

Het effect van de Fonologische en Leerpsychologische methode bij leerlingen met dyslexie

M. A. R. Gijssel, en A. M. T. Bosman

Samenvatting

In deze studie is het effect van de **Fonologische en Leerpsychologische methode (F&L-methode®)** onderzocht bij 392 leerlingen met dyslexie. Er deden 144 leerlingen in de onderbouw (gemiddelde leeftijd 8 jaar), 190 leerlingen in de middenbouw (gemiddelde leeftijd 9,6 jaar) en 58 leerlingen in de bovenbouw (gemiddelde leeftijd 11,1 jaar) van de basisschool mee aan het onderzoek. De effecten van de behandeling zijn onderzocht voor lezen (DMT, EMT, Klepel, AVI-toets) en spellen (PI-dictee). De leerlingen lieten na gemiddeld 8 en 16 individuele behandelingen een significante verbetering zien van hun lees- en spellingprestaties, waarbij de effecten voor het lezen van teksten groter waren dan voor het lezen van losse woorden. Effecten voor het spellen waren beduidend groter dan effecten voor het lezen.

1 Inleiding

Zowel in de internationale literatuur als in de Nederlandstalige literatuur zijn vele effectstudies gerapporteerd met betrekking tot leesproblemen en dyslexie. Vanuit verschillende theoretische invalshoeken zijn interventies met betrekking tot technische leesvaardigheid opgezet en geëvalueerd. Voorbeelden zijn interventies op basis van akoestisch gemodificeerde stimuli (zie bijvoorbeeld Habib et al., 1999), interventies gebaseerd op metacognitieve strategieën, zoals plannen en aandacht (zie bijvoorbeeld Parrila, Das, Kendrick, Papadopoulos, & Kirby, 1999), hemisfeerspecifieke specialisatie (zie bijvoorbeeld Dryer, Beale & Lambert, 1999; Goldstein & Obrzut, 2001; Kappers, 1997) en het expliciteren van de alfabetische code en training van het fonologisch decoderen. Van de laatst genoemde categorie zijn de meeste studies voorhanden. Hiertoe rekenen we interventies met betrekking tot het fonologische en fone-

mische bewustzijn (bijvoorbeeld Ehri et al., 2001), interventies op basis van het alfabetisch principe (bijvoorbeeld McGuinness, McGuinness, & McGuinness, 1996), de morfologie (zie Arnbak & Elbro, 2000) en het herhaald aanbieden van woorden (bijvoorbeeld Meyer & Felton, 1999). In een meta-studie naar behandel-effecten beschrijven Goetry, Nossent en Van Hecke (2006) deze behandelingsmethoden uitvoerig.

Interventies met betrekking tot het spellen zijn veel minder frequent geëvalueerd en gerapporteerd. We weten uit onderzoek dat fonologische verwerkingsproblemen niet alleen negatieve gevolgen hebben voor het lezen van woorden, maar ook voor het spellen ervan (o.a., Bosman, 2004; Snowling, 2000). Een slecht ontwikkeld fonologisch bewustzijn heeft vaak een zwakke segmenteervaardigheid tot gevolg, een vaardigheid die noodzakelijk is voor een goede spellingvaardigheid. We weten ook uit onderzoek dat alleen veelvuldig in aanraking komen met geschreven teksten niet voldoende is om een goede spellingvaardigheid te ontwikkelen. Directe instructie lijkt een essentieel element te zijn. Behalve effecten op de leesvaardigheid, zijn ook positieve effecten gevonden van expliciete instructie in fonologisch bewustzijn op spellingvaardigheid (zie bijv. Ball & Blachman, 1991; Graham, 1999; Lundberg, Frost, Peterson, 1988).

Veel methodieken zijn gebaseerd op psycholinguïstische modellen, waarbij de klankstructuur en de opbouw van woorden verduidelijkt wordt. In Nederland werkt een aantal behandelinstellingen volgens een psycholinguïstische methode voor het lezen en spellen. In een psycholinguïstische behandeling wordt de klankstructuur van woorden expliciet aangeleerd en leerlingen wordt geleerd om bepaalde algoritmes aan te wenden om tot de juiste grafeem-foneemkoppeling (lezen) of foneem-grafeemkoppelingen (spellen) te komen. Leerlingen leren effectieve leesstrategieën en het toepassen van spellingregels.

Verscheidende onderzoeken van behandelinstellingen in Nederland hebben het effect van een psycholinguïstische behandeling aangetoond voor lezen en spellen bij kinderen en volwassenen met dyslexie (bijv. Gerretsen, Vaessen, & Ekkebus, 2003; Tijms & Hoeks, 2005; Tijms, Hoeks, Paulussen-Hoogbeem, & Smolenaars, 2003). Een nadeel van het merendeel van bovenstaande studies is dat ze doorgaans zijn uitgevoerd door een onderzoeker die verbonden is aan het betreffende behandelinstituut. Dit staat een objectieve beoordeling in de weg. Daarnaast is het aantal proefpersonen in de studies vaak relatief klein. De onderhavige studie komt aan deze tekortkomingen tegemoet: De onderzoekers zijn op geen enkele wijze verbonden aan het behandelcentrum. Bovendien is er de beschikking over een groot aantal proefpersonen.

1.1 F&L-methode

In deze studie zal het effect van de Fono-logische en Leerpsychologische methode (F&L-methode®) bij leerlingen met dyslexie worden onderzocht. De F&L-methode is een psycholinguïstische methode. In deze methode wordt de training van fonologische vaardigheden gecombineerd met een cognitieve en taakgerichte training van strategieën voor het technisch lezen en de spelling. De F&L-methode is in de periode 1987 tot 1992 bij Stichting Taalhulp ontwikkeld door Boumans. Het is een hooggestructureerd systeem met een strakke opbouw in leerstappen op basis van de structuur van de Nederlandse taal.

De methode gaat uit van de klanken van de taal. Door de verschillende soorten klanken een kleur te geven, en door het gebruik van regelsymbolen wordt de klank- en regelstructuur van het taalsysteem zichtbaar gemaakt. Hierbij worden wetmatigheden behandeld die volgens de auteur van de methode in het basisonderwijs niet of onvoldoende aan bod komen, bijvoorbeeld de klinkerverkleuringen onder invloed van de -l en -r of de klanken en woordstructuren waaraan worden van buitenlands herkomst te herkennen zijn. In de F&L-methode worden de spraakklanken in schrift gevisualiseerd met behulp van kleuren. Er zijn in totaal zes kleuren voor de verschillende klankgroepen die in de metho-

de onderscheiden worden: lange klanken (aa, ee, oo, uu), korte klanken (a, e, i, u, o), tweetecken klanken (ij/ei, au/ou, ui, eu, oe, ie), stomme klinkers (e, 'i' in '-ig', 'ij' in '-lijk'), drietekenklanken (aai, ooi, oei,...) en medeklinkers. Daarnaast worden er leenwoordklanken onderscheiden (c, y,...). De methode kent acht fonologische regels die allemaal afhankelijk zijn van klinkers. Dit zorgt ervoor dat de klanken steeds het uitgangspunt en het beslissingscriterium zijn voor het al dan niet toepassen van een regel.

1.2 Lezen

De herkenning van de klinkers en het aantal klankgroepen (lettergrepen) wordt in de methode als belangrijke ondersteuning bij het technisch lezen beschouwd. Voor het ontsleutelen van meerlettergrepige woorden is het van groot belang dat de klinker waargenomen wordt als ankerpunt van de lettergreep en dat de segmentatie van de lettergrepen tussen de klinkers correct uitgevoerd wordt. Het waarnemen van het aantal klinkers in woorden geeft tevens informatie over de woordlengte zodat er beter geanticipeerd kan worden hoe een woord gelezen moet worden. Bij de aanvang van het technisch lezen worden in de F&L-methode twee fases onderscheiden: de eerste fase is het voorcoderen van de leestekst door de klinkers te kleuren en de tweede fase is het hardop lezen van de voorgecodeerde leestekst. Later wordt de kleurcodering weggelaten: de automatische koppeling tussen de grafemen/fonemen en de kleuren activeert het fonologisch bewustzijn en ondersteunt technisch lezen. Woorden worden altijd in een zins- of tekstcontext geoefend, waarbij tevens aandacht is voor het leesbegrip. Daarnaast is er aandacht voor leesbevordering en leespromotie.

1.3 Spellin

Ook bij de spelling nemen de klinkers als ankerpunten binnen de woorden een essentiële plaats in. De eigenschappen van de klinkers bepalen de basisregels van de spelling. Voor de didactiek wordt het in de F&L-methode belangrijk geacht om de naam van de spellingregel te laten aansluiten op de groep klinkers waar de regel betrekking op heeft. De lange klinkerregel heeft dus betrekking op

lange klinkers, de korte klinkerregel heeft betrekking op korte klinkers, de stomme klinkerregel heeft betrekking op stomme klinkers, enzovoorts. Termen als *letterdief*, *open lettergreep regel*, *jantje snoeper*, *dubbelzetter* en *gesloten lettergreep regel* zijn volgens Boumans, de auteur van de methode, verwarrend omdat zij geen enkele associatie opwekken met de klanken waarop de regels betrekking hebben. Alle regels hebben een maximaal generalisatievermogen doordat zij zowel op bestaande als niet bestaande woorden kunnen worden toegepast. Leerlingen die met dit systeem werken, leren om de spelling van woorden op het moment dat het woord aangeboden wordt, te (re)construeren vanuit de klankstructuur van de taal. In de F&L-methode wordt dan ook met een zeer grote diversiteit aan echte en pseudowoorden gewerkt die ingedeeld zijn op klinker-medeklinker structuren en op de benodigde regels. De schrijfwijze van een woord wordt daarvoor losgekoppeld van de betekenis en van eerder aangeboden woorden. Woorden worden in het algemeen slechts een- of tweemaal aangeboden, zodat bepaalde woordstructuren met telkens nieuwe woorden geoefend kunnen worden. In de F&L-methode wordt zowel voor lezen als voor spelling gebruik gemaakt van taakgerichte werkvormen. Deze werkvormen zijn zo gekozen dat de leerling impliciet een aantal belangrijke fonologische deelvaardigheden oefent.

1.4 F&L-methode in het kort

Specifieke kenmerken van de methode zijn: a) lezen en spellen worden aangeboden in een geïntegreerd systeem; b) er wordt gebruik gemaakt van een kleurcodering die de taalstructuur zichtbaar maakt; c) visualisatie staat centraal bij alle werkvormen; d) de training van het fonologisch bewustzijn wordt vanaf het begin geïntegreerd in alle werkvormen voor lezen en spelling; e) de spelling- en leesregels (bijvoorbeeld het verschil in spelling en uitspraak van *ramen* en *-rammen*) worden aangeleerd door middel van reeksen van klinker-medeklinkerstructuren; f) er wordt naast het oefenen met bestaande woorden geoefend met abstracte klankreeksen en pseudowoorden, g) de ondersteuning van de klank-tekenkoppeling door middel van de kleurcodering

wordt ook doorgevoerd bij leenwoordklanken en leenwoordregels.

In de behandeling wordt gewerkt volgens het Protocol Dyslexie Diagnostiek en Behandeling (Blomert, 2006). De behandeling is individueel en vindt meestal eenmaal per week plaats en duurt een uur. Daarnaast volgt de leerling een vrij intensief thuisprogramma dat gemiddeld 20 tot 40 minuten per dag in beslag neemt. Meer informatie over de methode is te vinden in Boumans (1997), Boumans en Karman (1993) en Heijerman en Karman (2004).

Eerder onderzoek heeft de effectiviteit van de F&L-methode aangetoond bij een beperkte groep kinderen (Van Geffen, Berends, & Franssens, 2008).

2 Onderzoeksvragen

In de onderhavige studie wilden we graag de resultaten van Van Geffen et al. (2008) repliceren bij een grotere steekproef. De eerste onderzoeksvraag is: Wat is het effect van de F&L-methode op zowel de lees- en spellingvaardigheid van leerlingen in het regulier basisonderwijs? Onze verwachting is dat behandeling volgens deze methode positieve effecten heeft op zowel het lezen als het spellen. Eerdere onderzoeken naar het effect van een psycholinguïstische behandeling hebben immers positieve effecten aangetoond (bijv. Van Geffen, Berends & Franssens, 2008; Gerretsen et al., 2003; Tijms & Hoeks, 2005; Tijms et al., 2003). Tevens zal worden nagegaan of het leerrendement tijdens het eerste deel van de behandeling (na gemiddeld acht lessen: testmoment 2, gemiddeld 4,7 maanden na de pretest) even groot is als het leerrendement tijdens het tweede deel van de behandeling (tussen de 8^{ste} les, het tweede testmoment, en de 16^{de} les, het derde testmoment 2). In de meeste studies worden alleen de scores op de pretest en de posttest gerapporteerd en is een vergelijking op verschillende tijdsintervallen niet mogelijk.

De tweede onderzoeksvraag is: Is het effect van de methode anders voor leerlingen van verschillende leeftijdsgroepen? Hiertoe zullen de resultaten van de onderbouw (groepen 3 en 4), middenbouw (groepen 5 en 6) en

bovenbouw (groepen 7 en 8) met elkaar worden vergeleken. Resultaten uit de internationale literatuur zijn niet eenduidig. Wise, Ring en Olson (2000) boden kinderen van groep 4 tot en met groep 7 twee typen leesinterventie aan en concludeerden dat jongere kinderen een grotere leerwinst boekten dan oudere kinderen in het lezen van woorden. Lovett en Steinbach (1997) vonden echter geen verschil tussen kinderen in groep 4/5, 6 of 7/8. Kinderen van groep 4 tot en met groep 8 met ernstige leesproblemen kregen intensieve systematische training op woord- en subwoordniveau. Daarbij werden twee typen interventie onderscheiden: één training richtte zich met name op fonologische vaardigheden en letter-klankkoppelingen, de andere training richtte zich met name op het aanleren van strategieën om tot goede woordidentificatie te komen. Zowel bij de fonologische taken als bij het lezen van woorden en non-woorden was er geen effect van groep en geen interactie-effect tussen groep en interventie. Ook Rashotte, MacPhee en Torgesen (2001) vonden geen effect van groep. Zij onderzochten het effect van een fonologisch georiënteerd trainingsprogramma bij 116 zwakke lezers. De kinderen waren onderverdeeld in drie groepen, namelijk: groep 3/4, groep 5/6, groep 7/8. Zowel bij de fonologische taken als bij de lees- en spellingtaken werd er geen hoofdeffect van groep gevonden en geen interactie tussen groep en interventie (interventiegroep versus controlegroep).

Om het effect van de methode te onderzoeken, is gebruik gemaakt van didactische leeftijdsequivalenten (dle's) als maat voor de lees- en spellingprestaties. We zijn ons ervan bewust dat het gebruik van dle's sterk ter discussie staat. Zowel in de internationale literatuur (bijv. Bracken, 1988; Maloney & Larrivee, 2007; Reynolds, 1981) als in de nationale literatuur (bijv. Evers & Resing, 2007; Oud & Mommers, 1990) zijn verschillende bezwaren geuit tegen het gebruik van dle's. Op basis van statistische, instrumentele en inhoudelijke kritiek komen Evers en Resing (2007) tot de conclusie: "Dle is ongeschikt om uitspraken te doen betreffende de achterstand (of voorsprong) t.o.v. andere leerlingen." Het dle lijkt niet geschikt om een diagnose te stellen of uitspraken te doen over de exacte

mate van achterstand. In het huidige onderzoek gaat het echter niet om dergelijke uitspraken, maar staat de mate van vooruitgang in lees- en spellingprestaties bij een groep dyslectische leerlingen centraal. We onderzoeken de vooruitgang van elke leerling ten opzichte van zichzelf. Bovendien wilden we ervoor zorgen dat de resultaten te vergelijken zijn met die van andere Nederlandse effectstudies, zoals die van Struiksma, Scheltinga, & Van Efferen-Wiersma (2006), Van den Bungelaar & Van der Schaft (2000), Van Gefen, Berends, & Franssens (2008), Van der Leij & Rolak (2002) en Van der Leij (2006).

3 Methode

3.1 Deelnemers

De onderzoeksgegevens van een groot deel van de leerlingen uit het regulier basisonderwijs die in behandeling zijn geweest van 1995 tot en met 2007 zijn gedigitaliseerd. Leerlingen met meer dan zeven testmomenten, leerlingen uit het speciaal basisonderwijs en leerlingen met zeer onvolledige dossiers zijn niet in het onderzoek opgenomen. Dit resulteerde in een databestand van 611 leerlingen. Voor de analyses zijn vervolgens extra selectiecriteria gehanteerd: 1) alleen leerlingen met drie of meer testmomenten worden opgenomen in het onderzoek, 2) het aantal behandelingen volgens de F&L-methode tussen het eerste en het tweede testmoment en tussen het tweede en het derde testmoment is groter dan 4 en kleiner dan 11. De gegevens van 392 leerlingen voldeden aan deze criteria. De onderzoeksgroep bestond uit 254 jongens (64,8%) en 138 meisjes (35,2%). Kenmerken van de onderzoeksgroep zijn weergegeven in Tabel 1. Het aantal leerlingen in de analyses verschilt per toets, omdat niet iedereen alle lees- en spellingtoetsen aangeboden heeft gekregen op alle onderzoeksmomenten.

3.2 Materiaal

Om het effect van de behandeling te meten is gebruik gemaakt van standaard lees- en spellingtoetsen. Het leesniveau van losse woorden is vastgesteld met behulp van de Drie-Minuten-Toets (DMT; Verhoeven, 1993) en

Tabel 1

Gemiddelde (en standaardafwijkingen) van een aantal kenmerken van de onderzoeksgroep

	Onderbouw	Middenbouw	Bovenbouw	Totaal
Aantal testmomenten	3,9 (1,2)	4,3 (1,1)	3,8 (,8)	4,1 (1,1)
Aantal F&L-behandelingen				
<i>Totaal</i>	22,4 (8,9)	24,3 (8,3)	20,4 (6,1)	23,0 (8,4)
<i>t1-t2</i>	8,4 (,9)	8,3 (,9)	8,5 (1,0)	8,3 (,9)
<i>t2-t3</i>	7,6 (1,3)	7,5 (1,4)	7,5 (1,3)	7,5 (1,3)
Leeftijd (in maanden)				
<i>t1</i>	96,1 (11,1)	115,5 (8,8)	132,6 (19,9)	110,9 (17,4)
<i>t2</i>	100,9 (11,1)	120,2 (9,2)	137,4 (19,7)	115,6 (17,4)
<i>t3</i>	105,3 (11,7)	124,3 (9,5)	141,8 (19,8)	119,9 (17,6)
DL (in maanden)*				
<i>t1</i>	17,4 (5,2)	32,4 (7,4)	49,2 (5,6)	29,4 (12,5)
<i>t2</i>	21,5 (5,3)	37,2 (10,7)	53,4 (5,6)	33,8 (13,7)
<i>t3</i>	25,3 (6,0)	40,9 (11,3)	57,1 (5,8)	37,6 (14,0)

Noot. * DL is didactische leeftijd

de Een-Minuuat Test (EMT; Brus & Voeten, 1973). De leesvaardigheid van pseudoworden is vastgesteld met behulp van de Klepel (Van den Bos, Lutje Spelberg, Scheepstra, & De Vries, 1994). Het lezen van teksten is gemeten met behulp van de AVI-niveaukaarten (Visser, Van Laarhoven & Ter Beek, 1994). Het spellingniveau ten slotte, is gemeten met behulp van het PI-dictee (Geelhoed & Reitsma, 1999). Deze toetsen worden veelvuldig gebruikt in de Nederlandse onderwijspraktijk. De COTAN (Evers, Van Vliet-Mulder & Groot, 2000, 2004) heeft de betrouwbaarheid van de Een-Minuuat Test en de Drie-Minuten Toets als goed beoordeeld en de betrouwbaarheid van het PI-dictee, de Klepel en AVI als voldoende. De begripsvaliditeit van de Een-Minuuat Test en Drie-Minuten Toets is als goed beoordeeld en de begripsvaliditeit van de Klepel en AVI als voldoende. De begripsvaliditeit van het PI-dictee is onvoldoende (geen onderzoek).

handelingen al waren uitgevoerd. Dit is een bekend probleem bij studies die zijn uitgevoerd in klinische settings. Om die reden zullen de leerlingen uitsluitend met zichzelf worden vergeleken. Om de vooruitgang tijdens de behandeling te meten, wordt er na gemiddeld acht behandelingen een onderzoek uitgevoerd door de behandelaars, waarbij dezelfde lees- en spellingtoetsen als in het diagnostisch onderzoek worden afgenomen. De behandelingen worden doorgaans uitgevoerd volgens de F&L-methode. Enkele behandelmomenten ($M = 0,18$, $SD = 0,99$ tussen $t1$ en $t2$ en $M = 0,39$, $SD = 1,32$ tussen $t2$ en $t3$) hadden een ander karakter: Deze behandelmomenten waren gericht op training van de grammatica en/of het aanleren van strategieën voor het begrijpend lezen en studerend lezen. Om de vooruitgang tijdens de behandeling te bepalen, zijn de ruwe scores op de lees- en spellingtoetsen op de eerste drie testmomenten omgezet naar didactische leeftijdsequivalenten (dle) en leerrendementen (lr). Het didactische leeftijdsequivalent geeft de didactische leeftijd aan waarop de bijbehorende score gemiddeld wordt behaald. De didactische leeftijd (dl) verwijst naar het aantal maanden dat een leerling formeel lees- en spellingonderwijs heeft genoten. Vervolgens is het leerrendement op de drie testmo-

3.3 Procedure

De studie heeft een quasi-experimenteel design: er is gebruik gemaakt van een voormeting, tussenmeting en nameting bij leerlingen die in behandeling zijn geweest bij Stichting Taalhulp. Er kon geen gebruik worden gemaakt van een controlegroep, omdat de be-

menten bepaald ($lr = dle/dl$). De gemiddelde (niet-dyslectische) leerling heeft een leerrendement van 1, het niveau (dle) komt dan dus overeen met de didactische leeftijd. Daarnaast is ook het leerrendement *tussen* twee testmomenten in berekend ($lr \text{ tijdens behandeling} = \text{delta } dle/\text{delta } dl$). Bij een gemiddelde (niet-dyslectische) leerling is het leerrendement *tussen* twee testmomenten 1.

Ten slotte dient opgemerkt te worden dat het derde testmoment voor slechts een deel van de leerlingen tevens een eindmeting en daarmee de afsluiting van de behandeling betekende. Voor de leerlingen met vier, vijf, zes of zeven testmomenten was het derde testmoment een tussenmeting.

4 Resultaten

Voor alle leerlingen is op het eerste testmoment (t1, diagnostisch onderzoek), tweede testmoment (t2, eerste voortgangsonderzoek) en derde testmoment (t3, tweede voortgangsonderzoek) de gemiddelde score op de leesen spellingtoetsen berekend, evenals de didactische leeftijd, het didactisch leeftijds-equivalent en het leerrendement. Daarnaast is het leerrendement tijdens de behandeling berekend, dus tussen het eerste en het tweede testmoment, $(dle2-dle1) / (dl2-dl1)$, tussen het tweede en het derde testmoment, $(dle3-dle2) / (dl3-dl2)$ en tussen het eerste en het derde testmoment, $(dle3-dle1) / (dl3-dl1)$.

Tabel 2

Resultaten op groepsniveau

	n	t1	t2	t3	lr t1- t2	lr t2- t3	lr t1- t3	Hoofdeffect testmoment
DMT1	340							
ruwe score		50,3 (22,0)	59,9 (21,3)	64,9 (21,1)	0,88	0,62	0,77	$F(2,336) = 30,06; p < 0,001$; partiële $\eta^2 = 0,15$ $lr1 < lr2 = lr3$
dl		29,3	33,4	37,1				
dle		13,6	17,2	19,6				
lr		0,48	0,53	0,54				
DMT2	333							
ruwe score		38,3 (22,2)	47,5 (22,6)	53,7 (22,7)	0,90	0,68	0,80	$F(2,329) = 34,21; p < 0,001$; partiële $\eta^2 = 0,17$ $lr1 < lr2 < lr3$
dl		29,6	33,8	37,5				
dle		13,7	17,5	20,0				
lr		0,47	0,52	0,54				
DMT3	318							
ruwe score		28,9 (18,7)	36,7 (18,8)	43,1 (18,9)	0,79	0,81	0,80	$F(2, 314) = 54,36; p < 0,001$; partiële $\eta^2 = 0,26$ $lr1 < lr2 < lr3$
dl		30,2	34,4	38,1				
dle		13,7	17,0	20,0				
lr		0,45	0,50	0,53				
EMT	257							
ruwe score		39,5 (11,7)	44,9 (11,7)	49,4 (12,0)	0,76	0,84	0,80	$F(2,253) = 14,27; p < 0,001$; partiële $\eta^2 = 0,10$ $lr1 < lr2 < lr3$
dl		33,5	37,7	41,4				
dle		17,6	20,8	23,9				
lr		0,55	0,57	0,59				
Klepel	174							
ruwe score		31,9 (9,8)	36,6 (9,8)	41,5 (10,7)	0,84	0,95	0,91	$F(2,170) = 6,94; p = 0,001$; partiële $\eta^2 = 0,08$ $lr1 = lr2 < lr3$
dl		36,7	41,0	44,8				
dle		18,6	22,4	26,0				
lr		0,54	0,57	0,61				
AVI	123							
behniveau		4,1 (2,1)	5,5 (2,1)	6,5 (2,0)	1,18	1,08	1,13	$F(2,119) = 33,48; p < 0,001$; partiële $\eta^2 = 0,36$ $lr1 < lr2 < lr3$
dl		30,3	34,8	38,5				
dle		16,1	21,4	25,4				
lr		0,55	0,63	0,68				
PI-dictee	366							
ruwe score		45,7 (24,2)	64,7 (23,5)	77,3 (22,9)	1,51	1,49	1,50	$F(2,362) = 337,07; p < 0,001$; partiële $\eta^2 = 0,65$ $lr1 < lr2 < lr3$
dl		29,6	34,1	37,8				
dle		13,2	20,0	25,5				
lr		0,43	0,59	0,69				

Een overzicht van de resultaten is te vinden in Tabel 2. Vervolgens zijn analyses uitgevoerd, waarbij alleen het leerrendement op de verschillende meetmomenten is meegenomen en dus niet het leerrendement *tussen* twee meetmomenten.

Om de onderzoeksvragen te beantwoorden, is een 3 (bouw: onderbouw vs. middenbouw vs. bovenbouw) x 3 (moment: diagnostiek vs. eerste voortgangsonderzoek vs. tweede voortgangsonderzoek) variantie-analyse uitgevoerd met leerrendement als afhankelijke variabele, moment als binnenproefpersonenfactor en bouw als tussenproefpersonenfactor. Op deze manier is onderzocht of de leereffecten afhankelijk zijn van de bouw waarin de leerlingen zich bevinden op het moment van het diagnostisch onderzoek.

De eerste onderzoeksvraag was: Wat is het effect van de F&L-methode op zowel de lees- en spellingvaardigheid van leerlingen in het regulier basisonderwijs? Onze verwachting was dat behandeling volgens deze methode positieve effecten heeft op zowel het lezen als het spellen. De hoofdeffecten van testmoment zijn weergegeven in Tabel 2.

Uit de resultaten blijkt dat de hypothese bevestigd wordt: bij alle tests steeg het leerrendement significant over de tijd. Alleen bij de DMT1 was het leerrendement op het tweede en derde testmoment niet significant verschillend van elkaar en bij de Klepel bleek het leerrendement op het eerste en het tweede testmoment niet significant verschillend.

De tweede onderzoeksvraag luidde: Is het effect van de methode anders voor leerlingen van verschillende leeftijdsgroepen? Hiertoe werden de resultaten van de onderbouw (groepen 3 en 4), middenbouw (groepen 5 en 6) en bovenbouw (groepen 7 en 8) met elkaar vergeleken en onderzocht of er effecten van bouw optreden. Hieronder worden de resultaten per toets weergegeven.

4.1 DMT1

Er is geen significant interactie-effect gevonden van moment * bouw ($F(4, 672) = 1,22$; $p = 0,30$). Het verschil in leerrendement op de drie testmomenten is dus niet verschillend voor de onderbouw, middenbouw en bovenbouw. Het hoofdeffect van bouw was even-

min significant ($F < 1$), wat aangeeft dat het leerrendement in de onderbouw dus niet afwijkend is van dat in de middenbouw of in de bovenbouw.

4.2 DMT2

Er is geen significant interactie-effect gevonden van moment * bouw ($F < 1$). Het verschil in leerrendement op de drie testmomenten is dus niet verschillend voor de onderbouw, middenbouw en bovenbouw. Ook nu was het hoofdeffect van bouw niet significant ($F(2, 330) = 2,26$; $p = 0,11$). Het leerrendement in de onderbouw was dus statistisch niet afwijkend van dat in de middenbouw of bovenbouw.

4.3 DMT3

In dit geval was er een significant interactie-effect tussen moment en bouw ($F(4, 628) = 5,49$; $p < 0,001$; partiële $\eta^2 = 0,03$). De gemiddelde ruwe scores en leerrendementen per bouw zijn weergegeven in Tabel 3. Afzonderlijke analyses voor de onderbouw, middenbouw en bovenbouw laten een significant hoofdeffect van moment zien: onderbouw ($F(2, 99) = 31,93$; $p < 0,001$; partiële $\eta^2 = 0,39$), middenbouw ($F(2, 172) = 34,64$; $p < 0,001$; partiële $\eta^2 = 0,29$), en bovenbouw ($F(2, 41) = 13,21$; $p < 0,001$; partiële $\eta^2 = 0,39$). In de onderbouw en in de middenbouw waren alle paarsgewijze vergelijkingen significant (alle p 's $< 0,001$). In de bovenbouw is echter enkel het verschil tussen t1 ($lr = 0,47$) en t3 ($lr = 0,52$) significant (p 's $< 0,001$). Deze interactie is geïllustreerd in Figuur 1. Er is geen significant hoofdeffect van bouw ($F(2, 315) = 2,62$; $p = 0,08$; partiële $\eta^2 = 0,02$).

4.4 EMT

Er is geen significant interactie-effect gevonden van moment * bouw ($F < 1$). Het verschil in leerrendement op de drie testmomenten was dus niet verschillend voor de onderbouw, middenbouw en bovenbouw. Het hoofdeffect van bouw was wel significant ($F(2, 254) = 8,99$; $p < 0,001$; partiële $\eta^2 = 0,07$). Het leerrendement in de onderbouw ($lr = 0,66$) was significant hoger dan het leerrendement in de middenbouw ($lr = 0,56$, $p = 0,001$) en in de bovenbouw ($lr = 0,53$, beide p 's $< 0,001$). Het

leerrendement in de middenbouw en de bovenbouw verschilt niet significant van elkaar ($p = 0,78$). De gemiddelde ruwe scores en leerrendementen per bouw zijn weergegeven in Tabel 4.

Tabel 3

Resultaten op DMT3 per bouw

	Toetsmoment					
	1		2		3	
	M	SD	M	SD	M	SD
Onderbouw (n = 101)						
ruwe score	12,1	8,5	21,2	11,0	28,4	12,7
lr	0,39	0,19	0,47	0,18	0,51	0,19
Middenbouw (n = 174)						
ruwe score	33,7	15,6	41,4	16,3	47,3	16,8
lr	0,48	0,19	0,51	0,19	0,54	0,21
Bovenbouw (n = 43)						
ruwe score	49,0	17,0	54,2	17,1	60,3	16,3
lr	0,47	0,22	0,49	0,23	0,52	0,22

4.5 Klepel

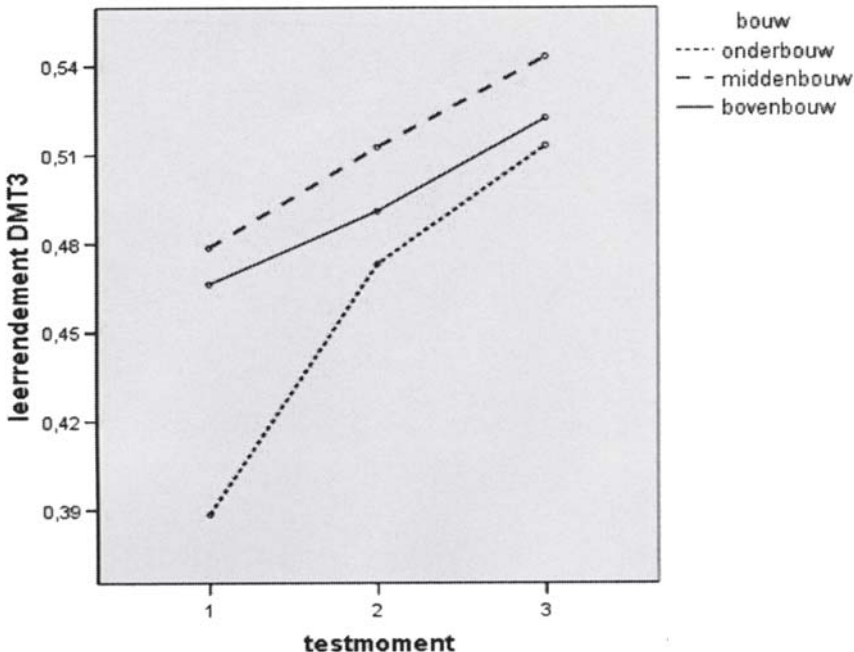
Er is geen significant interactie-effect gevonden van moment * bouw ($F < 1$). Het verschil in leerrendement op de drie testmomenten is dus niet verschillend voor de onderbouw,

Tabel 4

Resultaten op EMT per bouw

	Toetsmoment					
	1		2		3	
	M	SD	M	SD	M	SD
Onderbouw (n = 45)						
ruwe score	27,8	5,8	34,5	7,4	40,8	8,8
lr	0,63	0,15	0,67	0,17	0,69	0,22
Middenbouw (n = 161)						
ruwe score	39,3	9,8	44,8	9,8	48,8	10,5
lr	0,54	0,16	0,56	0,16	0,58	0,19
Bovenbouw (n = 51)						
ruwe score	50,5	11,2	54,4	12,3	58,7	12,7
lr	0,49	0,16	0,53	0,21	0,57	0,25

middenbouw en bovenbouw. Het hoofdeffect van bouw was significant ($F(2, 171) = 13,94$; $p < 0,001$; partiële $\eta^2 = 0,14$). Het leerrendement in de onderbouw ($lr = 0,81$) is significant hoger dan het leerrendement in de middenbouw ($lr = 0,59$, $p = 0,001$) en in de bovenbouw ($lr = 0,48$, $p < 0,001$). Het verschil in leerrendement tussen de middenbouw en bovenbouw is ook significant ($p = 0,004$). De gemiddelde ruwe scores en leerrendementen per bouw zijn weergegeven in Tabel 5.



Figuur 1. Leerrendement op de DMT3 voor de onderbouw, middenbouw en bovenbouw.

Tabel 5

Resultaten op Klepel per bouw

	Toetsmoment					
	1		2		3	
	M	SD	M	SD	M	SD
Onderbouw (n= 10)						
ruwe score	25,3	5,9	29,6	4,6	36,3	5,5
lr	0,81	0,21	0,77	0,18	0,85	0,19
Middenbouw (n= 117)						
ruwe score	30,8	8,7	36,1	9,1	40,5	10,2
lr	0,55	0,20	0,59	0,18	0,62	0,23
Bovenbouw (n= 47)						
ruwe score	36,0	11,5	39,2	11,5	45,1	11,8
lr	0,44	0,20	0,48	0,20	0,52	0,23

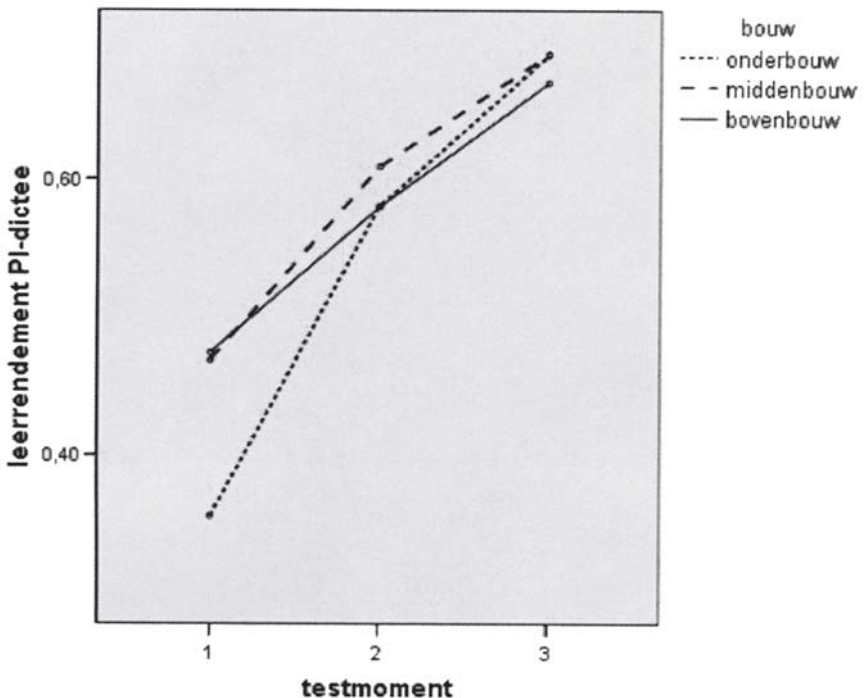
4.6 AVI

Er is geen significant interactie-effect gevonden van moment * bouw ($F < 1$). Het verschil in leerrendement op de drie testmomenten is dus niet verschillend voor de onderbouw, middenbouw en bovenbouw. Er was evenmin een significant hoofdeffect van bouw ($F(2, 120) = 2,80$; $p = 0,07$). Het leerrendement in de onderbouw is dus niet afwijkend van dat in de middenbouw of bovenbouw.

4.7 PI

Er is een significant interactie-effect gevonden van moment * bouw ($F(4, 724) = 12,38$; $p < 0,001$; partiële $\eta^2 = 0,06$). De gemiddelde ruwe scores en leerrendementen per bouw zijn weergegeven in Tabel 6. Afzonderlijke analyses voor de onderbouw, middenbouw en bovenbouw laten een significant hoofdeffect van moment zien: onderbouw ($F(2, 126) = 190,68$; $p < 0,001$; partiële $\eta^2 = 0,75$), middenbouw ($F(2, 184) = 229,06$; $p < 0,001$; partiële $\eta^2 = 0,71$), en bovenbouw ($F(2, 50) = 66,91$; $p < 0,001$; partiële $\eta^2 = 0,73$). Daarnaast zijn zowel in de onderbouw, middenbouw en bovenbouw alle paarsgewijze vergelijkingen significant ($p < 0,001$). Het verschil in leerrendement tussen het eerste en tweede moment is echter bij de onderbouw groter dan bij de middenbouw en bovenbouw. Deze interactie is geïllustreerd in Figuur 2. Er is geen significant hoofdeffect van bouw ($F(2, 363) = 2,45$; $p = 0,09$).

Samengevat kunnen we concluderen dat bij alle tests het leerrendement significant stijgt over de tijd. Alleen bij de DMT1 is het leerrendement op het tweede en derde testmoment niet significant verschillend van el-



Figuur 2. Leerrendement op het PI-dictee voor de onderbouw, middenbouw en bovenbouw.

Tabel 6

Gemiddelde scores op het PI-dictee per bouw

	Toetsmoment					
	1		2		3	
	M	SD	M	SD	M	SD
Onderbouw (n=128)						
ruwe score	25,0	14,8	45,3	15,2	58,6	16,9
lr	0,36	0,23	0,58	0,22	0,69	0,23
Middenbouw (n=186)						
ruwe score	52,5	18,2	71,1	18,7	83,6	17,7
lr	0,47	0,19	0,61	0,20	0,69	0,19
Bovenbouw (n=52)						
ruwe score	72,8	21,6	89,2	19,8	100,9	18,0
lr	0,47	0,19	0,58	0,19	0,67	0,21

kaar en bij de Klepel is het leerrendement op het eerste en het tweede testmoment niet significant verschillend. Daarnaast blijkt bij de EMT en de Klepel het leerrendement in de onderbouw hoger is dan het leerrendement in de middenbouw en bovenbouw. Bij de Klepel moet echter worden opgemerkt dat er in de onderbouw slechts 10 kinderen werden meegenomen in de analyse. De Klepel is voor de jonge kinderen nog erg moeilijk en bij een te lage score kan geen leerrendement worden vastgesteld. Bij de DMT3 en het PI-dictee laat de onderbouw een sterkere groei zien in het leerrendement tussen het eerste en het tweede testmoment dan de middenbouw en de bovenbouw.

5 Conclusie

De behandeling volgens de F&L-methode resulteerde na gemiddeld acht behandelingen (testmoment 2) en 16 behandelingen (testmoment 3) in een significante vooruitgang op het lezen van losse woorden, onzinwoorden en teksten. Alleen bij de DMT1 bleek het leerrendement op het tweede en derde testmoment niet verschillend van elkaar en bij de Klepel weken het leerrendement op het eerste en het tweede testmoment niet van elkaar af. De vooruitgang is bepaald door het leerrendement op drie verschillende tijdstippen met elkaar te vergelijken. Daarnaast is het leerrendement tijdens de behandeling berekend. Zo konden we onderzoeken of de leerlingen tijdens de behandeling hun achterstand inlo-

pen. Leerlingen kunnen immers vorderingen maken in hun lees- en spellingprestaties, maar wanneer het leerrendement tijdens de behandeling kleiner is dan 1, zullen ze hun achterstand ten opzichte van de gemiddelde leerling niet inlopen. Het leerrendement tussen het eerste testmoment en het derde testmoment lag rond de 0,80 voor het lezen van losse woorden, was 0,91 voor het lezen van onzinwoorden, 1,13 voor het lezen van teksten en 1,50 voor het spellen van woorden. Dit betekent dat voor het lezen van teksten en het spellen van woorden een inhaalslag werd gemaakt ten opzichte van de gemiddelde leerling. Voor een vergelijking van deze leerrendementen met die van andere Nederlandse studies verwijzen we naar Van der Leij (2006). Verschillen in de ernst van de leesproblemen, de duur en frequentie van de behandelingen, de keuze van de toetsen, en de maat voor vooruitgang maken structurele vergelijkingen tussen de resultaten van onze studie en de resultaten van anderen echter moeilijk. Van der Leij kwam tot de conclusie dat verschillen in effecten tussen de behandelinstututen minder lijken samen te hangen met de inhoud en methodiek dan met de selectiecriteria van de instituten. Een dergelijke vergelijking was dan ook niet het doel van onze studie.

Het leerrendement voor het lezen van teksten, gemeten met de AVI-toets, is dus groter dan het leerrendement voor het lezen van losse woorden. De berekende effectgrootte in de analyses bevestigden dit in het oog springende resultaat. Bij het lezen van teksten was het leerrendement tijdens de behandeling zelfs groter dan 1, wat betekent dat de achterstand wordt ingelopen. Dit resultaat, een groter behandelingseffect voor het lezen van teksten dan voor het lezen van losse woorden, is in overeenstemming met eerdere studies (Van Daal & Reitsma, 1999; Van der Leij & Rolak, 2002). Het grote verschil in effect tussen het lezen van losse woorden en het lezen van tekst is op verschillende manieren te verklaren. Voor het snel en correct lezen van losse woorden is automatisering van de foneem-grafeemkoppeling van groot belang. Juist die automatisering van de klank-tekenkoppelingen is het kernprobleem van leerlingen met dyslexie. Bij het lezen van een tekst spe-

len echter ook andere vaardigheden een rol, zoals woordenschat, het kunnen voorspellen van een tekst en het gebruik maken van de context. Deze vaardigheden kunnen bij leerlingen met dyslexie benut en getraind worden om het automatiseringsprobleem voor een deel te compenseren. De F&L-behandelingen hebben hier mogelijk een bijdrage aan geleverd. Tijdens de behandeling met de F&L-methode wordt altijd gebruik gemaakt van teksten en is er zowel aandacht voor remediëren als voor compenseren. Hoewel compensatievaardigheden niet expliciet worden aangeleerd, worden ze wel benoemd als een leerling hier gebruik van maakt. Daarnaast wordt er gewerkt aan het vergroten van het zelfvertrouwen van de leerlingen en is er aandacht voor leesbevordering. Deze factoren zouden ervoor kunnen zorgen dat kinderen weer vaker boeken gaan lezen en op die manier hun leesvaardigheid van teksten vergroten. Een tweede observatie is dat het behandelingseffect voor het spellen groter is dan het behandelingseffect voor het lezen van woorden en teksten. Voor het spellen van woorden werden significante verschillen in leerrendementen behaald. Het leerrendement tijdens de behandeling was 1,50. Hierbij was dus duidelijk sprake van een inhaalslag die gemaakt werd. Deze bevinding, grotere effecten voor spellen dan voor lezen, is consistent met andere studies (bijv. Van der Leij & Rolak, 2002). Een belangrijke verklaring voor dit verschil zou gevonden kunnen worden in de aard van de toetsen, omdat aan de spellingtoets (PI-dictee) geen tijdslimiet verbonden is. Leerlingen hebben dus voldoende tijd om de aangeleerde regels en strategieën toe te passen op de omzetting van fonemen naar grafemen. Voor het lezen is dat niet geval. Bij alle leestoetsen is immers de opdracht om zo snel mogelijk de woorden of tekst goed te lezen en de score wordt bepaald door een combinatie van snelheid en accuratesse. Leerlingen hebben dus weinig of geen tijd om de geleerde regels bewust toe te passen. De leestoetsen doen dus een (veel sterker) beroep op de automatisering van de klank-tekenkoppelingen dan de spellingtoets.

Het effect van bouw blijkt niet eenduidig. Bij de Een-Minut-Test en de Klepel was er een effect van bouw. Het leerrendement in de

onderbouw bleek groter te zijn dan het leerrendement in de middenbouw en bovenbouw. Bij de Klepel was er ook een verschil tussen de middenbouw en de bovenbouw, in het voordeel van de middenbouw. Hieruit blijkt dus dat de jongere leerlingen meer profiteerden van de behandeling dan de oudere leerlingen. Bij de DMT3 en het PI-dictee was er een significante interactie tussen moment en bouw: het verschil in leerrendement tussen twee opeenvolgende meetmomenten bleek bij beide toetsen groter in de onderbouw dan in de bovenbouw. Op bovengenoemde toetsen was het leerrendement voor de lagere groepen dus groter (EMT en Klepel) of steeg het leerrendement sneller in de tijd (DMT3 en PI-dictee) dan in de hogere groepen. De onderbouw profiteerde dus meer van de behandeling. Bij de DMT1, DMT2 en AVI-toets was dit niet het geval. De moeilijkheidsgraad van de toetsen zou deze verschillen kunnen verklaren. Bij relatief moeilijke toetsen, zoals de EMT, Klepel en de DMT3 blijkt het voor de oudere kinderen moeilijker om grote leerwinst te boeken in beperkte tijd. Wellicht hebben zij hun eigen plafond bereikt en hebben ze meer tijd nodig om toch nog vooruitgang te kunnen boeken.

Aparte analyses voor de onderbouw, middenbouw en bovenbouw lieten wel zien dat alle groepen vooruitgang boekten tussen het eerste en het derde testmoment. Dus, hoewel de leerlingen uit de onderbouw op enkele toetsen (EMT en Klepel) grotere effecten behaalden dan de middenbouw of bovenbouw, en hoewel op andere toetsen (DMT3, PI-dictee) de stijging van het effect tussen de toetsmomenten niet gelijk was voor alle groepen, bleek de interventie ook voor de hogere groepen effectief. Dit komt overeen met de conclusies van Lovett en Steinbach (1997) en Rashotte, MacPhee en Torgesen (2001).

We willen benadrukken dat we deze positieve effecten hebben aangetoond in een klinische setting met een retrospectieve onderzoeksopzet. Een voordeel van een retrospectieve opzet is dat het relatief goedkoop is omdat de gegevens al verzameld zijn. Relatief veel leerlingen kunnen op die manier over een langere tijd gevolgd worden. De klinische setting heeft als voordeel dat de ecologische validiteit groot is: doordat het

onderzoek is uitgevoerd in dezelfde omstandigheden als die van de dagelijkse praktijk, kunnen de bevindingen eerder worden generaliseerd. Dit komt ten goede aan de externe validiteit. De interne validiteit is echter wel in het geding. Omdat er geen sprake is van een experimenteel design met een controlegroep, is het moeilijk om verschillende validiteitsbedreigers volledig onder controle te krijgen. Hieronder bespreken we de belangrijkste validiteitsbedreigers en de consequenties daarvan voor het huidige onderzoek.

Een eerste validiteitsbedreiger is maturatie, oftewel rijping. Dat wil zeggen dat het verschil tussen de voor- en de nameting vanzelf ontstaat en niet als gevolg van de interventie. Bovendien zou het spontane rijpingsproces van de onderzoeksgroep anders kunnen zijn dan dat van een controlegroep. Doordat we in dit onderzoek gebruik hebben gemaakt van didactische leeftijdsequivalenten (dle) gekoppeld aan de didactische leeftijd (dl) van leerlingen, is er rekening gehouden met spontane rijping over de tijd: Een leerling met een dl van 20 maanden en een dle van 20 heeft een leerrendement (lr) van $20/20 = 1$. Tien maanden later heeft de leerling een dl van 30 maanden. Stel dat dan de dle is gestegen naar 30, dan is het lr $30/30 = 1$. Oftewel: het dle is dan wel met 10 gestegen, maar omdat de leerling ook 10 maanden ouder is, blijft het leerrendement constant. Om een significante vooruitgang te boeken, moet het didactische leeftijdsequivalent dus relatief sneller stijgen dan de didactische leeftijd. Het ligt niet voor de hand dat het spontane rijpingsproces bij de groep dyslectische leerlingen sneller verloopt dan bij kinderen uit een controlegroep die geen dyslexie hebben. Een criterium voor dyslexie is immers dat de vooruitgang in lees- en spellingprestaties zeer klein is, ondanks extra interventie.

Een tweede validiteitsbedreiger is het test-effect, dat wil zeggen, de metingen veroorzaken bepaalde effecten bij de leerlingen. De kans dat de leerlingen woord- of zinsmateriaal uit de test hebben onthouden en kunnen inzetten bij een volgende testmoment is echter zeer klein: Deze leerlingen maakten in het algemeen niet meer tests dan anderen. Wanneer de test die werd afgenomen bij Stichting

Taalhulp dezelfde was als die op school (bijvoorbeeld de DMT), werd vaak bij deze leerlingen de test niet meer op school afgenomen. Verder zat er minimaal drie maanden tijd tussen twee metingen en bij alle tests waar dit mogelijk was, is bij opeenvolgende testmomenten een parallelversie van de toets afgenomen, om bekendheid met de woorden of teksten zo veel mogelijk uit te sluiten. Bovendien is het aannemelijk dat dyslectische leerlingen juist door hun automatiseringsproblemen minder leren van een test dan andere leerlingen. Immers, dyslexie is een stoornis die gekenmerkt wordt door een hardnekkig probleem met het aanleren en/of vlot toepassen van het lezen en/of het spellen op woordniveau (SDN, 2008). Terwijl leerlingen zonder dyslexie woorden automatisch decoderen en opslaan in het geheugen, moeten leerlingen met dyslexie als ze een woord vaker tegenkomen, toch steeds opnieuw weer het woord decoderen; zij kunnen minder profiteren van de verbindingen tussen klanken en letters die zijn opgeslagen in het geheugen.

Een derde validiteitsbedreiger is de geschiedenis van de leerlingen, oftewel wat is er nog meer gebeurd tussen de voor- en de nameting, behalve de interventie met de F&L-methode? Omdat de behandelaars niet systematisch gegevens hadden verzameld over de vorm en mate van interventie op school tijdens de behandeling van Stichting Taalhulp, kunnen we niet uitsluiten dat bij een deel van de kinderen de vooruitgang te danken is aan interventie op school. Hoewel we het niet aannemelijk achten – het ligt niet voor de hand dat als een leerling wordt doorverwezen naar een extern behandelinstituut, school een andere en meer intensieve interventie zal inzetten bij deze leerling dan voorheen – zou dit aspect de validiteit negatief kunnen beïnvloeden.

Een vierde factor die de validiteit negatief kan beïnvloeden is statistische regressie. Dit houdt in dat groepen die zijn geselecteerd op basis van extreem lage scores bij een volgende afname minder extreem laag zullen scoren; regressie naar het midden. Recent onderzoek heeft echter aangetoond dat juist leerlingen met een relatief hogere beginscore meer vooruitgang boeken dan leerlingen met een relatief lage beginscore (Gijsel, 2009).

Gijssel toonde aan dat hoe zwakker de leerlingen waren in het lezen van losse woorden en pseudoworden bij aanvang van de behandeling, hoe kleiner de vooruitgang was op het lezen van losse woorden. Deze resultaten zijn ook in overeenstemming met resultaten gevonden in de nationale literatuur (zie voor een overzicht Van der Leij, 2006) en de internationale literatuur (bijv. Berninger et al., 2002).

Een vijfde validiteitsbedreiger is selectie. Dat betekent dat er geen sprake is van een representatieve steekproef, maar dat de steekproef al op voorhand is geselecteerd. Dit is niet aannemelijk; er zijn slechts beperkte selectiecriteria gehanteerd om de data-analyse goed te kunnen uitvoeren. Er zijn leerlingen opgenomen met uiteenlopende leeftijden, afkomstig van verschillende regio's in het land, met uiteenlopende intelligentie en met – hoe wel uiteraard laag- differentiatie in beginscores.

Een zesde aspect, dat enigszins samenhangt met de voorgaande validiteitsbedreiger, is uitval. Uitval van proefpersonen kan een vertekend beeld geven van de resultaten. Doordat er in het huidige onderzoek sprake was van een retrospectieve opzet, was er weliswaar geen uitval, maar wel een selectie op basis hiervan: leerlingen met veel ontbrekende scores zijn niet opgenomen in het onderzoek. De onvolledigheid was in de meeste gevallen te wijten aan slordigheden van de behandelaars, niet ingevulde oudervragenlijsten of kinderen die emotioneel zeer slecht reageerden op testafnames. We kunnen met redelijke zekerheid zeggen dat het hier niet om selectieve uitval gaat.

De laatste validiteitsbedreiger die we aan de orde willen stellen is de instrumentatie. De mogelijkheid dat de resultaten verstoord kunnen zijn doordat de toetsen en interventies door verschillende personen zijn uitgevoerd. Dit is niet aannemelijk, omdat zowel de toetsen als de interventies volgens een strakke procedure worden uitgevoerd. Bovendien is dit aspect in het huidige onderzoek minder van belang, omdat we de groep als geheel hebben onderzocht, en geen subgroepen hebben onderscheiden.

Uit onze studie kan dus met enige voorzichtigheid geconcludeerd worden dat behan-

deling volgens de F&L-methode voor zowel het lezen als het spellen effectief is. Het lijkt dus verantwoord om leerlingen een behandeling volgens de F&L methode aan te bieden. Bovendien verdient het aanbeveling om de behandeling zo vroeg mogelijk te starten. Dit onderzoek liet op twee tests (EMT en Klepel) een grotere leerwinst zien voor leerlingen in de onderbouw dan voor leerlingen in de middenbouw en bovenbouw. Dit vraagt om een vroegtijdige signalering en begeleiding van leesproblemen in de klas. Wij willen tot slot benadrukken dat het voorkomen van leesproblemen begint bij goed leesonderwijs in de klas, waarbij er voldoende ruimte is voor expliciete instructie en verlengde inoefening en begeleiding. De Protocollen Leesproblemen en Dyslexie (Wentink en Verhoeven, 2003, 2004) zijn een bruikbaar hulpmiddel hierbij. Voorbeelden van effectief gebleken methodieken zijn het SLIM-programma (Wentink, Wouters, & Van Hertum, 2006), gerapporteerd in Wentink, Wouters, Wennekers, Van Hertum, & Reuvekamp, 2006) en Zo leer je kinderen lezen en spellen (ZLKLS, ontwikkeld door Schraven), gerapporteerd in Bosman (2007). Wanneer er goed wordt lesgegeven in de klas, zijn er nog maar een paar kinderen (gemiddeld niet meer dan drie leerlingen per klas) die extra hulp nodig hebben. Deze kinderen zijn gebaat bij intensieve interventie van minimaal drie keer 20 minuten per week, gedurende vier tot zes maanden (zie bijv. Struiksmā, 2005; Struiksmā, Scheltinga, & Van Efferen-Wiersma, 2006). Wanneer ook die hulp niet voldoende is, is er mogelijk sprake van dyslexie en is doorverwijzing naar een extern behandelinstituut te adviseren. Behandeling volgens de F&L-methode kan bij deze kinderen de lees- en spellingprestaties verbeteren. We willen echter waarschuwen voor te hoge verwachtingen: na behandeling zal slechts een heel klein deel van de kinderen, ondanks een significante vooruitgang, op een leeftijdsadequaat lees- en spellingniveau functioneren. De vooruitgang kan echter wel betekenen dat leerlingen een *aanvaardbaar* niveau hebben bereikt waardoor ze zich beter in de klas en in de maatschappij kunnen handhaven.

Noot

Het onderzoek is financieel mogelijk gemaakt door Stichting Dyslexie Fonds, Stichting Lakeland Foundation en Stichting Iuvemus. We zijn deze stichtingen zeer erkentelijk hiervoor. Voor dit onderzoek is gebruik gemaakt van diagnostiek- en behandelingsgegevens van leerlingen die een dyslexiebehandeling hebben gekregen bij Stichting Taalhelp. Wij willen graag de kinderen, ouders en behandelaars bedanken voor het beschikbaar stellen van de gegevens.

Literatuur

- Arnbak, E., & Elbro, C. (2000). The effects of morphological awareness training on the reading and spelling skills of young dyslexics. *Scandinavian Journal of Educational Research, 44*, 229-251.
- Berninger, V. W., Abbott, R. D., Vermeulen, K., Ogiers, S., Brooksher, R., Zook, D., & Lemos, Z. (2002). Comparison of faster and slower responders to early intervention in reading: Differentiating features of their language profiles. *Learning Disability Quarterly, 25*, 59-76.
- Blomert, L. (2006). Protocol Dyslexie Diagnostiek en Behandeling. <http://www.nvo.nl/docs/Protocol%20Dyslexie%20Diagnostiek%20en%20Behandeling.pdf>.
- Ball, E. W., & Blachman, B. A. (1991). Does phoneme awareness training in kindergarten make a difference in early word recognition and developmental spelling? *Reading Research Quarterly, 26*, 49-66.
- Bos, K. P., van den, Lutje Spelberg, H. C., Scheepstra, A. J. M., & de Vries, J. (1994). *De Klepel. Vorm A en B. Een test voor de leesvaardigheid van pseudowoorden*. Nijmegen, Nederland: Berkhout.
- Bosman, A. M. T. (2004). Spellingvaardigheid en leren spellen. In A. Vyt, M.A.G. van Aken, J.D. Bosch, R.J. van der Gaag, & A.J.J.M. Ruijsse-naars (red.), *Jaarboek Ontwikkelingspsychologie, orthopedagogiek en kinderpsychiatrie 6, 2004-2005* (pp. 155-188). Houten, Nederland: Bohn Stafleu Van Loghum.
- Bosman, A. M. T. (2007). Zo leer je kinderen lezen en spellen. *Tijdschrift voor Orthopedagogiek, 46*, 451-465.
- Boumans, T. S. Y. (1997). Taal in blokjes. *Van Horen Zeggen, 37(2)*, 16-25.
- Boumans, T. S. Y. & Karman, S. (1993). Psycholinguïstische methode voor lees- en spellingproblemen. *Logopedie en Foniatrie, 4*, 113-116.
- Bracken, B. A. (1988). Ten psychometric reasons why similar tests produce dissimilar results. *Journal of School Psychology, 26*, 155-166.
- Bungelaar, H. C. I., van den & Schaft, A. J., van der. (2000). Effecten van neuropsychologische behandeling van kinderen met een ernstige dyslexie. *Tijdschrift voor Orthopedagogiek, 39*, 475-483.
- Brus, B. T., & Voeten, M. J. M. (1973). *Een-Minuu-Test*. Nijmegen, Nederland: Berkhout.
- Dryer, R., Beale, I. L., & Lambert, A. J. (1999). The balance model of dyslexia and remedial training: An evaluative study. *Journal of Learning Disabilities, 32*, 174-186.
- Daal, V. H. P. van, & Reitsma, P. (1999). Effects of outpatient treatment of dyslexia. *Journal of Learning Disabilities, 32*, 447-456.
- Ehri, L. C., Nunes, S. R., Willows, D. M., Schuster B. V., Yaghoub-Zadeh, Z., & Shanahan, T. (2001). Phonemic awareness instruction helps children learn to read: Evidence from the National Reading Panel's meta-analysis. *Reading Research Quarterly, 36*, 250-287.
- Evers, A., & Resing, W. (2007). Het drijfzand van didactische leeftijdsequivalenten. *De Psycholoog, 466*-472.
- Evers, A., Vliet-Mulder, J. C. van, & Groot, C. J. (2000). *Documentatie van tests en test-research in Nederland, deel I en II (COTAN)*. Assen, Nederland: Van Gorcum.
- Evers, A., Vliet-Mulder, J. C. van, & Groot, C. J. (2004). *Documentatie van tests en test-research in Nederland, aanvulling 2004/02 (COTAN)*. Amsterdam: Boom test uitgevers.
- Geelhoed, J. W., & Reitsma, P. (1999). *PI-dictee*. Lisse, Nederland: Swets & Zeitlinger.
- Geffen, E. C., van, Berends, M., & Franssens, J. (2008). Effectonderzoek naar de Fonologische en Leerpsychologische methode voor behandeling van dyslexie. *Tijdschrift voor Orthopedagogiek, 47*, 365-375.
- Gerretsen, P., Vaessen, A., & Ekkebus, M. (2003). Het effect van een psycholinguïstische behandeling bij kinderen en volwassenen met dyslexie. *Tijdschrift voor Remedial Teaching, 2*, 4-11.
- Gijsel, M. A.R. (2009). Lees- en spellingprestaties

- met de F&L-methode en de voorspellende factoren voor succes. *Tijdschrift voor Orthopedagogiek*, 48, 307-320.
- Goetry, V., Nossent, Ch., & Van Hecke, P. (2006). Taalontwikkelings- en Leerstoornissen. *Een literatuuronderzoek in opdracht van het Rijksinstituut voor ziekte- en Invaliditeitsverzekering*. www.riziv.fgov.be/care/nl/revalidatie/studies/study_language_learning/index.htm.
- Goldstein, B. H., Obrzut, J. E. (2001). Neuropsychological treatment of dyslexia in the classroom setting. *Journal of Learning Disabilities*, 34, 276-285.
- Graham, S. (1999). Handwriting and spelling instruction for students with learning disabilities: A review. *Learning Disability Quarterly*, 22, 78-98.
- Habib, M., Essesper, R., Rey, V., Giraud, K., Bruas, P., & Gres, C. (1999). Training dyslexics with acoustically modified speech: Evidence of improved phonological performance. *Brain and Cognition*, 40, 143-146.
- Heijerman, M. & Karman, S. (2004). Taalhelp bij dyslexie. *Logopedie en Foniatrie*, 6, 602-607.
- Kappers, E. J. (1997). Outpatient treatment of dyslexia through the stimulation of the cerebral hemispheres. *Journal of Learning Disabilities*, 30, 100-125.
- Lundberg, I., Frost, J., & Peterson, O. (1988). Effects of an extensive program for stimulating phonological awareness in preschool children. *Reading Research Quarterly*, 23, 1263-1284.
- Maloney, E. S., & Larrivee, L. S. (2007). Limitations of age-equivalent scores in reporting the results of norm-referenced tests. *Contemporary Issues in Communication Science and Disorders*, 34, 86-93.
- McGuinness, C., McGuinness, D., & McGuinness, G. (1996). Phonographix™: A new method for remediating reading difficulties. *Annals of Dyslexia*, 46, 73-96.
- Meyer, M. S., & Felton, R. H. (1999). Repeated reading to enhance fluency: Old approaches and new directions. *Annals of Dyslexia*, 49, 283-306.
- Leij, A. van der. (2006). Dyslexie: vergelijking van behandelingsstudies. *Tijdschrift voor Orthopedagogiek*, 45, 313-338.
- Leij, A. van der, & Rolak, M. (2002). Behandeling van dyslexie in een klinische setting. *Tijdschrift voor Orthopedagogiek*, 41, 181-195.
- Lovett, M. W., & Steinbach, K. A. (1997). The effectiveness of remedial programs for reading disabled children of different ages: Does the benefit decrease for older children? *Learning Disability Quarterly*, Vol. 20, No. 3, Second Special Issue on Intervention. (Summer, 1997), pp. 189-210.
- Oud, J. H. L., & Mommers, M. J. C. (1990). De valkuil van het didactisch leeftijdsequivalent. *Tijdschrift voor Orthopedagogiek*, 29, 445-459.
- Parrila, R. K., Das, J. P., Kendrick, M. E., Papadopoulos, T. C., & Kirby, J. R. (1999). Efficacy of a cognitive reading remediation program for at-risk children in Grade 1. *Developmental Disabilities Bulletin*, 27, 1-36.
- Rashotte, C. A., MacPhee, K., & Torgesen, J. K. (2001). The effectiveness of a group reading instruction program with poor readers in multiple grades. *Learning Disabilities Quarterly*, 24, 119-134.
- Reynolds, C. R. (1981). The fallacy of "two years below grade level for age" as a diagnostic criterion for reading disorders. *Journal of School Psychology*, 19, 350-358.
- SDN. (2008). *Dyslexie. Diagnose en behandeling van dyslexie. Brochure van de Stichting Dyslexie Nederland*. Geheel herziene versie, 2008. <http://www.stichtingdyslexienederland.nl>.
- Snowling, M. J. (2000). *Dyslexie*. Malden, MA: Blackwell Publishing.
- Struiksmas, C. (2005). *Organisatorisch continuüm voor de zorgroute van leerlingen met leesproblemen en dyslexie*. Notitie rond 'Tussenvormen' in het kader van het Masterplan Dyslexie. www.masterplandyslexie.nl.
- Struiksmas, A. J. C., Scheltinga F, & Efferen-Wiermas, E. van. (2006). De Rotterdamse aanpak Dyslexie, evaluatie van een project. *Tijdschrift voor Orthopedagogiek*, 45, 170-180.
- Tijms, J., & Hoeks, J. (2005). A computerized treatment of dyslexia: Benefits from treating lexico-phonological processing problems. *Dyslexia*, 11, 22-40.
- Tijms, J., Hoeks, J. J. W. M., Paulussen-Hoogboom, M. C., & Smolenaars, A. J. (2003). Long-term effects of a psycholinguistic treatment for dyslexia. *Journal of Research in Reading*, 26, 121-140.
- Verhoeven, L. (1993). *Drie-Minuten-Toets*. Arnhem, Nederland: Cito.
- Visser, J., Laarhoven, A., van, & Beek, A. ter. (1994). *AVI-toetspakket*. 's-Hertogenbosch, Nederland: Katholiek Pedagogisch Centrum.

Wentink, H., & Verhoeven, L. (2003). *Protocol leesproblemen en dyslexie* (herziene vierde druk). Nijmegen, Nederland: Expertisecentrum Nederlands.

Wentink, H., & Verhoeven, L. (2004). *Protocol leesproblemen en dyslexie voor groep 5-8*. Nijmegen, Nederland: Expertisecentrum Nederlands.

Wentink, H., Wouters, E., & Hertum, A. van. (2006). *Stimuleringsprogramma Lezen uitgaande van Instructie en Motivatie. Het SLIM-programma*. Nijmegen, Nederland: Expertisecentrum Nederlands.

Wentink, H., Wouters, E., Wennekers, M.C., Hertum, A. van, & Reuvekamp, E. (2006). Hoge verwachtingen leiden tot betere leesprestaties in het Speciaal Basisonderwijs. *Tijdschrift voor Orthopedagogiek*, 45, 182-191.

Wise, B. W., Ring, J., & Olson, R. K. (2000). Individual differences in gains from computer-assisted remedial teaching. *Journal of Experimental Child Psychology*, 77, 197-235.

Manuscript aanvaard: 27 oktober 2009

Auteurs

Dr. Martine A.R. Gijssel is werkzaam als senior projectmedewerker bij het Expertisecentrum Nederlands, verbonden aan de Radboud Universiteit Nijmegen. **Prof. dr. Anna M.T. Bosman**, is werkzaam als hoogleraar bij de vakgroep Orthopedagogiek Leren en Ontwikkeling aan de Radboud Universiteit Nijmegen.

Correspondentieadres: Martine Gijssel, Expertisecentrum Nederlands, Postbus 6610, 6503 GC Nijmegen. E-mail: m.gijssel@taalonderwijs.nl

Abstract

Effects of a psycholinguistic intervention program for children with dyslexia

In this study, the effects of a psycholinguistic intervention program for reading and spelling were investigated in 392 children with dyslexia. About 144 children in Grades 1 and 2 (mean age 8.0), 190 children in Grades 3 and 4 (mean age 9.6) and 58 children in Grades 5 and 6 (mean age 11.1) participated. Standardized tests measured reading and spelling performance at intake, after 8 and after 16 interventions. Children were trained individually during one hour a week and they received a home training program. Results showed a significant improvement of their reading and spelling scores after 8 and after 16 interventions. In reading, effects were larger for reading text than for reading isolated words. Moreover, training effects were larger for spelling than for reading.