

Effecten van een leergang begrijpend lezen

E. BOL en M. A. GRESNIGT
*Vakgroep Psychonomie Rijksuniversiteit
Utrecht*

Samenvatting

Op grond van theoretische inzichten ontleend aan de cultuurhistorische theorie van Vygotskij en modellen van Lompscher en Bol is in nauwe samenwerking met scholen een experimentele leergang voor het begrijpend lezen van zakelijke teksten ontwikkeld. Het experimentele onderwijs van deze leergang is gericht op de optimalisering van met het lezen samenhangende vaardigheden op het gebied van communicatie en denken. Om het effect van de leergang te meten hebben in totaal 4 klassen van 2 verschillende scholen het experimentele onderwijs van het derde t/m het zesde leerjaar systematisch gevolgd. In dit artikel brengen we verslag uit van de belangrijkste kwantitatieve resultaten van dit onderzoek. Het blijkt dat op beide experimentele scholen na vier jaar experimenteel onderwijs aan de hand van door ons zelf ontwikkelde toetsen significante positieve effecten kunnen worden gemeten – in vergelijking met een controleschool. Resultaten die door de experimentele klassen aan het eind van de zesde klas op diverse eindtoetsen basisonderwijs (CITO en Algemeen Proefwerk) zijn behaald, maken het aannemelijk dat de gemeten effecten in de voor het experimentele programma relevante onderdelen van die toetsen doorwerken.

1 Inleiding

In het onderwijs wordt vaak de klacht gehoord dat veel leerlingen er niet in slagen een aanvaardbaar leesniveau te bereiken. Een van de voornaamste problemen is dat leerlingen maar weinig relevante informatie weten te putten uit leerboekteksten, zodat zij er maar zeer ten dele in slagen kennis en denken verder te ontwikkelen door middel van het

lezen van teksten. En daarmee raakt één van de belangrijkste waarden van het kunnen lezen in de knel.

De grootste knelpunten liggen doorgaans niet op het terrein van het technisch lezen of de uitvoering van de noodzakelijke taalkundige operaties, al spelen deze aspecten van het leesproces bij de problematiek zeker een rol. Het kernprobleem is volgens ons dat een aantal voor het lezen onontbeerlijke cognitieve en sociale basisvaardigheden onvoldoende tot ontwikkeling komt. Het gevolg is dat leerlingen er veelal niet toe komen om grotere verbanden binnen teksten te leggen. Zo bleek uit oriënterend onderzoek (Bol, Carpay en Simons, 1971) dat nogal wat leerlingen van brugklassen als het ware zin voor zin lezen, dus zonder dat zij veel verband zien tussen de verschillende onderdelen van een tekst.

In het onderwijs worden de lessen in het begrijpend lezen vaak gekenmerkt door leesopdrachten, zoals het beantwoorden van vragen, het aangeven van kernzinnen, het maken van een samenvatting, het bedenken van een geschikte titel e.d. Het is aannemelijk dat deze opdrachten zijn bedoeld om de leerlingen te oefenen in het leggen van grotere verbanden binnen teksten. Het is daarbij niet altijd duidelijk op grond waarvan de opdrachten bij zouden dragen tot het met inzicht leren lezen. Een ander bezwaar is dat het leesonderwijs vaak gericht is op het rechtstreeks ontlokken van leesprestaties, zonder dat systematische aandacht wordt besteed aan de meer fundamentele cognitieve operaties die aan die prestaties ten grondslag liggen. Het is daarom verklaarbaar dat leerlingen de grond onder hun voeten dreigen te verliezen en niet in staat zijn de opdrachten op een bevredigende wijze uit te voeren.

Deze tekortkomingen zijn wellicht mede het gevolg van het feit dat niet precies bekend is wat er geleerd moet worden, dus om welke cognitieve vaardigheden het gaat. Om hierop meer zicht te krijgen hebben we op basis van theoretische inzichten ontleend aan o.a. de taalhandelingstheorie, de cultuurhistorische theorie van Vygotskij en modellen van Bol (1982) en Lompscher (1975) een

analyse gemaakt van de mentale operaties die noodzakelijk zijn om de tekstgegevens in hun verband te kunnen begrijpen. Het centrale uitgangspunt van deze theoretische analyse is dat het lezen van een zakelijke tekst een activiteit is waarbij de lezer tracht te volgen wat de ander (de schrijver, of degene naar wie de schrijver verwijst) heeft waargenomen, gedacht of gedaan. Dit houdt in dat de lezer in staat moet zijn de tekst te beschouwen als het product van een doelbewust denkende en handelende persoon. Bovendien moet de lezer vertrouwd zijn met allerlei methoden van onderzoek en beschrijving die door een schrijver worden gehanteerd en die corresponderen met de aard van tekstverbanden. Dit betreft b.v. het beheersen van logische relaties (causaliteit, middeldoelrelaties), redeneervormen, het maken van 'voorstellingen' van hetgeen wordt beschreven (o.a. analyse), het groeperen en structureren van gegevens en deze in verband brengen met bestaande kennis (vergelijking, classificatie, definitie).

Op grond van deze theoretische analyse is een experimentele leesmethode voor de basisschool ontwikkeld. Daarin staat de opvatting centraal dat leesopdrachten pas tot optimale resultaten kunnen leiden als eerst de relevantie basisvaardigheden op het gebied van communicatie en denken op een voldoende hoog peil zijn gebracht aan de hand van meer praktische taken, waarbij het uitwendig handelen als uitgangspunt wordt gekozen. Dit sluit aan bij de opvattingen van Galperin (vgl. Van Parreren en Carpay, 1980). De kinderen leren de betreffende vaardigheden beheersen door middel van het verrichten van onderzoek, waarbij gebruik wordt gemaakt van schemata.

Kinderen van een tweetal basisscholen hebben de gehele experimentele leergang van het derde t/m het zesde leerjaar systematisch gevolgd om de effecten van de leergang in de praktijk te kunnen meten. In dit artikel zullen we verslag uitbrengen van de voornaamste kwantitatieve resultaten van dit onderzoek. We zullen hier geen uitgebreide beschrijving geven van de theoretische grondslagen en de didactische uitwerking van het experimentele programma (zie hiervoor Bol, 1982; Bol, Gresnigt en Vroege, 1982; Bol en Gresnigt, 1983).

2 Experimentele gegevens

De leergang bestaat uit 5 'blokken'. Elk blok bestaat uit een handleiding voor de docent en een leerlingenboek met allerlei opdrachten, spelletjes, teksten e.d. De lessen zijn gedetailleerd uitgewerkt. Ze vereisen veel voorbereiding, zeker als een docent nog niet vertrouwd is met het materiaal.

Blok 1 (gepland voor klas 3) bevat een oriëntering op de functie van gesproken en geschreven taal, waarbij de kinderen o.a. hun eigen taalgebruik onderzoeken. De nadruk ligt daarbij op een verkenning van het communicatieve kader waarbinnen geschreven teksten functioneren. In blok 2 (gepland klas 3/4) leren de kinderen een analysemethode hanteren met behulp waarvan allerlei zaken systematisch kunnen worden onderzocht en beschreven. Het model moet de kinderen helpen een duidelijk beeld van in teksten beschreven zaken op te bouwen. In blok 3 (gepland klas 4) wordt op basis van het voorafgaande geleerd objecten systematisch met elkaar te vergelijken. Daarbij komen tevens denkoperaties aan de orde die op het vergelijken gebaseerd zijn, zoals het ordenen van gegevens, het classificeren, het definiëren en bepaalde redeneervormen. In blok 4 (gepland klas 5) gaan de leerlingen zich bezighouden met een nader onderzoek van veranderingsprocessen. Aan de orde komen diverse patronen van verandering (zoals b.v. fasen in een ontwikkeling) die in zakelijke teksten vaak terugkeren. In blok 5 (gepland klas 5/6) ten slotte leren de kinderen omgaan met verklaringen voor verschijnselen. Daarbij wordt de aard van causale en redengevende relaties verkend, alsmede de logische regels die daarbij van toepassing zijn.

Voor elk van deze 5 blokken is een parallelle voor- en natoets ontworpen, dus in totaal 10 toetsen. Iedere-toets is opgebouwd uit 20 meerkeuzevragen met 4 alternatieven. In de toetsitems moeten opdrachten worden uitgevoerd aan de hand van korte tekstjes. Daarbij wordt niet rechtstreeks gevraagd naar begrippen die in de blokken aan de orde komen. Het gaat steeds om nieuwe toepassingen van het geleerde. De opgaven zijn zodanig gemaakt dat ze ook foutloos kunnen worden beantwoord door leerlingen die de leergang niet hebben gevolgd¹.

De toetsen zijn bedoeld om na te gaan in

hoeverre de leerlingen de schemata, die in de leergang aan de orde komen, beheersen op mentaal niveau (zie Bol, 1982). Langs deze weg kan worden bepaald of het mogelijk is door middel van de leergang de beheersing van de schemata door leerlingen te manipuleren. Indien blijkt dat experimentele leerlingen tot een hogere graad van beheersing gebracht kunnen worden dan bij controleleerlingen het geval is, kan vervolgens worden nagegaan of dit effect heeft op leesprestaties gemeten volgens externe leestoetsen (zie ook Bol en Verhelst, 1985). Hierbij wordt gebruik gemaakt van de CITO-toets en het Algemeen Proefwerk van het SAC te Utrecht. In de opzet van het onderzoek vormt de beheersing van schemata de onafhankelijke variabele en de prestaties op externe leestoetsen de afhankelijke variabele.

De eerste lessen van het experimentele programma werden in het schooljaar 1978-1979 gegeven in twee derde klassen van verschillende scholen. Deze leerlingen zullen wij, samen met een groep van twee derde klassen van een controleschool, aanduiden met groep 1. De twee experimentele klassen van groep 1 geven we aan met 'ex. 1.1' resp. 'ex. 1.2'. In het daaropvolgende schooljaar 1979-1980 werd gestart met een tweede groep leerlingen van het derde leerjaar. Deze leerlingen zullen we aanduiden met 'groep 2'. Deze groep was op overeenkomstige wijze samengesteld en afkomstig van dezelfde scholen als groep 1, met dit verschil dat er 3 controleklassen waren in plaats van twee. De beide experimentele klassen van groep 2 zullen wij aanduiden met 'ex. 2.1' en 'ex. 2.2'. De experimentele klassen 1.1 en 2.1 waren opvolgende klassen van dezelfde school. Dit was ook het geval met de klassen 1.2 en 2.2.

Het programma werd dus in totaal door 4 klassen gevolgd. Deze klassen hebben gedurende een periode van 4 jaar (van de derde t/m de zesde klas) experimenteel onderwijs gehad. Er werd in principe één lesuur per week gespendeerd aan de leergang. Dit is 2 à 3 % van de totale schooltijd.

Voordat met het experiment werd gestart werden enkele bijeenkomsten georganiseerd waarbij het gehele experimentele programma werd besproken met de betreffende docenten. Daarnaast was er enkele malen per jaar een bespreking, in elk geval bij de start van

een nieuw blok. In deze gesprekken werden de grote lijnen van het programma nader toegelicht. Het materiaal werd tevoren door alle betrokkenen kritisch doorgenomen.

Bij de experimentele klassen van groep 1 werden alle experimentele lessen door een medewerker van het project bijgewoond. Indien nodig of gewenst was er een korte voor- en nabespreking met de docent. Bij deze gesprekken was er de gelegenheid positieve en negatieve punten van het programma te bespreken en werden eventueel voorstellen tot een bepaalde aanpak gedaan, mede gezien de specifieke omstandigheden in de klas. Daarbij ging het evenwel alleen om de soorten van voorbeelden, de hoeveelheid oefening e.d., de 'harde kern' van het programma mocht niet worden aangetast. Op deze manier werd de aard en de volgorde van de leerstappen in de verschillende experimentele groepen gelijk gehouden.

Bij de aanvang van het programma voor de experimentele klassen van groep 2 (in het daarop volgende jaar) werd besloten in deze klassen niet meer te observeren en slechts incidenteel contact te onderhouden met de leerkrachten. Dit gaf de mogelijkheid na te gaan of het programma ook zou werken zonder intensieve begeleiding.

Tabel 1 geeft weer in welke klassen de diverse blokken zijn behandeld. Tevens is aangegeven de tijdsperiode in maanden tussen de eerste en de laatste les van een bepaald blok. Over het algemeen verliepen de blokken iets langzamer dan verwacht.

Tabel 1 *Overzicht van planning en feitelijke uitvoeringstijd van het experiment*

Blok:	Uitvoering gr. 1 (1978-1982)		Uitvoering gr. 2 (1979-1983)	
	klas:	duur:	klas:	duur:
1	3	6 mnd.	3	7 mnd.
2	3/4	14 mnd.	3/4	13 mnd.
3	4/5	11 mnd.	4/5/6	19 mnd.
4	5/6	8 mnd.	—	—
5	6	6 mnd.	6	6 mnd.

Uit Tabel 1 blijkt dat de groepen 1 en 2 ongeveer evenveel tijd nodig hadden voor de blokken 1, 2 en 5. Bij blok 3 was er echter een groot tijdsverschil van meer dan een half jaar. Dit was het gevolg van vertragingen in het programma bij groep 2.

Vanwege deze vertraging was het niet meer haalbaar de resterende blokken 4 en 5 bij groep 2 in hun geheel te behandelen. In overleg met de scholen werd daarom besloten om blok 4 grotendeels te laten vervallen en verder te gaan met blok 5. Wat groep 2 betreft zijn – met uitzondering van blok 4 – alle toetsen afgenomen. Bij groep 1 werden bij de blokken 1 en 2 alleen de natoetsen afgenomen, omdat de voortoetsen nog niet beschikbaar waren. Bovendien bleek bij groep 1 niet mogelijk bij de controle-klassen een natoets af te nemen bij blok 5.

3 Resultaten

3.1 Interne toetsen

In Tabel 2 staan de betrouwbaarheidsgegevens (Lambda-2) van de door ons zelf ontwikkelde voor- en natoetsen vermeld. De betrouwbaarheid is niet hoog, maar acceptabel als in aanmerking wordt genomen dat de toetsen uit slechts 20 items bestaan. Een uitzondering vormt de zeer lage betrouwbaarheid van de voortoets van blok 3, wat groep 2 betreft, terwijl ook de betrouwbaarheid van de voortoets van blok 1 vrij laag uitvalt.

Tabel 2 *Overzicht van de betrouwbaarheid van de toetsen*

blok	groep	voortoets			natoets		
		klas	N	betr.	klas	N	betr.
1	1	3	–	–	3	95	.70
	2	2	144	.64	3	134	.72
2	1	3	–	–	4	107	.81
	2	3	140	.79	4	138	.79
3	1	4	109	.75	5	104	.71
	2	4	153	.45	6	161	.77
4	1	5	106	.72	6	102	.71
	2	–	–	–	–	–	–
5	1	6	108	.72	6	62	.79
	2	6	154	.75	6	157	.79

De oorzaak van de lage betrouwbaarheid van de voortoets van blok 3 is waarschijnlijk dat de toets te moeilijk is (zie Tabel 3). Omdat de voortoets van blok 3 door een aantal leerlingen van groep 1 toch al zeer goed werd gemaakt waren we bang voor een plafond-effect op de natoets. Daarom is de toets moeilijker gemaakt. Het blijkt evenwel dat

de toets te moeilijk is uitgevallen, waardoor de scores van de minder goede leerlingen van het toeval afhankelijk zijn. Over de validiteit van de toetsen merken we hier slechts op dat de toetsen niet programmaspecifiek zijn. Voor de behandeling van deze problematiek zie Bol en Verhelst (1985).

Tabel 3 geeft de gemiddelden en standaarddeviaties van groep 1 voor elke toets. De scores op de natoetsen zijn geanalyseerd door middel van een covariantieanalyse (SPSS- Anova) met (waar mogelijk) de scores op de voortoetsen als covariaat. Er zijn alleen scores opgenomen van kinderen die alle toetsen hebben gemaakt en dus het gehele programma hebben doorlopen, zodat de berekeningen niet zijn beïnvloed door een eventuele selectieve uitval van proefpersonen. Bij groep 1 bestaat de experimentele groep dan uit 29 kinderen en de controlegroep uit 31 kinderen. Voor groep 2 is dit 27 resp. 34. Andere berekeningswijzen geven geen fundamenteel ander beeld te zien.

Tabel 3 laat zien dat het verschil in gemiddelde tussen de experimentele en de controlegroep bij blok 1 statistisch significant is. Omdat een voortoets ontbreekt is het niet duidelijk in hoeverre dit verschil kan worden toegeschreven aan het experimentele programma. Ook bij blok 2 ontbreken, wat groep 1 betreft, voortoetsgegevens. Hier kunnen evenwel de scores op de natoets van blok 1 als covariaat dienen. Dan blijkt dat de experimentele groep significant hoger scoort. Bij de blokken 3 en 4 is wel een voortoets afgenomen. Bij beide blokken wordt een significant hoofdeffect gemeten in het voordeel van de experimentele groep. Bij blok 5 beschikken we wat groep 1 betreft niet over de natoetsgegevens van de controlegroep. Het blijkt dat de experimentele groep op deze natoets hoog scoort (het percentage goed gescoorde items bedraagt bijna 87). Wanneer de gegevens van de controlegroep van groep 2 worden benut kan een significant hoofdeffect worden gemeten in het voordeel van de experimentele groep 1.

De natoetsen werden gelijktijdig afgenomen met de voortoetsen van het volgende blok. Zoals te verwachten is levert een variantieanalyse van de voortoetsen met de gelijktijdig afgenomen natoetsen als covariaat in geen enkel geval een significant hoofdeffect op. Dit geldt ook voor groep 2.

Tabel 3 *Overzicht van de resultaten van groep 1 (n exp = 29; n contr. = 31)*

blok	cond	voortoets		natoets		covariaat		hoofdeffect		R ²
		m	sd*	m	sd*	F	P	F	P	
1	ex	-	-	13,6	3,0	-	-	10,2	.002	-
	co	-	-	11,0	3,3					
2	ex	-	-	17,3	2,7	25,6	.001	21,3	.001	.45
	co	-	-	12,6	3,5					
3**	ex	13,9	3,0	11,3	2,8	39,8	.001	27,0	.001	.54
	co	10,4	3,9	6,8	2,0					
4	ex	13,0	2,9	15,6	2,1	36,2	.001	20,2	.001	.50
	co	10,3	3,5	11,8	2,8					
5	ex	15,0	2,6	17,3	2,0	-	-	-	-	-
	co	12,1	3,2	-	-					

* varianties zijn homogeen

** van dit blok is de natoets moeilijker gemaakt.

In Tabel 4 staan de gemiddelde somscores van de voor- en natoetsen van groep 1. We gebruiken hierbij alleen toetsen van blokken waarvan zowel de voor- als natoetsen bekend zijn, dus wat groep 1 betreft de voor- en natoetsen van de blokken 3 en 4, dus 40 items. De betrouwbaarheid (Lambda-2) van de gesommeerde voortoetsen is .79 en van de gesommeerde natoetsen .80 (60 ppn).

Tabel 4 *Gemiddelde somscores van de voor- en na-toetsen in de experimentele en controleklassen van groep 1, uitgedrukt in percentages goed gemaakte items*

	exp. (n = 29)	contr. (n = 31)	P
gem. som voortoetsen	67,3	51,8	.001
gem. som natoetsen	67,3	46,5	.001
verschil	0	- 5,3	

Het blijkt dat de scores op de natoetsen voor de controlegroep iets lager uitvallen dan die van de voortoets. Dit komt omdat de natoets van blok 3 moeilijker is gemaakt. De experimentele leerlingen scoren zowel op de voor- als op de natoetsen hoger. Maar op de natoetsen is het verschil groter, hetgeen wijst op een leereffect bij althans 2 van de 5 blokken van groep 1.

Over de data samengevat in Tabel 4 is een covariantieanalyse uitgevoerd (zie Tabel 5).

Er blijkt een goede samenhang tussen de toetsen en er is een zeer significant hoofdeffect.

Tabel 5 *Covariantieanalyse van de somscores voor- en natoetsen van groep 1*

	SS	F	P
covar. som vt.	1051,2	90,4	.001
hoofdeffect	284,5	24,5	.001
residu	662,9		
totaal	1998,6		

Toets homogeniteit varianties R = .82
 Bartlett-Box F = .000; P = .989 R² = .67

Tabel 6 geeft een overzicht van de resultaten per blok van groep 2. De scores op de natoetsen zijn geanalyseerd d.m.v. een covariantieanalyse, met de voortoetsen als covariaat. Bij groep 2 is de moeilijkheidsgraad van de voor- en natoetsen zoveel mogelijk gelijk gehouden.

Als we kijken naar het begin-niveau van de kinderen vóórdat het experimentele onderwijs startte dan blijkt dat de experimentele groep niet significant hoger scoort (F = .110; p = .741). Op de natoets is er wel een hoofdeffect. Ook bij blok 2 wordt een beter resultaat op de natoetsen behaald door de experimentele groep. Bij de blokken 3 en 5 wordt geen significant hoofdeffect gemeten.

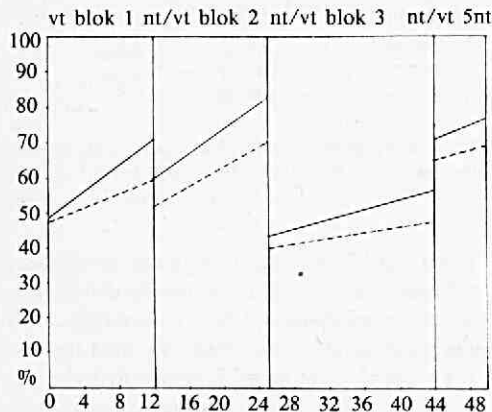
Figuur 1 geeft de gemiddelden van de experimentele en controlegroep van groep 2 in

Tabel 6 *Overzicht van de resultaten per blok van groep 2 (n ex=27; n co=34)*

Blok	cond.	voortoets		natoets		covariant		hoofdeffect		R ²
		m	sd*	m	sd*	F	P	F	P	
1	ex	9,7	3,1	14,0	3,0	17,2	.001	6,9	.011	.29
	co	9,4	2,9	11,9	3,4					
2	ex	11,9	4,3	16,6	3,1	26,0	.001	5,0	.029	.35
	co	10,3	3,9	14,0	3,7					
3	ex	8,7	3,0	11,3	4,0	13,6	.001	2,1	.154	.21
	co	8,0	2,2	9,5	3,9					
5	ex	14,2	2,6	15,4	3,2	48,4	.001	0,9	.348	.46
	co	13,0	3,1	13,9	3,5					

* alle varianties zijn homogeen per toets.

een grafiek weer, horizontaal het aantal maanden tussen de meetmomenten, vertikaal het percentage goed gemaakte items per toets. We zien dat de blokken 3 en 5 niet zo goed passen in het beeld van de eerste 2 blokken. Mede gezien de ervaringen in de klassen lijkt het verstandig de volgorde van de blokken te wijzigen.



tijd tussen de toetsen in maanden

opm.: De voortoets van blok 1 is enkele maanden vóór het begin van de lessen afgenomen.

Figuur 1 *Percentage gemiddeld goed beantwoorde items op de voor- en natoetsen van groep 2*

— = experimentele groep (n = 27)
 - - - = controlegroep (n = 34)

In Tabel 7 staan de gemiddelde somscores

voor groep 2 over 4 voor- en natoetsen. Dit betreft dus 80 items elk. De betrouwbaarheid (Lambda-2) bedraagt voor de gecombineerde voortoetsen .86 en voor de natoetsen .92 (n = 61).

Tabel 7 *Gemiddelde somscores van de voor- en natoetsen van de experimentele en controlegroep, in percentages goed beantwoorde items*

	exp. (n = 27)	contr. (n = 34)	P
gem. som voortoetsen	55,6	50,9	.125
gem. som natoetsen	71,6	61,6	.009
verschil	16,0	10,7	

Op de voortoetsen scoren de experimentele leerlingen wel hoger, maar niet significant. Dat doen ze wel op de natoetsen. Ook de controleleerlingen scoren op de natoetsen hoger, wat ontwikkelingspsychologisch te verwachten is. Maar de experimentele leerlingen gaan meer vooruit. Als we ons realiseren dat een deel van het leereffect ook in de voortoetsen zit dan is het geconstateerde verschil in toename zeker niet geflatteerd.

In Tabel 8 is een covariantieanalyse uitgevoerd over deze gegevens. Er blijkt wederom een goede samenhang tussen voor- en natoetsen. Het hoofdeffect is significant. Zowel bij groep 1 als groep 2 kan dus een leereffect worden gemeten.

Tabel 8 *Covariantieanalyse van de somscores voor- en natoetsen van groep 2*

	SS	F	P
covar. som vt.	5425,99	109,1	.001
hoofdeffect	275,91	275,9	.002
residu	2885,12		
totaal	8587,02		

Toets homogeniteit variantie $R = .82$
 Bartlett-Box $F = .08$ $P = .78$ $R^2 = .66$

3.2 Eindtoetsen van het basisonderwijs

In deze paragraaf gaan we na of er aanwijzingen te vinden zijn dat de experimentele leerlingen beter presteren op de leesonderdelen van de externe eindtoetsen van het basisonderwijs. Op de drie betrokken scholen worden helaas verschillende eindtoetsen gebruikt.

In de experimentele klassen ex. 1.1 en ex. 2.1 is de CITO-toets afgenomen in 1982 resp. 1983. Van klas ex. 2.1 beschikken we niet over de relevante gegevens. In de overige experimentele klassen ex. 1.2 en ex. 2.2 is het Algemeen Proefwerk (SAC- Utrecht) gemaakt in 1982 en 1983. Op de controleschool zijn ISI-toetsen afgenomen.

Omdat vergelijking met de controle klassen niet mogelijk is wegens deze verschillende toetsen hebben we vergeleken met voorgaande jaarklassen van de experimentele scholen, die de leergang niet hebben gehad. Bij de eindtoetsen zijn die onderdelen waarvan kan worden aangenomen dat zij nauw samenhangen met hetgeen in de experimentele leergang aan de orde komt bij elkaar opgeteld tot een totaalscore. Wat de CITO-toets betreft zijn dit de onderdelen 'lezen van teksten: reproductie', 'lezen van teksten, conclusies', 'betekenis en betekenis-relatie' en 'vraagstukjes'. De scores van de 'experimentele' klas 1.1 (1982) zijn vergeleken met die van de jaarklas van 1981 (de toetsen van vóór die tijd zijn niet geheel vergelijkbaar). Om voor een mogelijk niveauverschil dat reeds vóór het experimentele onderwijs bestond te compenseren zijn onderdelen waarvan kan worden aangenomen dat ze niet zijn beïnvloed door de leergang als covariaat genomen. Dit waren rekenen (het totaal van de onderdelen 'getallen', 'hoofdrekenen',

'bewerkingen', 'procenten', 'meten' en 'verhoudingen'), grammatica (het totaal van de onderdelen 'woordvorming' en 'grammaticale bouw van de zin'), spelling (bestaande uit 'spellen van werkwoorden' en 'spellen van niet-werkwoorden') en informatieverwerking (het totaal van 'kaartlezen' en 'lezen van tabellen en grafieken'). Uit een hiërarchische covariantieanalyse bleek dat de covariaten spelling en informatieverwerking geen extra bijdrage geven bij hetgeen reeds wordt verklaard door rekenen en grammatica. Er blijft een significant hoofdeffect in het voordeel van de experimentele groep over (zie Tabel 9). De gemiddelden van klas 1.1 en de controlegroep van 1981 staan in Tabel 10. De gemiddelden zijn uitgedrukt in z-scores, waarbij alle CITO-gegevens van 1981 en 1982 van de regio Gooi zijn betrokken. De voor de covariaat gecorrigeerde waarde is gemiddeld één-derde standaarddeviatie hoger bij de experimentele leerlingen van klas 1.1 in vergelijking met de controleklas.

Tabel 9 *Covariantieanalyse van de scores van experimentele klas 1.1 (1982) en de controleklas (1981, van dezelfde school) op de CITO-eindtoets*

Bron	SS	DF	F	P
covariaat rekenen	11,0	1	58,3	.001
covariaat grammatica	2,7	1	14,1	.001
hoofdeffect	1,5	1	7,8	.007
error	9,85	52		

De covariaten en het hoofdeffect verklaren 61% van de variantie.

In de experimentele klassen 1.2 en 2.2 is in de jaren 1982 resp. 1983 het Algemeen Proefwerk afgelegd. De gegevens van 1981 zijn niet gebruikt omdat in deze klas een groot deel van de leergang is uitprobeerd. In plaats daarvan werden de gegevens van 1979 en 1980 gebruikt. Het Algemeen Proefwerk bestaat uit de onderdelen 'begrijpend lezen', 'vraagstukjes', 'hoofdrekenen', 'cijferen', 'spelling' en 'taal'. Het onderdeel taal bevat zeer diverse opgaven, die soms aansluiten bij het experimentele programma (b.v. betekenisrelaties) en soms niet (b.v. grammatica). Omdat slechts één totaalscore bestaat is tevoren besloten dit onderdeel buiten beschouwing te laten. De onderdelen 'begrij-

pend lezen' en 'vraagstukjes' konden worden geacht nauw samen te hangen met het experimentele programma. Hiervan werden gecombineerde scores berekend. De scores van de leerlingen die in 1982 en 1983 de toets hadden gemaakt (ex. 1.2 en ex. 2.2) en die van de leerlingen van 1979-1980 werden geanalyseerd door middel van een covariantie-analyse (zie Tabel 11). Het Algemeen Proefwerk werkt met standaardcores met 20 als gemiddelde (zie Tabel 10). Als covariaat werd gebruikt rekenen (= 'hoofdrekenen' plus 'cijferen') en spelling.

De analyse laat een significante bijdrage van de beide covariaten en een hoofdeffect ten gunste van de experimentele groep zien. Het percentage verklaarde variantie bedraagt 63.

Tabel 10 *Gemiddelden en voor covariaten aangepaste gemiddelden van experimentele en controleklassen op relevante onderdelen van schooltoetsen (zie tekst)*

groepen	N	M	M aangepast voor covariaten
ex 1.1	28	.91 (z scores)	.91
co 1981	28	.57	.58
ex 1.2 + 2.2	83	42,1	43,3
co 1979 + 1980	84	42,3	41,2

Tabel 11 *Covariantie-analyse van de scores van de experimentele klassen 1.2 en 2.2 en de controleklassen (1979 en 1980, van dezelfde school) op Algemeen Proefwerk eindtoets (zie tekst)*

bron	SS	df	F	P
covariaat rekenen	9396,0	1	134,9	.001
covariaat spelling	376,6	1	10,8	.001
hoofdeffect	178,7	1	5,1	.025
error	5675,7	163		

3.3 Sterkte van de effecten

In 1983 is op een zestal mavo-scholen een toets afgenomen, bestaande uit een selectie van 40 items uit alle door ons gebruikte toetsen. Deze toets is ook afgenomen bij de experimentele en controle-leerlingen van groep 2 (toen klas 6) aan het einde van het experimentele programma. We kunnen de scores

van deze zesdeklassers vergelijken met die van de diverse mavo-klassen. Uit Figuur 2 blijkt dat de experimentele zesdeklassers ongeveer op het niveau van klas 3 mavo scoren. Het niveau van de mavo-klassen is misschien niet helemaal vergelijkbaar. De controlezesde-klassen scoren immers ook hoger dan klas 1 mavo. Verder kan hebben meegespeeld dat de zesdeklassers vertrouwd waren met de afnameprocedure. Ze hadden de items ook al een keer gemaakt.

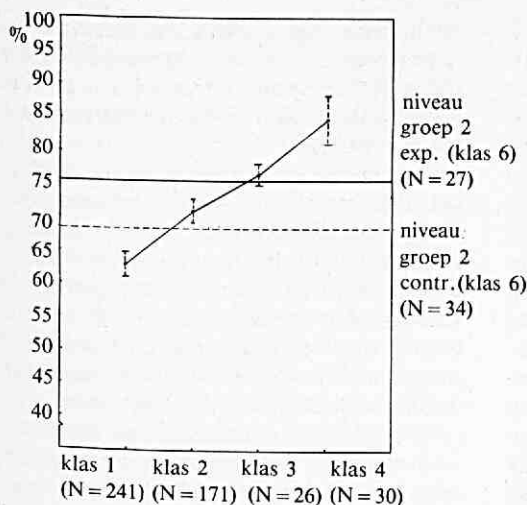
Het verschil tussen de gemiddelde scores van de experimentele en controlezesdeklassers op deze 40 item toets geeft een schatting van het totale effect van de leerang over 4 jaar (klas 3-6). Ter compensatie van aanvangsverschillen gebruiken we de scores op de voortoets van blok 1 (klas 3) als covariaat. Dan blijkt dat de experimentele zesdeklassers significant hoger scoren. De voor de covariaat aangepaste leerwinst bedraagt 8,4%, (d.w.z. de experimentele groep maakt 8,4 % méér items goed). Een overeenkomstige vergelijking van de laatst gemaakte natoets geeft een geschatte leerwinst van 7,5%, van dezelfde orde van grootte dus. De vraag is natuurlijk wat deze verbetering van de scores van de experimentele groep t.o.v. de controlegroep ontwikkelingspsychologisch betekent.

In het verleden is op de controleschool de voortoets van blok 1 gelijktijdig afgenomen in de tweede t/m de zesde klas (zie Bol, 1979). De resultaten staan afgebeeld in Figuur 3. Er blijkt een vrij regelmatige jaarlijkse stijging van de scores in de orde van 7 à 8% op te treden. Kijken we naar de scores in de diverse mavo-klassen, dan blijkt een regelmatige jaarlijkse stijging van ongeveer 7% op de 40-item toets. Omdat noch in de controleschool, noch in de mavo-scholen experimenteel onderwijs is gegeven kan de stijging in prestaties op deze toetsen over de leerjaren worden toegeschreven aan het gewone onderwijs en de cognitieve ontwikkeling van de kinderen. Dit geeft de indicatie dat de ontwikkeling die de kinderen doormaken zonder experimenteel onderwijs verantwoordelijk is voor een verbetering van de toetsprestaties op onze toetsen van 7-8 % per jaar. We zullen deze gegevens in de discussie bespreken.

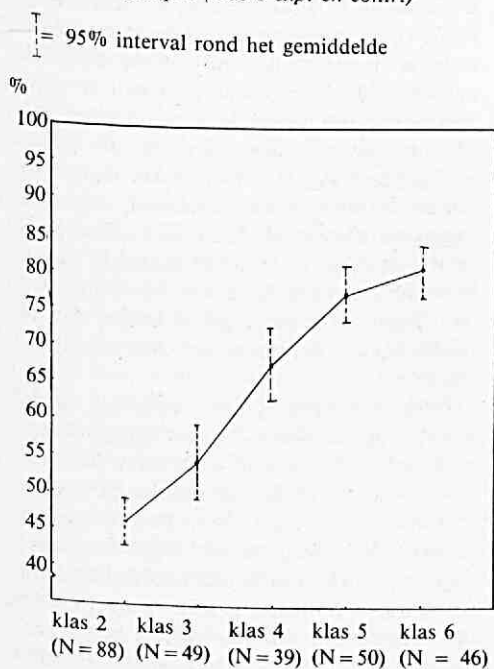
4 Discussie

De eerste kwestie die we aan de orde willen stellen heeft betrekking op het effect van het experimentele programma op de kennis en beheersing van de in de cursus behandelde schemata. Uit de Tabellen 3 en 6 blijkt dat alle blokken een positief effect opleveren. Bij het merendeel van de blokken is het hoofdeffect statistisch significant, de blokken 3 en 5 bij de experimentele groep 2 vormen hierop een uitzondering. In het algemeen lijkt groep 1 het wat beter te doen dan groep 2 wat het effect van het experimentele programma betreft. Dit kan samenhangen met het feit dat bij groep 1 een intensieve begeleiding heeft bestaan. Daarnaast kan er sprake zijn van een Hawthorne-effect in groep 1. Dit lijkt ons echter niet zo waarschijnlijk, omdat een Hawthorne-effect nu juist een vrij kortdurend effect is. In dit onderzoek zou het effect zich moeten uitstrekken over vier jaren. Waarschijnlijker lijkt ons dat het verschil begrepen kan worden als een uitvloeisel van de omstandigheid dat door een aantal oorzaken het programma in groep 2 veel minder soepel verliep dan in groep 1 het geval was. Er was nogal wat ziekte van leerkrachten, één van de klassen was een extreem lastige klas, in beide experimentele scholen was er tijdens het laatste jaar van groep 2 wat onrust in verband met de interne organisatie etc. Hierdoor zijn ook veel lessen en zelfs een heel blok uitgevallen. Desondanks vinden we, over het gehele programma gezien, bij beide experimentele groepen een significant hoofdeffect (zie Tabellen 5 en 8). Deze uitkomst wordt bevestigd door de resultaten op de selectie van 40 items die aan het eind van de zesde klas bij de groepen 2 is afgenomen. De experimentele leerlingen scoren daar significant hoger dan de controle leerlingen (zie onder 3.3). We mogen met enige voorzichtigheid concluderen dat de beheersing van de schemata op mentaal niveau door leerlingen met behulp van het programma op een hoger plan kan worden gebracht. Dit houdt in dat leerlingen door het programma leren systematischer te observeren en te lezen en dat zij ook beter toegerust worden op het leggen van verbanden tussen gegevens, resp. het plaatsen van gegevens in algemene kaders.

Het volgende punt betreft de betekenis van het effect. Absoluut gezien lijkt een leerwinst



Figuur 2 Percentage goed beantwoorde items van een selectie van 40 toetsvragen, afgenomen in de klassen 1 t/m 4 van het voortgezet onderwijs (mavo) en de klassen van groep 2 (klas 6 exp. en contr.)



Figuur 3 Percentage goed beantwoorde items van de voortoets van blok 1 bij gelijktijdige afname in de klassen 2 t/m 6 van een lagere school (1979 controleschool)

95% interval rond het gemiddelde

van 7 à 8 procent over vier schooljaren niet groot. Bekijken we evenwel de groei van de scores op de programmatoetsen bij leerlingen die het gebruikelijke leesonderwijs krijgen, waar geen systematische en expliciete aandacht wordt geschonken aan in beschrijvingen verborgen schemata, dan zien we een vrij constant beeld. Per leerjaar nemen de scores ongeveer 7 procent toe, zowel op een lagere school als op een aantal mavo-scholen (zie Fig. 2 en 3). Deze gegevens wijzen op de mogelijkheid dat het eindniveau van de experimentele leerlingen door de controleleerlingen pas ongeveer een jaar later behaald zal worden. De leerwinst bedraagt dan zowat een jaar. We moeten evenwel voorzichtig zijn met de toepassing van dwarsdoorsnede gegevens bij de prognose van een ontwikkelingsproces, al dan niet beïnvloed door onderwijs.

Kijken we naar Tabel 6 en Fig. 1 dan zien we de groei van scores in een longitudinaal onderzoek. Zowel bij de experimentele als bij de controle-groep zien we een vrij sterke groei in de beginjaren en een tragere groei in de latere jaren. We moeten hier bedenken dat bij dit onderzoek niet steeds dezelfde toets is gebruikt. Er zijn acht verschillende (4 maal 2 parallele) toetsen gebruikt. We kunnen de toename van de scores van leerlingen die geen experimenteel onderwijs krijgen per blok bekijken. De groei van de scores over de achtereenvolgende blokken is bij de controle-groep resp. 12.5%, 18.5%, 7.5% en 4.5%. Dit beeld verschilt nogal van het dwarsdoorsnede onderzoek. Verschillen in moeilijkheidsgraad van de toetsen (de toetsen van blok 3 lijken te moeilijk en die van het laatste blok te gemakkelijk) en bekendheid met de toetsprocedure en de soort items kunnen hier een rol spelen. We kunnen een gemiddeld cijfer voor de groei van de scores bij de controle-groep bepalen door het gemiddelde te nemen van bovenstaande getallen, dan wel de groei per blok om te rekenen op jaarbasis en daar het gemiddelde van te nemen. In beide gevallen krijgen we vrijwel dezelfde uitkomst, namelijk een groei van 10.8 procent per jaar. Op deze basis berekend bedraagt de voorsprong van de experimentele leerlingen op de controle-leerlingen acht à negen maanden. Nemen we de tragere groei van de scores gedurende de latere jaren in aanmerking, dan lijkt deze schatting zeker

niet te hoog. Dit resultaat lijkt niet onbevredigend wanneer we ook nog bedenken dat tijdens het onderzoek slechts 2 à 3 procent van de totale onderwijstijd aan het programma is besteed.

Er valt nog iets anders te zeggen over de omvang van het effect van het experimentele programma. Op basis van onze theoretische analyse (zie Bol, 1982) nemen we aan dat de door het experimentele programma beoogde vaardigheden nauw samenhangen met de cognitieve ontwikkeling. Dit houdt in dat het programma de cognitieve ontwikkeling van leerlingen positief moet beïnvloeden. Het spreekt dat deze doelstelling van ontwikkelend onderwijs geen simpele zaak is en dat men ook geen spectaculaire effecten mag verwachten. Onze theoretische analyse wordt ondersteund door het gegeven dat er een sterke samenhang bestaat tussen prestaties van leerlingen op onze interne toetsen en hun prestaties op toetsen voor verbale intelligentie (Otis en ISI). Onze toetsen verklaren 55 tot 75 procent van de variantie van dergelijke tests (zie Bol, Gresnigt en Vroege, 1982; Bol en Verhelst, 1985). En dit terwijl er soms meerdere jaren tussen de verschillende toetsafnames liggen. Ook bij de in dit artikel weergegeven statistische analyses vinden we dat de bijdragen van de covariaten veel hoger liggen dan die van hoofdeffecten. Dit wijst er op dat de interne toetsen betrekkelijk stabiele cognitieve kenmerken van leerlingen meten. Tegen deze achtergrond bezien zijn de resultaten van het onderzoek zeer redelijk te noemen.

Onze conclusie is dat, ondanks allerlei smetten op de uitvoering, het experimentele programma de beoogde effecten redelijk bereikt. We hebben hier de resultaten besproken van leerlingen die alle toetsen hebben gemaakt. We hebben evenwel dezelfde analyse uitgevoerd over de groepen van leerlingen die ten hoogste één toets hadden gemist en die ten hoogste twee toetsen hadden gemist. De aantallen in de analyses betrokken leerlingen stijgen dan sterk. De resultaten die gevonden werden wijken nauwelijks af van de hier gepresenteerde. We vinden dus geen aanwijzingen dat leerlingen select uitvallen. Voorts hadden de in de discussie betrokken data betrekking op groep 2, omdat daarvan de meeste gegevens bekend zijn. We hebben echter geen aanwijzingen dat de effecten van

het programma bij groep 1 sterk verschillen van die bij groep 2. We zitten natuurlijk nog wel met de vraag of de gevonden effecten niet aan andere factoren toegeschreven kunnen worden. Bijvoorbeeld doordat op de controle-school in een aantal vakken systematisch ander onderwijs wordt gegeven dan op de experimentele scholen. Echter in het vooronderzoek waren dergelijke systematische verschillen niet aanwijsbaar (zie Bol, 1979). Geen van de scholen wordt gekenmerkt door een bijzondere vorm van onderwijs (zoals Jenaplan) en alle drie de scholen staan bekend als goede scholen waar prima resultaten worden bereikt. Bovendien dekken de experimentele scholen enerzijds en de controle-school anderzijds een zeer brede leerlingpopulatie.

We stellen thans de kwestie aan de orde of het effect van het programma inderdaad samenhangt met de prestaties van de leerlingen op externe leestoetsen. We hebben reeds aangegeven dat het niet mogelijk is de scholen met elkaar te vergelijken, aangezien ze verschillende schooltoetsen gebruiken. We hebben er niet voor gekozen één van de schooltoetsen (bijv. de Cito-toets) op de andere scholen af te nemen, omdat de examensfeer waarmee de schooltoetsen gepaard gaan dan ontbreekt. Daardoor zijn de resultaten toch niet goed vergelijkbaar. Door experimentele klassen te vergelijken met voorgaande klassen van dezelfde school spelen eventuele verschillen tussen scholen geen rol. Het nadeel is dat we moeten aannemen dat vorige zesde klassen van dezelfde school lager gescoord zouden hebben op de interne toetsen. Strikt genomen weten we dat niet. We hebben daarvoor wel een indirecte aanwijzing. We kennen van de voorgaande zesde klassen de scores op de voortoets van blok 1 toen deze leerlingen resp. in de vierde, de vijfde en de zesde klas zaten (vgl. Fig. 2). Welnu, de experimentele leerlingen blijken op het moment dat ze achtereenvolgens in die klassen zaten hoger te scoren op de toetsen van latere blokken (die dus moeilijker zijn). Maar hier speelt natuurlijk wel verschil in toetservaring. Uit de Tabellen 9 en 11 blijkt dat de experimentele leerlingen op de externe leestoetsen een significant hoofdeffect vertonen. Dus volgens verwachting presteren ze in vergelijking met de controle-klassen relatief beter op de leesonderdelen van de schooltoet-

sen dan op de overige onderdelen daarvan. In groep 2 hebben we daarnaast in alle klassen nog een leestoets afgenomen, gemaakt door enige vakleerkrachten Nederlands van een mavo. Hoewel deze leestoets wat gemakkelijker bleek, scoorden de experimentele leerlingen significant hoger dan de controleleerlingen. Het verschil kon geheel worden verklaard op basis van verschillen tussen leerlingen gemeten met behulp van de interne toetsen (zie Bol en Verhelst, 1985).

Wanneer we het geheel in ogenschouw nemen, dan vinden we in de empirische gegevens een redelijke onderbouwing van de theoretische aanname dat het op mentaal niveau kunnen schematiseren van gegevens een belangrijke factor vormt in het proces van het begrijpend lezen. Daarnaast hebben we aannemelijk kunnen maken dat deze vaardigheid door middel van systematisch onderwijs verhoogd kan worden. Aan welke voorwaarden een school zal moeten voldoen om met het experimentele programma vergelijkbare resultaten te kunnen behalen is op grond van dit onderzoek niet uit te maken. Daarvoor is een anders geaard onderzoek noodzakelijk.

We maken nog enkele slotopmerkingen. Ten eerste is bij observatie in de klassen en uit gesprekken met leerkrachten en leerlingen gebleken dat het programma moeilijk is, maar dat het niet als saai of schools wordt ervaren. In de meeste lessen waren de leerlingen zeer actief en ontstonden vaak levendige discussies. Veel leerkrachten zijn na het onderzoek doorgegaan met het programma. Bovendien is het materiaal thans in tientallen scholen en schooladviesdiensten in gebruik. Ten tweede zal de leerang in de toekomst verder worden uitgebreid met voorbereidende programma's voor de onderbouw en vervolgpogramma's voor de brugklassen. Dit is nu in voorbereiding. Daarbij zal worden geprobeerd dwarsverbanden te leggen met de zaakvakken en de praktijk van de omgang met zakelijke teksten op school en daar buiten. Ook zal gestreefd worden naar een vergroting van de effectiviteit van het programma door meer variatie in de didactische aanpak, meer oefenmateriaal en meer inzicht in de specifieke problemen waarmee leerlingen te kampen hebben. Wat dit laatste betreft zal het bestaande toetsmateriaal nader worden bestudeerd. Bovendien hopen we door middel van gericht onderzoek een gede-

tailleerder inzicht te verkrijgen in de aard en de ernst van de problemen. Daartoe zullen individuele leestoetsen worden afgenomen, waarbij speciaal aandacht zal worden geschonken aan verschillen tussen leerlingen.

Noten

1 Het lesmateriaal en de toetsen kunnen na afspraak worden ingezien. Vakgroep Psychonomie, Heidelberglaan 2 te Utrecht (M.A. Gresnigt), tel 030-533022, kamer 17.06.

Literatuur

- Bol, E., Onderzoekopzet voor begrijpend lezen op de basisschool (derde t/m zesde leerjaar). In: M.J.C. Mommers en B.W.G.M. Smits (red), *Leesonderwijs in de Basisschool*. Den Haag: Staatsuitgeverij, 1979.
- Bol, E., *Leespsychologie*. Groningen: Wolters-Noordhoff, 1982.
- Bol, E., J.A.M. Carpay en R. Simons, Praktische taalkunde in het brugjaar. *Pedagogische Studiën*, 1971, 48, 119-128.
- Bol, E. en M.A. Gresnigt, Worden leerlingen in het basisonderwijs voldoende voorbereid op het

lezen van leerboekteksten in het vervolgonderwijs? *Moer*, 1983, 3-4-5, 176-190.

- Bol, E., M.A. Gresnigt en A.P.J. Vroege, Begrijpend lezen: communicatie en denken. *Pedagogische Studiën*, 1982, 59, 261-276.
- Bol, E., en N.D. Verhelst, Inhoudelijke en statistische analyse van een leestoets. *Tijdschrift voor onderwijsresearch*, 1985, 2, 49-68.
- Lompscher, J., *Theoretische und experimentelle Untersuchungen zur Entwicklung geistiger Fähigkeiten*. Berlin: Volk und Wissen, 1972.
- Parreren, C.F. Van en J.A.M. Carpay, *Sovjetpsychologen over onderwijs en cognitieve ontwikkeling*, Groningen: Wolters-Noordhoff, 1980.

Curricula vitae

E. Bol studeerde psychologie aan de Rijksuniversiteit te Utrecht (doctoraal 1968); in 1974 promoveerde hij op het proefschrift 'Psychologie en cybernetica: een leertheoretische verkenning'.

M.A. Gresnigt studeerde psychologie aan de Rijksuniversiteit te Utrecht, hoofdvak psychologische functieleer (doctoraal 1978).

Adres: Vakgroep Psychonomie Rijksuniversiteit Utrecht, Heidelberglaan 2, 3584 CS Utrecht

Manuscript aanvaard 10-9-'85

Summary

Bol, E. & M.A. Gresnigt. 'Effects of an experimental program for reading comprehension'. *Pedagogische Studiën*, 1986, 63, 49-60.

An experimental program for reading comprehension was constructed for elementary education, aimed to the systematic development of cognitive factors. The course is based upon the socio-historical theory of Vygotskij and models developed by Lompscher and Bol.

The program ran for a period of four years (3rd to 6th grade) in four classes of two different schools. Weekly about one hour was spent on these lessons, which were given by teachers of the experimental classes.

Significant positive effects, measured by a series of specially developed pre- and posttests were found in comparison with control groups. Data indicate that it is possible to speed up the acquisition of reading comprehension with about one year. Results of experimental and control groups on school tests constructed by others suggest a broader positive effect upon educational achievement.