

Effecten van begeleid hardop lezen van teksten op technisch lezen, begrijpend lezen, woordenschat en leesplezier bij zwakke lezers in de leerjaren vier tot en met zes¹

H. Blok, R. Oostdam en C. Boendermaker

Samenvatting

Bijna elke basisschool kent leerlingen die in de fase van het voortgezet technisch leesonderwijs ernstig achterblijven. Zulke leerlingen kenmerken zich door een traag leestempo of door het maken van veel fouten. In dit onderzoek is nagegaan of de methodiek begeleid hardop lezen van teksten een effectieve remedie is. Begeleid hardop lezen houdt in dat een leerling hardop een tekst leest, terwijl een ervaren lezer de voortgang controleert en zo nodig hulp of instructie geeft. Aan het onderzoek hebben 126 zwakke lezers meegedaan, verdeeld over drie leerjaren (4, 5 en 6) en acht scholen. De leeshulp is één-op-één aangeboden, ofwel in de variant Opnieuw Lezen, ofwel in de variant Verder Lezen. Een derde groep leerlingen, de controlegroep, kreeg de leeshulp niet. De duur van de leeshulp bedroeg 48 leessessies van 20 minuten, in principe vier sessies per week. De onderzoeksoorzak voorzorg in een voormeting, twee tussenmetingen, een nameting en een retentiemeting. De uitkomsten laten zien dat begeleid hardop lezen effectief is ten aanzien van technisch lezen (lezen van woordrijen en van tekst) en leesplezier. Effecten op begrijpend lezen en woordenschat konden niet worden aangetoond. Tussen de twee varianten Opnieuw Lezen en Verder Lezen is geen verschil in effectiviteit, evenmin als tussen leerlingen uit verschillende leerjaren.

1 Inleiding

De Inspectie van het Onderwijs heeft de laatste jaren veel aandacht besteed aan de opbrengsten van het leesonderwijs op basisscholen. De aanleiding daarvoor is dat een grote groep leerlingen in het basisonderwijs problemen ondervindt met technisch en begrijpend lezen. In de brochure 'Iedereen kan leren lezen' doet de Inspectie voorstellen om

de kwaliteit van het leesonderwijs op basisscholen te verbeteren (Inspectie van het Onderwijs, 2006). De voorstellen hebben met name betrekking op het onderwijs in technisch lezen, omdat een goede beheersing van de leesteknik als een belangrijke conditie wordt beschouwd voor begrijpend en studerend lezen (zie bijvoorbeeld LaBerge & Samuels, 1974; Perfetti, 1985; Walczyk, 2000). Het belangrijkste voorstel van de Inspectie is dat scholen zwakke lezers meer instructie- en oefentijd geven. Daarmee wordt beoogd dat zo veel mogelijk zwakke lezers in de eigen groep kunnen blijven. Een kansrijke methodiek in dit verband is begeleid hardop lezen.

Begeleid hardop lezen van teksten (*guided oral reading*) is een methodiek om met name zwakke lezers beter te leren lezen. De methodiek komt er in essentie op neer dat een leerling hardop een tekst leest, terwijl een ervaren lezer de leerling aanmoedigt en controleert of deze de tekst correct leest en begrijpt. Begeleid hardop lezen wordt vaak ingezet in de fase van het voortgezet technisch leesonderwijs. Voortgezet technisch leesonderwijs is in de leesdidactiek de fase die volgt op het aanvankelijk leesonderwijs. De fase van het aanvankelijk lezen eindigt als leerlingen de elementaire leeshandeling beheersen (voor veel leerlingen is dat rond de kerst in leerjaar 3). Op dat moment kunnen leerlingen teksten op niveau AVI-M3 zelfstandig lezen, zij het nog langzaam of moeizaam spellend. Voor het vervolg van hun leesontwikkeling is het van belang dat leerlingen vloeiend leren lezen. Voor het merendeel van de leerlingen verloopt de fase van voortgezet technisch leesonderwijs zonder problemen. Door regelmatig zelfstandig lezen worden veel leerlingen steeds vaardiger in het vlot en foutloos lezen. Echter, ongeveer 10 tot 20 procent van de leerlingen ondervindt problemen in deze fase. Zulke leerlingen blijven traag lezen of veel fouten maken.

In dit artikel doen we verslag van een on-

derzoek waarin twee varianten van begeleid hardop lezen onderzocht zijn op effectiviteit bij zwakke lezers in de groepen vier, vijf en zes van het basisonderwijs. De varianten worden aangeduid als Verder Lezen (*continuous reading*) en Opnieuw Lezen (*repeated reading*). Bij Verder Lezen leest de leerling steeds nieuwe teksten, bij Opnieuw Lezen oefent de leerling een aantal malen met dezelfde tekst. Bij beide varianten oefenen leerlingen alleen met integrale teksten en niet met woordrijen. Het oefenen met integrale tekst ligt dicht bij het uiteindelijke criteriumgedrag, namelijk het zelfstandig kunnen lezen van teksten op een leeftijdsadequaat niveau. Het biedt tevens mogelijkheden om leerlingen ondersteuning te bieden op het vlak van begrijpend lezen.

2 Probleemstelling en onderzoeksvragen

Hoofddoel van het onderwijs in voortgezet technisch lezen is dat leerlingen vloeiend leren lezen. Vloeiend lezen (*reading fluency*) is een vaardigheid die drie aspecten omvat: vlot, foutloos en met een passende expressie (Kuhn, Schwanenflugel & Meisinger, 2010). Onderzoekers zijn het erover eens dat vloeiend lezen alleen kan worden verworven door veel en regelmatig oefenen. De vraag is hoe dat oefenen het beste kan gebeuren.

Een belangrijke review rond dit vraagstuk is die van het National Reading Panel (2000). In dit overzicht is een onderscheid gemaakt tussen twee typen didactieken. De eerste is aangeduid als begeleid hardop lezen van teksten (*guided oral reading*), de tweede als leesbevordering (*encouraging children to read on their own*). Een voorbeeld van de eerste methodiek is het in Nederland bekende AVI-lezen, waarbij leerlingen in groepjes van een ongeveer gelijk leesniveau aan elkaar voorlezen, begeleid door een ervaren lezer, bijvoorbeeld een leesouder. Ook het RALFI-lezen (Smits & Braams, 2006) is een voorbeeld van de methodiek begeleid hardop lezen. Uit het overzicht van het National Reading Panel komt naar voren dat begeleid hardop lezen een effectieve didactiek is, zowel voor leerlingen met een modale als

met een vertraagde leesontwikkeling. De gemiddelde effectgrootte op basis van veertien afzonderlijke onderzoeken bedroeg 0.50. Opgemerkt zij dat in Nederland naar AVI- en RALFI-lezen nooit experimenteel onderzoek is gedaan. Deze aanpakken kunnen dus niet gelden als *evidence-based* methodieken. Voor de effectiviteit van het tweede type didactiek, gebaseerd op leesbevordering, vond het National Reading Panel geen aanwijzingen.

We concentreren ons hier op het onderscheid tussen de twee eerder genoemde varianten, namelijk Verder Lezen en Opnieuw Lezen. De vraag is of de ene variant effectiever is dan de andere. Kuhn en Stahl (2003) vinden in hun review op basis van 58 onderzoeken geen duidelijk voordeel voor de ene of de andere variant. Ze stellen vast dat een effectieve interventie uit twee kernelementen moet bestaan, namelijk *practice* en *support*. Met andere woorden, leerlingen moeten oefenen in begeleide omstandigheden. Of ze bij het oefenen een tekst al of niet vaker lezen, maakt voor de leesontwikkeling weinig uit. Ook Wexler, Vaughn, Edmonds en Reutebuch (2008) maken een vergelijking tussen Verder Lezen en Opnieuw Lezen. Zij vergelijken de uitkomsten van negentien onderzoeken, in dit geval bij zwakke lezers in de leeftijd tussen 11 en 21 jaar. Op grond van de gevorderde leeftijd van deze zwakke lezers mag men aannemen dat zich onder deze groep veel leerlingen bevinden die ernstig gedemotiveerd zijn geraakt. Ook in deze groep treedt tussen Verder Lezen en Opnieuw Lezen geen systematisch verschil op; beide varianten hebben een gunstig effect op de vloeiendheid van lezen. Therrien (2004) vergelijkt de uitkomsten van achttien onderzoeken die exclusief waren gericht op de effectiviteit van Opnieuw Lezen. Deze review resulteert in een bevestiging van de effectiviteit van deze variant. Een verdere conclusie uit Therriens review is dat de effectiviteit afhankelijk is van wie tijdens het lezen de feedback geeft. Een volwassen feedbackgever leidt tot een grotere effectiviteit dan wanneer een medeleerling de feedback verzorgt.

Het is begrijpelijk dat veel onderzoek rond vloeiend lezen zich beperkt tot aspecten van technisch lezen. Toch is het relevant om

daarnaast nog te weten of Opnieuw Lezen en Verder Lezen een doorwerking hebben naar andere doelen van het leesonderwijs, met name naar de vaardigheid in begrijpend lezen en naar woordenschat. Er zijn redenen om zo'n doorwerking of transfereffect inderdaad te verwachten. Immers, naarmate leerlingen minder moeite hebben met het ontcijferen van tekst, kunnen ze meer aandacht besteden aan de betekenis ervan (LaBerge & Samuels, 1974; Perfetti, 1985). Ook van de toegenomen leessnelheid en nauwkeurigheid van lezen kan een positief effect uitgaan op de vaardigheid in begrijpend lezen en woordenschat (Klauda & Guthrie, 2008). Er zijn enkele reviews waarin aan deze kwestie aandacht is gegeven. Zo merken Kuhn en Stahl (o.c.) hierover het volgende op: "Generally, where an increase in fluency was found, there was also an increase in comprehension" (p. 9). Ook Therrien (o.c.) besteedt in zijn review aandacht aan transfereffecten naar begrijpend lezen. Hij vindt een gemiddelde effectgrootte van 0.25 voor tekstbegrip, naast een gemiddelde effectgrootte van 0.50 voor vloeiendheidsmaten. Deze schattingen zijn overigens gebaseerd op *within designs*, derhalve op onderzoeken zonder een controlegroep. Ze zijn niet zonder meer te generaliseren naar situaties waarin de effectiviteit geschat wordt op basis van een vergelijking tussen een behandelde en een onbehandelde groep.

Het experiment van O'Connor, White en Swanson (2007) biedt een relevante aanvulling op de hiervoor besproken reviews, mede omdat de uitkomsten ervan nog niet in reviews zijn verwerkt. Zij bedienden zich van een *between design* met drie groepen leerlingen: een Verder Lezengroep, een Opnieuw Lezengroep en een onbehandelde controlegroep. Hun verwachtingen omtrent de effectiviteit van de twee varianten waren verschillend. Verder Lezen zou gunstiger zijn voor de ontwikkeling van begrijpend lezen en woordenschat dan Opnieuw Lezen. Als reden daarvoor geven ze aan, dat leerlingen in de eerste conditie meer ervaring opdoen met verschillende teksten en vaker nieuwe woorden tegenkomen. Het lezen van telkens nieuwe teksten zou betere mogelijkheden bieden voor het verwerven van gevarieerde leesstra-

tegieën en nieuwe woordbetekenissen dan het herlezen van teksten. Voor het technisch lezen zouden beide varianten ongeveer even effectief zijn. De uitkomsten van dit onderzoek komen maar ten dele overeen met de verwachtingen. Inderdaad blijkt er wat betreft technisch lezen geen verschil te zijn tussen de beide varianten: leerlingen in beide condities gaan sterker vooruit dan leerlingen in de controlegroep (effectgroottes bedragen tussen de 0.84 en de 1.03). Maar ook wat betreft begrijpend lezen en woordenschat ontwijken de twee varianten elkaar weinig; beide zijn effectief (effectgroottes bedragen omstreeks 1.00 voor begrijpend lezen en 0.50 voor woordenschat). Opvallend zijn de grote effectgroottes. Ze zijn beduidend hoger dan men op grond van de hier besproken reviews zou verwachten. De auteurs geven hiervoor geen verklaring. Duidelijk is in elk geval dat de periode van leeshulp niet excessief lang is geweest. De leeshulp is gegeven gedurende een periode van 14 weken op basis van drie keer per week een kwartier (in totaal ongeveer 10 uur). Mogelijk heeft wel een rol gespeeld dat de leeshulp individueel in een setting van één-op-één is aangeboden.

Een opvallend afwezige factor in de genoemde reviews is het plezier dat leerlingen in lezen ervaren. Over het algemeen hebben zwakke lezers door opgedane negatieve ervaringen minder plezier in lezen gekregen. Ook hun inzet tijdens leesinstructies en -oefeningen kan daaronder lijden. Het is een relevante vraag of de varianten Opnieuw Lezen en Verder Lezen hierop ingrijpen. Opnieuw Lezen kan zwakke lezers het voordeel bieden dat ze tijdens het oefenen successen ervaren, en wel meer successen naarmate ze een tekst vaker geoefend hebben. Daar staat tegenover dat herhaald oefenen ook tot verveling kan leiden. Bij Verder Lezen kan sprake zijn van omgekeerde mechanismen: frequente faal-ervaringen maar ook de continue uitdaging van iedere keer weer nieuwe teksten. Welke mechanismen de overhand hebben, is door gebrek aan empirisch onderzoek niet aan te geven.

Op grond van de hier besproken overwegingen hebben we een vijftal onderzoeksvragen geformuleerd.

a. Vormen de varianten Verder Lezen en Op-

nieuw Lezen een effectieve methodiek voor leerlingen die een achterstand hebben ten aanzien van technisch lezen?

- b. In hoeverre dragen de varianten bij aan de ontwikkeling van de vaardigheid in begrijpend lezen, aan de omvang van de woordenschat en aan leesplezier?
- c. Verschillen de varianten van elkaar in effectiviteit ten aanzien van de vaardigheid in technisch en begrijpend lezen, woordenschat en leesplezier?
- d. Zijn eventuele verschillen in effectiviteit (wat betreft technisch en begrijpend lezen, woordenschat en leesplezier) afhankelijk van het leerjaar waarin de hulp wordt gegeven?
- e. Hoe is de relatie tussen de duur van de behandeling en de effectiviteit ervan voor de vaardigheid in technisch lezen?

3 Methode

Het onderzoek is opgezet als een *true experimental design* met een voormeting, twee tussenmetingen, een na- en een retentiemeting (Campbell & Stanley, 1963). Afhankelijke variabelen zijn technisch lezen (lezen van woordrijen, lezen van tekst), begrijpend lezen, woordenschat en leesplezier. In de opzet onderscheiden we twee experimentele groepen, namelijk Opnieuw Lezen (OL) en Verder Lezen (VL), naast een controlegroep.

3.1 Proefpersonen

Het onderzoek is in het schooljaar 2009-2010 uitgevoerd op scholen in de provincie Flevoland. Vanuit de schoolbesturen zijn scholen aangemeld voor deelname aan het onderzoek. Alleen scholen met tenminste zes zwakke lezers in elk van de leerjaren vier, vijf en zes kwamen voor deelname in aanmerking. Het betreft leerlingen die volgens de richtlijnen van het 'Protocol leesproblemen en Dyslexie' op grond van geringe vorderingen als risicoleerling worden aangemerkt. In totaal hadden zich twaalf scholen aangemeld. Vier scholen zijn uiteindelijk afgewezen, omdat ze niet in alle leerjaren voldoende zwakke lezers hadden, acht scholen hebben aan het onderzoek meegedaan.

Per school hebben achttien zwakke lezers

kunnen deelnemen aan het experiment, zes uit elk van de genoemde leerjaren. Een vereiste is geweest dat een leerling de elementaire leeshandeling diende te beheersen. Leerlingen die niet aan deze eis voldoen, kunnen naar verwachting nog onvoldoende profiteren van begeleid hardop lezen. Toewijzing van de leerlingen aan een van de drie condities (OL, VL, controlegroep) heeft plaatsgevonden via *blocking*. Leerlingen uit één jaargroep zijn per school triogewijs gematcht op hun leesniveau (aan de hand van hun score op de toets Technisch Lezen 345678; zie hierna) en vervolgens via toeval toegewezen aan een van de drie condities. Matching vergroot de statistische precisie en randomisatie is een essentiële bijdrage aan de interne validiteit van het experiment.

Bij aanvang van het experiment namen in totaal 143 leerlingen aan het onderzoek deel, zeven scholen met 18 leerlingen, één school met 17 leerlingen. Gedurende het experiment heeft er zich geen uitval voorgedaan. Wel hebben we van 17 leerlingen (verspreid over zes scholen) de gegevens alsnog buiten de analyses gehouden. Reden is geweest dat deze leerlingen op de voormeting een te geringe achterstand bleken te hebben. Als criterium hebben we de score op de toets Technisch Lezen 345678 gehanteerd; de gehanteerde grensscore was een leerrendement van 68 procent.² Van de resterende 126 leerlingen die in de hier te rapporteren uitkomsten zijn betrokken, zaten er 43 in de conditie OL, eveneens 43 in de conditie VL en 40 in de controlegroep.

Op een aantal variabelen is nagegaan of er tussen de drie condities op het moment van de voormeting statistisch significante verschillen bestonden (zie Tabel 1). Bij toetsing blijkt zulks niet het geval te zijn. Op de toets Technisch Lezen 345678 varieert de gemiddelde score in de drie condities tussen 10.1 en 10.6. Aangezien het hier een dle-score betreft en de leerjaren 4 tot en met 6 zijn samengenomen, kan men hieruit afleiden dat het gemiddelde leerrendement omstreeks 50 procent moet bedragen. Gemiddeld genomen hebben de deelnemende leerlingen dus een forse leesachterstand.

Tabel 1

Gemiddelden en standaarddeviaties (tussen haakjes) voor de drie groepen leerlingen

Variabelen	Opnieuw Lezen (n = 43)	Verder Lezen (n = 43)	Controlegroep (n = 40)
Proportie in leerjaar 4	.33 (.47)	.28 (.45)	.27 (.44)
Proportie in leerjaar 5	.33 (.47)	.33 (.47)	.37 (.48)
Proportie in leerjaar 6	.35 (.48)	.38 (.49)	.37 (.48)
Proportie jongens	.49 (.51)	.64 (.49)	.56 (.50)
Proportie zittenblijvers	.23 (.43)	.10 (.30)	.22 (.42)
Proportie met dyslexie in familie	.26 (.44)	.19 (.40)	.15 (.36)
Proportie voor wie Nederlands niet de eerste taal is	.12 (.32)	.19 (.40)	.07 (.26)
Leeftijd bij voormeting (aantal maanden)	105 (11)	105 (10)	105 (13)
CBVL, voormeting (aantal seconden)	38.3 (9.1)	38.0 (10.8)	36.2 (10.7)
Technisch Lezen 345678, voormeting (dle-score)	10.1 (5.7)	10.1 (5.2)	10.6 (5.3)

Noot. Op geen van de variabelen zijn de gemiddelden statistisch significant verschillend van elkaar ($\alpha = .05$)

3.2 Meetinstrumenten

Zoals al vermeld onderscheiden we vijf afhankelijke variabelen: technisch lezen (lezen van woorden), technisch lezen (lezen van teksten), begrijpend lezen, woordenschat en leesplezier. De gekozen toetsen hebben een voldoende breed schaalbereik (minimaal vanaf groep 4 tot en met groep 6). Als controlevariabele hebben we een toets voor *rapid naming* toegevoegd. Van deze variabele is bekend dat deze een matig tot sterke samenhang vertoont met de groei van de technisch leesvaardigheid (Van den Bos, Zijlstra & Lutje Spelberg, 2002).

Technisch lezen (woorden)

Afgenomen zijn twee toetsen, de toets Technisch Lezen 345678 (Aarnoutse & Kapinga, 2005) en de Drie-Minuten-Toets (versie 2009), verder aangeduid als de DMT.

Van de toets Technisch Lezen 345678 is kaart 90A afgenomen. De leerling leest zo snel en nauwkeurig mogelijk de woorden die de kaart biedt. De ruwe score is het aantal correct gelezen woorden in 90 seconden. Via de beschikbare normtabel kan de ruwe score worden omgezet in een dle-score.

Evenals de vorige toets meet de DMT hoe snel en nauwkeurig een leerling losse woor-

den leest. Afgenomen zijn alle drie de beschikbare kaarten. Per kaart bedraagt de leestijd 60 seconden. Van de kaarten bestaan drie versies, aangeduid als A, B en C, die alternerend afdgenomen zijn, te beginnen met op de voormeting de A-versie. De score is de vaardigheidsscore, bepaald op basis van de somscore over de drie kaarten.

De correlatie tussen de DMT (vaardigheidsscore) en de toets Technisch Lezen 345678 (ruwe score) bedraagt .96 (bepaald op basis van de voormeting).

Technisch lezen (tekst)

Afgenomen is de toets AVI (versie 2009). Deze versie bestaat uit elf toetskaarten, waarmee bepaald wordt hoe snel en nauwkeurig een leerling teksten van oplopende moeilijkheidsgraad kan lezen. Van de toetskaarten bestaan twee parallelle versies. Op de voor- en de retentiemeting is de A-versie afgenomen, op de nameting de B-versie. De hier gebruikte score representeert de moeilijkste kaart die een leerling op het zogenoemde instructieniveau leest. De moeilijkheidsgraad van de toetskaarten is gecodeerd van 0 (voor het startniveau) tot en met 11 (het plusniveau). De correlaties van de aldus gevormde schaal met de toets Technisch Lezen 345678 en de

DMT bedragen respectievelijk .90 en .93 (bepaald op basis van de voormeting).

Begrijpend lezen

Voor het meten van de vaardigheid in begrijpend lezen is de toets Begrijpend Lezen uit het Cito-leerlingvolgsysteem gebruikt (versie 2006-2011). Voor de verschillende leerjaren zijn afzonderlijke opgavenboekjes beschikbaar. In de voormeting zijn de volgende opgavenboekjes afgenomen: leerjaar 4 (Opgavenboekje E3 Start), leerjaar 5 (Opgavenboekje E4 Start), leerjaar 6 (Opgavenboekje M5 Start). In de nameting zijn de volgende opgavenboekjes afgenomen: leerjaar 4 (Opgavenboekje M4 Start + Vervolg 1/2), leerjaar 5 (Opgavenboekje M5 Start + Vervolg 1/2) en leerjaar 6 (Opgavenboekje M6 Start + Vervolg 1/2). Hier presenteren we de zogenoemde schaa scores, die over de leerjaren heen door Cito geëquivalenciert zijn.

Woordenschat

Voor het meten van de receptieve woordenschat is de in het Nederlands vertaalde Peabody (PPVT-III-NL) gebruikt. De toets bevat 204 items, verdeeld over 17 sets. Een set wordt altijd in zijn geheel afgenomen en op grond van het aantal gemaakte fouten wordt bepaald of een eerdere (makkelijker) of volgende (moeilijker) set moet worden afgenomen. De hier gepresenteerde scores betreffen het woordbegripquotient. Gemiddelde en standaarddeviatie in de populatie bedragen 100, respectievelijk 15.

Leesplezier

Naast de gestandaardiseerde toetsen is ook de vragenlijst Leesplezier afgenomen om vast te stellen hoe de houding van een leerling is ten opzichte van lezen. Deze specifiek voor het onderzoek samengestelde vragenlijst bevat zeven items met een driepuntschaal (1 = niet waar, 2 = weet niet, 3 = waar). De items luiden: a. Er zijn veel boeken die ik graag wil lezen; b. Ik vind lezen fijn; c. Ik ken veel leuke boeken; d. Als ik thuis ben, pak ik nooit een boek; e. Ik vind het leuk om veel boeken te hebben; f. Ik vind het leuk om over boeken te praten; g. Ik vind lezen saai. Na spiegeling van de negatieve items is de somscore gebruikt als een algemene indicator voor het

leesplezier dat leerlingen ervaren. De homogeniteit van de schaal varieert tussen .70 (op de voor- en retentiemeting) en .74 (op de nameting). De test-hertestbetrouwbaarheid, geschat als de correlatie tussen de voor- en de nameting, bedraagt .66. Omdat er van de conditie waaraan de leerlingen zijn toegewezen een effect kan uitgaan, is hier mogelijk sprake van een onderschatting.

Rapid naming

Afgenomen is de subtest Continu Benoemen Van Letters uit de test Continu Benoemen En Woorden Lezen (Van den Bos & Lutje Spelberg, 2007). Met de subtest, verder aangeduid als de CBVL, kan men vaststellen hoe snel een leerling een eenvoudige coderingstaak kan uitvoeren. Op de testkaart worden slechts vijf letters gebruikt, die in wisselende volgorde worden aangeboden. De gebruikte score is de benodigde tijd in seconden.

Afnameschema toetsen

In Tabel 2 is het afnameschema voor de toetsen weergegeven. De DMT is vaker afgenomen dan de andere toetsen, omdat deze blijkens normtabellen gevoeliger is voor groei. Aan de retentiemeting is niet meer deelgenomen door de leerlingen in de controlegroep. Reden is dat deze leerlingen vanuit onderzoeksethische overwegingen direct na de nameting alsnog leeshulp is aangeboden. De toetsen zijn als regel afgenomen door de onderwijsassistenten die ook de leeshulp hebben aangeboden. Alle toetsen zijn individueel afgenomen, behalve Begrijpend Lezen die in groepjes is afgenomen.

3.3 De condities

We onderscheiden drie condities, respectievelijk Verder Lezen, Opnieuw Lezen en de controleconditie waarin leerlingen géén leeshulp gekregen hebben. Op elk van de acht scholen is gedurende de looptijd van het experiment voor vier dagen per week een onderwijsassistent aangesteld.³ Condities en scholen zijn volledig gekruist (elke conditie is op elke school gerealiseerd). Hetzelfde geldt bijgevolg ook voor condities en onderwijsassistenten (elke onderwijsassistent heeft beide interventies aangeboden).

Tabel 2

Afnameschema toetsen

Meetmoment (maand)	Conditie	TL345678 en CBVL	DMT	AVI	Begr. lezen	PPVT III-NL	Lees- plezier
Voormeting (september)	OL, VL, CO	X	X	X	X	X	X
Tussenmeting 1, na leessessie 16 (oktober)	OL, VL, CO		X				
Tussenmeting 2, na leessessie 32 (november)	OL, VL, CO		X				
Nameting, na leessessie 48 (januari)	OL, VL, CO		X	X	X	X	X
Retentie, twee lesmaanden na nameting (maart/april)	OL, VL		X	X			X

Noot. Zie de tekst voor de gebruikte acroniemen.

Conditie Verder Lezen

Leerlingen in deze conditie oefenden gedurende twaalf aaneengesloten weken één-op-één met de onderwijsassistent. Het streven was om per week vier leessessies te realiseren, met een gemiddelde duur van 20 minuten per sessie. In verband met het hardop lezen vonden de sessies buiten het groepslokaal plaats. Leeshulpsessies werden zo georganiseerd dat leerlingen mee konden blijven doen aan het reguliere programma voor technisch lezen zoals verzorgd door de groepsleerkracht. Eventuele remedial teaching is stopgezet. De leerlingen kozen zelf hun leesboeken, onder de randvoorwaarde dat de tekstmoeilijkheid geen onoverkomelijke leesproblemen mocht opleveren. De onderwijsassistenten beschikten over een gevarieerde boekenvoorraad, waarin expliciet plaats was ingeruimd voor informatieve jeugdboeken. De leerlingen oefenden elke sessie met nieuwe tekst. De onderwijsassistenten zorgden voor afwisseling in werkvormen, waaronder simultaan lezen (ook wel koorlezen genoemd), om de beurt lezen en 'voorlezen aan de onderwijsassistent' (de onderwijsassistent leest niet mee, maar stelt zich op als luisteraar). Een gevarieerd boekenaanbod en afwisselende werkvormen dragen naar wij verwachten eraan bij, dat het oefenen voor de leerling gevarieerder en daardoor boeiender wordt. De onderwijsassistent bewaakte de *time-on-task* en bood hulp en feedback volgens de wacht-hint-prijsmethodiek (Struiksmá, 2003, p. 31). De onderwijsassistent bewaakte tevens het leesbegrip door op passende momenten met de

leerling over de inhoud van de teksten te spreken. Als vuistregel is gehanteerd dat er per leessessie vijftien minuten daadwerkelijk gelezen werd en vijf minuten over de inhoud van het gelezene kon worden gesproken. De totale duur van de leeshulp is per leerling zestien uur geweest (48 leessessies met een duur van 20 minuten).

Conditie Opnieuw Lezen

Leerlingen in deze conditie zijn op dezelfde manier begeleid als leerlingen in de conditie VL, met één belangrijk verschil: ze oefenden pas met nieuwe tekst als ze de vorige tekst voldoende vloeiend konden voorlezen. Per leessessie werd een fragment met een leestijd van omstreeks vier minuten gekozen, zodanig dat het fragment per leessessie drie tot vier keer kon worden gelezen. Het oefenen met het fragment werd in volgende leessessies voortgezet tot maximaal twaalf maal, tenzij de onderwijsassistent oordeelde dat er al eerder van voldoende beheersing sprake was.

Controleconditie

Leerlingen in deze conditie kregen géén leeshulp van de onderwijsassistent. Zij kregen ten aanzien van het voortgezet technisch lezen alleen het reguliere programma van hun groepsleerkracht aangeboden. Dat kon per school verschillend zijn. Wanneer het bijvoorbeeld gebruikelijk was om aan zwakke lezers een verlengde instructie aan te bieden, dan kon een leerling hier gewoon aan deelnemen. Ook deelname aan duolezen of tutorlezen kon gewoon doorgang vinden, evenals een eventuele computertraining, zoals het

werken met geflitste woorden. Het was namelijk geenszins de bedoeling dat leerlingen in de controleconditie didactisch verwaarloosd zouden worden. Wat daarentegen niet was toegestaan, was het aanbieden van individuele hulp buiten de groep door een ander dan de eigen leerkracht: geen hulp door leesouders, remedial teachers, stageaires en dergelijke. Hetzelfde gold voor buitenschoolse gespecialiseerde leeshulp, bijvoorbeeld van een orthopedagoog of van een instituut voor dyslexie.

Verdere procedures

De onderwijsassistenten beschikten over een gedetailleerd protocol met aanwijzingen voor de te bieden leeshulp. Onderwerpen waren onder andere: het toepassen van de wachthint-prijsmethodiek, het belang van variëren de werkvormen, leesbegrip en woordenschat, aandacht voor de prosodie, boekkeuze en motivatieproblemen. Als hulpmiddel bij het bewaken van de *time-on-task* hebben de onderwijsassistenten een wekker gekregen. Voorafgaande aan het experiment hebben ze gedurende drie dagen een training gehad. In de training is ingegaan op de achtergrond van het experiment, het afnemen van de toetsen en het nauwgezet toepassen van het leeshulp-protocol. Aanvullend hebben de onderwijsassistenten tweemaal *coaching-on-the-job* gekregen in de vorm van een observatiebezoek met nabespreking door hun trainers.

De onderwijsassistenten hebben voor elke leerling een logboek bijgehouden. Leessessies werden genummerd van 1 tot en met 48 en gemiste sessies moesten worden ingehaald. Deze procedure garandeerde dat elke leerling het aantal van 48 geplande leessessies daadwerkelijk gekregen heeft, en ook dat de tussenmetingen op de juiste momenten gebeurd zijn, namelijk na de 16^e en de 32^e leessessie. In het logboek noteerden de onderwijsassistenten gegevens over de gelezen teksten, hoe vaak een tekst gelezen is (alleen in de conditie OL) en of zich afwijkingen van het leeshulp-protocol hebben voorgedaan. Alle logboeken zijn door de onderzoekers ingenomen en gezien op afwijkingen van het leeshulp-protocol. De conclusie is dat de beide experimentele condities met een grote mate van getrouwheid zijn uitgevoerd.

Na afloop van elke leessessie hebben leerlingen in de vorm van een zelfbeoordeling aangegeven, hoe zeer ze zich bij het oefenen hebben ingezet. De antwoordschaal bevatte vijf schaalpunten (1 = ik heb helemaal mijn best niet gedaan; 2 = ik heb niet zo goed mijn best gedaan; 3 = ik weet het niet; 4 = ik heb gewoon mijn best gedaan; 5 = ik heb heel erg goed mijn best gedaan). Uit een analyse van deze inzetscores is gebleken dat de leerlingen gedurende de gehele leeshulpfase niet alleen een hoog niveau van inzet hebben gerealiseerd (het gemiddelde over alle leessessies en leerlingen bedraagt 4.65 met een standaarddeviatie van 0.60), maar ook dat het niveau van inzet is toegenomen, van gemiddeld 4.15 (standaarddeviatie 0.78) op de eerste leessessie tot gemiddeld 4.82 (standaarddeviatie 0.45) op de laatste leessessie. Overigens is gebleken dat er tussen de experimentele condities een verschil in niveau van inzet is geweest (het gemiddelde in de conditie OL bedraagt 4.60, standaarddeviatie 0.62; in de conditie VL is dit 4.66, standaarddeviatie 0.59). Het verschil is statistisch significant ($p < .01$), maar de effectgrootte uitgedrukt als Cohens d is gering ($d = 0.10$).

3.4 Gegevensanalyse

We rapporteren hier twee typen analyses. Het eerste type is voornamelijk beschrijvend van aard. Met dit type analyses beantwoorden we de vraag of er tussen de drie experimentele condities niveauverschillen optreden en hoe groot deze verschillen zijn. Daartoe berekenen we de gemiddelden per groep en de verschillen daartussen drukken we uit in een maat voor de effectgrootte, namelijk Cohens d , waarbij de standaardisatie is uitgevoerd op basis van de *gepoolde* standaarddeviatie. We presenteren steeds twee uitkomsten, één voor het contrast tussen de condities Opnieuw Lezen en Controle, de ander voor het contrast tussen de condities Verder Lezen en Controle. We berekenen bovendien van elke effectgrootte de standaardfout. Indien de effectgrootte in absolute zin minstens twee maal zo groot is als de bijbehorende standaardfout, spreken we van een statistisch significant verschil. De gehanteerde grenswaarde komt overeen met een p -waarde van iets minder dan .05.

Het tweede type analyses is gedaan op basis van longitudinale meerniveaumodellen, geformuleerd en geschat met MLwiN, versie 2.10 (Rasbash, Steele, Browne & Goldstein, 2009). Met dit type analyses beantwoorden we de vraag of er tussen de drie condities een verschil in groei of ontwikkeling optreedt. Waar in het eerste type analyses het accent ligt op verschillen per meetmoment, komt in het tweede type analyse het *repeated measures* karakter van de gegevens tot zijn recht. In beginsel kunnen we vier hiërarchische niveaus onderscheiden, van laag naar hoog: het niveau van de meetmomenten (vier meetmomenten op de DMT, twee meetmomenten op de andere toetsen), het niveau van de leerlingen ($n = 126$), het niveau van de leerjaren ($n = 3$) en het schoolniveau ($n = 8$). Omdat de variantiecomponent die betrekking heeft op het schoolniveau in geen van de geschatte modellen significant van nul blijkt te verschillen, rapporteren we hier alleen over modellen met drie niveaus (meetmoment, leerling, leerjaar). In het *fixed* deel van de modellen zijn de volgende variabelen opgenomen: meetmoment (geschaald van 0 voor de voormeting tot en met 3 voor de naming), twee *dummies* die het effect van de experimentele condities aanduiden ten opzichte van de controlegroep (resp. voor Opnieuw Lezen en Verder Lezen), twee *dummies* voor het effect van de leerjaren ten opzichte van leerjaar 5 (resp. voor leerjaar 4 en leerjaar 6) en verder nog twee *dummies* voor de interacties tussen meetmoment en conditie. Aan het *fixed* deel zijn in eerste aanleg vervolgens nog drie covariaten toegevoegd, namelijk a. het geslacht van de leerling, b. een variabele die aangeeft of er dyslexie in de familie voorkomt en c. de score op de toets Continu Benoemen Van Letters (CBVL). In de uiteindelijke modellen zijn alleen de laatste twee variabelen gehandhaafd, omdat alleen deze twee covariaten in een significant betere passing resulteren. Beide covariaten blijken stelselmatig een negatief effect te hebben op de leesontwikkeling, een bevinding die overeenstemt met bevindingen in de literatuur ter zake. In het *random* deel van de modellen is de variabele meetmoment als *random* beschouwd op leerling- en leerjaarniveau. Dit houdt in dat hellingshoeken van de regressie-

lijnen op de genoemde niveaus kunnen variëren.

In de meerniveaumodellen zijn twee hypothesen getoetst. De eerste hypothese behelst de vraag of er op het eerste meetmoment een niveauverschil is tussen de drie experimentele condities. De tweede hypothese behelst de vraag of de hellingshoek van de regressielijn, die de groeisnelheid representeert, voor de drie experimentele condities gelijk dan wel ongelijk is. Toetsing is steeds gedaan door de parameterschattingen te vergelijken met de bijbehorende standaardfouten. Indien een schatting in absolute zin minstens twee maal zo groot is als de geschatte standaardfout, beslissen we dat het bijbehorende effect statistisch significant van nul verschilt. In het voorkomende geval is er dus sprake van een verschil in startniveau, dan wel in groeisnelheid.

Een belangrijke opmerking is nog dat we in de resultatensectie geen onderscheid maken tussen de drie betrokken leerjaren 4, 5 en 6. Uit analyses is namelijk gebleken dat deze factor weliswaar een direct effect heeft op het leesniveau van de leerlingen, maar dat interactie-effecten van leerjaar met de experimentele conditie of met het meetmoment statistisch niet significant zijn. Dit geldt ook voor de driewegsinteractie tussen leerjaar, conditie en meetmoment. In de meerniveaumodellen is het leerjaar, zoals hiervoor al gemeld, overigens wel steeds als *fixed* effect opgenomen, omdat deze variabele in de meeste modellen een groot deel van de variantie tussen leerlingsscores wegneemt.

4 Resultaten

4.1 Beschrijvende analyses: verschillen per meetmoment

Tabel 3 bevat de uitkomsten van de beschrijvende analyses, zowel gemiddelden en standaarddeviaties als effectgroottes. Een negatieve effectgrootte wil zeggen dat de leerlingen in de controlegroep een hogere gemiddelde score hebben behaald. De effectgroottes variëren tussen -0.39 (het contrast VL-CO op de voormeting Leesplezier) en 0.33 (het contrast VL-CO op de naming PPVT-III-NL). Geen van de effectgroottes

Tabel 3

Behaalde vaardigheidsscores op de DMT, AVI, Begrijpend Lezen, PPVT-III-NL en Leesplezier, gemiddelden, standaarddeviaties en effectgroottes

Toets/ Meetmoment	Leeshulpgroep						Effectgrootte	
	Opnieuw Lezen (n = 43)		Verder Lezen (n = 43)		Controlegroep (n = 40)			
	Gem.	S.d.	Gem.	S.d.	Gem.	S.d.	OL-CO	VL-CO
<i>DMT</i>								
Voormeting	31.67	16.04	31.65	16.06	33.53	16.24	- 0.12	- 0.11
Tussen 1	37.40	17.28	37.40	16.06	37.50	16.71	- 0.01	- 0.01
Tussen 2	42.53	17.60	42.35	17.05	41.20	16.83	0.08	0.07
Nameting	47.19	19.03	47.47	17.71	44.40	16.82	0.16	0.17
Retentie	51.21	18.39	51.23	18.13	–	–	–	–
<i>AVI</i>								
Voormeting	2.81	2.19	2.79	2.02	3.05	2.03	- 0.11	- 0.13
Nameting	4.14	1.93	4.42	2.02	4.00	1.85	0.07	0.22
Retentie	4.98	2.11	5.16	2.32	–	–	–	–
<i>Begrijpend Lezen</i>								
Voormeting	2.73	17.75	3.74	19.44	4.09	17.80	- 0.08	- 0.02
Nameting	11.10	14.69	11.98	18.16	13.27	17.97	- 0.14	- 0.07
<i>PPVT-III-NL</i>								
Voormeting	95.65	13.82	98.69	11.15	95.82	10.89	- 0.01	0.26
Nameting	98.76	11.94	102.60	11.88	98.90	10.40	- 0.01	0.33
<i>Leesplezier</i>								
Voormeting	16.09	3.76	15.63	3.93	16.95	3.25	- 0.25	- 0.39
Nameting	16.95	3.54	16.93	3.71	16.23	3.81	0.22	0.21
Retentie	17.77	2.81	16.91	3.52	–	–	–	–

Noot. De effectgroottes zijn geschat op basis van de standaarddeviatie in de totale steekproef. Geen van de vermelde effectgroottes verschilt statistisch significant van nul ($\alpha = .05$). De retentiemeting is in de controlegroep niet gedaan.

verschilt significant van nul (de standaardfout bedraagt 0.22).

4.2 Uitkomsten van de meerniveau-analyses

Zoals in paragraaf 3.4 is aangegeven, hebben we twee typen effecten getoetst. Het eerste type zijn effecten die de verschillen op het eerste meetmoment, de voormeting, representeren. Aangezien de controlegroep steeds als referentiegroep is gedefinieerd, hebben we te maken met twee effecten, waarvan één het verschil tussen de groep OL en de controlegroep vertegenwoordigt, en de ander het verschil tussen de groep VL en de controlegroep. In de middelste kolom van Tabel 4 zien we dat de geschatte effecten op geen van de vijf afhankelijke variabelen significant van nul verschillen. Ook blijkt, indien per toets beschouwd, dat de schattingen voor OL en VL dicht bij elkaar liggen en in geen enkel geval significant van elkaar verschillen. Met andere woorden, op de voormeting is er op geen van de vijf toetsen een verschil

tussen de drie groepen waar te nemen.

Bezien we vervolgens de laatste kolom van de tabel, dan vinden we daar de schattingen voor de interactie-effecten tussen conditie en meetmoment. Deze effecten representeren de hellingshoek van de regressielijn, ofwel de snelheid van de leesontwikkeling tussen de voor- en eindmeting. Zowel op de DMT als op de AVI blijkt het tempo van de leesontwikkeling afhankelijk van de conditie te zijn. Leerlingen in OL en VL ontwikkelen zich sneller dan de leerlingen in de controlegroep. Tussen OL en VL daarentegen zijn de verschillen niet significant. Op de toets Begrijpend Lezen en op de toets voor woordenschat PPVT is de groeisnelheid in alle drie de condities min of meer gelijk. Op de vragenlijst Leesplezier zien we, ten slotte, dat de ontwikkeling in beide groepen OL en VL positiever is dan in de controlegroep. Wie de moeite neemt om nog even naar Tabel 3 terug te blikken, kan zien dat in de groepen OL en VL het leesplezier is gestegen, maar in de controlegroep is gedaald.

Tabel 4

Uitkomsten van de meerniveau-analyses, afzonderlijk voor de condities Opnieuw Lezen (OL) en Verder Lezen (VL); weergegeven worden de parameterschattingen en tussen haakjes de bijbehorende standaardfouten; de controlegroep vormt steeds de referentiegroep

Toets/Conditie	Effecten van experimentele conditie op startniveau	Effecten van experimentele conditie op groeisnelheid
DMT		
OL	- 0.06 (± 1.52)	1.43 (± 0.42)*
VL	- 1.81 (± 1.52)	1.44 (± 0.42)*
AVI		
OL	- 0.41 (± 0.32)	0.37 (± 0.18)*
VL	- 0.93 (± 0.32)	0.66 (± 0.19)*
Begrijpend Lezen		
OL	1.30 (± 3.86)	0.83 (± 2.39)
VL	2.33 (± 3.86)	1.86 (± 2.40)
PPVT-III-NL		
OL	0.09 (± 3.76)	0.44 (± 2.02)
VL	1.37 (± 3.79)	1.29 (± 2.02)
Leesplezier		
OL	- 0.87 (± 0.80)	1.59 (± 0.64)*
VL	- 1.35 (± 0.80)	2.03 (± 0.64)*

* $p \leq 0.05$

4.3 Uitkomsten op de retentiemetingen

De retentiemeting omvatte drie toetsen (DMT, AVI en Leesplezier). De meting is, zoals in par. 3.2 vermeld, alleen verricht in de groepen OL en VL. Daarom kunnen geen contrasten berekend worden met de controlegroep. We beperken ons hier tot de leesontwikkeling binnen de groepen OL en VL, de zogenoemde intra-individuele groei. Om de ontwikkeling van enige context te voorzien, bevat Tabel 5 effectgroottes voor twee periodes: vanaf de voormeting tot de eindmeting en vanaf de eindmeting tot de retentiemeting. Opgemerkt zij dat de voor- en eindmeting drie maanden uit elkaar liggen, de eind- en retentiemeting twee maanden. De standaardfout voor de in de tabel vermelde effectgroottes bedraagt 0.22. Tussen eind- en retentiemeting valt over het algemeen geen statistisch significante terugval of leerwinst te constateren. Enige uitzondering zijn de leerlingen in de OL-conditie, die in de laatste twee maanden op AVI een significante groei laten zien van bijna een halve standaarddeviatie.

4.4 Uitkomsten in termen van leerrendement

De tot nu toe besproken uitkomsten geven géén inzicht in de omvang van de leesachterstanden en in de vraag hoezeer ze vermin-

Tabel 5

Effectgroottes tussen de voor-, eind- en retentiemeting op drie toetsen (DMT, AVI, Leesplezier), afzonderlijk voor de condities Opnieuw Lezen (OL) en Verder Lezen (VL)

Toets/Conditie	Tussen voormeting en eindmeting	Tussen eindmeting en retentiemeting
DMT		
OL	0.89*	0.21
VL	0.93*	0.21
AVI		
OL	0.63*	0.45*
VL	0.80*	0.37
Leesplezier		
OL	0.24	0.25
VL	0.37	- 0.03

Noot. Effectgroottes zijn geschat op basis van de gepoolde standaarddeviaties.

* $p \leq 0.05$

derd zijn. Toch is het een relevante kwestie of de leerlingen daadwerkelijk een inhaalslag hebben gemaakt. Om op deze kwestie enig licht te werpen hebben we de vaardigheidscores op de DMT uitgedrukt in een score voor het leerrendement. Daartoe hebben we de verhouding bepaald tussen de dle-score en de didactische leeftijd van de leerling (zie noot 2). De leerrendementscore is een aanduiding van het vorderingentempo van een leerling: hoe hoger het leerrendement, des te sneller vordert een leerling. Voor een leerling met een gemiddeld vorderingentempo be-

Tabel 6

Gemiddelde leerrendement voor de drie leeshulpgroepen op de DMT, uitgedrukt als een percentage (leerrendement = $dle/dl \times 100$), op vijf meetmomenten

Leeshulpgroep	Meetmoment				
	Voormeting	Tussenmeting 1	Tussenmeting 2	Nameting	Retentiemeting
Opnieuw Lezen	42	50	57	63	65
Verder Lezen	42	50	55	61	65
Controlegroep	45	50	53	56	--

draagt het leerrendement 100 procent. Aan de hand van de leerrendementscores kan men beoordelen in welke mate de achterstand van leerlingen is verminderd. Uit de gegevens in Tabel 6 blijkt dat de drie leeshulpgroepen op de voormeting een heel laag leerrendement hebben, net boven de 40 procent. In de groepen OL en VL neemt het leerrendement toe tot ruim 60 procent op de nameting, en nog iets verder op de retentiemeting. Opvallend is dat óók in de controlegroep het leerrendement stijgt, zij het minder dan in OL en VL. Dit is een onverwacht resultaat, waarop we in de discussie nog terug zullen komen.

5 Conclusies en discussie

5.1 Conclusies

We structureren de conclusies aan de hand van de eerder beschreven onderzoeksvragen.

- Vormen de varianten VL en OL een effectieve methodiek voor leerlingen die een achterstand hebben ten aanzien van technisch lezen? Zowel op de DMT als op de AVI laten OL- en VL-leerlingen een snellere groei zien dan leerlingen in de controlegroep.
- In hoeverre dragen de varianten bij aan de ontwikkeling van de vaardigheid in begrijpend lezen, aan de omvang van de woordenschat en aan leesplezier? Er is niet aangetoond dat een of beide varianten hebben geleid tot een versnelde groei wat betreft begrijpend lezen of woordenschat. Ten aanzien van leesplezier hebben leerlingen in de OL- en VL-condities wél een positievere ontwikkeling doorgemaakt dan leerlingen in de controlegroep.
- Verschillen de varianten van elkaar in effectiviteit ten aanzien van de vaardigheid

in technisch en begrijpend lezen, woordenschat en leesplezier? Tussen OL en VL is op geen van de genoemde maten een verschil in effectiviteit aangetroffen.

- Zijn eventuele verschillen in effectiviteit (wat betreft technisch en begrijpend lezen, woordenschat en leesplezier) afhankelijk van het leerjaar waarin de hulp wordt gegeven? De factor leerjaar hangt op geen van de genoemde maten samen met het groeitempo.
- Hoe is de relatie tussen de duur van de behandeling en de effectiviteit ervan voor de vaardigheid in technisch lezen? In de periode tussen de voor- en nameting is de groei van de OL- en VL-leerlingen op de DMT nagenoeg lineair geweest. Per 16 leessessies, overeenkomend met een maand leeshulp, bedroeg de groei ruim vijf punten op de DMT-vaardigheidsschaal. Omgerekend naar dle-scores komt dit overeen met een maandelijks groei van ongeveer twee dle's. In de twee maanden tussen de nameting en de retentiemeting, waarin geen leeshulp is gegeven, bedroeg de maandelijks toename ongeveer één dle-punt.

5.2 Discussie

De vraag kan opkomen in hoeverre de uitkomsten van het experiment intern consistent zijn. Uit de beschrijvende analyses is namelijk gebleken dat er op geen van de meetmomenten en op geen van de toetsen een statistisch significant verschil is tussen de drie groepen (OL, VL en controlegroep). Daarentegen resulteren de meerniveau-analyses in de conclusie dat er op drie van de vijf toetsen sprake is van een statistisch significant verschillende groei. Als groepen zich in een verschillend tempo ontwikkelen, dan moet er toch op enig moment een waarneembaar

verschil ontstaan, kan men opwerpen. Ter verklaring wijzen wij erop dat de leerlingen zowel in OL als in VL op de voormeting enigszins achter liggen op leerlingen in de controlegroep, maar bij de nameting een voorsprong hebben ontwikkeld. Beide, de achterstand aan het begin en de voorsprong aan het eind, zijn gering en niet statistisch significant gebleken, maar laten tegelijkertijd wél zien dat er een verschil in groeitempo is.

Hoe verhouden de hier beschreven uitkomsten zich tot die in de onderzoeksliteratuur? Dat begeleid hardop lezen een effectieve methodiek is, althans ten aanzien van de technische leesvaardigheid, is al in veel ander onderzoek vastgesteld, zij het nog niet in Nederlands onderzoek. En ook de conclusie dat er tussen OL en VL geen noemenswaardige verschillen in effectiviteit zijn, ligt in lijn met eerder onderzoek. Nieuw is wel de constatering dat de gelijke effectiviteit van OL en VL voor alle drie de leerjaren geldt. Met andere woorden, wij troffen geen interactie-effect aan tussen het type leeshulp en het leerjaar. Nieuw is voorts dat beide methodieken óók effectief zijn gebleken ten aanzien van leesplezier. De onderwijsassistenten zijn er kennelijk in geslaagd – misschien door de afwisseling in werkvormen of het gevarieerde boekenaanbod, misschien ook door de toepassing van de wacht-hint-prijsmethodiek – de negatieve spiraal om te buigen. Deze uitkomst is een stimulans om in vervolgonderzoek preciezer vast te stellen welke componenten van leesplezier hier aan de orde zijn. Gaat het om de beleving van de eigen competentie, door Pintrich (2000) aangeduid als *mastery orientation* of als *performance orientation*? Of beleven leerlingen meer plezier in de leestaak zelf? Door de eenvoud van de vragenlijst Leesplezier kunnen zulke vragen nu nog niet beantwoord worden.

In de probleemstelling is naar voren gekomen dat begeleid hardop lezen vaak óók tot transfereffecten leidt ten aanzien van begrijpend lezen en woordenschat. O'Connor, White en Swanson (2007) bijvoorbeeld rapporteerden effectgroottes van omstreeks 1.00 voor begrijpend en omstreeks 0.50 voor woordenschat. Echter, noch bij begrijpend lezen noch bij woordenschat vonden wij een verschil tussen de drie condities. Hier is geen

sluitende verklaring voor te geven. Enerzijds is bekend dat het in dit geval om vaardigheden gaat die zich, zeker in termen van de beschikbare meetinstrumenten, tamelijk langzaam ontwikkelen. Het Cito adviseert, mede om deze reden, de toets Begrijpend Lezen maar eenmaal per jaar af te nemen (behalve in leerjaar 4). Maar die langzame groei in combinatie met niet zeer gevoelige onderwijskundige toetsen bestaat in andere landen waarschijnlijk ook, terwijl men daar toch middelgrote tot grote effecten aantreft. We kunnen daarom niet uitsluiten dat onze uitwerking van de methodiek op het domein van begrijpend lezen en woordenschat beduidend minder effectief is geweest dan uitwerkingen elders.

Het meervoudige karakter van onze metingen met de DMT – deze toets is vijf maal afgenomen – stelt ons in staat een schatting te maken van de dosis-effectrelatie, een relatie waaraan in de literatuur nog nauwelijks aandacht is gegeven. Om de interpretatie te vergemakkelijken stappen we over naar de dle-schaal, ook al is bekend dat deze schaal beperkingen heeft (Evers & Resing, 2007). Voor elke maand leeshulp vorderden de leerlingen gemiddeld met twee dle-punten. Men vergelijkte dit met het vorderingentempo voor de modale leerling dat één punt per maand bedraagt. Bedenk daarbij dat dit ‘dubbele tempo’ is gerealiseerd door leerlingen die tot aan de aanvang van het experiment gemiddeld nog niet eens een halve dle-punt per maand vorderden, zoals af te leiden is aan de in Tabel 6 vermelde leerrendementen. De tijdens de periode van de leeshulp gerealiseerde leerrendementen zijn in het licht hiervan indrukwekkend. De leeshulp heeft tot een verdubbeld ontwikkelingstempo geleid in vergelijking met het modale ontwikkelings-tempo, en zelfs tot een verviervoudiging van het oorspronkelijke tempo. Opmerkelijk is dat ook de controlegroep een versnelde ontwikkeling heeft laten zien van ongeveer één dle-punt per maand. Wat hiervan de achtergrond is, kunnen we niet met zekerheid aangeven. We vermoeden dat twee mechanismen een rol hebben kunnen spelen. De ene is dat groepsleraren, en mogelijk ook ouders – tegen onze instructies in – aan controleleerlingen extra leeshulp hebben geboden, mis-

schien gemotiveerd door een als oneerlijk ervaren achterstelling van deze leerlingen. Een ander mechanisme is misschien geweest dat controleleerlingen door het intensieve meetprogramma waaraan ze deelgenomen hebben, een meer effectieve taakoriëntatie hebben ontwikkeld. In algemene zin is het verleidelijk in dit verband te spreken van het zogenoemde Hawthorne effect, ware het niet dat dit begrip zo weinig zeggend is (Chiesa & Hobbs, 2008; Merrett, 2006).

Wat betekenen de onderzoeksuitkomsten voor de praktijk van het leesonderwijs op school? Voorop dient te blijven staan dat de aandacht primair moet uitgaan naar effectief groepsonderwijs. Dat betekent onder meer dat leraren ambitieuze doelen moeten stellen, goede instructies moeten geven, voldoende tijd voor instructie en oefening moeten inroosteren en in moeten spelen op verschillen tussen leerlingen (Houtveen, Smits, Koekebacker & Kuijpers, 2006; Vernooy, 2006). Wanneer het leesonderwijs in de groep kwalitatief in orde is, hoeven er nog maar weinig uitvallers te zijn. Ons onderzoek laat zien dat begeleid hardop lezen van teksten, aangeboden in de één-op-één setting op basis van vier begeleidingsmomenten per week, een effectieve vorm van remediëring is voor leerlingen die bij het technisch lezen uitvallen. De varianten OL en VL zijn beide effectief. Dat biedt de mogelijkheid een keuze te maken die overeenstemt met de voorkeur van de leerling. Ook is het mogelijk om beide varianten met elkaar af te wisselen. De leeshulp zou gegeven kunnen worden door een remedial teacher of een onderwijsassistent. Een naar omvang bescheiden training volstaat om onderwijsassistenten voor te bereiden op hun remediërende taak. Niet onbelangrijk voor de praktijk is de bevinding dat oefenen met tekst lezen óók op de DMT tot aantoonbare leerwinst leidt. Deze bevinding is in strijd met het in de praktijk wel gehoorde argument, dat zwakke lezers aparte oefeningen met losse woorden nodig hebben om betere decodeerders te worden.

De uitkomsten van het onderzoek vormen een pleidooi voor vroegtijdige interventie. Hoe langer met leeshulp gewacht wordt, des te groter wordt de achterstand en des te langer zal leeshulp geboden moeten worden om

de achterstand weg te werken. Daar komt nog bij dat achterstanden voor de leerlingen erg demotiverend kunnen zijn. Liever dus al leeshulp bieden in leerjaar 4 dan in leerjaar 5 of 6. Als we de cijfers uit ons onderzoek als uitgangspunt nemen, dan heeft een vierdegroeper gedurende in doorsnee een half jaar leeshulp nodig om weer bij te komen. Voor een vijfdegroeper geldt een termijn van een jaar, voor een zesdegroeper is de termijn nog weer langer.⁴ Deze rekensom, zo zij nog opgemerkt, is gebaseerd op de 10 tot 20 procent zwakste lezers per cohort, de populatie waaruit onze proefpersonen afkomstig zijn. De minder zwakke lezers uit deze groep hebben waarschijnlijk voldoende aan een kortere begeleidingstijd. De zwaksten, waaronder zich natuurlijk dyslectische leerlingen bevinden, vergen een langere begeleidingstijd, en heel waarschijnlijk ook meer gespecialiseerde hulp, waaronder hulp met spelling.

Noten

- 1 Het onderzoek is gefinancierd door het Ministerie van Onderwijs, Cultuur en Wetenschap in het kader van de prijsvraag Onderwijsbewijs, een programma om het evidence-based werken in het onderwijs te versterken (project ODB08079). Het is uitgevoerd door een consortium bestaande uit het Kohnstamm Instituut UvA BV, het lectoraat Maatwerk Primair van Pabo Almere, vijf schoolbesturen in de provincie Flevoland en de gemeente Almere.
- 2 Het zogenoemde leerrendement wordt bepaald als de verhouding tussen de didactische leeftijd (het aantal maanden leesonderwijs dat een leerling heeft gehad) en het didactische leeftijdsequivalent (het leesniveau dat een leerling heeft bereikt). De gehanteerde grenzen, uitgedrukt als een die-score, waren respectievelijk 7 (in leerjaar 4), 14 (in leerjaar 5) en 21 (in leerjaar 6). Bij de berekening van het leerrendement is verdisconteerd dat sommige leerlingen als gevolg van zittenblijven een extra jaar leesonderwijs hebben gehad. Een leerrendement van minder dan 100 procent betekent dat een leerling trager vordert dan leeftijdgenoten.
- 3 Eén school koos ervoor twee onderwijsassistenten aan te stellen, elk voor twee dagen.

- 4 Deze rekensom is gebaseerd op een leerrendement van 50 procent bij aanvang van de leeshulp, een leerwinst van 2 dle-punten per maand en een gewenst leerrendement van 100 procent bij de beëindiging van de leeshulp.

Literatuur

- Aarnoutse, C. A. J., & Kapinga, T. J. (2005). *Handleiding Technisch Lezen 345678*. Ridderkerk, Nederland: 678 Onderwijs Advisering.
- Campbell, D. T., & Stanley, J. C. (1963). *Experimental and quasi-experimental designs for research*. Chicago, IL: Rand McNally & Company.
- Chiesa, M., & Hobbs, S. (2008). Making sense of social research: How useful is the Hawthorne effect? *European Journal of Social Psychology, 38*, 67-74.
- Evers, A., & Resing, W. (2007). Het drijfzand van didactische leeftijdsequivalenten. *De Psycholoog, 42*, 466-472.
- Houtveen, A. A. M., Smits, A. E. H., Koekebacker, E. A., & Kuijpers, J. M. (2006). Vlot lezen in het speciaal basisonderwijs. *Tijdschrift voor Orthopedagogiek, 45*, 339-353.
- Inspectie van het Onderwijs. (2006). *Iedereen kan leren lezen*. Utrecht, Nederland: Inspectie van het Onderwijs.
- Klauda, S. L., & Guthrie, J. T. (2008). Relationships of three components of reading fluency to reading comprehension. *Journal of Educational Psychology, 100*, 310-321.
- Kuhn, M. R., Schwanenflugel, P. J., & Meisinger, E. B. (2010). Aligning theory and assessment of reading fluency: Automaticity, prosody, and definitions of fluency. *Reading Research Quarterly, 45*, 230-251.
- Kuhn, M. R., & Stahl, S. A. (2003). Fluency: A review of developmental and remedial practices. *Journal of Educational Psychology, 95*, 3-21.
- LaBerge, D., & Samuels, J. (1974). Toward a theory of automatic information processing in reading. *Cognitive Psychology, 6*, 293-323.
- Merrett, F. (2006). Reflections on the Hawthorne effect. *Educational Psychology, 26*, 143-146.
- National Reading Panel. (2000). *Teaching children to read: An evidence-based assessment of the scientific research literature on reading and its implications for reading instruction* (Reports of the subgroups). Washington, DC: National Institute for Child Health and Human Development.
- O'Connor, R. E., White, A., & Swanson, H. L. (2007). Repeated reading versus continuous reading: Influences on reading fluency and comprehension. *Exceptional Children, 74*, 31-46.
- Perfetti, C. A. (1985). *Reading ability*. New York, NY: Oxford University Press.
- Pintrich, P. R. (2000). An achievement goal theory perspective on issues in motivation terminology, theory and research. *Contemporary Educational Psychology, 25*, 92-104.
- Rasbash, J., Steele, F., Browne, W. J., & Goldstein, H. (2009). *A user's guide to MLwiN* (version 2.10). Bristol, United Kindom: University of Bristol.
- Smits, A. & Braams, T. (2006). *Dyslectische kinderen leren lezen*. Amsterdam, Nederland: Boom.
- Struiksma, A. J. C. (2003). *Lezen gaat voor*. Dissertatie. Amsterdam, Nederland: VU Uitgeverij.
- Therrien, W. J. (2004). Fluency and comprehension gains as a result of repeated reading; a meta-analysis. *Remedial and Special Education, 24*, 252-261.
- Van den Bos, K. P., & Lutje Spelberg, H. C. (2007). *Continu benoemen en woorden lezen*. Amsterdam, Nederland: Boom Test Uitgevers.
- Van den Bos, K. P., Zijlstra, B. J. H., & Lutje Spelberg, H. C. (2002). Life-span data on continuous-naming speeds of numbers, letters, colors, and pictured objects, and word-reading speed. *Scientific Studies of Reading, 6*, 25-49.
- Vernooy, K. (2006). Het LISBO- en VLOT-project: leerkracht gestuurde leesinterventieprojecten voor het special basisonderwijs. *Tijdschrift voor Orthopedagogiek, 45*, 162-169.
- Walczyk, J. J. (2000). The interplay between automatic and control processes in reading. *Reading Research Quarterly, 35*, 554-566.
- Wexler, J., Vaughn, S., Edmonds, M., & Reutebuch, C. K. (2008). A synthesis of fluency interventions for secondary struggling readers. *Reading and Writing, 21*, 317-347.

Manuscript aanvaard op: 11-11-2011

Auteurs

Henk Blok is als senior onderzoeker verbonden aan het Kohnstamm Instituut UvA BV; tevens lid van de Kenniskring Maatwerk Primair van Pabo Almere. **Ron Oostdam** is als lector Maatwerk Primair verbonden aan Pabo Almere, tevens senior onderzoeker bij het Kohnstamm Instituut UvA BV. C. Boendermaker is als docent Nederlands verbonden aan Pabo Almere; tevens lid van de Kenniskring Maatwerk Primair van Pabo Almere.

Correspondentieadres: Dr. H. Blok, Kohnstamm Instituut UvA BV, Postbus 94208, 1090 GE Amsterdam, h.blok@uva.nl

Abstract

Effects of guided oral reading interventions on reading fluency, comprehension, vocabulary, and reading attitude of poor readers in grades 2 through 4

In nearly every primary school, there are students that lag behind in reading fluency. These students read slowly or with many errors. We wanted to know if guided oral reading would be an effective remedy for low achieving readers in the lower grades. Guided oral reading implies that a student reads a text aloud while an experienced reader supervises the reading process, providing feedback or help when necessary. A total of 126 poor reading students from grades 2 to 4 participated in the study. They were randomly divided between two treatment groups and a non-treated control group. In the treatment groups, the intervention was delivered one-to-one by a paraprofessional, either in a repeated reading or in a continuous reading format, depending on how often students practiced with the same text. Each student went through 48 reading sessions of 20 minutes, with a maximum of four sessions each week. Measures included tests for fluency, reading comprehension, vocabulary, and reading attitude. The results demonstrate that guided oral reading is effective for improving fluency and reading attitude. Within and between grades the two versions (continuous or repeated reading) proved to be equally effective. Transfer effects on reading comprehension and vocabulary could not be established.