijze waarop de docent de methode gebruikt. Immers, wat in het algemeen onder methode-effecten verstaan wordt, zijn in feite effecten van het gebruik van methoden, en dat verschillende docenten eenzelfde methode op een totaal andere manier gebruiken is zowel in het basis- als in het voortgezet onderwijs meermalen aangetoond.

Hoewel een causale interpretatie van de verschillen tussen de drie typen leesmethoden, zeker gezien de inhoud van deze methoden voor de hand ligt, willen wij erop wijzen dat op grond van deze gegevens daar geen directe uitspraak over gedaan kan worden. Wij weten bijvoorbeeld niets over selectiemechanismen, zoals zittenblijven, doorverwijzingen naar het speciaal onderwijs etc., die verschillend tussen scholen zijn. Bovendien zouden dergelijke selectiemechanismen kunnen verschillen tussen de drie categorieën scholen – wat op zich niet zo vreemd zou zijn, gezien de verschillen in schoolgewicht en voedingsgebied.

## Noten

1. Het gevolg van deze transformatie is dat géén uitspraak gedaan kan worden over toename of afname in het niveau van het begrijpend lezen in de onderzochte periode; wél blijven relatieve verschillen tussen scholen en leerlingen in stand.
2. Het is jammer dat er géén informatie beschikbaar is over de tijdspanne waarin de leerlingen onderwijs op basis van één van de methoden genoten hebben, of dat zij gedurende hun hele schoolse carrière op grond van dezelfde methode les gekregen hebben.
3. Gezien het geringe aantal scholen dat in de loop van het onderzoek van type methode wisselde, is het onmogelijk een complete covariantiematrix op schoolniveau te schatten; de correlatie tussen typen methoden is helaas niet schatbaar. Aan het model is derhalve de restrictie opgelegd dat de residuele scores op schoolniveau ongecorrigeerd zijn. Eenzelfde restrictie is opgelegd aan de covariantiematrix op het jaarniveau. Merk op dat de residuele scores op leerling- en taak-

niveau niet gecorreleerd kunnen zijn; er wordt niet gewisseld van methode gedurende een schooljaar, waardoor elke leerling slechts aan één (type) begrijpend leesmethode is blootgesteld. Hetzelfde geldt voor de residuele scores op taakniveau, aangezien die genest zijn binnen leerlingen.

4. De parameters zijn geschat met behulp van MLN.

## Literatuur

- Aarnoutse, C.A.J. (1991). Onderwijs en leren. In A.J.W.M. Thomasen, L.G.M. Noordman & P.A.T.M. Eling (Eds.), *Lezen en begrijpen. De psychologie van het leesproces*. Lisse: Swets & Zeitlinger.
- Baker, L. & Brown, A.L. (1984). Metacognitive skills in reading. In D. Pearson (Ed.), *Handbook of reading research*. New York: Longman.
- Bosker, R. (1991). De consistentie van schooleffecten in het basisonderwijs. *Tijdschrift voor Onderwijsresearch*, 16, 206-218.
- Brand-Gruwel, S. (1995). *Onderwijs in tekstbegrip. Een onderzoek naar het effect van strategisch lees- en luisteronderwijs bij zwakke lezers*. Academisch proefschrift. Ubbergen: Tandem Felix.
- Chall, J.S. (1983). *Learning to read: The great debate (updated edition)*. New York: McGraw-Hill Book Company.
- Clark, H.H. (1973). The language as a fixed effect fallacy: A critique of language statistics in psychological research. *Journal of Verbal Learning and Behaviour*, 12, 335-359.
- Goldstein, H. (1987). *Multilevel models in educational and social research*. London: Charles Griffin & Company Ltd.
- Goldstein, H. (1995). *Multilevel statistical models*. Kendall's library of statistics. Arnold Publishers: London.
- Hout-Wolters, B.H.A.M. van (1992). *Cognitieve strategieën als onderwijsdoel*. Groningen: Wolters-Noordhoff.
- Kuhlemeier, H. & Bergh, H. van den (1998). Relationships between language skills and task effects. *Perceptual and Motor Skills*, 86, 443-463.
- Inspectie van het onderwijs (1996). *Begrijpend en studerend lezen in het basisonderwijs*. Inspectierapport 1996-1.
- Palincsar, A.S. & Brown, A.L. (1984). Reciprocal teaching of comprehension forstoring and

comprehension monitoring activities. *Cognition and Instruction*, 1(2), 117-175.

- Paris, S.G., Wasik, B.A. & Turner, J.C. (1991). The development of strategic readers. In R.Barr, M.L. Kamil, P.B. Mosenthal & P.D. Pearson (Eds.), *Handbook of reading research vol 2*. New-York: Longman.

Sijstra, J. (red.) (1997). *Balans van het taalonderwijs aan het einde van de basisschool 2. Uitkomsten van de tweede taalpeiling einde basisonderwijs*. Arnhem: Cito (PPON-reeks nr. 10a).

Manuscript aanvaard: 30 maart 2000

## Auteurs

**Huub van den Bergh** is verbonden aan de onderzoeksgroep 'language learning' van het Utrechts Instituut voor Linguïstiek.

Adres: Universiteit Utrecht, Utrechts Instituut voor Linguïstiek (OTS), Trans 10, 3512JK Utrecht.

E-mail: Huub.vandenBergh@let.uu.nl

**Jan Rijkers** is coördinerend inspecteur basisonderwijs. Hij is nauw betrokken bij lees- en taalonderwijs in het primaire onderwijs.

**Michel Zwarts** is staf-inspecteur bij de sector voortgezet onderwijs van de Inspectie van het Onderwijs.

Adres: Inspectie van het Onderwijs, Park Voorn 4, Postbus 2725, 3500 GS Utrecht.

E-mail: M.Zwarts@owinsp.nl

## Abstract

### Effects of reading method on reading achievement

**H. van den Bergh, J. Rijkers and M. Zwarts.** *Pedagogische Studiën*, 2000, 77, 152-165.

In this study differences in reading achievements of eight grade students from six consecutive years (with on average 310 primary schools and 6767 students per year) are studied in relation to type of reading method used in a school. Three types of reading method were distinguished: new methods that cover the so-called attainment targets of



primary education (type 1), old methods in which neither reading strategies nor attainment targets are main aspects of the teaching of reading (type 3), and an intermediate type of method in which only limited attention is paid to target attainments and/or reading strategies (type 2).

It was shown that the mean reading achievement did not vary between types of reading method. However, differences between methods in variances were assessed. It was shown that both the between-student-variance and the between-reading-task-variance are larger for the third type of reading method than for the other two types of reading method.

The school-weight, as an indication of the social economic composition of a school proved a relative good predictor of the differences between schools. Conditionally on both the school-weight and urbanization of the region of a school, the differences between schools are relatively small compared to differences between years. So, it was concluded that school characteristics that are generally expected to be related to school effectiveness can only play a very minor role.