

Vergelijkend onderzoek naar methoden voor aanvankelijk leesonderwijs

Henk Blok , Martha E. Otter & Kees de Glopper

Henk Blok, Martha E. Otter en Kees de Glopper zijn onderzoeker bij het SCO-Kohnstamm Instituut van de Universiteit van Amsterdam

Correspondentieadres: SCO-Kohnstamm Instituut, Universiteit van Amsterdam, Wibautstraat 4, 1091 GM Amsterdam

Inleiding

Veel leerkrachten maken bij het onderwijs in aanvankelijk lezen gebruik van een methode. In het PRIMA-cohort blijkt dat ongeveer 90 procent van de leerkrachten in groep 3 het aanvankelijk leesonderwijs vormgeeft aan de hand van een methode (Overmaat & Ledoux, 1998). Een methode verlicht het dagelijkse werk van leerkrachten aanzienlijk. Een methode functioneert namelijk als een concrete invulling van het schoolwerkplan. Dankzij een methode hoeven leerkrachten weinig of geen tijd te besteden aan het ontwikkelen van leermaterialen. Een methode vergemakkelijkt bovendien de afstemming van werkzaamheden tussen leerkrachten onderling. Afspraken, onder meer over de verdeling van de leerstof en over de te volgen didactiek, kunnen dankzij een methode gemakkelijker tot stand komen.

Diverse educatieve uitgeverijen brengen een methode voor aanvankelijk lezen op de markt. De beschikbaarheid van vergelijkbare methoden stelt scholen voor een keuzeprobleem. Weliswaar kunnen scholen zich uitvoerig informeren op basis van voorlichtingsmateriaal van de uitgevers. Maar scholen willen ook graag weten hoe een methode in de praktijk blijkt te bevallen. Het gaat bijvoorbeeld om vragen naar de gebruiksmogelijkheden en naar de effectiviteit. Zulke vragen zijn alleen te beantwoorden op basis van empirisch onderzoek.

Niet alleen schoolteams hebben belang bij empirisch methodeonderzoek. Ook andere bij het onderwijs betrokken actoren, zoals ouders, schoolbesturen, schoolbegeleiders, lerarenopleiders en de Inspectie van het Onderwijs tonen interesse in verschillen tussen methoden. Beoordeling van methoden en publicatie van de resultaten daarvan vormt voorts onderdeel van het beleid van het Ministerie van Onderwijs, Cultuur en Wetenschappen (1995, p. 3).

Dit artikel beschrijft een vergelijkend onderzoek naar methoden voor aanvankelijk lezen. De vergelijking is gebaseerd op twee onderdelen: een onderzoek naar effecten op de lees- en spellingvaardigheid van leerlingen en een vragenlijstonderzoek naar het gebruik en de waardering door leerkrachten.

Probleemstelling

Lezen leren

Vertrekpunt voor deze paragraaf is het model van technisch lezen van Struiksma, Van der Leij en Vieijra (1997). Het lezen van een woord houdt in dat de lezer het woord herkent. Woordherkenning

kan langs drie wegen tot stand komen: via directe herkenning op grond van onder meer orthografische informatie, via indirecte herkenning door omzetting van het geschreven woord in zijn klankvorm, en door een mengvorm waarbij naast directe herkenning van woorddelen sprake is van gedeeltelijke verklanking.

Tijdens de fase van het leren lezen domineert het proces van indirecte herkenning. De essentiële deelvaardigheden daarbij zijn de grafeem-foneemkoppeling en de auditieve synthese. Bij onvoldoende beheersing van een van deze deelvaardigheden is het snel en accuraat decoderen van woorden niet mogelijk. Tijdens de aanleerfase vindt een ontwikkeling plaats van indirecte naar directe woordherkenning, met directe herkenning van woorddelen als tussenstation. In deze periode is een flexibele toepassing van indirecte en directe woordherkenningsstrategieën van belang. Volledige verklanking is met name geschikt voor korte, klankzuivere woorden. De combinatie van gedeeltelijke verklanking en directe herkenning van woorddelen is juist geschikt voor niet-klankzuivere woorden. De ontwikkeling van de technische leesvaardigheid staat niet op zichzelf. De ontwikkeling van het technisch lezen hangt nauw samen met de ontwikkeling van het spellen en het begrijpend lezen (Aarnoutse et al., 1986). Hoewel er tussen lezen en spellen verschillen zijn in het ontwikkelingstempo - bij het spellen blijft de indirecte weg via de foneem-grafeemkoppeling veel langer de belangrijkste dan bij het lezen - vinden de auditieve synthese en grafeem-foneemkoppeling bij het spellen hun logische tegenhangers in de auditieve analyse en de foneem-grafeemkoppeling. In veel methoden voor aanvankelijk lezen wordt van meet af aan aandacht aan spelling geschonken. Wat betreft de relatie met begrijpend lezen moet worden gesteld dat het technisch lezen vanzelfsprekend een voorwaarde is voor het kunnen lezen met begrip. Maar omgekeerd geldt dat aandacht voor het begrip van de tekst ten goede komt aan de ontwikkeling van de leesteknik. Daarom wordt in de hedendaagse methoden voor aanvankelijk lezen vanaf het begin aandacht besteed aan het begrip van gelezen tekst.

Vergelijkend onderzoek naar methoden voor aanvankelijk lezen kan dus niet volstaan met alleen technisch lezen als afhankelijke variabele. Ook vaardigheden in begrijpend lezen en spellen dienen bestreken te worden.

Effectieve instructie

Voor methodevergelijkend onderzoek is naast een vakdidactische basis ook een instructietheoretisch fundament nodig. De mate waarin een methode voor aanvankelijk lezen effectief is, hangt immers niet alleen af van de leesdidactiek die in de methode gevolgd wordt. Ook andere, niet-domeinspecifieke didactische kenmerken doen ertoe. Overzichten van kenmerken van effectieve instructie zijn bijvoorbeeld te vinden bij Creemers (1991) en Veenman (1992).

Hol, De Haan en Kok (1995) ontwikkelden - specifiek voor methoden voor aanvankelijk lezen - een lijst met negen effectiviteitskenmerken. Op deze lijst vinden we onder meer de volgende punten: weergave van didactische uitgangspunten, duidelijk omschreven doelstellingen, herkenbare structurering van de leerstof in eenheden en duidelijk omschreven onderwijsactiviteiten voor de leerkracht en leerling.

Aan deze bruikbare lijst van kenmerken vallen in elk geval nog verdere kenmerken inzake differentiatie toe te voegen. Het lijkt onder meer van belang na te gaan welke

differentiatiemogelijkheden methoden bieden, met name waar het gaat om uiteenlopende groeperingsvormen, zoals homogeen of juist heterogeen samengestelde groepen.

Overwegingen van effectiviteit gelden niet alleen de methode, maar ook de wijze waarop deze wordt gebruikt. Een leesmethode heeft immers op zichzelf geen effecten. Eerst door het gebruik in de klas wordt een meer of minder effectieve instructie gerealiseerd. Ook de context waarbinnen de instructie plaatsvindt kan van belang zijn. Toegespitst op het leesonderwijs geeft Vernooij (1993) een overzicht van kenmerken van effectief onderwijs, onderscheiden naar het niveau van de leerkracht, de schoolorganisatie en de schoolleiding. Bij het maken van een methodevergelijking moet derhalve aandacht geschonken worden aan de manier waarop leerkrachten hun methode gebruiken en een aantal schoolorganisatorische aspecten.

Vergelijkend onderzoek naar methoden voor aanvankelijk lezen

Vergelijkend onderzoek naar methoden voor aanvankelijk lezen is in ons land nog tamelijk schaars. De uitkomsten lijken desalniettemin een duidelijk antwoord op te leveren. De dominerende conclusie is dat er - wat betreft de leesvorderingen van de leerlingen - weinig of geen verschillen tussen methoden aantoonbaar zijn (Creemers, 1974; Van Dongen, 1984; Harskamp & Deinum, 1995; Hol et al., 1995; Van Oorsouw, 1971; Reitsma et al., 1981). De enige uitzondering vormt het onderzoek van Kooreman (1974), die laat zien dat een onder zijn leiding geconstrueerde methode op acht proefscholen aanzienlijk betere resultaten opleverde dan andere leesmethoden op andere scholen. Op een toets voor technisch lezen bedroeg de voorsprong 0,8 standaarddeviatie.

In de genoemde onderzoeken is weinig aandacht besteed aan andere vergelijkingscriteria, zoals de wijze waarop methoden gebruikt worden of de waardering die leerkrachten voor hun methode hebben. De verwaarlozing van deze aspecten valt te begrijpen vanuit de gedachte dat bereikte leerresultaten het primaire vergelijkingscriterium dienen te zijn. Maar naarmate de indruk sterker wordt dat methoden zich op dit primaire criterium niet van elkaar onderscheiden, groeit de behoefte aan een vergelijking ten aanzien van secundaire criteria.

Onderzoeksvragen

Het in de vorige paragraaf genoemde onderzoek geeft ons inziens steun aan de verwachting dat ook tussen recente methoden, na 1990 verschenen, geen verschillen, of in elk geval geen grote verschillen in leervorderingen aantoonbaar zullen zijn. Maar er zijn voor deze verwachting ook andere argumenten aan te voeren.

Zo blijkt uit methodeanalyses dat methoden voor aanvankelijk lezen in veel opzichten op elkaar lijken (Filipiak & Van Laarhoven, 1992; Groeneveld et al., 1992; Hol et al., 1995). De overeenkomsten betreffen belangrijke aspecten van de leesdidactische aanpak. Zo is er gelijkenis op het vlak van de leerstof: vrijwel alle methoden bieden leerstof voor aanvankelijk lezen (inclusief begrijpend lezen) en aanvankelijk spellen. Zo is er ook gelijkenis ten aanzien van de methodiek. In alle methoden wordt de elementaire leeshandeling centraal gesteld: kinderen leren lezen via het letter voor letter verklanken van klankzuivere woorden. In samenhang daarmee speelt het gebruik van eigen teksten - in de Verenigde Staten bekend onder de naam 'Whole language approach' - in alle methoden een heel bescheiden rol.

Vanzelfsprekend zijn er toch ook verschillen tussen de methoden. De meest in het oog lopende verschillen betreffen de onderwijskundige uitgangspunten, waaronder de manier waarop gedifferentieerd wordt. Omdat ook wij de door ons onderzochte methoden inhoudelijk hebben geanalyseerd, wijzen wij voor een verdere bespreking van de verschillen naar de volgende paragraaf. Daar worden de bij het onderzoek betrokken methoden gekarakteriseerd.

De betrekkelijk geringe verschillen tussen de leesdidactische en onderwijskundige uitgangspunten, plus het meerdimensionale karakter van deze verschillen, verhinderen het formuleren van specifieke hypothesen over een mogelijk verschillende leeropbrengst. Dit impliceert dat wij voor het onderhavige onderzoek alleen van open onderzoeksvragen konden uitgaan.

Zijn er verschillen tussen methoden wat betreft de leeropbrengst voor technisch en begrijpend lezen en voor spelling?

Zijn er verschillen tussen methodegebruikers wat betreft de manier waarop ze hun methode gebruiken?

Zijn er verschillen tussen methodegebruikers wat betreft de waardering voor hun methode?

Methodebeschrijving

Bij ons onderzoek waren vijf methoden betrokken, alle uitgegeven na 1990 (Tabel 1). De keuze voor deze vijf was opgelegd door de opdrachtgever van het onderzoek [1]. Uit opgaven van de uitgevers omtrent het gebruik kan worden afgeleid dat de methoden samen 60 à 65 procent van de markt dekken. Dit percentage neemt in de komende jaren wellicht nog toe, onder meer omdat mag worden verwacht dat de circa 1500 scholen die nu nog gebruik maken van de oude versie van Veilig leren lezen, hun methode zullen gaan vervangen.

Tabel 1 maakt duidelijk dat de methoden qua leerstofdomein niet gemakkelijk onder één noemer te brengen zijn. Weliswaar bieden ze alle vijf leerstof voor het voorbereidend en aanvankelijk lezen, maar het aanvankelijk spellen wordt maar door drie van de vijf gedekt. 'Alle kinderen leren lezen' en 'Leeslijn' bieden op dit gebied geen leerstof. Bij de interpretatie van de uitkomsten zullen we hiermee dan ook rekening houden. In het navolgende geven we van elke methode een korte karakteristiek. Een uitgebreidere analyse van de leesdidactische en onderwijskundige uitgangspunten vindt men in het onderzoeksrapport (Blok & Otter, 1997).

Tabel 1: Overzicht van de in het onderzoek betrokken methoden

Naam	Jaar van uitgave	Domein	Leerjaren	Eindniveau	Aantal technische gebruikers*
lezen					
Alle kinderen leren	1995	Aanvankelijk lezen (met een	2, 3	AVI-2	50

lezen (AK)		voorbereidende auditieve leergang)			
De Leesbus (LB)	1992-1993	Voorbereidend en aanvankelijk lezen, aanvankelijk spellen	1, 2, 3	AVI-4	30
Leeslijn (LL)	1991-1995	Voorbereidend, aanvankelijk en technisch lezen	1, 2, 3, 4, 5, 6	AVI-9	800
De Leessleutel (LS)	1991	Voorbereidend en aanvankelijk lezen, aanvankelijk spellen, spreken, luisteren, stellen, taalbeschouwing	1, 2, 3	AVI-2 à 3	600
Veilig leren lezen (maanversie), Schatkist (VL)	1991	Voorbereidend en aanvankelijk lezen, aanvankelijk spellen	1, 2, 3	AVI-3 à 4	ruim 3000

* Aantal scholen in Nederland dat de methode heeft ingevoerd, volgens schatting van de uitgever, najaar 1996.

Alle kinderen leren lezen (AK)

AK is een bewust eenvoudig gehouden methode voor aanvankelijk leesonderwijs met een voorbereidende leergang, gericht op de stimulering van de auditieve ontwikkeling. De methode bestrijkt de periode vanaf medio leerjaar 2 tot april leerjaar 3. De eenvoud van de methode blijkt onder meer uit de uitvoering van het leerlingmateriaal (alleen in zwart/wit), uit de bondigheid van de handleiding voor leerkrachten (circa 30 pagina's), het ontbreken van leesseries en uit het feit dat er weinig materiaal voor differentiatie bij de methode hoort. De lesbegeleiding voor leerkrachten is tot een minimum beperkt. De methode gaat uit van een voornamelijk klassikale werkwijze.

De Leesbus (LB)

LB is een methode voor voorbereidend en aanvankelijk leesonderwijs en aanvankelijk spellingonderwijs. In de praktijk bestrijkt de methode de leerjaren 2 en 3. De vooropgezette bedoeling van de auteur was een aantrekkelijke leesmethode te maken: goede teksten, mooie illustraties en zinvolle opdrachten. De doelstellingen van de methode liggen op het terrein van technisch lezen (streefniveau AVI- 4), begrijpend lezen (inclusief aandacht voor leesstrategieën) en de leeshouding. Voor spelling is geen expliciete doelstelling opgenomen, maar spelling vormt wel onderdeel van de leerstof. In LB wordt een klassikale werkwijze aanbevolen.

Leeslijn (LL)

LL is een methode voor voorbereidend, aanvankelijk en voortgezet technisch lezen. De methode is bestemd voor de leerjaren 1 tot en met 6. Het beoogde eindniveau is AVI-9. LL stimuleert individualiserend onderwijs in tegenstelling tot klassikaal onderwijs. Om dit ideaal zo goed mogelijk te realiseren, wordt de leerkracht aanbevolen de klas in niveaugroepen te verdelen, afhankelijk van de behoefte die leerlingen hebben aan expliciete instructie. De LL biedt twee leergangen: de

Aktiviteitenlijn (bedoeld voor kinderen die zichzelf leren lezen via vormen van ontdekkend leren) en De Leesweg (voor kinderen die in sterkere mate instructie nodig hebben). De Activiteitenlijn biedt vergaande mogelijkheden tot zelfontdekkend leren, onder meer via een omvangrijke verzameling leesspelletjes. De Leesweg biedt een sterk gestructureerde leergang, waarin het leren lezen met kleine stapjes opgebouwd wordt.

De Leessleutel (LS)

LS is een zogenoemde geïntegreerde lees- en taalmethode. De methode bevat leerstof voor voorbereidend en aanvankelijk lezen, voor spelling en voor taal (spreken, luisteren, stellen en taalbeschouwing). De methode is bestemd voor de leerjaren 1 tot en met 3. Het beoogde eindniveau voor technisch lezen is AVI-2 à AVI- 3. LS biedt de keuze het proces van het leren lezen langs drie verschillende didactische lijnen te laten plaatsvinden: via een methodische door de leerkracht gestuurde lijn, via de didactiek van zelfontdekkend leren en via een orthodidactische lijn. LS besteedt ruime aandacht aan de functies die geschreven taal in het dagelijks leven heeft en aan teksten met een non-fictie karakter.

Veilig leren lezen, maanversie (VL)

VL vormt in combinatie met Schatkist een methode voor voorbereidend en aanvankelijk leesonderwijs en aanvankelijk spellingonderwijs. Schatkist is bedoeld voor leerjaar 1 en 2. VL bestrijkt leerjaar 3. Het beoogde eindniveau voor technisch lezen komt overeen met AVI-3 à AVI-4. De beide onderdelen, hier als één methode beschreven, pretenderen een doorgaande lijn aan te bieden. De didactische aanpak is evenwel verschillend. Schatkist biedt een open benadering waarbij het accent ligt op taalverrijking in brede zin. In VL wordt een methodische lijn gevolgd met veel klassikale instructiemomenten. Een karakteristiek punt is dat bij de methode een omvangrijke verzameling differentiatiematerialen is ontwikkeld. De materialen vergroten de mogelijkheden voor de leerlingen om zelfstandig te werken.

Opzet van het onderzoek

Het uitgevoerde onderzoek bestond uit twee deelstudies, de deelstudie 'Leeropbrengst' naar de eerste onderzoeksvraag en de deelstudie 'Gebruik en waardering' naar de beide andere onderzoeksvragen. Van beide deelstudies bespreken we kort de opzet.

In de deelstudie 'Leeropbrengst' hebben metingen plaatsgevonden in leerjaar 3 en leerjaar 4. Maar omdat vier van de vijf methoden voor leerjaar 4 geen leerstof bieden (zie Tabel 1), rapporteren we hier alleen de resultaten voor leerjaar 3. De metingen hebben op twee momenten in het schooljaar 1996/97 plaatsgehad. Direct na de herfstvakantie zijn van de leerlingen enkele persoonsgegevens bepaald - waaronder het onderwijsvoorrangsgewicht - en is een woordenschattoets afgenomen. Het onderwijsvoorrangsgewicht en de woordenschatsscore bieden de mogelijkheid de vergelijkbaarheid van de vijf methodesteekproeven achteraf te controleren en zo nodig langs statistische weg te vergroten. In de laatste maand van het schooljaar zijn lees- en spellingtoetsen afgenomen (zie verder *Variabelen*).

Steekproef

Op verzoek van de opdrachtgever is het onderzoek uitsluitend uitgevoerd onder basisscholen, met uitsluiting derhalve van scholen voor speciaal onderwijs. Wij hebben de betrokken uitgevers verzocht een gebruikerslijst ter beschikking te stellen. Uit de daarop vermelde scholen hebben wij aselekt scholen getrokken. Deze hebben we vervolgens benaderd met het verzoek aan het onderzoek deel te nemen. In de uitnodigingsbrief zijn de scholen op de hoogte gesteld van het doel en de opzet van het onderzoek. Voorwaarde onzerzijds was dat de leerkracht van groep 3 de methode tenminste één jaar in gebruik moest hebben. Voor de deelstudie 'Leeropbrengst' streefden wij per methode naar 20 deelnemende scholen en per school naar tien deelnemende leerlingen. De beperking van het aantal leerlingen per school is ingegeven door de wens de toetslast voor de leerkracht te beperken. Hierbij speelde mee dat niet alle geplande toetsen klassikaal afgenomen kunnen worden. De keuze van de leerlingen is op basis van een toevalsprocedure tot stand gekomen. Tabel 2 laat zien dat we voor drie methoden (LL, LS, VL) in deze opzet redelijk geslaagd zijn. Voor de methoden AK en LB kon het beoogde aantal deelnemende scholen bij lange na niet worden gerealiseerd, omdat deze methoden erg weinig gebruikers kennen. Uit tabel 2 valt verder op te maken dat niet op alle deelnemende scholen ook daadwerkelijk tien leerlingen bij het onderzoek betrokken zijn geweest. Dat vindt zijn verklaring in het feit dat onder meer in combinatieklassen soms minder dan tien leerlingen uit leerjaar 3 aanwezig waren. Het percentage leerlingen met leerlinggewicht 1,90 (de allochtone leerlingen) blijkt bij de meeste methoden weinig af te wijken van het landelijke percentage, dat in het schooljaar 1996/97 ongeveer twaalf procent bedroeg. Alleen in de LB-steekproef is het percentage allochtone leerlingen afwijkend hoog.

Ten behoeve van de deelstudie 'Gebruik en waardering' is de 'Leeropbrengststeekproef' uitgebreid door middel van een verdere steekproeftrekking. De 'Gebruik en waardering-steekproef' overlapt de eerste steekproef derhalve volledig. We streefden naar 50 leerkrachten per methode. Bij drie methoden is dit streven bij benadering gerealiseerd, bij de methoden AK en LB, wederom in verband met het geringe aantal gebruikers, niet. Het gemiddelde aantal ervaringsjaren met de methode - of met de proefversie daarvan - bedraagt minimaal bijna drie jaar. Dat lijkt ons voldoende lang om een gefundeerd oordeel te kunnen geven over gebruik en waardering.

Variabelen

Voor de deelstudie 'Leeropbrengst' hebben we gebruik gemaakt van bestaande toetsen die onderdeel zijn van het leerlingvolgsysteem van het Cito. Voordeel hiervan is dat bij deze toetsen landelijke normscores zijn vastgesteld. Dit is door het Cito gedaan in een afzonderlijk normeringsonderzoek. De toetsen bestrijken vier domeinen:

1. woordenschat; afgenomen is Woordenschattoets 1;
2. technisch lezen; afgenomen is de Drie-Minuten-Toets (de kaarten 1c, 2c en 3c);
3. begrijpend lezen; afgenomen is de toets Lezen met Begrip 1 (twee toetsen, namelijk de Schaal Betekenisrelaties, toets E3 en de Schaal Verwijsrelaties, toets 1);
4. spelling; afgenomen is de Schaal Vorderingen in Spellingvaardigheid 1 (Dictee e3b).

Tabel 2: Opbouw van de steekproef in de deelstudies 'Leeropbrengst' en 'Gebruik en waardering'

Methode	'Leeropbrengst'		% 1,90 leerlingen	'Gebruik en waardering'	
	Aantal scholen	Aantal leerlingen		Aantal leerkrachten	Gemiddelde ervaring*
AK	10	93	10	9	4,2
LB	11	109	17	11	2,7
LL	18	174	10	40	5,3
LS	17	167	10	37	2,9
VL	21	196	9	54	3,1

* Vermeld wordt het gemiddeld aantal ervaringsjaren met de methode of met een proefversie daarvan.

Voor de deelstudie 'Gebruik en waardering' hebben we een eigen vragenlijst ontwikkeld. De vragenlijst bevatte 32 vragen, merendeels over het gebruik dat leerkrachten van hun methode maken. Daarnaast waren er 37 vragen die de waardering voor de methode betroffen. Deze waarderingsvragen hadden de vorm van stellingen over de methode. De leerkrachten hebben per stelling de waardering voor hun methode op een vijfpuntsschaal aangegeven. Vragen die inhoudelijk dicht bij elkaar liggen, zijn samengevoegd tot één schaal. In totaal zijn zes schalen onderscheiden. De betrouwbaarheid van de schalen varieerde tussen .76 en .91 (Cronbach's alpha).

Gegevensverzameling en -analyse

In de laatste maand van het schooljaar hebben de leerkrachten de toetsen zelf afgenomen en gescoord. De ruwe scores zijn vervolgens opgestuurd, samen met de ingevulde vragenlijsten. Ter vermijding van absolute scores - die moeilijk rechtstreeks te interpreteren zijn - is in beide deelstudies gebruik gemaakt van een norm- of referentiegroep (zie hierna).

In de deelstudie '*Leeropbrengst*' zijn allereerst analyses uitgevoerd met het oog op de vergelijkbaarheid van de vijf methodesteekproeven. In het voorafgaande is al gebleken dat het percentage allochtone leerlingen niet gelijk verdeeld is: de LBsteekproef telt meer van zulke leerlingen. Ten aanzien van de tweede controlevariabele, woordenschat, blijken de gemiddelden van de vijf groepen niet statistisch significant van elkaar te verschillen (eenweg-variantieanalyse; $F(4,734) = 1,90$; $p = .11$). Desondanks valt het gemiddelde in de LB-steekproef net iets lager uit - ongeveer 0,2 standaarddeviatie - dan in de andere vier groepen. Het vermoeden is daarom dat de LB-steekproef in vergelijking met de vier andere steekproeven licht in het nadeel is. Om voor dit nadeel te corrigeren zijn de definitieve analyses niet uitgevoerd op de geobserveerde scores, maar op scores die gecorrigeerd zijn op basis van covariantie-analyses met twee covariaten, namelijk het onderwijsvoorrangsgewicht en de score op de woordenschattoets. Voor elk van de zes toetsen (zie Tabel 3) is een dergelijke analyse uitgevoerd. De analyses zijn steeds verricht op de totale steekproef van ruim 700 leerlingen, omdat bleek dat de regressielijnen in de vijf methodesteekproeven bij

benadering parallel liepen. Overigens vertonen de twee covariaten slechts een geringe samenhang met de afhankelijke variabelen. Het percentage gedeelde variantie (multiple R²) varieert tussen de 2 procent (voor DMT, kaart 1c) en 14 procent (voor Lezen met Begrip, toets E3). Als gevolg hiervan hebben de covariantie-analyses dan ook een bescheiden effect op de verdeling van de afhankelijke variabelen. Het grootste effect, ter grootte van 0,06 standaarddeviatie, trad op bij de methode LS, toets E3.

In de volgende stap zijn de gecorrigeerde toetsscores vergeleken met de scores van de normgroep van het Cito. De vergelijking is op twee punten uitgevoerd.

Het gemiddelde in de deelsteekproeven is vergeleken met het gemiddelde van de Cito-normgroep. Dit is gedaan door de effectgrootte te berekenen (het verschil tussen de beide gemiddelden, gedeeld door de standaardafwijking in de normgroep). Ook is de standaardfout van de effectgrootte geschat.

Het percentage leerlingen in Cito-niveaugroep A (de bij benadering 25 procent hoogst scorende leerlingen) en in Cito-niveaugroep E (de bij benadering 10 procent laagst scorende leerlingen) is bepaald.

In de deelstudie '*Gebruik en waardering*' hebben we bij de waarderingsschalen gebruik gemaakt van een eigen referentiegroep. Voor elke methode wordt de referentiegroep gevormd door de gebruikers van de andere vier methoden². Nadeel hiervan is dat de referentiegroep voor elke methode enigszins anders is samengesteld. Maar zo wordt in elk geval vermeden dat gebruikers van een methode vergeleken worden met een referentiegroep waarvan ze zelf deel uitmaken, hetgeen de vergelijking tautologisch zou maken. Voor elke methode en elke dimensie is het gemiddelde vergeleken met het gemiddelde van de referentiegroep. Dit is wederom gedaan door de effectgrootte te berekenen (het verschil tussen de beide gemiddelden, gedeeld door de standaardafwijking in de referentiegroep). Ook is de standaardfout van de effectgrootte geschat.

In de volgende sectie bespreken we alleen effectgroottes die met 95 procent zekerheid afwijken van nul. Daartoe is steeds het 95-procent betrouwbaarheidsinterval rondom de effectgroottes bepaald. Door deze handelwijze minimaliseren we het kapitaliseren op steekproeffluctuatie. De standaardfouten bij de effectgroottes voor de lees- en spellingtoetsen variëren tussen de 0,08 en 0,12. Voor de waarderingsschalen variëren ze tussen de 0,16 en 0,35. Het kleinst aantoonbare effect bedraagt derhalve 0,16 (voor de lees- en spellingtoetsen), respectievelijk 0,32 (voor de waarderingsschalen). Overigens bevatten de tabellen alle effectgroottes, ongeacht of deze statistisch significant van nul verschillen.

Resultaten

De leeropbrengst

Allereerst bespreken we de *gemiddelde* leervorderingen, uitgedrukt als een effectgrootte (Tabel 3). Voor het domein 'technisch lezen' blijkt dat twee methoden geen of bijna geen verschil laten zien met de Cito-normgroep. Het betreft de methoden AK en LB, waarvoor dus geconcludeerd kan worden dat de leerlingen die met een van deze methoden onderwijs gekregen hebben, gemiddeld genomen na één jaar even ver gekomen zijn als de gemiddelde Cito-normleerling. Voor de drie

andere methoden geldt dat de leerlingen aan het eind van leerjaar 3 voor technisch lezen gemiddeld een voorsprong hebben gerealiseerd. De omvang van de voorsprong varieert van 0,2 standaarddeviatie tot 0,4 standaarddeviatie.

Voor het domein 'begrijpend lezen' blijkt de uitkomst iets complexer te zijn. Twee methoden laten een negatieve effectgrootte zien, twee andere methoden realiseren een positieve effectgrootte. Van belang is dat de geconstateerde effectgroottes zich niet op beide toetsen tegelijk manifesteren. De achterstand of voorsprong blijkt steeds maar op één van beide toetsen zichtbaar.

Op de toets voor 'spelling' komt slechts één negatieve effectgrootte aan het licht. AK-leerlingen scoren gemiddeld 0,2 standaarddeviatie onder de gemiddelde normleerling. De andere methoden laten geen verschillen met de Cito-normgroep zien.

Ter afsluiting van deze paragraaf besteden we aandacht aan de *verdelingsstaarten*, het percentage leerlingen in Cito-niveaugroep A (de zeer goede leerlingen) en in Cito-niveaugroep E (de zeer zwakke leerlingen). Voor de methoden AK en LB geldt dat het percentage A-leerlingen - bezien over het geheel van de zes toetsen - in de buurt ligt van de verwachte 25 procent. Bij de andere drie methoden blijkt het aantal A-leerlingen overwegend iets groter, met name op de toetsen voor technisch lezen. Het percentage E-leerlingen, de uitvallers, laat tussen de methoden weinig verschillen zien. Methode LL levert weliswaar de minste uitvallers op, maar om te bepalen of dit verschil statistisch significant is, is geen statistische toets voorhanden.

Tabel 3: Overzicht van de leervorderingen einde leerjaar 3; op basis van gecorrigeerde scores (na covariantie- analyse)

Methode	Domein	Toets	Effectgrootte	% leerlingen in	
				niveau Ac	niveau Ed
AK	TL	DMT-1c	- 0,06	28	15
	TL	DMT-2c	0,12	39	9
	TL	DMT-3c	0,07	32	6
	BL	Toets E3	- 0,23	17	8
	BL	Toets 1	- 0,04	30	16
	SP	SVS-e3b	- 0,23	17	11
LB	TL	DMT-1c	0,10	31	7
	TL	DMT-2c	0,23	35	6
	TL	DMT-3c	0,15	35	5
	BL	Toets E3	- 0,30	13	7
	BL	Toets 1	0,17	31	10
	SP	SVS-e3b	- 0,10	26	8
LL	TL	DMT-1c	0,22	33	3
	TL	DMT-2c	0,33	36	5
	TL	DMT-3c	0,24	36	6

	BL	Toets E3	- 0,02	29	4
	BL	Toets 1	0,08	29	9
	SP	SVS-e3b	- 0,09	23	5
LS	TL	DMT-1c	0,26	39	8
	TL	DMT-2c	0,44	42	8
	TL	DMT-3c	0,42	42	8
	BL	Toets E3	0,08	33	9
	BL	Toets 1	0,37	37	7
	SP	SVS-e3b	- 0,14	24	13
VL	TL	DMT-1c	0,30	37	6
	TL	DMT-2c	0,44	39	8
	TL	DMT-3c	0,34	38	7
	BL	Toets E3	-0,01	34	7
	BL	Toets 1	0,36	41	9
	SP	SVS-e3b	0,09	31	9

a TL = Technisch Lezen; BL = Begrijpend Lezen; SP = Spelling.

b De effectgrootte is uitgedrukt als het verschil tussen het gemiddelde van de steekproef en het gemiddelde van de normgroep van het Cito. Het verschil is gestandaardiseerd (op basis van de standaardafwijking in de normgroep). Een negatieve effectgrootte wil zeggen dat de leerlingen in de steekproef gemiddeld een lagere score hebben gehaald, een positieve effectgrootte duidt op een gemiddeld hogere score. Effectgrootten die met 95 procent zekerheid van nul verschillen, worden cursief weergegeven.

c Niveau A: goed tot zeer goed (overeenkomend met de bij benadering 25 procent hoogst scorende leerlingen).

d Niveau E: zwak tot zeer zwak (overeenkomend met de bij benadering 10 procent laagst scorende leerlingen).

Gebruik van de methode

Om te beginnen vermelden we dat het methodegebruik niet los staat van de schoolorganisatie. De methoden AK en LL worden vaker in een combinatieklas gebruikt dan de drie andere methoden. Voor AK en LL gaat het om ongeveer 70 procent van de steekproef, voor de drie andere methoden om 20 tot 30 procent van de steekproef. Het betreft met name de combinatie 3/4, maar ook de combinaties 2/3 en 3/4/5 komen een aantal malen in de steekproef voor.

We hebben de leerkrachten gevraagd of er in leerjaar 2 gebruik wordt gemaakt van de 'voorloper', het deel van de methode dat betrekking heeft op het voorbereidende leesonderwijs. Ook ten aanzien van deze vraag blijken er verschillen te bestaan. De gebruikers van de methoden LB en LL geven bijna allen aan gebruik te maken van het voorbereidende leesprogramma. Voor de drie andere methoden geldt dat 55 tot 70 procent van de leerkrachten aangeeft dat de 'voorloper' in het tweede leerjaar in

het geheel niet gebruikt wordt. Dit verschil heeft aantoonbare gevolgen voor de activiteiten die in leerjaar 2 aan leerlingen worden aangeboden. Met name het 'bevorderen van de letterkennis' en het 'woorden schrijven of stempelen' komt op scholen waar de methode LL wordt gebruikt vaker voor. Enkele andere activiteiten - rijmen, het oefenen van de auditieve analyse en synthese, het namen schrijven of stempelen - bleken bij alle vijf de methoden ongeveer even vaak voor te komen. Dit zijn activiteiten die volgens bijna alle leerkrachten, ongeacht de gebruikte methode, in leerjaar 2 worden aangeboden.

De mate waarin de methode 'bepalend is voor het aanvankelijk leesonderwijs' blijkt hoog. Vrijwel alle leerkrachten, ongeacht de methode die ze gebruiken, geven aan dat de methode hun onderwijs vrijwel geheel of in belangrijke mate bepaalt. Dat neemt niet weg dat leerkrachten - naast de methode die zij als hoofdmethode beschouwen - nog andere methoden gebruiken. Voor aanvankelijk lezen wordt door 12 procent van de leerkrachten naar een andere methode gegrepen. Voor begrijpend lezen en voor spelling bedraagt dit aantal 24, respectievelijk 49 procent. Vergelijken we de methoden in dit opzicht, dan treden er enkele verschillen aan het licht. Zo grijpen gebruikers van AK voor alle drie de domeinen vaker naar een andere methode. Gebruikers van LB en LL gebruiken alleen voor spelling vaker een additionele methode. Gebruikers van LS en VL hebben de behoefte aan een additionele methode het minst.

'Hoeveel tijd ontvangen de leerlingen in leerjaar 3 in een doorsnee week leesen spellingonderwijs?' was een van de in de vragenlijst gestelde vragen. Uit de antwoorden valt op te maken dat er wat betreft de tijdsbesteding sprake is van een tweedeling tussen de methoden. Gebruikers van AK, LB en LL besteden wekelijks gemiddeld 315 minuten aan lezen en spellen. Voor gebruikers van LS en VL bedraagt het gemiddelde 370, respectievelijk 400 minuten. Het verschil komt neer op één à anderhalf uur per week.

De manier waarop leerkrachten hun instructie organiseren blijkt tussen methoden duidelijk te verschillen. Het zijn met name de gebruikers van LL die instructies vaak in subgroepen of individueel aanbieden. Bij de vier andere methoden betreft het hoogstens 10 procent van de leerkrachten die zo werken, maar bij de LL gaat het om 70 procent van de gebruikers. Ook wat betreft de verdeling van de leerstof blijkt er een pregnant verschil te zijn. Ruim 80 procent van de LL-gebruikers laat de leerlingen in eigen tempo de leerstof doorwerken. Bij de andere methoden bedraagt dit aantal minder dan 15 procent. De niet-LL gebruikers hanteren in overgrote meerderheid een model met basisstof voor iedereen en differentiatiestof voor snelle en langzame leerlingen.

Tabel 4: De waardering voor de methoden (uitgedrukt op een vijfpuntsschaal: 1 = sterk negatief; 2 = negatief; 3 = neutraal; 4 = positief; 5 = sterk positief); gemiddelden, standaardafwijkingen en effectgrootten

Methode	Dimensie	Gem.	Stand.afw.	Eff.gr.*
AK	1. Duidelijke doelen	3,94	0,68	- 0,09
	2. Opbouw van de leerstof	4,02	0,40	- 0,08
	3. Zorgverbreding	3,51	0,71	- 0,70
	4. Duidelijke lesbeschrijvingen	3,96	0,69	- 0,08

	5. Handleiding: algemene deel	3,46	0,69	- 0,44
	6. Aantrekkelijkheid leerlingenmateriaal	3,60	0,53	- 1,32
LB	1. Duidelijke doelen	3,98	0,47	- 0,02
	2. Opbouw van de leerstof	4,14	0,49	0,28
	3. Zorgverbreding	3,91	0,34	0,11
	4. Duidelijke lesbeschrijvingen	4,27	0,45	0,44
	5. Handleiding: algemene deel	4,17	0,32	0,85
	6. Aantrekkelijkheid leerlingenmateriaal	4,42	0,40	0,63
LL	1. Duidelijke doelen	3,87	0,60	- 0,24
	2. Opbouw van de leerstof	4,05	0,51	- 0,02
	3. Zorgverbreding	3,68	0,65	- 0,32
	4. Duidelijke lesbeschrijvingen	3,34	0,79	- 1,45
	5. Handleiding: algemene deel	3,12	0,64	- 1,17
	6. Aantrekkelijkheid leerlingenmateriaal	4,21	0,51	0,18
LS	1. Duidelijke doelen	4,03	0,72	0,10
	2. Opbouw van de leerstof	4,10	0,60	0,13
	3. Zorgverbreding	4,41	0,64	1,22
	4. Duidelijke lesbeschrijvingen	4,25	0,70	0,43
	5. Handleiding: algemene deel	3,91	0,68	0,40
	6. Aantrekkelijkheid leerlingenmateriaal	4,15	0,63	0,06
VL	1. Duidelijke doelen	4,11	0,55	0,24
	2. Opbouw van de leerstof	3,96	0,47	- 0,23
	3. Zorgverbreding	3,73	0,51	- 0,21
	4. Duidelijke lesbeschrijvingen	4,23	0,43	0,36
	5. Handleiding: algemene deel	3,82	0,55	0,22
	6. Aantrekkelijkheid leerlingenmateriaal	4,26	0,45	0,27

a De effectgrootte is uitgedrukt als het verschil met de waardering voor de groep 'Gebruikers overige methoden'. Het verschil is gestandaardiseerd (op basis van de standaardafwijking in de referentiegroep 'Gebruikers overige methoden'). Een negatieve effectgrootte betekent dat de methode minder wordt gewaardeerd, een positieve effectgrootte duidt op een grotere waardering. Effectgrootten die met 95 procent zekerheid van nul verschillen, zijn cursief weergegeven.

Het gebruik van een methodeoverstijgend leerlingvolgsysteem - in de vragenlijst omschreven als 'een systeem waarbij leerlingen meermalen per jaar een toets maken, zodat hun voortgang door de leerstof vergeleken kan worden met een landelijke normgroep' - is beslist niet algemeen. Voor de drie domeinen - aanvankelijk lezen, begrijpend lezen en spelling - bedraagt het invoeringspercentage 85, respectievelijk 46 en 77. Gebruikers van AK en LL maken relatief wat minder gebruik van zo'n methodeoverstijgend systeem. Tussen de gebruikers van andere methoden zijn geen grote verschillen gebleken.

Waardering voor de methode

Tot slot bezien we dan nog de waardering van de leerkrachten voor hun methode. We wijzen er allereerst op dat die waardering, afgaande op de in Tabel 4 vermelde gemiddelden, overwegend positief is. De meeste gemiddelden bewegen zich rond de vier, hetgeen op de antwoordschaal omschreven is als 'positief'. De laagste gemiddelde score betreft 3,1, de gemiddelde waardering op de dimensie 'Handleiding: algemene deel' voor de methode LL.

Vergelijken we elke methode met de referentiegroep - bestaande uit de gebruikers van de andere vier methoden - dan komen er enkele verschillen aan het licht. De methoden AK en LL worden elk op twee van de zes dimensies minder gewaardeerd. De mindere waardering voor LL - op de dimensies 'Duidelijke lesbeschrijvingen' en 'Handleiding: algemene deel' - is zelfs fors te noemen. Voor de methoden LB en LS is de waardering het meest positief. LB wordt onder meer extra gewaardeerd op de dimensie 'Aantrekkelijkheid leerlingmateriaal', een aspect dat in de methode expliciete aandacht heeft gekregen. LS wordt onder meer extra gewaardeerd op de dimensie 'Zorgverbreding'. Kennelijk hebben de gebruikers grote waardering voor de orthotheek die bij de methode hoort.

Conclusies en discussie

Conclusies

Ten behoeve van de conclusies grijpen we allereerst terug op de onderzoeksvragen.

De *eerste* onderzoeksvraag handelt over de mogelijke verschillen tussen methoden wat betreft de leeropbrengst voor technisch en begrijpend lezen en voor spelling. Onze conclusie is dat zulke verschillen inderdaad aangetoond zijn: methoden voor aanvankelijk lezen verschillen aan het eind van het derde leerjaar wat betreft de leeropbrengst. Voor technisch lezen concluderen wij dat drie methoden (LL, LS en VL) tot een leeropbrengst leiden die hoger is dan de in het Cito-leerlingvolgsysteem vastgestelde normopbrengst. Voor begrijpend lezen realiseren twee methoden (LS en VL) een hogere leeropbrengst, die overigens maar op één van beide toetsen zichtbaar wordt. Voor spelling valt aangaande één methode, namelijk AK, te constateren dat de leeropbrengst onder de normopbrengst ligt. Hierbij tekenen we aan dat het aanvankelijk spellen door de methode *AKniet* gedekt wordt. Overigens dekt ook de methode LL het aanvankelijk spellen niet, maar dat leidt bij deze methode niet tot een achterstand op de Citonormgroep. Een tweede kanttekening is dat alle geconstateerde verschillen klein zijn, maximaal 0,4 standaarddeviatie. Verschillen van deze omvang blijken - na interpolatie op basis van de normeringsgegevens van het Cito - bij benadering equivalent met één à twee maanden leesonderwijs.

Naast conclusies over de gemiddelde leeropbrengst kunnen we ook conclusies trekken over het aantal hele goede lezers (Cito-niveau A), dan wel het aantal hele zwakke lezers (Cito-niveau E). Het percentage hele goede lezers - we beperken ons hier tot het technisch en begrijpend lezen - overstijgt bij drie methoden stelselmatig de verwachte 25 procent. Dat zijn de methoden LL, LS en VL. Het percentage hele zwakke lezers is bij vier methoden stelselmatig ten hoogste tien procent. Dat zijn wederom de methoden LL, LS en VL, aangevuld met LB.

De *tweede* onderzoeksvraag betreft de verschillen in het gebruik. Zulke verschillen blijken er inderdaad te zijn. Zo worden de methoden AK en LL beduidend vaker in combinatieklassen gebruikt. Wat betreft de inrichting van de onderwijsleersituatie lijken vier van de vijf methoden op elkaar: het accent ligt op klassikale instructiemomenten, afgewisseld met aparte leerstof voor zwakke en

snellere leerlingen. Uitzondering vormen de gebruikers van LL. Ongeveer 70 procent van hen biedt instructies vooral in kleine groepen of individueel aan. LLgebruikers kiezen er in grote meerderheid voor dat leerlingen in eigen tempo de leerstof doorwerken. Deze gebruikswijze is consistent met de door de methode aanbevolen werkwijze. Verdere verschillen ten aanzien van het gebruik betreffen onder meer het gebruik van additioneel materiaal (dit gebeurt vaker bij AK) en de tijdbesteding (gebruikers van LS en VL besteden wekelijks één à anderhalf uur extra aan lees- en spellingonderwijs).

Verschillen wat betreft de waardering van leerkrachten voor hun methode vormen het onderwerp van de *derde* onderzoeksvraag. Allereerst wijzen wij erop dat de waardering voor de meeste methoden op de meeste onderscheiden dimensies overwegend positief is. Het hoge niveau van waardering neemt evenwel niet weg dat er tussen methoden in dit opzicht verschillen blijken te bestaan. Twee methoden worden op enkele dimensies minder dan gemiddeld gewaardeerd, namelijk AK en LL. Drie andere methoden hebben daarentegen een relatief hoge waardering ontvangen, namelijk LB, LS en VL, de laatste overigens slechts marginaal hoger. De omvang van de verschillen overstijgt in vier gevallen de waarde van één standaarddeviatie.

Discussie

De validiteit van de voorgaande conclusies is vanzelfsprekend afhankelijk van de kwaliteit van de onderzoeksopzet. In dat verband brengen we enkele overwegingen naar voren, waarvan sommige tot voorzichtigheid nopen.

Een eerste groep overwegingen heeft te maken met selectie. Er zijn verscheidene selectiemechanismen die de interpretatie van de onderzoeksuitkomsten bemoeilijken. Zo kan er selectie optreden op het moment van methodekeuze. Scholen die voor de ene methode kiezen, kunnen systematisch verschillen van scholen die voor een andere methode kiezen. Selectie kan ook optreden in de vorm van zelfselectie. De deelnemers aan het onderzoek zijn immers geworven op basis van vrijwilligheid. De scholen die 'namens methode A' aan het onderzoek hebben deelgenomen, hoeven niet representatief te zijn voor de populatie van scholen die methode A hebben ingevoerd. Het is bijzonder lastig om op beide selectiemechanismen greep te krijgen, niet in het minst omdat bij gebrek aan populatiegegevens beide mechanismen niet te ontwarren zijn. In feite kan men slechts achteraf een vergelijking maken op variabelen die een zuivere vergelijking mogelijksterwijs verstoren. Eén zo'n variabele is de achterstandssituatie van een school. Een school met veel achterstandsleerlingen heeft nu eenmaal, ondanks ruimere onderwijsfaciliteiten, minder mogelijkheden bij zijn leerlingen goede leervorderingen te realiseren dan een school met weinig achterstandsleerlingen. Het is in het onderhavige onderzoek een gelukkige omstandigheid dat het percentage 1,90-leerlingen in vier van de vijf steekproeven weinig bleek af te wijken van het landelijke percentage. Alleen in de LB-steekproef troffen we een oververtegenwoordiging van 1,90-leerlingen aan, namelijk zeventien procent van het aantal leerlingen. Landelijk wordt twaalf procent van de leerlingen tot deze categorie gerekend. In dezelfde LB-steekproef bleek ook de omvang van de woordenschat van de leerlingen iets lager te zijn (circa 0,2 standaarddeviatie). Het ligt voor de hand dat de leeropbrengst in deze groep hierdoor enigszins onder druk is komen te staan en het is niet zeker dat onze analyses op basis van (na covariantie-analyse) gecorrigeerde scores het nadelige effect voldoende compenseren. Leerlinggewicht en woordenschat zijn immers slechts 'proxies' voor onderwijsachterstanden. Maar er kunnen ook in andere opzichten tussen methodegroepen

verschillen bestaan die een eerlijke vergelijking verhinderen. Voor een selectie van school- en klaskenmerken hebben wij achteraf een vergelijking tussen de methodegroepen gemaakt. Op geen van de in de analyse betrokken kenmerken stuiten we op een statistisch significant verschil. Het betreft de volgende variabelen: het aantal jaren dat de leerkracht in leerjaar 3 gestaan heeft, de mate waarin de leerkracht de nadruk legt op basisvaardigheden, de beschikbare hoeveelheid extra hulp in de vorm van een extra leerkracht of 'remedial teacher' en het schoolbeleid wat betreft de bevordering van leerjaar 2 naar leerjaar 3. Uit schooleffectiviteitsonderzoek in Nederland is aangaande deze variabelen gebleken dat ze een samenhang vertonen met de leerprestaties van leerlingen (Ledoux et al., 1997; Volman et al., 1997). Gezien deze resultaten trekken we de voorzichtige conclusie dat er geen voor de hand liggende verschillen op school- of klasniveau lijken te bestaan die de onderhavige methodevergelijking kan verstoren.

Aparte aandacht vragen we voor de aard van de onafhankelijke variabele. Weliswaar laat deze zich simpelweg benoemen als de 'gehanteerde methode', maar in werkelijkheid hebben we natuurlijk met een complex van factoren te maken. Het gaat niet alleen om de materialen en de verschillen daartussen, het gaat ook om de manier waarop de materialen worden ingezet. In feite bestaan er zelfs over de gebruikte materialen al onzekerheden. Een aantal leerkrachten geeft bijvoorbeeld aan dat ze additionele materialen inzetten. Een andere materiële factor is dat drie van de vijf methoden een aanzienlijke hoeveelheid keuzematerialen kennen. Mogelijk bepaalt het inzetten daarvan mede de effectiviteit van de methode. In feite hebben we dus te maken met een vijftal niet erg duidelijk gedefinieerde 'treatments'. Waar dringend behoefte aan bestaat, is een theorie die de 'black box' tussen de gehanteerde methode en de leerresultaten openlegt. Voor het onderhavige onderzoek, dat van meet af aan is opgezet als een methodevergelijkend onderzoek met een toegepast karakter, is de ontwikkeling van zo'n theorie een stap te ver.

Het toegepaste karakter van het hier gerapporteerde onderzoek komt tot uiting in de relevantie van de uitkomsten voor de schoolpraktijk. Scholen maken over het algemeen veel werk van de aanschaf van een nieuwe methode. Methoden vormen voor leerkrachten immers een heel belangrijk hulpmiddel bij de inrichting van onderwijsleersituaties. De aanschaf van een methode is bovendien een zware aanslag op het schoolbudget. Het is niet ongewoon een afschrijvingstermijn van tien jaar te hanteren. Een verkeerde beslissing kan daarom langdurig vervelende gevolgen hebben. Ons onderzoek laat zien dat een uitvoerige oriëntatie op de beschikbare methoden voor aanvankelijk lezen inderdaad de moeite loont. De vijf onderzochte methoden verschillen wat betreft de leeropbrengst, de gebruiksmogelijkheden en de waardering die gebruikers voor hun methode hebben. Het moge een geruststellende conclusie zijn dat leerkrachten met alle vijf methoden een leeropbrengst weten te realiseren die - in elk geval wat betreft technisch lezen - dicht in de buurt van het landelijke gemiddelde ligt. Deze uitkomst biedt scholen ruimte bij de aanschaf ook overwegingen te hanteren op het vlak van gebruik en waardering. Eén zo'n overweging is de mate waarin een school het leesonderwijs wil individualiseren, zoals in de methode LL gebeurt, dan wel juist een voorkeur heeft voor een meer klassikale aanpak, zoals in de andere vier methoden is uitgewerkt. Onze resultaten laten zien dat met beide aanpakken goede resultaten bereikt worden.

Ter afsluiting schenken we nog kort aandacht aan de vraag of de betrokken methoden de zelf gestelde streefdoelen (zoals aangegeven in Tabel 1) weten te realiseren. Vertaalt men de scores op de Drie-Minuten-Toets (op basis van Gillijns & Visser, 1995) in AVI-scores, dan blijkt voor alle vijf

methoden het eindniveau dat feitelijk door de gemiddelde leerling wordt behaald, ongeveer AVI 2 à 3 te zijn. Dit impliceert dat twee methoden, LB en VL, net iets hoger mikken dan in de praktijk realiseerbaar blijkt. Voor de andere drie methoden geldt dat het beoogde eindniveau vrijwel exact gelijk is aan het feitelijk gerealiseerde eindniveau.

Literatuur

Aarnoutse, C.A.J., Mommers, M., Smits, B. & Leeuwe, J. van (1986). De ontwikkeling en samenhang van technisch lezen, begrijpend lezen en spellen. *Pedagogische Studiën*, 63, 97-110.

Blok, H. & Otter, M.E. (1997). *Vijf methoden voor aanvankelijk lezen onderzocht* (vijf delen). Amsterdam: SCO-Kohnstamm Instituut.

Creemers, H.P.M. (1974). *Evaluatie van onderwijsstijlen binnen het aanvankelijk lezen* (diss.) Utrecht: Universiteit Utrecht.

Creemers, B.P.M. (1991). *Effectieve instructie: een empirische bijdrage aan de verbetering van het onderwijs in de klas*. 's-Gravenhage: SVO.

Dongen, D. van (1984). *Leesmoeilijkheden: naar diagnostiserend onderwijzen bij het leren lezen*. Tilburg: Zwijsen.

Filipiak, P. & Laarhoven, A. van (red.) (1992). *Analyse van methoden voor aanvankelijk lezen* (diverse delen). 's-Hertogenbosch: Katholiek Pedagogisch Centrum.

Gillijns, P. & Visser, J. (1995). De relatie tussen de Drie-Minuten-Toets en de AVI-toetskaarten. *School & Begeleiding*, (december), 9-12.

Groeneveld, K., Smout, W., Stegeman, H. & Vernooy, K. (1992). *Zes methoden voor aanvankelijk lezen nader bekeken*. Hoevelaken: Christelijk Pedagogisch Studiecentrum.

Harskamp, E. & Deinum, J.F. (1995). *Leerresultaten met twee methoden voor aanvankelijk lezen*. Groningen: GION.

Hol, G.G.J.M., Haan, M. de & Kok, W.A.M. (1995). *De effectiviteit van methodes voor aanvankelijk leesonderwijs*. Utrecht: ISOR/Universiteit Utrecht.

Ledoux, G., Overmaat, M. & Koopman, P. (1997). *Kwaliteitszorg in het primair onderwijs*. Amsterdam: SCO-Kohnstamm Instituut.

Kooreman, H.J. (1974). Konstruktie en resultaten van een onderwijsleerpakket voor het technisch leren lezen. *Pedagogische Studiën*, 51, 398-412.

Ministerie van Onderwijs, Cultuur en Wetenschappen (1995). *Een impuls voor het basisonderwijs*. Zoetermeer: Ministerie van Onderwijs, Cultuur en Wetenschappen.

Oorsouw, G.A. van (1971). *Een vergelijking tussen twee methoden voor aanvankelijk lezen* (doctoraalscriptie Didactiek) Amsterdam: Vrije Universiteit.

Overmaat, M. & Ledoux, G. (1998). *School- en klaskenmerken basisonderwijs; technische rapportage PRIMA-cohortonderzoek*. Amsterdam: SCO-Kohnstamm Instituut.

Reitsma, P., Komen, N. & Kapinga, T. (1981). Methoden voor aanvankelijk lezen: een vergelijking van leerresultaten na één jaar. *Pedagogische Studiën*, 58, 174-189.

Struiksma, A.J.C., Leij, A. van der & Vieijra, J.P.M. (1997). *Diagnostiek van technisch lezen en aanvankelijk spellen* (zesde herziene druk). Amsterdam: VU Uitgeverij.

Veenman, S. (1992). Effectieve instructie volgens het directe instructiemodel. *Pedagogische Studiën* 69, 242-269.

Vernooy, C.G.Th. (1993). *Leesproblemen voorkomen en er effectief mee omgaan*. Hoevelaken: CPS.

Volman, M., Deckers, P. & Roeleveld, J. (1997). *De invloed van organisatiefactoren op het functioneren van leerkrachten en leerlingen*. Amsterdam: SCO-Kohnstamm Instituut.

Noten

[1] Het onderzoek is uitgevoerd in opdracht van het Instituut voor Onderzoek van het Onderwijs (SVO), project 96030.

[2] Om te bereiken dat alle methoden in de referentiegroep even zwaar tellen is een weging toegepast, omgekeerd evenredig aan het aantal leerkrachten in de steekproef.