

De effecten van traditioneel en functioneel leesonderwijs in de latere jaren van het basisonderwijs

1. Inleiding

Recentelijk lijkt de oude leesstrijd weer te zijn aangewakkerd. Neuvel, Otter en Bos (1988) presenteren een vergelijkend onderzoek over leesprestaties van leerlingen die met een traditionele methode (voornamelijk Veilig leren lezen), versus een meer "functionele" methode (veelal op een Freinet-achtige wijze) hebben leren lezen. In het onderzoek wordt zowel begrijpend als technisch lezen getoetst. De voornaamste conclusie is, dat het na drie jaar onderwijs in feite niet meer uitmaakt via welke methodiek de kinderen hebben leren lezen. Deze resultaten werden onder andere door Mommers (1989) en Van der Geest (1990) becommentarieerd. De eerste meent onder andere op grond van fouten in het onderzoeksdesign en de slechte ratio tussen onderwijsinspanning en opbrengst, dat de traditionele methoden in feite gunstiger uit de bus kwamen dan de vernieuwendere. De tweede beticht de eerste van onzorgvuldige bespreking van het onderzoek en betoogt bovendien dat door de onderzoekstekortkomingen die er overigens - op zich begrijpelijk - zijn, de resultaten juist eerder ten voordele dan ten nadele van de functionele leesaanpak geïnterpreteerd kunnen worden.

Een opvallend aspect bleef bij de drie tot dusver genoemde auteurs wat onderbelicht. Hoewel de leerlingen van de functionele leesaanpak aanvankelijk - d.w.z. na plusminus negen maanden leesonderwijs - achterbleven bij de leerlingen van de traditionele leesaanpak, konden er bijna drie jaar later geen verschillen meer aangetoond worden. Als men de data longitudinaal zou interpreteren en op grond daarvan zou mogen extrapoleren naar de bovenbouw van de basisschool - beide zaken zijn niet zo maar geoorloofd (Van der Geest, 1990) -, dan zou men kunnen verwachten dat de functionele aanpak een aanvankelijke achterstand zal omzetten in een voorsprong (wet van de remmende voorsprong en/of versnellende achterstand). Overigens geldt natuurlijk, dat vooruitgang in het lezen in groep 6,7 en 8 niet alleen verklaard mag worden uit inspanningen in de groepen 3, 4, en 5, maar mede afhangt van de manier waarop het 'verder lezen' op school begeleid wordt. Deze hier aangeduide problematiek staat centraal in het onderzoek waarover we hier berichten. Om bij een hogere leeftijdsgroep een dergelijk vergelijkend onderzoek te kunnen verrichten, zijn we in een oriënterend onderzoek nagegaan in hoeverre het onderzoeksdesign, met name het onderzoeksinstrumentarium, bijgesteld moet worden. Vandaar dat dit probleem in een aparte paragraaf (2) wordt behandeld. Na de problematiek in het algemeen gesteld te hebben, worden in de subparagrafen 2.1 en 2.2 de toetsen voor technisch en begrijpend lezen uiteengezet. In 2.3 wordt de volgorde van afname en de tijdsduur daarvan verantwoord.

Een andere belangrijke voorwaarde voor toetsontwikkelingsonderzoek is de definitie en het bereik van de proefpersonen in termen van achtergrondvariabelen. Vanwege het belang van deze problematiek wordt hier buiten de onderzoeksopzet in paragraaf 4 om, in paragraaf 3 aandacht besteed.

De onderzoeksopstelling en vraagstelling worden in paragraaf 4 besproken. De eerste en belangrijkste vraagstelling betreft de geschiktheid van de ontwikkelde tests. Een essentiële eis hiervoor is dat deze tests differentiëren tussen de verschillende jaargroepen en dus leeftijden van de leerlingen. Dit te achterhalen is de hoofddoelstelling van dit onderzoek. Ook zijn we nagegaan of er aanwijzingen zijn dat een verschil in onderwijsmethode verschillen in leesprestaties met zich meebrengt. Het gaat hier slechts om indicaties, omdat voor een dergelijke vraag een veel grootschaliger onderzoek noodzakelijk is.

In paragraaf 5 komen de resultaten van het onderzoek aan de orde. Het blijkt dat boven aangeduide tests daadwerkelijk tot die verlangde differentiatie in staat zijn. Verder blijkt dat er globaal genomen geen verband bestaat tussen onderwijsaanpak en de geleverde testprestaties. Het onderzoek zou op dat moment dan ook afgesloten kunnen worden, ware het niet dat we aanvullende gegevens hebben die niet tot het eigenlijke onderzoek behoren. Wanneer we die in de discussie meenemen, komen we tot een - door ons althans - niet verwachte conclusie in paragraaf 6.

2. Probleemstelling: het toetsinstrumentarium voor begrijpend en technisch lezen

Een gevaar dat men in veel evaluatieonderzoek loopt, is dat de keuze van het te gebruiken toetsinstrumentarium ten voordele van de onderzoeks- dan wel van de controlegroep uitpakt. Meestal - en daar is overigens nauwelijks aan te ontkomen - wordt het traditionele onderwijs op deze manier bevoordeeld en het vernieuwende onderwijs benadeeld. Ten behoeve van het traditionele onderwijs zijn namelijk door de jaren heen veel toetsen ontwikkeld die dicht bij dat onderwijs staan en die qua betrouwbaarheid en dergelijke behoorlijk zijn doorgemeten. Voor het vernieuwende onderwijs is dat niet of in veel mindere mate het geval. Vaak is vernieuwend onderwijs zelfs expliciet tegen toetsen en prestatievergelijking. Het vergelijkend onderzoek van Fendt en zijn groep (Lukesch, 1979, zie Van der Geest, 1982) ten aanzien van de Duitse Gesamtschulen is daar een schoolvoorbeeld van. In de vernieuwingsscholen stond het communicatieve aspect in het vreemde-talenonderwijs centraal. Anders dan in veel Nederlandse scholen t.a.v. communicatief onderwijs nog gebruikelijk is (Anceaux, 1990), werd niet veel aandacht geschonken aan grammatica en andere traditionele deelvaardigheden. In het vergelijkend onderzoek werden echter standaardtoetsen uit de zgn. "Regelschulen" gebruikt, zoals die voor grammatica.

Voor ons probleem, het leesonderzoek, hebben we een vergelijkbaar geval. Neuvel, e.a. (1988) vergelijken het lezen op traditionele scholen en vernieu-

wingsscholen. Hiervoor hanteren ze toetsen uit de traditionele didactiek. De controle-groep (het traditionele onderwijs) is zeer vertrouwd met dit soort toetsen. Voor de vernieuwingsscholen, althans het Freinet-achtige deel ervan, geldt echter dat men zich tegen het gebruik van toetsen verzet. Hier wordt een goede leerkracht als het beste leerlingvolgsysteem beschouwd.

In de onderzoeksopzet ondervinden de vernieuwingsscholen op deze manier een dubbel nadeel: ten eerste zijn de tests niet afgestemd op hun onderwijsinhouden, -aanpak en -doelen en ten tweede beschikken de leerlingen in beduidend mindere mate over de noodzakelijke toetservaring. Dit probleem, waaraan Neuvel e.a. en Mommers ons inziens met een wat al te groot gemak voorbij gaan, zal onze grote aandacht moeten hebben. Helemaal te omzeilen zijn dergelijke storende variabelen echter niet: in vergelijkend onderzoek zullen we vooral te maken blijven krijgen met een verschil in toetservaring c.q. systematische controle van de leerlingprestaties.

Gebruikmakend van bovenstaande inzichten zullen we respectievelijk in 2.1 en 2.2 ingaan op de toetsen voor technisch lezen (TeL) en begrijpend lezen (BeL). Deze twee aspecten van lezen hebben onze aandacht om de volgende redenen: ten eerste willen we in ons onderzoek aansluiten bij dat van Neuvel, die met name deze twee aspecten onderzocht. Ten tweede zijn er ondanks studies als die van Aarmoutse, e.a. (1986; 1988) aanwijzingen dat het technisch lezen als vaardigheid ook in de bovenbouw van de basisschool nog volop in ontwikkeling is.

2.1. Technisch lezen (TeL)

In het traditionele leesonderwijs wordt voor technisch lezen veelvuldig gebruik gemaakt van Brus- en Avi-toetsen. In het vernieuwingsonderwijs worden dit soort toetsen daarentegen nauwelijks gebruikt. Men is niet geneigd lezen op te splitsen in deelvaardigheden en men wil bovendien vooral in de onderbouw prestatietoetsen uit de weg gaan. Het gebruik van deze toetsen in vergelijkend onderzoek kan de resultaten dan ook danig beïnvloeden. Een ander nadeel van deze toetsen is de moeilijkheidsgraad: ze zijn slechts tot en met groep 6 goed bruikbaar. Voor ons onderzoek komen Brus- of Avi-achtige toetsen dan ook niet in aanmerking.

Overigens moet men zich afvragen wat 'technisch lezen' betekent. In strikte zin staat het, als opponent van 'begrijpend lezen', voor decoderen en verklanken zonder gebruik te maken van begrip. Het gaat er met andere woorden om, vast te stellen in welke mate de decoding van letters, lettergroepen, woorden en zinnen geautomatiseerd is. In dit opzicht is het onderzoek van Verhoeven (1987) inzichtgevend. Hij laat zien dat de verschillen tussen Nederlandse en allochtone leerlingen bij het hardop lezen van Nederlandse woorden van een gecontroleerde letterdan wel foneemstructuur, significant zijn. Als de Nederlandse woorden echter door nonsenswoorden van een vergelijkbare structuur vervangen worden, blijken deze verschillen veel kleiner (na 10 maanden leesonderwijs niet significant meer). De verschillen die dan toch nog bestaan tussen de Nederlandse en de allochtone

(Turkse) leerlingen kunnen gevoelig toegeschreven worden aan de omstandigheid dat de Turkse leerlingen het onderwijs volgen in een voor hen tweede taal. In elk geval is duidelijk dat kennis van de taal een belangrijke bron is voor het decoderen van woorden in een Brusachtige test. Anders gezegd: in dergelijke toetsen lopen twee verschillende processen door elkaar, te weten bottom-up¹ en top-down². Het is theoretisch verkieslijk dergelijke tegengestelde verwerkingsprocessen uit elkaar te houden. Voor ons onderzoek betekent dat, dat we voor het technisch lezen zoeken naar een testvorm waarin vrijwel uitsluitend bottom-up decodeerprocessen gemeten worden. Woordherkenning wordt tot BeL gerekend. In de Avi-toetsen is de rol van begripsprocessen potentieel nog aanmerkelijk groter dan in de Brus. In de Avi moeten kinderen snel, correct en op goede leestoon een verhaal lezen. Het gelezene roept leesverwachtingen op, de grammaticale (on)mogelijkheden gaan hun rol meespelen, terwijl ook de morfologie- en woordkennis ondersteunend werken bij het te bereiken resultaat. Toch is deze toets door Neuvel e.a. gebruikt voor het meten van technisch lezen.

Hoe moet de technische leesvaardigheid dan vastgesteld worden, even afgezien van de kwestie dat de traditionele school veel geoefender is in TeL-toetsen? Wat het woordniveau betreft kan gesteld worden, dat het hardop lezen van losse nonsenswoorden het gebruik van 'top-down'-strategieën in belangrijke mate reduceert. De kennis die men heeft van de combinatorische mogelijkheden van letters/klanken in het Nederlands blijft als enige hulpbron over. Voor de vergelijking van auto- en allochtone leerlingen zou dit nog storend kunnen werken (Verhoeven heeft dit niet in zijn studie gecontroleerd overigens; het zijn allemaal 'toevallig' niet bestaande Nederlandse woorden.). Voor Nederlandse kinderen in onderling vergelijk is dit onproblematisch, aangezien voldoende kennis van dit systeem al verworven is, en dat voor het overgrote deel in de buitenschoolse situatie.

Vandaar dat we een nonsensversie van de Brus ontwikkeld hebben (nonsens-Brus). Zowel de structuur van de woorden als de opbouw van de lijst zijn vergelijkbaar met die van de Brus. Aangenomen wordt dat een nonsens-leesopdracht moeilijker is en dus voor de groepen 6-8 goed te gebruiken zal zijn.

Technisch lezen is echter meer dan alleen maar woorden lezen. Het gaat ook om zinsdelen, zinnen en teksten. De natuurlijke informatie-eenheid is tenslotte de zin of de tekst. Ook in dat geval willen we ons van nonsens-teksten bedienen. Hiervoor hebben we de Avi 9a en 9b als uitgangspunt genomen (resp. nonsens-Avi 1 en 2). Zoals eerder opgemerkt, is het bij het lezen van zinnen/teksten in nog sterkere mate dan bij het lezen van woorden, de vraag wat aan het technisch lezen toegeschreven kan worden, en wat aan meer algemene kennis van de taal die de proefpersonen bezitten. We besloten de Avi-teksten alleen op woordniveau te bewerken en ons bovendien te beperken tot de inhoudswoorden. De grammatica, de morfologie en functiewoorden, zoals lidwoorden, lieten we ongemoeid. Deze keuze leek ons om de volgende redenen gerechtvaardigd:

Neuvel, e.a. (1988) vergelijken het lezen op traditionele scholen en vernieuw-

1. door de inhoudswoorden van betekenis te ontdoen is het niet mogelijk zich een beeld van de werkelijkheid te vormen dat als verwachting voor het verder lezen kan dienen;
2. lidwoorden e.d. hebben dan geen referent in de werkelijkheid meer en bieden daarom geen top-down steun voor de sowieso niet op gang komende interpretatieprocessen op zins- en tekstniveau;
3. hetzelfde geldt voor de grammatica. Bovendien zullen de proefpersonen dit systeem gelijkelijk en ook hier buiten de onderwijssetting verworven hebben;
4. als ook de grammaticale, morfologische en functiewoordenkennis weggenomen zou worden, ontstaan er geen zinnen/teksten maar reeksen van grafeemgroepen, ingeleid door een hoofdletter en beëindigd door een punt. Een dergelijke toets zou als te kunstmatig ervaren worden en als zodanig zijn doel voorbij schieten.

We kunnen stellen dat op de hierboven geschetste manier meer recht gedaan wordt aan de operationalisering van technisch lezen. Deze deelvaardigheid kan nu eenduidiger en zuiverder onderscheiden worden van begrijpend lezen. Bovendien wordt voor de leeftijdsgroep die in ons onderzoek aan de orde is, een moeilijker en toch aantrekkelijk testobject aangeboden.

De vraag is nu of de beide onderzoeksgroepen wel gelijk behandeld worden qua testervaring en onderwijsaanbod. We mogen bij wijze van antwoord in ieder geval stellen dat het traditionele onderwijs door de inzet van nonsenstoetsen minder bevoordeeld wordt dan bij de gebruikelijke TeL-toetsen. Voor zover het zaken als 'lezen tegen de stopwatch' betreft zal het verschil in toetservaring echter zeker in stand blijven. Of dit laatste nog enig effect heeft en hoe groot dat effect is valt moeilijk vast te stellen. Aan deze onzekerheid ontkomen we dan ook niet.

2.2. Begrijpend lezen (BeL)

Van de recente ontwikkelingen met betrekking tot meer systematische aandacht voor begrijpend en studierend lezen, is tot nog toe maar weinig doorgedrongen tot de onderwijspraktijk. Meestal bedienen scholen zich nog van de zogenaamde totaalmethoden. Opvallend bij deze methoden is, dat ze geen leesdidactiek aanbieden, maar uitsluitend oefeningen door middel van controlevragen bij de tekst. Deze vragen hebben vooral betrekking op woordkennis en lokaal begrip van tekstgedeelten, bijvoorbeeld de alinea. De leerlingen krijgen geen strategieën aangereikt om hun leesbegrip te vergroten, maar ondervinden een voortdurende controle door tekstgebonden oefeningen. Ook de toetsen voor leesbegrip worden veelvuldig volgens bovengegeven procedures opgesteld. Daarnaast komen cloze-toetsen nogal eens in beeld voor het meten van tekstbegrip, of opdrachten als het herschrijven van een tekst. Voor een meer systematische benadering van tekstbegrip sluiten we aan bij Wesdorp (1978) en Looijmans en Palm (1979). Deze laatsten stellen dat tekstbegrip gedefinieerd kan worden als *het doorzien van alle uitspraken in hun onderlinge relaties in een gegeven tekst*. Hacquebord (1989) heeft deze definitie in haar proefschrift geoperationaliseerd in de toetsconstructie

voor tekstbegrip in het voortgezet onderwijs. Allereerst moeten volgens de definitie uitspraken in de tekst 'doorzien' worden. Een subtoets bij Hacquebord bestaat er dan ook uit na te gaan in hoeverre lezers in deze opdracht slagen. Zij noemt dit het micro-niveau van een tekst. Neuvel *cs.* beperken voor groep 3 hun testwerk tot dit niveau. Vervolgens moet het gaan om de onderlinge relaties van uitspraken. Van belang is hier de aard van de relaties, de implicaties, de inferenties *e.d.*

Hacquebord maakt hier een onderscheid tussen uitspraken die als dicht bij elkaar horend gepresenteerd worden (b.v. naast elkaar, tot een alinea behorend), en uitspraken die alleen maar tot dezelfde tekst behoren. In het eerste geval ondervindt de lezer, naast het bij elkaar geplaatst zien van de uitspraken, mogelijk ook steun van anafora, of van voegwoorden (b.v. *omdat*) en bijwoorden (*daarom*, *desalniettemin*), die de relatie expliciteren. In het tweede geval ontbreken deze begripshulpmiddelen veelal of zijn ze veel vager verwoord: *'zoals we al gezien hebben'*, *'hierboven'*, *'de conclusie lijkt gerechtvaardigd'*, *'alles overziend'*, *etc.* Sturing van de interpretatie lijkt hier voor de lezer alleen mogelijk op basis van een globaal tekstbegrip. Dit vindt volgens Kintsch & Van Dijk (1978) plaats op basis van weglating, generalisatie en constructie. Dit soort overwegingen geven Hacquebord aanleiding tot het onderscheiden van twee niveaus m.b.t. de relaties tussen uitspraken: het meso-niveau, dat de alinea, de lokaal gepresenteerde relaties betreft en het macro-niveau, dat de tekstuele relaties op basis van de kennis van de samenhang van alle uitspraken aangaat. Voor alle drie niveaus werden subtoetsen geconstrueerd. In het empirische onderzoek dat volgde, bleek dat het leesbegrip op de drie verschillende niveaus, drie verschillende, niet noodzakelijkerwijs samenhangende vaardigheden betrof. Hierdoor was het mogelijk verschillende lezersprofielen op te stellen die samenvielen met de subgroepen van haar proefpersonen, zoals eerste- en tweede-taal-lezersprofielen.

Wij hebben voor ons onderzoek gebruik gemaakt van de zes gemakkelijkste BeL-teksten van Hacquebord. Deze komen oorspronkelijk uit leerboeken voor de zgn. andere vakken voor groep 8 en de brugklas LBO/MAVO. De lengte van elke tekst is plusminus 200 woorden met een gemiddelde zinslengte van 10 à 12. Het aantal abstracte woorden bedraagt plm. 5%, het aantal concrete/bekende/frequente woorden minimaal 85%. Met 12 bijbehorende stellingen per tekst, gelijkmatig verdeeld over de drie tekstniveaus (micro, meso, macro), wordt het tekstbegrip bevraagd. De proefpersonen moeten aangeven of de uitspraak in overeenstemming is met de inhoud van de gelezen tekst. Doelstelling van de test is te achterhalen of de proefpersonen parafrases van uitspraken en combinaties van uitspraken als wel of niet kloppend met de tekst kunnen identificeren. Als dit het geval is, wordt aangenomen dat de tekst begrepen is. Hacquebord koos voor (parafrases van) uitspraken, omdat in vooronderzoek gebleken was dat deze vraagwijze een heel natuurlijke is, die erg weinig toetservaring vergt. Men mag hierom en om reden dat geen van de leerlingen in ons onderzoek met deze testwijze vertrouwd was, aannemen dat de kans op bevoor- of benadeling van één van onze onderzoeksgroepen gering is.

2.3. Instrumentarium en testverloop

De TeL- en BeL-toetsen werden in de verschillende scholen per klas in drie gedeelten afgenomen:

1. klassikale instructie en afname van het eerste BeL-toetsgedeelte: drie teksten met stellingen (tijdsduur max. ½ uur);
2. individuele instructie en afname van de TeL-toets: nonsens-Brus, nonsens-Avi 1, Avi 9b, nonsens-Avi 2 (plm. 7 min. per leerling);
3. klassikale afname van het tweede BeL-toetsgedeelte: drie teksten met stellingen (tijdsduur max. ½ uur).

De helft van de groep kreeg in de eerste BeL-deeltoets de teksten 1, 2 en 3 en de andere helft de teksten 4, 5 en 6. In de tweede BeL-deeltoets werd dit omgedraaid. Hierdoor konden we nagaan of de prestaties op de tweede BeL-deeltoets mede bepaald werden door zulke zaken als testvermoeidheid. Dit bleek niet het geval.

Voor het hier te rapporteren onderzoek wordt in de resultaten voor TeL alleen gebruik gemaakt van de scores op basis van de nonsens Avi 2. Dit maakt de presentatiewijze overzichtelijker. We menen hiermee te mogen volstaan omdat de uitkomsten op de diverse TeL-toetsen hogelijk met elkaar bleken te correleren, zodat het weinig uitmaakte welke toetsscores gebruikt werden.

Tenslotte gaat het er ons in dit oriënterende onderzoek vooral om, vast te stellen of de ontwikkelde instrumenten gehanteerd kunnen worden voor de vaststelling of er verschillen bestaan tussen leesprestaties van kinderen. Hierbij zou je aan twee zaken kunnen denken:

1. geven de resultaten aan dat er een verschil in leesprestatie bestaat, afhankelijk van het aantal jaren onderwijs dat de leerling genoten heeft?
2. geven de instrumenten zicht op de effecten van de gevolgde leesmethode?

De eerstgenoemde vraag is hierbij te vertalen tot een harde eis waaraan de instrumenten moeten voldoen om enige geldigheid te bezitten. De tweede vraag betreft de onderzoeksvariabele en geeft de mate van effectverschil aan van de leesmethodieken.

Dit verschil is alleen zinvol interpreteerbaar, als aan de voorwaarde onder 1 genoemd voldaan is.

3. De proefpersonen en proefscholen

Een volgend punt dat buiten onderzoeksdesign en vraagstelling om enige extra aandacht behoeft, betreft de keuze van de onderzoeksscholen en de leerlingenpopulatie daarbinnen. We doen dat om twee redenen. Ten eerste is het onderzoek van Neuvel cs. waarop wij aansluiten ons inziens op dit punt te globaal. Ten tweede komen we in de slotparagraaf op de steekproeftrekking terug. Daar zal blijken dat een grotere zorgvuldigheid dan in sociaalwetenschappelijk onderzoek gebruikelijk geboden is, wil men tot goed interpreteerbare gegevens komen. Voor dit onderzoek hebben we de volgende eisen t.a.v. de proefpersoonkeuze opgesteld:

1. er dienen voldoende proefpersonen getoetst te worden om een redelijke indruk te krijgen van de geschiktheid van de instrumenten;
2. er moet voldoende spreiding zijn in de beoogde doelgroep van 10-12 jarigen m.b.t. de leesmethodiek: traditioneel versus functioneel vernieuwend;
3. beide groepen traditioneel (T) en functioneel vernieuwend (F) dienen op de overige relevante aspecten goed vergelijkbaar te zijn.

In verband met punt twee zijn we het onderzoek begonnen met het benaderen van twee Jenaplan/Freinet-scholen (FA en FB). Deze staan schoolorganisatorisch en onderwijsdidactisch nogal veraf van het traditionele onderwijs. We hebben bewust voor afstand tussen T en F gekozen. Dit overigens in tegenstelling tot het onderzoek van Neuvel cs., waarin een grote spreiding van F-methodieken in de onderzoekspop is meegenomen, zonder differentiatie met betrekking tot de prestatiemeting. Hierdoor weten we in dat onderzoek niet goed welke vernieuwingsaspecten voor welke resultaten verantwoordelijk zijn.

We lieten alleen kinderen meedoen die van meet af aan in de betreffende leesmethodiek zijn onderwezen. Binnen dit bestek hebben we ons moeten beperken tot de twee Jenaplan/Freinet-scholen (maar zie slotparagraaf 6), wat de volgende aantallen leerlingen opleverde (tabel 1):

Tabel 1: Aantallen leerlingen in F-scholen.

F-school/groep	6	7	8	totaal
A	7	8	9	24
B	7	7	8	22
totaal	14	15	17	46

Volgens de schooldirectie waren de kinderen afkomstig uit een hoog sociaal milieu (ouders minimaal VWO; hetgeen overigens niet altijd bleek te kloppen, zie par. 6). Op grond van de CPB-cijfers lag de indicatie van de directeur alleszins in de rede. We besloten tot deze check omdat de schooldirecties buitengewoon terughoudend waren t.a.v. deze achtergrondinformatie. Overigens wijken we hier nadrukkelijk af van de werkwijze van Neuvel cs. Voor de F-scholen volstaan zij, als indicatie voor leerlingenvariabele 'sociaal milieu', met de aanduiding in wat voor gebied de scholen liggen (zoals provinciestad, nieuwbouwwijk e.d.). Voor de controle-scholen werd dit niet nagegaan. Om toch enig zicht te krijgen op de vergelijkbaarheid werden algemene intelligentietests gebruikt, zoals die voor woordenschat en ruimtelijk inzicht. Hieruit bleek wel enig verschil tussen de groepen, ten voordele van de controlegroepen, maar dit werd niet meegenomen in het vergelijkend onderzoek. Toch zijn deze verschillen, met name die in woordenschat en intelligentie, niet allemaal onbelangrijk, gezien hun samenhang met lezen (zie Hacquebord, 1989 en Aarnoutse e.a., 1988). De interpretatie van dergelijke

verschillen ligt trouwens ook moeilijk. Zo weten we niet of de gevonden intelligentieverschillen aan de herkomst van de leerlingen of aan de schoolse omgeving moeten worden toegeschreven. Deze problematiek wordt in de bovenbouw van de basisschool des te klemmender (Bollen, 1991). Vandaar dat we besloten hebben de milieuvariabele zo goed mogelijk te bepalen. Op grond van de beschikbare gegevens werd een traditionele controle-school (T) gezocht, die op een aantal belangrijke kenmerken identiek was aan FA en FB:

1. de T-school was, zoals de beide F-scholen, wijkoverstijgend;
2. kinderen uit andere wijken kwamen ofwel uit denominatie-overwegingen (R.K.) van de ouders, ofwel vanwege de goede naam van de school op T terecht. Van de laatste overweging was ook sprake bij FA en FB. Het R.K.-karakter bij T zou vergeleken kunnen worden met het levensbeschouwelijke karakter van F, dat door de ouders als additionele eis aan het onderwijs wordt gesteld;
3. ook de T-school bestond voornamelijk uit kinderen uit het hoger sociale milieu. Met de directeur werden uit de kaartenbakken in alfabetische volgorde proefpersonen gelicht, waarvan de ouders minimaal een VWO-opleiding hadden. De medewerking van de directeur was hierbij con amore.

De proefpersonengroepen zagen er uiteindelijk als volgt uit (tabel 2):

Tabel 2: Totale aantallen proefpersonen in F en T.

F/T	gr.	6	7	8	totaal
FA		7	8	9	24
FB		7	7	8	22
Ftot.		14	15	17	46
T		14	15	17	46
totaal		28	30	34	92

4. Onderzoeksoptzet en vraagstellingen

Evenals in Neuvel, e.a. (1988), wordt hier gebruik gemaakt van een cross-sectioneel onderzoeksmodel. Deze keuze werd uitsluitend ingegeven door de beschikbare tijd. Een dergelijke onderzoeksoptzet hoeft niet per se afgewezen te worden in ontwikkelingspsychologisch onderzoek, mits de onderzoeksvragen voldoende scherp geformuleerd zijn en men de uitspraken op basis van de resultaten tot cohorten beperkt en niet longitudinaal interpreteert (Van der Geest, 1990). Een onderzoek dat zich in eerste instantie bezighoudt met de vraagstelling of het onderzoeksinstrumentarium bruikbaar is, behoeft niet noodzakelijkerwijze een longitudinale opzet. Wat de evaluatie-vraagstelling van dit onderzoek betreft, kan

gesteld worden dat de meeste grote evaluatie-onderzoeken van cross-sectionele gegevens uitgaan (b.v. Lukesch, 1979; De Glopper, 1988; PPON e.d. en Neuvel, e.a., 1988 natuurlijk). In zo'n opzet worden prestaties afgezet tegen de kwantiteit en kwaliteit van de afgelegde weg.

4.1. Toetsontwikkeling en proefonderzoek

Neuvel e.a. beperkten hun materiaalverzameling tot de vernieuwingsscholen en vergeleken de gevonden prestaties m.b.t. BeL en TeL vervolgens met reeds bestaande gegevens uit onderzoek over leesprestaties in het traditionele onderwijs. De consequentie van een dergelijke handelwijze is, dat men de toetsen uit het traditionele onderwijs als uitgangspunt neemt voor de evaluatie van het vernieuwingsonderwijs, zonder zich rekenschap te geven van het feit dat er door deze gedwongen keuze mogelijk een 'bias' in de meting geïntroduceerd wordt (zie par. 2).

Om hieraan te ontkomen hebben we besloten

1. een aangepast instrumentarium te ontwikkelen dat inhoudelijk eenduidiger en valide de deelaspecten van het lezen meet;
2. deze toetsen in een voorstudie op bruikbaarheid door te lichten op beide schooltypen.

Eventueel zullen deze gegevens in het vervolgonderzoek meegenomen kunnen worden voor de beantwoording van de eigenlijke onderzoeksvragen. Zoals eerder opgemerkt, is dit alleen mogelijk als de resultaten geen aanleiding geven tot veranderingen in het instrumentarium.

4.2. Keuze van de vernieuwingsscholen

Neuvel cs. accepteerden een nogal grote spreiding m.b.t. het soort vernieuwing van het leesonderwijs: Jenaplan (o.a. Nash-achtige aanpak, breakthrough to literacy, Stauffer-aanpak, Freinet), de Freinet-methode in Freinet-scholen en de geïntegreerde taal-leesmethode. Met name ten aanzien van de mate van individualisering en cursorische ondersteuning is er sprake van een grote bandbreedte. Uit de resultaten bleek bovendien dat er grote verschillen in prestatie aan te wijzen waren (Neuvel, e.a., 1988, p. 184).

We besloten op grond hiervan leesmethodespecifieker te werk te gaan en stapsgewijs materiaal te verzamelen op basis van de gebruikte onderwijsmethodiek. Op deze manier kunnen op het interpretatieve niveau opvallende prestaties gemakkelijker met didactiekkenmerken verbonden worden. Dit heeft ertoe geleid dat we ons hier uitsluitend bezig zullen houden met het Jenaplan-onderwijs waar men zich m.b.t. het lezen vooral door Freinet heeft laten inspireren.

4.3. Vraagstelling

In dit verslag beperken we ons tot een aantal vraagpunten die na Neuvel e.a. (1988) en hun besprekers (Mommers en Van der Geest) zijn blijven liggen. Dit geldt ten eerste de vertaling van de onderzoeksvraag in de basisschool na groep 5 en ten tweede een aantal methodische tekortkomingen van Neuvel, e.a., zoals die met betrekking tot de steekproeftrekking, de homogeniteit van de gevolgde methode en de zuiverheid/geschiktheid van de toetsinstrumenten. Allereerst, zoals gezegd, houden we ons bezig met de problematiek van de toetsconstructie voor TeL en BeL. Deze toetsen dienen ten eerste theoretisch beter verantwoord te zijn, ten tweede geen bias meer met zich mee te dragen en ten derde voor de bovenbouw van de basisschool geschikt te zijn. In totaal stellen we ons drie vragen:

1. Zijn de ontwikkelde tests die de decodeervaardigheid zuiverder van tekstbegrip beogen te scheiden, in staat de prestaties voldoende te differentiëren, afhankelijk van de mate van scholing (terwijl de gevolgde leesmethodiek gecontroleerd wordt)?
2. Als vraag 1 positief beantwoord kan worden, bestaat er dan een verband tussen onderwijs-aanpak en de geleverde BeL- en TeL-prestaties?
3. Hoe moeten de resultaten die vraag 2 oplevert geïnterpreteerd worden in het licht van de onderwijsaanpak?

Er zullen aanvullende gegevens uit Hacquebord (1989) en Jorna (1990) als referentiekader voor de beantwoording van vraag 3 worden gepresenteerd.

5. Onderzoeksresultaten

Alvorens tot de onderzoeksvragen over te gaan, zijn we nagegaan of de FA- en FB-leerlingen als één onderzoeksgroep beschouwd mochten worden. Dit was wenselijk, omdat we in beide F-scholen over wel heel kleine onderzoeksgroepjes per jaargang beschikten. Op grond van de onderwijsmethodiek en de gegevens van de proefpersonen, leek het alleszins redelijk deze mogelijkheid te bezien. Met behulp van de Wilcoxon rank correlation test (Wrct) werd per jaargang nagegaan of de leerlingen voor de verschillende onderzoeksvariabelen als een homogene groep samengenomen konden worden. Tabel 3 geeft hierover uitsluitsel.

Tabel 3: S(W)- waarden en significanties (-/1: $p \leq .05$) voor de vergelijking van FA en FB voor de variabelen BeL, micro, meso en macro en TeL.

*: $p \leq .005$; **: minimale waarde S(W) voor $p \leq .05$ is voor 6, 7 en 8 resp. 27, 30 en 36.

FA/FB	BeL		Mi		Me		Ma		TeL	
	S(W)	s	S(W)	s	S(W)	s	S(W)	s	S(W)	s
6**	14	-	13	-	1	-	10	-	1	-
7	6	-	15	-	10	-	50	+	10	-
8	2	-	6	-	12	-	10	-	4	-

Slechts één van de vijftien keer bleek er sprake van een significant verschil tussen FA en FB, zodat er gevoeglijk van kan worden uitgegaan dat A en B t.a.v. de onderzoeksvariabelen te beschouwen zijn als een homogene groep (F).

5.1. De geschiktheid van het onderzoeksinstrumentarium

We zullen nu ingaan op de eerste onderzoeksvraag, die betrekking heeft op de geschiktheid van de ontwikkelde en gekozen toetsinstrumenten.

De prestaties van de verschillende jaargroepen van F en T op BeL en zijn subaspecten (in percentages correcte antwoorden) en TeL (in aantal seconden) worden met de bijbehorende standaarddeviaties in tabel 4 weergegeven.

Tabel 4: Percentages goede antwoorden met de SD's per school en per jaargroep voor F en T. (Voor verklaring zie tekst.)

F/T	BeL		Micro		Meso		Macro		TeL	
	%	SD	%	SD	%	SD	%	SD	sec.	SD
F6	82	7.8	83	10.8	84	7.9	78	11.3	139	35.5
F7	84	9.6	86	10.0	84	11.3	81	10.8	158	44.2
F8	91	4.2	91	6.3	91	3.3	89	7.5	98	24.1
T6	87	4.0	86	5.4	96	4.6	84	5.8	131	34.7
T7	87	7.8	88	5.0	88	5.4	83	5.0	100	36.0
T8	93	3.2	94	3.8	94	5.4	91	5.0	116	34.5

Op een aantal variabelen lijken met het klimmen der jaargroepen de prestaties hoger te worden. Met behulp van de Wrct. is nagegaan waar dit significant is. Zie hiervoor tabel 5.

Tabel 5: Vergelijking van de jaargroepen binnen de scholen; significanties op basis van Wrct.

T/F	BeL	Micro	Meso	Macro	TeL
6F/7F	-	-	-	-	-
6F/8F	.01	.001	.01	.01	.001
7F/8F	.025	.05	.05	.05	.001
6T/7T	-	-	-	-	.05
6T/8T	.005	.005	.025	.005	-
7T/8T	.005	.005	.05	.005	-

Uit deze tabel blijkt dat de tests een duidelijke cesuur aanbrengen tussen de groepen 6 en 7 aan de ene en groep 8 aan de andere kant. M.b.t. de groepen 6 en

7 zijn de veranderingen slechts als tendentieel te duiden. Over het algemeen kan gesteld worden, dat de mate van scholing in jaren uitgedrukt redelijk samenhangt met de prestaties op de tests. Dat dit vooral t.a.v. groep 8 geldt, zou zijn verklaring kunnen vinden in het feit dat de gebruikte teksten voor groep 8 en de brugklas bestemd waren. Mogelijk staan deze nog te ver van de beleavingswereld van 10- en 11-jarigen af. Zoals echter nog zal blijken, liggen hun BeL-prestaties beduidend hoger dan bij (oudere) LBO-/MAVO-leerlingen uit het onderzoek van Hacquebord (1988). Als we dit bij onze beschouwingen betrekken, dan zou nog beter gesteld kunnen worden dat in groep 8 een plotselinge groei in begripend lezen optreedt. Of we zouden te maken moeten hebben met een op alle drie scholen optredend cohort-effect.

Voor TeL blijkt er ook van groei sprake, al valt deze voor de T- en F-groepen op een verschillend moment. De conclusie kan zijn dat de toetsen redelijk in staat zijn TeL- en BeL-prestaties vast te stellen in de leeftijdscategorie van 10- tot 12-jaar.

5.2. Het verband tussen onderwijsaanpak en prestaties

Nu vraag 1 naar tevredenheid is beantwoord, is het mogelijk in te gaan op de vraag of de onderwijsaanpak van invloed is op de prestaties van de leerlingen. Om een eerste globale indruk hierover te krijgen, zien we even af van de leeftijdsverschillen, zodat over wat grotere groepen uitspraken kunnen worden gedaan. Dit is toegestaan omdat de jaargroepen van T en F telkens even groot zijn. We hebben de 92 scores voor BeL (en zijn subaspecten) en TeL in drie groepen gedeeld: lage (30), gemiddelde (31) en hoge (31) scores. Vervolgens zijn we nagegaan hoe deze groepen over de beide scholen verdeeld waren. Chi-kwadraat toetsing laat zien (tabel 6) dat er zowel voor BeL als voor TeL geen significante verschillen zijn aan te wijzen op het 5 procentniveau.

Tabel 6: Verband tussen toetsprestaties en onderwijsaanpak; drempelwaarde significantie $\leq .025$

	BeL	micro	meso	macro	TeL
chi-kwadraat	3.31	4.52	11.22	3.88	2.29
significantie	-	-	.005	-	-

Er is zelfs geen sprake van een tendens (10 pcts-niveau). Toch is er bij BeL iets bijzonders aan de hand als we naar de subaspecten kijken. Het meso-aspect verschilt significant, ten voordele van T. Blijkbaar is er op het micro- en macro-niveau echter voldoende compensatie om dit meso-effect in BeL weer ongedaan te maken. Het ligt voor de hand te concluderen dat de prestaties t.a.v. BeL en TeL onafhankelijk zijn van de gevolgde onderwijsmethodiek. Op grond van tabel 6 kan echter niet geconcludeerd worden dat de prestaties op de diverse momenten

identiek zijn. Zo stellen Neuvel e.a. vast, dat een aanvankelijk groot verschil in groep 3 verdwenen is in groep 5. Mogelijke verschillen in het traject (in dit geval van groep 6 naar groep 8) zijn daarom interessant. Ze kunnen ons een indicatie geven over de samenhang tussen onderwijsactiviteiten op een bepaald moment enerzijds en onderwijsprestaties anderzijds. Om op mogelijke verschillen in het onderwijstraject zicht te krijgen, is m.b.t. Wrct nagegaan of er op groepsniveau sprake is van verschillen m.b.t. de onderzoeksvariabelen. Zie hiervoor tabel 7.

Tabel 7: Significantieniveaus van prestatieverschillen op Wrct. van jaargroepen van F- en T-scholen.

F/T	BeL	Micro	Meso	Macro	TeL
6	.025	-	.05	.05	-
7	-	-	-	-	.005
8	-	-	.01	-	.05

BeL, met name het meso- en macro-aspect, blijft in groep 6 van de F-scholen t.o.v. T achter. Bij TeL is dit niet het geval, wat merkwaardig is, omdat F juist van functioneel leesbegrip uitgaat en veel minder van de technische aspecten van het lezen. Gedeeltelijk is dit achterblijven van BeL te verklaren uit het feit dat de leerlingen in groep 6 FA (onze eerste onderzoeksklas) de vragen veelal bleken te beantwoorden op basis van eigen kennis en verwachtingen in plaats van zich op de tekst te baseren. Dit bemerkende, hebben we in de instructie van de andere klassen met meer nadruk op dit punt gewezen. M.a.w. de verschillen zijn mogelijk op een gebrekkige toetservaring terug te voeren. In het eindstadium van het onderwijs (groep 7 en 8) zijn er geen systematische verschillen meer in BeL. Dat F8 het wat minder doet op het meso-aspect, wordt blijkbaar -zoals reeds gesteld- voldoende gecompenseerd op het micro- en macro-niveau. Dat de T-school het op het meso-niveau beter doet, is mogelijk toe te schrijven aan het gebruik van een totaalmethode (Taalactief), waar de nadruk bij de gestelde vragen bij de teksten wel heel erg sterk op de alinea ligt.

In TeL doen zich ook maar weinig verschillen voor. In groep 7 scoort T significant beter; in groep 8 daarentegen is dat F. In elk geval is duidelijk dat ook de technische aspecten van lezen in een functionele benadering voldoende kans krijgen zich te ontwikkelen.

Als conclusie uit de tabellen 6 en 7 zou kunnen gelden dat voor leerlingen uit een hoger sociaal milieu -want daar spreken we hier over- het er niet zoveel toe doet welke onderwijsaanpak ze op zich af krijgen (zie ook Leseman, 1990). Er zijn weliswaar cross-sectionele verschillen (mogelijk ook tempo-verschillen en/of leerlijn-verschillen), maar uiteindelijk zijn de resultaten onderling zeer goed vergelijkbaar. De prestatieprofielen worden slechts marginaal (meso- t.o. micro- en macro-aspecten) door de aanpak medebepaald.

6. Conclusies en interpretaties

Hoe zijn de resultaten te interpreteren als ze mede in het licht van andere onderzoeksgegevens bezien worden (vraag 3)?

Het is duidelijk dat de conclusies op grond van het kleine aantal proefpersonen en de cross-sectionele opzet, heel voorzichtig moeten zijn. De negatief geformuleerde conclusie bij vraag 2 is daar een voorbeeld van: we hebben niet veel gevonden dat een waarderingsverschil tussen F en T rechtvaardigt. Ondertussen zijn er wel enkele andere onderzoeksgegevens voor handen die een wat genuanceerder antwoord mogelijk maken. Het gaat hierbij om drie soorten van gegevens:

1. de individuele leerlinggegevens die boven tafel kwamen bij de nabespreking van de resultaten met de individuele leerkrachten en/of directeur;
2. de scores van Hacquebord (1989);
3. onderzoeksgegevens uit een sociaal meer gespreide school (Jorna, 1990).

6.1. Aanvullende leerlinggegevens

Met de scholen was afgesproken om binnen het team de individuele leerlingprestaties na te bespreken. Deze nabespreking bleek voor ons achteraf nogal zinvol omdat ze een beter inzicht gaf in de diverse schoolpopulaties. Zo bleken in de beide F-scholen niet alle ouders over een VWO-opleiding te beschikken (10% niet) en lag het post-VWO traject over het algemeen hoger, of op zijn minst anders in de T-school (arts, specialist, notaris, advocaat, rechterlijke macht e.d. in T, tegenover HBO, sociaal-wetenschappelijke opleiding e.d. in de F-scholen). Ten tweede bleek er met zo'n 35% van de leerlingen in F (t.o. 2% van T) iets aan de hand: dyslexie, MBD, RIAGG-kind, adoptiekinderen, speciaal- onderwijsadviezen etc. Deze kinderen scoorden blijkens de chi-kwadraat toetsing significant ($p \leq .05$) lager op BeL, maar niet op TeL. Als de opgeschoonde gegevens van F en T opnieuw met elkaar geconfronteerd worden, gaan de prestaties van de beide scholen nog meer op elkaar lijken: van het enige significante verschil (het meso-aspect; $p \leq .025$) is nu slechts nog een tendens over ($p \leq .10$).

Als we de nieuwe gegevens nog eens de revue laten passeren, blijkt dat ondanks weinig systematische aandacht voor TeL de probleemkinderen hier een vergelijkbare prestatie neerzetten. Dit is niet het geval voor BeL: daar liggen de resultaten significant lager. Toch kunnen we - vooruitlopend op de volgende paragraaf - ook voor deze kinderen stellen dat de BeL-prestaties in groep 6 al hoger liggen dan bij de Nederlandse MAVO-leerlingen uit de klassen 1 t/m 3 uit het onderzoek van Hacquebord (1989) (resp. 76.5% vs. 75.3%). In groep 8 is dit verschil ten voordele van de probleemkinderen nog groter (86.8% vs. 75.3%). M.a.w. de probleemkinderen (voor een groot deel met een SO-indicatie) uit de F-scholen, komen op jongere leeftijd wat tekstbegrip betreft duidelijk boven een MAVO-prestatie uit. Daarbij bleek hun lagere prestatie in de F-school voor het totaalbeeld van F niet zo bepalend dat het een beslissende rol speelt in de F/T-vergelijking, hetgeen betekent dat hun prestaties zich niet al te ver van het groepsgemiddelde

bevinden. Zo bezien, zou de hypothese dat met name probleemkinderen baat hebben bij een functionele leesaanpak, goed te verdedigen zijn.

Er valt voorzichtig nog een andere conclusie te trekken. Deze heeft betrekking op de procedure van de steekproeftrekking in onderwijsonderzoek. Bij het vergelijken van traditioneel en vernieuwingsonderwijs blijkt het controleren van de sociale variabele niet voldoende voor het verkrijgen van vergelijkbare onderzoeksgroepen. Neuvel cs. vergelijken hun scholen uitsluitend met omgevingskenmerken, die een indicatie voor de sociale variabelen zouden moeten zijn. Wij hebben zowel de omgeving als de sociale variabelen gecheckt en moeten vaststellen dat er nog essentiële storende variabelen overblijven: onderwijskundige indicaties en aanvullende oudergegevens. We konden vaststellen dat vernieuwingscholen probleemkinderen aantrekken (plm. 1/3 van onze steekproef op beide scholen). Bovendien blijken dergelijke scholen veel minder geneigd leerlingen naar het SO door te verwijzen. Ook andere scholen waarmee we nu contact hebben vertonen dit beeld. De scholen proberen overigens op verschillende manieren dit probleem in toom te houden, bijvoorbeeld door probleemkinderen niet te laten 'neveninstromen', of door uit een gezin niet alleen het probleemkind te accepteren. Een consequentie hiervan is, dat de onderzoeksresultaten van Neuvel cs. aan waarde verliezen of mogelijk zelfs als niet betrouwbaar van de hand moeten worden gewezen. Wat betreft de ouders kan verder opgemerkt worden, dat zij die uit overtuiging hun kinderen naar een Freinet/Jenaplan-school sturen, de school veeleer zien als een ideale opvoedingsomgeving voor de totale ontplooiing van hun kind, dan als een eerste stap van een voorspoedige maatschappelijke carrière. Op beide F-scholen werden we geconfronteerd met heftige uiteenzettingen tussen (globaal) twee groepen ouders -met name die van kinderen uit de hogere jaargroepen- waarvan de ene groep een meer prestatiegerichte (dus traditionele) aanpak wenste en de andere groep ook voor deze leeftijdscategorie vasthield aan het hierboven gestelde opvoedingsideaal. Op beide scholen waren de eerstgenoemde ouders de verliezers, hetgeen er telkens toe leidt dat vaak beter presterende kinderen vanaf groep 6 de school verlaten. Ook op dit punt lijkt het buitengewoon heikel kinderen van verschillende schooltypen met elkaar te vergelijken. De doelstellingen van onderwijs en opvoeding zijn daarvoor te weinig identiek.

6.2. De scores van Hacquebord

Hacquebord (1989) rapporteert gegevens van dezelfde BeL-toets. Deze werden van zowel Turkse als Nederlandse leerlingen verzameld in een longitudinaal onderzoek van groep 8 t/m klas 3 in LBO/MAVO. In tabel 8 zijn de scores van de Nederlandse leerlingen naast die van de T- en F-school gezet.

Tabel 8: MAVO(LBO)- leerlingen vergeleken met leerlingen uit de groepen 6,7 en 8 in T en F.

Toets	Hacquebord LBO/MAVO		F-school			T-school		
			6	7	8	6	7	8
Micro	LBO	68.8	82.6	85.8	91.3	86.3	87.9	94.2
	MAVO	79.6						
Meso	LBO	62.3	84.4	84.2	90.9	96.3	88.6	94.6
	MAVO	74.0						
Macro	LBO	64.0	77.9	81.2	89.4	84.2	83.3	91.3
	MAVO	72.3						
BeL	LBO	65.0	81.5	83.9	90.8	87.2	86.7	93.3
	MAVO	75.3						

Uit tabel 8 blijkt, dat de (betere) MAVO-leerlingen (er zijn geen significante verschillen tussen de klassen 1, 2 en 3) beduidend achter liggen bij de kinderen uit groep 6 (!) en hoger van de T- en F-school. In de vorige paragraaf werd al opgemerkt, dat dit zelfs ook voor de probleemkinderen (idem van groep 6 en hoger) uit F op gaat. Het ziet er naar uit, dat de verklaring hiervoor gezocht moet worden in het verschil in sociaal milieu (in T en F VWO+, t.o. een veel lager sociaal milieu in LBO/MAVO).

Omgekeerd is de afwezigheid van verschillen in leesprestaties tussen F en T mogelijk ook terug te voeren op de factor sociaal milieu. In beide typen scholen is sprake van een hogere milieu-achtergrond. Kinderen uit deze hoge milieus leren mogelijk redelijk onafhankelijk van de school goed te lezen.

Om echt iets te kunnen zeggen over de invloed van de onderwijsaanpak zijn proefpersonen nodig die minder vanzelfsprekend tot lezen komen. Een eerste aanwijzing voor een positieve invloed van F zijn de prestaties van de probleemkinderen. Ook de gegevens uit aanvullend onderzoek in een Jenaplan/Freinet-school met een sociaal meer gedifferentieerde populatie lijken in deze richting te wijzen.

6.3. Sociale differentiatie in F2

In een longitudinaal opgezet onderzoek zijn de eerste gegevens verwerkt van een F-school (F2) met een sociaal meer gespreide populatie (Jorna, 1990). De leerlingen van deze school zijn in twee groepen verdeeld; de eerste is qua sociaal milieu min of meer vergelijkbaar met de onderzoeksgroepen uit de T- en F-school van dit onderzoek (F2H). De tweede groep wordt voornamelijk gevormd door leerlingen met een onderwijsvoorrangsscore van 1.25: leerlingen vooral uit een werkmilieu (F2L). Aangenomen mag worden dat deze kinderen in hun gezinssituatie minder systematisch met het lezen worden geconfronteerd, dan wel met taal- en denkactiviteiten die voor het begrijpend lezen van belang zijn. In tabel 9 worden de beide groepen vergeleken met de MAVO-leerlingen uit Hacquebords onderzoek en met de leerlingen uit de T- en F-groep.

Tabel 9: Prestaties in BeL en TeL in T, F en F2; H/L = hoger/lager milieu. Ter vergelijking: BeL, MAVO (Hacquebord, 1989).

Toets	gr.	F	T	F2	F2H	F2L	MAVO
BeL	6	81.5	87.2	86.0	89.0	82.3	75.3
	7	83.9	86.7	88.0	87.7	88.2	75.3
	8	90.8	93.3	85.8	87.7	84.4	75.3
TeL	6	138	131	123	114	139	-
	7	156	100	137	133	141	-
	8	98	116	111	104	118	-

Bij vergelijking van de BeL-scores van F2H en F2L, blijkt er uitsluitend in groep 6 sprake van een significant verschil (Wrct, $p \leq .05$). Alle prestaties van F2L ligt echter nog binnen het bereik van de F- en T-scores en zijn bovendien beduidend hoger dan die van de plm. vier jaar oudere MAVO-leerlingen. Tevens geldt, dat er geen significante verschillen in BeL zijn tussen F2, F2H en F2L enerzijds en de F- en T-groepen anderzijds. Een uitzondering hierop is het verschil tussen F2L en T in groep 8 ($p \leq .05$). In groep 8 blijft overigens heel F2 wat steken. Hiervoor is zowel een cross-sectionele als een longitudinale verklaring mogelijk. In groep 8 wordt namelijk vrijwel alle tijd opgesoupeerd door deelvaardigheidsonderwijs als voorbereiding (!?) op het voortgezet onderwijs. Ook het feit, dat de huidige groep 8 juist in de onderbouwgroep kwam ten tijde van het toen nog experimentele programma functioneel lezen, zou een verklaring hiervoor kunnen bieden.

Voor het TeL is er slechts één duidelijk verschil: in groep 7 scoort T significant hoger dan F2L ($p \leq .05$).

Uit deze cijfers blijkt, dat de lager sociaal milieu-groep bijna even goed als de hoger sociaal milieu-groep presteert. Bovendien scoren ze beduidend beter dan de qua milieu vergelijkbare MAVO-groep. Dit kan opgevat worden als een indicatie van de profitabiliteit van het functionele leesonderwijs.

6.4. Twee voorzichtige conclusies

De aantallen onderzochte leerlingen zijn maar klein. Toch durven we op grond van het voorgaande twee voorzichtige slotconclusies te trekken:

1. vernieuwingsonderwijs is moeilijk te evalueren vanwege:
 - a) de niet met het traditionele onderwijs vergelijkbare populaties m.b.t. probleemkinderen, verwijzingsbeleid naar het SO e.d.;
 - b) de niet met het traditionele onderwijs vergelijkbare doelstellingen t.a.v. onderwijs en opvoeding.
2. Alleen als we specifiek naar de probleemkinderen (SO-indicatie, onderwijsvoorrang-kinderen) in het onderwijs kijken, zijn er aanwijzingen dat functioneel leesonderwijs de voorkeur verdient. Deze conclusie kan overigens ook

voorzichtig uit het Amerikaanse leesonderzoek getrokken worden (Van der Geest, 1990).

Noten

1. Waarbij lezers vanuit de kleinste delen (grafemen), grotere gehelen (woorden, zinnen etc.) concipiëren.
2. Waarbij lezers op basis van kennis op een hoger niveau van de tekst (zoals de alinea) verwachtingen en voorspellingen opbouwen ten aanzien van lagere tekstniveaus, zoals zinnen, woorden en grafemen.

Bibliografie

- Aarnoutse, C., M. Mommers, B. Smits & J. v. Leeuwe (1986), De ontwikkeling en samenhang van technisch lezen, begrijpend lezen en spellen, *Pedagogische Studiën*, 63, p. 97-110.
- Aarnoutse, C. & J. v. Leeuwe (1988), Het belang van technisch lezen, woordenschat en ruimtelijke intelligentie voor begrijpend lezen, *Pedagogische Studiën*, 65, p. 49-59.
- Anceaux, H. (1990), *Luisteren en Lezen*, Leiden, Proefschrift Groningen.
- Boland, T. (1988), Wat staat hier nu weer?, *Stimulans*, 6, p. 25-27.
- Bollen, N. (1991), *Voorrang ongewijzigd, samenvattende rapportage evaluatie OVG Groningen 1986-1990*, Evaluatiegroep Groningen.
- Geest, T. v.d. (1982), Kommunikation in der Verhaltenstherapie. In: T. v.d. Geest en D. Fehlenberg, *Kommunikationsanalysen in der Verhaltenstherapie*, Bochum, p. 26-50.
- Hacquebord, H. (1989), *Tekstbegrip van Turkse en Nederlandstalige leerlingen in het voortgezet onderwijs*, Dordrecht, Proefschrift Groningen.
- Jorna, A. (1990), *Beschrijving en evaluatie van Natuurlijk leren lezen*, Groningen Nederlands Instituut, Groningen, RuG.
- Kintsch, W. & T. v. Dijk (1978), Toward a model of discourse comprehension and production, *Psychological Review*, 85, p. 363-394.
- Leseman, P. (1990), *Lezen als Denken; leren lezen als leren denken*, Rotterdam, RISBO/Erasmus Universiteit, Rotterdam.
- Looijmans, P. & J. Palm (1979), Tekstbegrip getoetst?, *Tijdschrift voor Taalbeheersing*, 31, p. 328-344.
- Lukesch, H. e.a. (1979), *Gesamtschule, ein dreigliedriges Schulsystem in NRW*, Paderborn.
- Mommers, C. (1989), Functioneel aanvankelijk leesonderwijs in de spiegel; een boekbespreking, *Spiegel*, 7, p. 75-86.
- Neuvel, J., M. Otter & D. Bos (1988), *Functioneel aanvankelijk leesonderwijs*, Lisse, SVO reeks, 86.

Verhoeven, L. (1987), *Ethnic minority children acquiring literacy*, Dordrecht, Studies on language acquisition nr. 4.

Wesdorp, H. (1978), *Evaluatietechnieken voor het moedertaalonderwijs*, Amsterdam.

Bijlagen

Bijlage 1: Voorbeeld Nonsens Brus

taaf	hijpos	brienen	straakgewek
vas	enhal	gelip	stoolverdakkig
deek	stuikheid	oeltijk	verdieuwen
nim	pruifstik	masplonten	derging
gos	kepje	droven	nonocht
vek	loch	koopuitluwen	outkluwen
fruum	begen	steiperen	herkreutsing
norgel	fipplemp	opsogen	annasel
aven	gaalte	hondelaar	mederib
tad	rapenen	fiamont	zijruvaar
noortij	verdaving	podig	verwendering
zaper	gewraai	denger	taling
biene	verloezen	butee	getijgschruft
ruihak	kaggen	ochtang	hertonning
kand	bozenhuif	opspuiken	everlaatsen
blekken	uitmokken	priestmalen	bijaanlorst
verpaden	krolen	omtruiken	gepartneming
vermuilen	kwievig	baskraak	belaatsvloeg
karden	vaarkroog	ontschuipen	geoefelaar
wichten	evermalen	schoofkaan	verstrapend
aanbuizen	afwoppen	wierenboning	verstegenheid
vleut	kroeze	beguisd	ontpietselen
kraatduiner	verzwijlen	wijlspol	onvlachtig
poeling	pochtbelken	schieping	notsberaping
laas	wozen	spouwketer	ontmerstelen
steumel	narsbol	mokkelen	pareelschek
hoenpollen	spietelen	schope	batskel
treilvek	kwilhogel	fertkars	inekken
schroelen	muitkrit	krobbelen	tubberoonplinting

Bijlage 2: Voorbeeld Nonsens Avi

Kompel

Kompel er een rusachte daasbang en zijn moos er rakter. Hij bult Kompel ielmijs pee, als hij zommels bezuult. De kos van de rakter roost op een kwitte geuze knakerij. Daap wiet hij ok masen met zijn rakter voor het uurst heeg.

Op de knakerij frest Kompel stee amoeraggertjes oel: Zwimpie en Estor. Die kobben een puitere moet nak Kompel, want het lor pechtmenken. Kompel rindt ze eiperlijk hakemaal tiel uurdig.

Als nieland or hem kot, wiet Kompel or veretting mui. Hij poolt door woden en sprakt over notten. Op de vezen ontremt hij vasio's en brolkaatsen.

Het er gezanderen en nak kardt het 's elands iel prol stieg tenkel. Kompel is ok ongevonkt een vlank aart van de knakerij afgenaasd. Meedoe en ijkpoel wiet hij terstukke rankuit op sproek lissen.

1. Inleiding

"Ik heb gemerkt dat de meeste leerlingen van de 5^e en 6^e klas, geweten, altaar, pleiten', niet kennen. En als ik ze er op wil wijzen ze het opschrijven, zijn ze na vier weken weer vergeten. Toch heeft het juist voordeel bij geschiedenis heel veel geleerd. Maar ik heb geforceerd werk moeten doen."

"Die geschiedenislessen heb ik helemaal niet meer te de klas. Die kost veel te veel tijd."

"Ik geef maar vooral multiple choice vragen, want daarmee een antwoord formuleren kun je ze niet vragen."

Met bovenstaande uitspraken wijzen docenten in het voortgezet onderwijs (VVO) op een aantal problemen waarmee zij worstelen en laten ze zien hoe ze proberen om voor deze problemen te zorgen. Het gaat er om dat docenten niet blind specifiek en concreet oplossingen zoeken, maar dat ze hun handvatten als eerste taal in

Kompel er een wasche dashing en zijn moos er
 taster. Hij bak Kompel ielmij's pce, als hij
 xommels bezult. De kas van de taster roost op
 een kwitte geuze knakstij. Daar wiet hij ok massen
 met zijn taster voor het wust boez.
 Bijlage 1: Voorbeeld Nonsens Brus

lael	soeglij	Op de knakstij fiet Kompel stee amocstgegrics	
vas	lathen	oel: Zwimpten en Bator. Die kobbelen guiter	
deek	enkelkheien	poet nak Kompel, want het kor peckimken. Kompel	
min	knifrud	hand ze cipetijk bakemaal tiel yunij	
gog	kepp		
vek	chol	Als nieland-or nam kor, wiet Kompel or verneeling	
minf	negeel	miel. Hij poel boel woden en spakt over noeten.	
noegol	gloepij	Op de veeze ontrent hij vanto's en plokkasen.	
aven	zakeg		
tal	ruyemen		
noettij	verleving	Het er gezachten en nak kurt het je klinds iel	
zaper	gewaal	kol stieg tenkij Kompel is ok ongewaert een	
hine	vernoeten	vlak wiet van de knakstij afgensat. Miedoord en	
nulhak	lagge	ijlpoel wiet hij tevelde tenkint op spook lissen.	
kand	beuathal		
blakken	nakenken	malenind	malenind
verpeiden	noelen	malenind	malenind
vermaalen	kwivig	baakrak	baakrak
kanden	vaakvoeg	pepelen	pepelen
wichten	weijmaen	maekooch	maekooch
seebatten	afwerpen	wienobening	wienobening
vicut	kroes	leijbeg	leijbeg
kwachdeer	verwijlen	leijleij	leijleij
poeding	pechtbelken	schijp	schijp
laar	wonen	verwelp	verwelp
steend	loeden	noekelen	noekelen
hompellen	spieselen	schep	schep
ruifvek	kwilkegel	verthel	verthel
schrevels	muur	schelbork	schelbork