

## Schrijfvaardigheid en schooleffectiviteit

### 1. Inleiding

In dit artikel doe ik verslag van een onderzoek naar schrijfvaardigheid en schooleffectiviteit in het voortgezet onderwijs<sup>1</sup>. Onderzoek naar de effectiviteit van scholen staat de laatste jaren sterk in de belangstelling. De aanwijzingen hiervoor zijn legio. In Cardiff, Wales, vindt in januari 1991 het vierde internationale congres voor schooleffectiviteit plaats<sup>2</sup>. Sinds kort verschijnt het internationale tijdschrift *School Effectiveness*<sup>3</sup>. En in Nederland verscheen onder auspiciën van SVO onlangs de overzichtsstudie *Wat maakt scholen effectief?* (Scheerens 1989). Aan de hand van deze laatste publikatie geef ik hieronder een zeer beknopte schets van het schooleffectiviteitsonderzoek.

Onderzoek naar de effectiviteit van scholen is in de eerste plaats gericht op de output van scholen: de leeropbrengsten van leerlingen. De kenmerkende vraag in zulk onderzoek is in welke mate verschillen in prestaties tussen leerlingen toegeschreven kunnen worden aan verschillen tussen scholen. Het kan daarbij gaan om verschillen op allerlei niveaus: de schoolorganisatie, de deskundigheid en ervaring van leerkrachten, de groepsgrootte, vormen van leerlingdifferentiatie, instructiekenmerken, etcetera.

Ook in het onderzoek waarvan ik hieronder verslag doe, vormen leeropbrengsten het effectiviteitscriterium. Dat hiermee een bepaalde invulling van het begrip effectiviteit gegeven wordt, moge duidelijk zijn. Ook andere effectiviteitscriteria zijn immers denkbaar. Zo onderscheidt Scheerens (1989) als alternatieve effectiviteitscriteria het aanpassingsvermogen van de schoolorganisatie, de betrokkenheid van de individuele leden van de schoolorganisatie, de continuïteit van de schoolorganisatie en de responsbereidheid van de schoolorganisatie ten opzichte van externe belanghebbende instanties.

De concentratie op leeropbrengsten valt goed te verdedigen, vanuit de opvatting dat onderwijs een doelgericht proces is: een proces dat mikt op het aanbrengen van kennis, inzichten, vaardigheden en houdingen. Andere waarden dan leeropbrengsten vormen in mijn ogen geen eigenlijk effectiviteitscriterium. Ik schuif ze daarmee niet als onbeduidend ter zijde; met Scheerens (o.c., p.17) acht ik ze juist belangrijk, als condities waaronder leeropbrengsten tot stand moeten komen. Inzicht in zulke condities kan het begrip van verschillen in effectiviteit tussen scholen bevorderen.

Binnen het op leeropbrengsten gerichte onderzoek naar schooleffectiviteit kunnen verschillende onderzoekstradities onderscheiden worden. Scheerens (o.c., p.19) onderscheidt vijf stromingen, die in de onderstaande tabel worden weergegeven:

Tabel 1 Typering van stromingen in op leeropbrengsten gericht effectiviteitsonderzoek

stroming	type onafh. variabele	type afh. variabele	discipline	type onderzoek
ongelijke kansen	-SES -IQ leerlingen -materiële schoolkenmerken	-leerprestaties -schoolloopbanen	-sociologie	-survey
productiefuncties	-materiële schoolkenmerken	-leerprestaties	-economie	-survey
compensatieprogramma	-specifieke curricula	-leerprestaties	-interdisciplinaire onderwijskunde	-quasi-experiment
effectieve scholen	-proceskenmerken scholen	-leerprestaties	-interdisciplinaire onderwijskunde	-case-studie
effectieve instructie	-kenm. leerkr. instructie -klasorganisatie	-leerprestaties	-interdisciplinaire onderwijskunde	-experiment -observatie

(Naar Scheerens 1989, p.20)

Het onderzoek naar kansenongelijkheid (met name Coleman e.a. 1966; Jencks e.a. 1972; Thorndike 1973) mondde uit in weinig optimistische conclusies omtrent de mate waarin de school iets af kan doen aan de milieubepaaldheid van leerprestaties. Ook uit het economisch onderzoek naar productiefuncties en uit de evaluatie van compensatieprogramma's komen geen overtuigende schooleffecten naar voren.

Veel meer optimisme omtrent de mate waarin scholen 'verschil maken', is ontleend aan het effectieve-scholenonderzoek dat gericht was op kenmerken van de organisatie en de inhoud van het onderwijs op vaak bijzondere scholen<sup>4</sup> (met name Brookover e.a. 1979; Mortimore e.a. 1988). Dit onderzoek heeft een aantal factoren geïdentificeerd die van belang lijken<sup>5</sup> voor het verklaren van verschillen in schooleffectiviteit. Deze factoren worden vaak samengevat aan de hand van het 5-factor model van Edmonds (1979). Effectieve scholen zouden gekenmerkt worden door: (1) de aanwezigheid van sterk onderwijskundig leiderschap; (2) het benadrukken van leerprestaties met betrekking tot basisvaardigheden; (3) een veilig en ordelijk leerklimaat; (4) hoge verwachtingen ten aanzien van de leerlingprestaties; en (5) geregelde evaluatie van individuele leervorderingen.

Ook uit onderzoek naar kenmerken van effectieve instructie komt naar voren dat bepaalde school- of onderwijskenmerken er wel degelijk toe kunnen doen. Omvangrijke meta-analyses van duizenden afzonderlijke empirische studies (met name Walberg 1986) laten zien dat bepaalde instructiekenmerken onmiskenbare effecten hebben. Voorbeelden hiervan zijn effectieve leertijd (de tijd die leerlingen aan het leren in een vakgebied besteden), de duidelijkheid en aanvaardbaarheid van doelstellingen, het frequente gebruik van feedback en evaluatie en de kwaliteit van het leerklimaat.



In een samenvattende interpretatie van de uitkomsten van verschillende typen schooleffectiviteitsonderzoek bepleit Scheerens (o.c.) een integratie van verschillende onderzoekstradities. Hij postuleert drie kernfactoren die bepalend zijn voor schooleffectiviteit: (a) de politieke wil van scholen om hoge onderwijsopbrengsten te realiseren; (b) maximalisering van de daadwerkelijke leertijd; en (c) gestructureerd onderwijs. Daarnaast onderscheidt Scheerens verschillende niveaus van handelen: de schoolcontext, de school, de klas en haar leerkracht en de individuele leerling. Toekomstig onderzoek dient, volgens Scheerens, zoveel mogelijk rekening te houden met genoemde kernfactoren en met de verschillende niveaus.

In het vervolg van dit artikel bespreek ik een onderzoek dat in een aantal opzichten tegemoet komt aan deze aanbevelingen. Het onderzoek betreft een heranalyse van gegevens waarover ik eerder gerapporteerd heb (De Gloppe 1988).

## 2. Schrijfvaardigheid en schooleffectiviteit

Mijn onderzoek naar schrijfvaardigheid en schooleffectiviteit is verwant aan verschillende stromingen binnen het schooleffectiviteitsonderzoek. Het sluit enerzijds aan bij het onderzoek naar kansenongelijkheid en naar schooleffectiviteit, omdat de aandacht (mede) gericht is op de bijdrage die schoolkenmerken, na verdiscontering van milieukenmerken van de leerlingen, leveren aan de leerprestaties van leerlingen. Het onderzoek staat tevens in de traditie van het onderzoek naar effectieve instructiekenmerken: de keuze van een belangrijk deel van de variabelen was gebaseerd op een review van onderzoek naar de effecten van instructievariabelen op de schrijfvaardigheid (Wesdorp 1983). Het onderzoek was gericht op de inhoud en opbrengsten van het schrijfvaardigheidsonderwijs in het voortgezet onderwijs en vormde het Nederlandse aandeel in de IEA Written Composition Study (Gorman, Purves & Degenhart 1988; De Gloppe 1988).

In het onderzoek naar schooleffectiviteit is bij mijn weten niet eerder aandacht besteed aan schrijfvaardigheid. Dat is begrijpelijk, want het meten van schrijfvaardigheid is moeilijk en kostbaar. Toch is schrijven voor zulk onderzoek een zeer interessante afhankelijke variabele. Schrijfvaardigheid maakt (in Nederland en in veel andere landen) deel uit van het moedertaalonderwijs, een leergebied waar de invloeden van het thuismilieu zich in de regel sterk doen gelden. Schrijfvaardigheid is echter een onderdeel van het taalonderwijs waarvoor dit niet zozeer lijkt te gelden: (bijna) alle leerlingen ontwikkelen hun schrijfvaardigheid overwegend op school en niet thuis. Schrijven mag in dit licht met recht een echt *schoolvak* genoemd worden. Daarom en omdat schrijfvaardigheid een zeer centrale onderwijsdoelstelling is, vormt het een interessant criterium voor onderzoek naar schooleffecten.

Aandacht voor één specifieke onderwijsopbrengst als schrijfvaardigheid is voor onderzoek naar schooleffectiviteit om nog een andere reden interessant. Veel onderzoek naar schooleffecten heeft gebruik gemaakt van nogal brede en algemene effectiviteitscriteria. Dit zou één van de redenen kunnen zijn waarom in veel studies, ook in Nederland, de netto-schooleffecten klein zijn. Iedereen weet uit eigen ervaring dat er op scholen goede en slechte leerkrachten

rondlopen en dat individuele leerkrachten in bepaalde onderdelen van hun vak beter en in andere onderdelen slechter lesgeven. Door het gebruik van brede en algemene effectiviteitscriteria zou een niet onaanzienlijk deel van de bestaande, specifieke verschillen wel eens uitgemiddeld kunnen worden. Zo bezien kan het gebruik van specifieke criteria het inzicht in schooleffecten bevorderen.

### 3. Onderzoeksvragen en -opzet

In mijn onderzoek stonden twee vragen centraal:

1. Hoe groot zijn de verschillen in schrijffprestaties tussen scholen, in verhouding tot de verschillen op het niveau van individuele leerlingen?
2. Welk deel van de prestatieverschillen op schoolniveau valt toe te schrijven aan kenmerken van scholen en van het schrijfonderwijs en welk deel aan achtergrondkenmerken van leerlingen?

Op 107 scholen voor voortgezet onderwijs werden schrijfprodukten van leerlingen uit het derde leerjaar verzameld<sup>6</sup>. Per school nam één klas aan het onderzoek deel: een LTO-, een LHNO-, een MAVO-, een HAVO- of een VWO-klas. Uit een totaal van acht schrijfopdrachten maakte iedere leerling er drie. Per klas werd het schrijfwerk van 12 leerlingen beoordeeld, hetgeen neerkomt op een totaal van 36 opstellen per klas. De opstellen werden op uiteenlopende kwaliteitsaspecten beoordeeld. In de onderhavige analyses wordt één samenvattende prestatievariabele gebruikt: het gemiddelde van oordelen over globale kwaliteit, inhoud, opbouw en stijl voor drie verschillende schrijfopdrachten. De validiteit van deze gemiddelde schrijfvaardigheidsscore lijkt voldoende gewaarborgd (De Gloppe 1988, p.73-77). De betrouwbaarheid van de individuele scores bedroeg .75 (Cronbachs alpha).

Met behulp van schriftelijke vragenlijsten zijn bij de leerkrachten Nederlands en de leerlingen van de 107 klassen gegevens verzameld over een grote hoeveelheid variabelen: achtergrondkenmerken van leerlingen en leerkrachten, instructiekenmerken en schoolkenmerken.

### 4. De omvang van de verschillen in schrijffprestaties op schoolniveau

Er bestaat geen absolute standaard waaraan de grootte van verschillen tussen scholen afgemeten kan worden. Een voor de hand liggende relatieve standaard is er wel: de grootte van verschillen tussen individuele leerlingen. Veel gebruikt is de intra-klasse correlatiecoëfficiënt rho. Rho is gebaseerd op de verhouding van de variantie van de klasgemiddelden en de variantie van de individuele scores. Onder de aanname van een in twee trappen aselechte steekproef (namelijk een aselechte steekproef van klassen en binnen klassen een aselechte steekproef van leerlingen) geeft de volgende vergelijking een zuivere schatter van rho:

$$\rho = ((b)/(b-1))(S_s^2/S_i^2) - ((1)/(b-1)) \quad (\text{Kish 1965})$$



De kern van de formule is de verhouding tussen  $S_s^2$  en  $S_i^2$ : de variantie van de klasgemiddelden en de variantie van de individuele scores. Hoe groter de verschillen tussen klassen zijn, *in verhouding tot de individuele verschillen*, hoe groter rho. De waarde b in de formule staat voor het aantal leerlingen dat uit iedere klas getrokken wordt<sup>7</sup> en bepaalt de correctiefactor voor de ratio van de twee varianties. Deze correctiefactor moet voorkomen dat bij kleine leerlingsteekproeven per klas de tussen-klassen variantie overschat wordt. Hoe groter b, hoe dichter rho ligt bij de ratio  $S_s^2/S_i^2$ . Immers, bij toename van b nadert  $(b)/(b-1)$  naar 1 en  $(1)/(b-1)$  naar 0.

Een eenvoudig rekenvoorbeeld kan de betekenis van rho verhelderen. Stel dat we te maken hebben met een steekproef van 1000 leerlingen, afkomstig uit 100 verschillende klassen. Aan deze leerlingen is een woordenschattest afgenomen van 100 items. De gemiddelde leerlingsscore bedraagt 50, de leerlingenstandaarddeviatie 15, en de tussen-leerlingen-variantie 225<sup>8</sup>. Als binnen iedere klas alle geteste leerlingen dezelfde score behalen (in klas 1 tien maal 20, in klas 2 tien maal 30, etc.), dan zullen het klasgemiddelde, de klasse-standaarddeviatie en de tussen-klassen-variantie eveneens 50, 15 en 225 bedragen. De variantie tussen klassen is in dit geval even groot als de variantie tussen leerlingen en rho neemt (voor  $b=10$ ) zijn maximale waarde aan. We hebben hier een extreem geval: alle verschillen op het individuele niveau zijn herleidbaar tot verschillen op klasniveau.

In een tweede voorbeeld hebben we te maken met een steekproef van 1000 andere leerlingen, afkomstig uit een ander honderdtal klassen. Aan deze leerlingen is dezelfde test afgenomen en op het leerlingenniveau bedragen het gemiddelde, de standaarddeviatie en de variantie opnieuw 50, 15 en 225. Als in iedere klas de verdeling van de scores hetzelfde is (in alle klassen bijvoorbeeld ongeveer 25, 35, 40, 45, 50, 50, 55, 60, 65, 75) dan zullen het klasgemiddelde, de klasse-standaarddeviatie en de tussen-klassen-variantie 100, 0 en 0 bedragen. De variantie tussen klassen is nihil, want het gemiddelde van iedere klas bedraagt 50. Nu neemt rho (voor  $b=10$ ) zijn minimale waarde aan. Ook dit is een extreem geval: niets van de verschillen op het individuele niveau is herleidbaar tot verschillen op klasniveau.

In werkelijkheid is de omvang van verschillen tussen klassen niet zo extreem als in de gegeven voorbeelden. Het is gebruikelijk om reeds bij betrekkelijk kleine waarden van rho, in de orde van grootte van .20, van aanzienlijke verschillen tussen klassen of scholen te spreken. In veel onderzoek worden lagere waarden van rho gevonden. Blok en Eiting (1987) rapporteren over de verschillen tussen basisscholen op taalprestatiemetingen een gemiddelde rho van .13. Scheerens (1989) rapporteert in een samenvatting van Nederlandse onderzoeksresultaten een gemiddelde tussen-scholen-variantie van dezelfde orde van grootte.

In tabel 2 zijn de waarden van rho uit mijn onderzoek weergegeven. Uit de tabel blijkt dat de verschillen tussen de klasgemiddelden in de totale steekproef zeer aanzienlijk zijn. Deze grote verschillen weerspiegelen het sterk selectieve karakter van het Nederlandse voortgezet onderwijs. Binnen de schooltypen, en dat is niet verwonderlijk, zijn de waarden van rho veel lager. Toch zijn ze nog steeds omvangrijk te noemen. Ook door andere variabelen dan het schooltype valt derhalve nog verschil te verklaren. We zijn daarmee toe aan de beantwoording van de tweede onderzoeksvraag.

Tabel 2 De grootte van de verschillen tussen scholen in de totale steekproef en per schooltype

Totaal (107)	LTO (20)	LHNO (14)	MAVO (40)	HAVO (14)	VWO (18)
.48	.20	.28	.19	.19	.18

## 5. Achtergrondkenmerken van leerlingen

Achtergrondkenmerken werden bij de leerlingen met behulp van vragenlijsten gemeten. De vragenlijsten bevatten vooral vragen die op de sociaal-economische status van het gezin betrekking hadden, en op taalgebruik en taalgewoonten. In tabel 3 zijn de achtergrondvariabelen weergegeven die in de hier gerapporteerde analyses gebruikt zijn.

Tabel 3 Geselecteerde achtergrondkenmerken van leerlingen

Variabele	Operationalisatie
Geslacht	-
Leeftijd	-
SES	Opleiding en beroepsniveau ouders/verzorgers
Bockenbezit	Aantal boeken thuis
Bibliotheekbezoek	Frequentie van bibliotheekbezoek
Leestijd Boeken	Tijd besteed aan het lezen van boeken
Leestijd Strips	Tijd besteed aan het lezen van strips
Tijd Huiswerk	Tijd besteed aan het maken van huiswerk
Tijd TV	Tijd besteed aan het kijken naar de t.v.
Tijd Radio	Tijd besteed aan het luisteren naar de radio
Mondelinge Interactie	Frequentie en inhoud van mondelinge interactie thuis

De variabele sociaal-economische status (SES) is een index van het opleidings- en beroepsniveau van de ouders of verzorgers van de leerling. Cronbachs alpha voor de vier items bedraagt .87. Tijd TV en Tijd Radio zijn ieder gebaseerd op twee vragen (kijk/luistertijd op doordeweekse dagen en kijk/luistertijd op weekenddagen). Cronbachs alpha bedraagt respectievelijk .70 en .66. Mondelinge interactie is een index van 11 vragen die betrekking hebben op de frequentie en inhoud van de mondelinge interactie in het gezin. Cronbachs alpha bedraagt .70. De andere variabelen zijn alle met afzonderlijke vragen gemeten.

## 6. Instructievariabelen

De vragenlijsten voor de leerkrachten en de leerlingen bevatten tal van vragen over het schrijfvaardigheidsonderwijs. De keuze van de vragen was voor een belangrijk deel gebaseerd op een review van 158 studies naar de effecten van instructievariabelen op de schrijfvaardigheid (Wesdorp 1983). Uit dit review en uit een later review van Hillocks (1986) komen bepaalde veelbelovende



instructie-aanpakken naar voren: pre-writing activiteiten (activiteiten die het schrijven voorbereiden door leerlingen na te laten denken over de opdracht en over het onderwerp); peer evaluation (beoordeling van schrijfprodukten door mede-leerlingen, vergelijk Rijlaarsdam 1986); zinscombinatie-onderwijs (vergelijk Schuurs 1990); revisie-onderwijs (onderwijs waarin leerlingen hun schrijfprodukten leren herschrijven); en het gebruik van specifieke kwaliteitscriteria en schalen met voorbeeldopstellen (als specificatie van de doelen van het schrijf-onderwijs). De vragenlijsten stemden verder overeen met meer algemene uitkomsten van onderzoek naar effectieve instructie (bijvoorbeeld Walberg 1986). Er werd voor gezorgd dat er variabelen betreffende de leertijd, de duidelijkheid van doelstellingen en de hoeveelheid feedback opgenomen waren.

De overgrote meerderheid van de instructievariabelen bleek niet of nauwelijks met de leerlingprestaties samen te hangen. Eén plausibele verklaring voor dit gegeven is dat er nauwelijks variantie was op de variabelen die betrekking hadden op de bovengenoemde effectieve steldidactieken. Uit dit resultaat en uit corresponderende resultaten van een grootschalige enquête onder 1000 leerkrachten Nederlands in het voortgezet onderwijs (Damhuis e.a. 1983) bleek overduidelijk dat deze specifieke aanpakken van het schrijfonderwijs in het Nederlandse voortgezet onderwijs vrijwel non-existent zijn. Vanuit het perspectief van schooleffectiviteit, en daarmee natuurlijk vanuit didactisch oogpunt, is dit gegeven op zichzelf hoogst belangrijk. Effectief onderwijs kan niet zonder effectieve instructie-aanpakken. Het gemis daaraan in het schrijfonderwijs geeft aan dat er een belangrijk potentieel is voor de verbetering van het schrijfvaardigheidscurriculum. Dat het, in principe, veel beter kan, geldt niet alleen voor het schrijfonderwijs en niet alleen in Nederland. Het onderwijs in schrijven en lezen, binnen en buiten Nederland, is voor belangrijke verbetering vatbaar (vgl. Durkin 1978, NAEP 1985, Applebee e.a. 1986, Weterings & Aarnoutse 1986).

Enkele variabelen bleken echter wel substantieel met de schrijfprestaties van de leerlingen samen te hangen. Drie daarvan zijn betrokken in de analyses waarover ik hieronder rapporteer. Het betreft de hoeveelheid directe instructie in tekstenmerken, het aantal oefeningen in tekstproductie en de hoeveelheid feedback (zie ook bijlage 1). Dat deze variabelen samenhangen met de schrijfprestaties, komt overeen met de meer algemene uitkomsten van onderzoek naar effectieve instructie (vergelijk Walberg 1986).

Tabel 4 Geselecteerde instructievariabelen

Variabele	Operationalisatie
Instructie Tekstenmerken	Instructie in kenmerken van verschillende typen teksten
Oefeningen Tekstproductie	Oefeningen in onderdelen van tekstproductie
Feedback	Frequentie, timing en toon van de feedback

Instructie Tekstenmerken is een index van acht variabelen die betrekking hebben op de benadrukking in het onderwijs van kennis over tekstenmerken. De betreffende vragen gaan bij voorbeeld over de aandacht voor narratief proza en de analyse van tekststructuren en stilistiek. Cronbachs alpha bedraagt .76.

Oefening Tekstproductie is een index van 11 vragen over oefeningen in verschillende onderdelen van het schrijfproces (bijvoorbeeld het kiezen of opzetten van een tekstschema). Het gaat hier uitdrukkelijk niet om oefeningen die betrekking hebben op spelling, interpunctie of grammatica. Cronbachs alpha bedraagt .83. Feedback is een index van 22 vragen over de frequentie waarmee, het moment (onmiddellijk of met meer of minder uitstel) en de toon (kritisch/positief) waarop leerkrachten feedback op de schrijfprodukten van hun leerlingen geven. Cronbachs alpha bedraagt ook hier .83. Alle genoemde vragen zijn gemeten met behulp van de leerkrachtvragenlijst.

## 7. Schooltype

De variabele schooltype is voor onderzoek naar schooleffectiviteit vanzelfsprekend van groot belang. Allereerst omdat binnen het voortgezet onderwijs selectie naar leerprestaties plaats vindt. Verschillen in schrijfprestaties tussen klassen zullen daarom voor een niet onaanzienlijk deel terug gaan op prestatieverschillen tussen schooltypen, en daarmee op verschillen tussen eerdere leerprestaties. Hetzelfde geldt tot op zekere hoogte ook voor het onderwijsaanbod. Het onderwijsaanbod verschilt van klas tot klas. Ook deze verschillen zijn voor een deel terug te voeren zijn op verschillen tussen schooltypen (De Gloppe 1988).

De vijf in de steekproef vertegenwoordigde schooltypen zijn met behulp van vier zogeheten 'dummy-variabelen' geschaald. De wijze van schaling is weergegeven in tabel 5. Met de vier dummy-variabelen wordt alle informatie over het schooltype weergegeven. De LTO-scholen hebben op alle dummy-variabelen de waarde nul; alleen de LHNO-scholen hebben de waarde één op Dummy 1; alleen de MAVO-scholen hebben de waarde één op Dummy 2; etcetera. Bij statistische analyses geven de dummy-variabelen daardoor de contrasten tussen de schooltypen aan.

Tabel 5 Codering dummy-variabelen schooltype. In de tabel de waarden (0 en 1) van de vijf schooltypen op de vier dummy-variabelen.

	Dummy1	Dummy2	Dummy3	Dummy4
LTO	0	0	0	0
LHNO	1	0	0	0
MAVO	0	1	0	0
HAVO	0	0	1	0
VWO	0	0	0	1

## 8. Resultaten

De onderzoeksgegevens hebben typisch een 'multilevel' of multi-niveau structuur. Op het individuele niveau hebben we metingen over leerlingen, op het groepsniveau hebben we metingen over klassen of scholen. De gegevens



zijn geanalyseerd met een daarbij passend analyseprogramma: Goldsteins programma voor multilevel analyse (Goldstein 1987). Bij de multiniveau analyse worden metingen op individueel niveau (hier: schrijfvaardigheid) gezien als een functie van andere individuele metingen (zie tabel 3) en van metingen op groepsniveau (zie tabellen 4 en 5). Voor ieder van de afzonderlijke schooltypen was het volgende model gespecificeerd:

$$y_{ij} = (\alpha + \alpha_j) + \beta x_{ij} + \delta z_j + e_{ij}$$

waarbij  $y_{ij}$  de schrijfvaardigheidsscore van leerling  $i$  in klas  $j$  is,  $\alpha$  de gemiddelde score van de totale steekproef is,  $\alpha_j$  de score is die de afwijking van het gemiddelde van klas  $j$  ten opzichte van de het totaal gemiddelde aangeeft,  $\beta$  de regressiecoëfficiënt<sup>9</sup> is voor  $x_{ij}$  (de score van leerling  $i$  in klas  $j$  op een individuele achtergrondvariabele),  $\delta$  de regressiecoëfficiënt is voor  $z_j$  (de score voor klas  $j$  op een instructievariabele) en  $e_{ij}$  een residuele scorecomponent is voor leerling  $i$  in klas  $j$  (een component die aangeeft welk deel van de individuele schrijfvaardigheidsscore niet door de rest van de vergelijking beschreven kan worden).

Voor de steekproef als geheel zijn de vier dummy-variabelen voor schooltype (als variabelen op het klasniveau) aan de bovenstaande vergelijking toegevoegd.

Bij de multilevel analyse hoeft, anders dan in klassieke regressie analyse, niet aangenomen te worden dat coëfficiënt  $\beta$  in iedere klas dezelfde waarde heeft. Ik ben daarom nagegaan of  $\beta$  voor een of meer individuele variabelen tussen klassen significant varieert<sup>10</sup>. Ik heb geen modellen gevonden waarin van significante variatie in  $\beta$  sprake was. Deze bevinding kan mede een gevolg zijn van de beperkte steekproefomvang binnen schooltypen.

De resultaten van de analyses zijn samengevat in de tabellen 6 en 7. In tabel 7 zijn schattingen van  $\rho$  weergegeven, onder verschillende modellen. Letten we alleen op de eerste kolom, dan valt op dat de dummy-variabelen voor schooltype alleen reeds een groot deel van de verschillen tussen klassen verklaren. De initiële  $\rho$  van .48 (zie tabel 2) daalt door de introductie van deze variabelen naar .31. De introductie van de achtergrondvariabelen doet  $\rho$  verder dalen tot .14. In de totale steekproef zorgt de introductie van de instructievariabelen niet voor een verdere daling van  $\rho$ .

Voor de afzonderlijke schooltypen zijn de resultaten zeer overeenkomstig. De individuele achtergrondvariabelen zorgen voor een flinke reductie van  $\rho$  (LTO van .20 tot .14; LHNO van .28 tot .17; MAVO van .19 tot .17; HAVO van .19 tot .11 en VWO van .18 tot .10). In twee van de vijf schooltypen, LTO en LHNO, wordt  $\rho$  door de introductie van de instructievariabelen verder verkleind, respectievelijk tot .07 en .08. Waarom de instructievariabelen er hier wel en elders niet toe doen, is niet duidelijk. Als verklaring kan de hypothese naar voren gebracht worden dat verschillen in instructie-kwaliteit zich het eerst in de prestaties van zwakke schrijvers (met name LBO-leerlingen) zullen openbaren. Net als zwakke lezers zijn zwakke schrijvers het meest gebaat bij goede instructie.

Tabel 6 Schattingen van rho onder verschillende modellen

	Totaal	lto	lhno	mavo	havo	vwo
Schooltype	.31					
Achtergrondvariabelen	.14	.14	.17	.17	.11	.10
Achtergrond- en instructievariabelen	.14	.07	.08	.17	.11	.10

Tabel 7 Schattingen van regressiegewichten en bijbehorende standard errors

Variabele	$\beta/\delta$	s.e.
Dummy1	.09	.12
Dummy2	.65	.09
Dummy3	.99	.12
Dummy4	1.17	.11
Geslacht	.35	.05
Leeftijd	.01	.01
SES	.01	.02
Bockenbezit	.03	.02
Bibliotheekbezoek	.03	.01
Leestijd Boeken	.02	.01
Leestijd Strips	.04	.02
Tijd Huiswerk	-.01	.01
Tijd TV	-.02	.02
Tijd Radio	.01	.02
Mondelinge Interactie	.13	.04
Instructie Tekstkenmerken	.05	.01
Oefeningen Tekstproductie	.01	.01
Feedback	.01	.01

Tabel 7 bevat de schattingen van de regressie-gewichten voor de variabelen op individueel ( $\beta$ ) en op groepsniveau ( $\delta$ ) en de bijbehorende standard errors. Significant op het .05 niveau zijn die coëfficiënten die in absoluut opzicht 1.96 maal groter zijn dan hun standard error. Uit de tabel kan opgemaakt worden dat de dummy-variabelen voor schooltype veel gewicht in de schaal leggen. Van de achtergrondvariabelen zijn geslacht, bibliotheekbezoek, leestijd boeken en strips en mondelinge interactie significant. SES draagt, in de aanwezigheid van andere achtergrondvariabelen, opmerkelijk genoeg niet bij aan de voorspelling van de individuele schrijfvvaardigheidsscores.

Van de instructievariabelen draagt Instructie Tekstkenmerken significant bij aan de voorspelling.

## 9. Discussie

In het bovenstaande zijn twee vragen beantwoord. In het Nederlandse voortgezet onderwijs bestaan er tussen scholen grote verschillen in schrijfpredaties. Het



grootste deel van de verschillen tussen scholen kan echter toegeschreven worden aan de variabele schooltype (een gevoelige indicator voor eerdere leerprestaties en intelligentie) en aan achtergrondkenmerken van de leerlingen. Binnen het LBO leggen instructievariabelen bij het verklaren van verschillen tussen scholen nog enig gewicht in de schaal. Dit laatste vormt een aanwijzing voor 'echte' verschillen in schooleffectiviteit. Het is interessant om op te merken dat de instructievariabelen er bij de zwakkere leerlingen het meest toe lijken te doen. Deze bevinding correspondeert met het gegeven dat veel van het onderzoek dat positieve schooleffecten demonstreerde, gericht was op laag presterende leerlingen. Hierboven is al opgemerkt dat verschillen in instructiekwaliteit zich wellicht het eerst in de prestaties van zwakke schrijvers zullen openbaren. Net als zwakke lezers zijn zij het meest gebaat bij goede instructie. Aan het lage uiteinde van de prestatieverdeling zou verder de variatie in aanpak en resultaten wel eens het grootst kunnen zijn.

Deze resultaten geven geen aanleiding tot hoge verwachtingen omtrent de grootte van schooleffecten in het Nederlandse schrijfvaardigheidsonderwijs. Toch moet gewaarschuwd worden tegen de interpretatie dat scholen geen verschil maken.

In de eerste plaats moet met nadruk opgemerkt worden dat allerlei effectieve vormen van schrijfdidactiek in het Nederlandse schrijfvaardigheidsonderwijs niet of nauwelijks voorkomen. Het schrijfcurriculum bestaat overwegend uit globale oefening, zonder veel instructie of gerichte oefening van deelprocessen en -vaardigheden. Onder deze omstandigheden is het niet verbazingwekkend dat scholen weinig verschil maken. Uit het feit dat ze *onder de huidige omstandigheden* geen verschil maken, mag echter niet afgeleid worden dat ze geen verschil *kunnen* maken.

In de tweede plaats moet een kanttekening geplaatst worden bij de kwaliteit van metingen van achtergrond- en instructievariabelen. In grootschalig survey-onderzoek (zoals de onderhavige IEA-studie en nationale peilingsstudies) laat de kwaliteit van metingen van onderwijs- en leerprocessen en van individuele achtergrondkenmerken te wensen over. De onbetrouwbaarheid van de meting van tijdbesteding is door Otter (1987) aangetoond. Voor een treffend voorbeeld van de onbetrouwbaarheid van metingen van buitenschoolse activiteiten kan verwezen worden naar de discrepantie tussen verschillende typen metingen van de tijd die kinderen besteden aan het kijken naar televisie. De Dienst Kijk- en Luisteronderzoek van de NOS hanteert al vele jaren de zogenaamde knoppenmethode. In een groot aantal gezinnen staat een televisietoestel waaraan een afzonderlijk kastje met knoppen verbonden is. Ieder gezinslid (kinderen vanaf 6 jaar en volwassenen) heeft een eigen knop die bij het kijken naar een of ander programma ingedrukt moet worden. Uit dit 'knoppenonderzoek', dat algemeen geldt als de meest valide en betrouwbare kijktijdmeting, blijkt dat 9- tot 12-jarigen in het tweede kwartaal van 1988 dagelijks gemiddeld minder dan vier kwartier televisie keken (NOS 1990). Uit gegevens van PPON, die betrekking hebben op de maanden mei en juni 1988, komt naar voren dat 81% van de getoetste 12-jarigen de dag voor de toetsafname gemiddeld zeven kwartier televisie heeft gekeken (Zwarts 1990, p.39). Dit geeft te denken over de kwaliteit van vragenlijstmetingen.

Ten derde: onderzoek naar schooleffecten concentreert zich op individuele en op groepsverschillen en op de daarmee samenhangende ongelijkheidsproblematiek. Hoe belangrijk dat ook is, er mag niet vergeten worden dat prestatieverdelingen ook andere belangrijke kenmerken bezitten. Niet alleen de spreiding van scores is interessant. Ook de hoogte van gemiddelde scores en de ontwikkeling daarvan in de tijd is belangrijk. Hetzelfde geldt voor de omvang van het lage uiteinde van de prestatieverdeling: het percentage leerlingen dat niet of nauwelijks in staat is tot het verrichten van belangrijk geachte taken. In onderzoek naar schooleffectiviteit worden deze aspecten nog teveel verwaarloosd.

## Noten

1. Deze tekst is een bewerking van een lezing voor het Third International Congress for School Effectiveness te Jeruzalem (3-5 Januari 1990).
2. Eerdere congressen vonden plaats in Londen (1988), Rotterdam (1989) en Jeruzalem (1990). Van de eerste twee congressen zijn bundels uitgebracht (resp. Creemers e.a. 1989; Reynolds e.a. 1989).
3. *School Effectiveness and School Improvement. An International Journal of Research, Policy and Practice*. Editors: David Reynolds & Bert P.M. Creemers. Uitgever: Swets & Zeitlinger, Amsterdam/Lisse.
4. Het ging in zulk onderzoek vaak om bijzondere scholen: scholen die opvallend goed presteren gelet op de aard van hun leerlingeninstroom (veel leerlingen uit anderstalige milieus en uit milieus met een laag SES).
5. Over de status van de resultaten (gebrek aan replicatie-onderzoek) en over de interpretatie (wat zijn grote/kleine effecten?) bestaat nog geen consensus (Scheerens 1989).
6. Het responspercentage bedroeg 75%. De representativiteit van de steekproef lijkt gewaarborgd: vergelijking van vragenlijstgegevens met corresponderende resultaten van een andere onafhankelijke steekproef bracht alleen zeer geringe discrepanties aan het licht.
7. In peilingsonderzoek worden per taak of toets vaak steekproeven uit klassen getrokken. Opdrachten worden vaak niet aan een intacte klas voorgelegd, omdat de informatieopbrengst bij het toetsen van een complete klas niet opweegt tegen de kosten. Vanwege de intra-klasse correlatiecoëfficiënt  $\rho$  is het bij een vast aantal te testen leerlingen wenselijk om meer klassen te onderzoeken en per klas minder leerlingen (voor een illustratie hiervan, zie Hoeksma e.a. 1985).
8. De standaarddeviatie is een maat voor de omvang van de afwijking van een verzameling scores ten opzichte van hun gemiddelde. De variantie van een verzameling scores is het kwadraat van hun standaarddeviatie.
9. De regressiecoëfficiënt geeft aan hoeveel 'invloed' een onafhankelijke variabele op een afhankelijke variabele heeft. Preciezer gezegd geeft de coëfficiënt aan hoeveel de score op de afhankelijke variabele mee-verandert met een verandering van de score op de onafhankelijke variabele.
10. Men moet zich hierbij voorstellen dat bijvoorbeeld de samenhang tussen SES en schrijfvaardigheid van klas tot klas kan variëren. Er kunnen klassen zijn waarin wel een duidelijke samenhang tussen beide variabelen is, en



klassen waarin die samenhang om een of andere reden ontbreekt (bijvoorbeeld door differentiatie-maatregelen).

## Bibliografie

- Applebee, A.N., Langer, J.A. & Mullis, I.V.S. (1986). *Writing. Trends Across the Decade, 1974-1984*. Princeton NJ: ETS/NAEP.
- Blok, H. & Eiting, M.H. (1987). De grootte van schooleffecten: Hoe verschillend presteren leerlingen van verschillende scholen? *Tijdschrift voor Onderwijsresearch*, 13, 16-30.
- Brookover, W.B., Beady, C. & Flood, P. (1979). *School Social Systems and Student Achievement - Schools Can Make a Difference*. New York: Prager Publishers.
- Coleman, J.S. et al. (1966). *Equality of Educational Opportunity*. Washington D.C.: U.S. Government Printing Office.
- Creemers, B.P.M., Peters, T. & Reynolds, D. (1989). *School Effectiveness and School Improvement. Proceedings of the Second International Congress, Rotterdam, 1989*. Lisse: Swets & Zeitlinger.
- Damhuis, R., Glopper, K. de & Wesdorp, H. (1983). *Het opstelonderwijs. Stelvaardigheid in het voortgezet onderwijs: theorie en praktijk*. Amsterdam, SCO, 1983.
- Durkin, D. (1978). What classroom observations reveal about reading comprehension instruction. *Reading Research Quarterly*, 14, 481-533.
- Edmonds, R.R. (1979). Effective schools for the urban poor. *Educational Leadership*, 37, 15-24.
- Goldstein, H. (1987). *Multilevel models in educational and sociological research*. London: Charles Griffin.
- Gorman, T.P., Purves, A.C. & Degenhart, R.E. (1988). *The IEA Study of Written Composition I: The International Writing Tasks and Scoring Scales*. Oxford: Pergamon Press.
- Glopper, K. de (1988). *Schrijven beschreven*. Den Haag, SVO: Selecta Reeks.
- Hillocks, G. Jr. (1986). *Research on Written Composition. New Directions for Teaching*. Urbana, Illinois, NCRE/ERIC.
- Jencks, C. et al. (1972). *Inequality*. New York: Basic Books.
- Kish, L. (1965). *Survey Sampling*. New York: John Wiley & Sons.
- Mortimore, P., Sammons, P., Stool, L., Lewis, D. & Ecob, R. (1988). *School Matters: the Junior Years*. Somerset: Open Books.
- NAEP (1986). *The Reading Report Card. Progress Toward Excellence in Our Schools. Trends in Reading over Four National Assessments, 1971-1984*. Princeton NJ: ETS/NAEP.
- NOS (1990). *Continu Kijk Onderzoek Jeugd. Tweede kwartaal 1990*. Hilversum: NOS, Afdeling Kijk- en Luisteronderzoek.
- Otter, M.E. (1987). *Tijdbesteding aan moedertaalonderwijs op vier manieren gemeten*. Amsterdam: SCO.
- Reynolds, D., Creemers, B.P.M. & Peters, T. (1989). *School Effectiveness and Improvement. Proceedings of the First International Congress, London, 1988*. Cardiff: School of Education/Groningen: RION.

- Rijlaarsdam, G.C.W. (1986). *Effecten van leerlingenrespons op aspecten van de stelvaardigheid*. Amsterdam: SCO.
- Scheerens, J. (1989). *Wat maakt scholen effectief? Samenvatting en analyse van onderzoeksresultaten*. Den Haag: SVO.
- Schuurs, U.R.I. (1990). *Leren schrijven voor lezers*. Universiteit Twente: Faculteit Wijsbegeerte en Maatschappijwetenschappen. WMW-publikatie 3.
- Thorndike, R.L. (1973). *Reading Comprehension Education in Fifteen Countries*. Stockholm: Almqvist & Wiksell.
- Walberg, H.J. (1986). Syntheses of Research on Teaching. In: M.C. Wittrock (ed.). *Handbook of Research on Teaching*. Third Edition. New York/London, Macmillan, 214-229.
- Wesdorp, H. (1983). *Schrijven in het voortgezet onderwijs; een overzicht van het onderzoek naar de effecten van diverse instructie-variabelen op de stelvaardigheid*. Harlingen: Flevodruk.
- Weterings, A.C.E.M. & Aarnoutse, C.A.J. (1986). De praktijk van het onderwijs in begrijpend lezen. *Pedagogische Studiën* 63, 387-400.
- Zwarts, M. (red.) (1990). *Balans van het taalonderwijs aan het einde van de basisschool. Uitkomsten van de eerste taalpeiling einde basisonderwijs*. Arnhem: Cito.



## Bijlage 1

### Instructie Tekstkenmerken

Welke nadruk krijgen de onderstaande vakonderdelen in uw onderwijs aan deze klas? Welke nadruk hebben deze vakonderdelen gekregen in het moedertaalonderwijs dat deze leerlingen in voorafgaande leerjaren ontvangen hebben, voorzover u daarvan op de hoogte bent?

	Veel nadruk in mijn onderwijs aan deze klas	Weinig nadruk in mijn onderwijs aan deze klas	Veel nadruk in de voorafgaande leerjaren	Weinig nadruk in voorafgaande leerjaren
--	---------------------------------------------	-----------------------------------------------	------------------------------------------	-----------------------------------------

De structuur van verschillende tekstsoorten (verhalend, argumenterend)

Esthetica en literatuur-theorie (kennis van genres en stromingen)

Fictioneel proza

Zakelijk proza

### Oefeningen Tekstproductie

Hoe vaak maakt u gebruik van de volgende oefeningen binnen uw stelonderwijs aan deze klas?

	Zelden of nooit, hoogstens één keer per jaar	Soms, één of enkele malen per kwartaal	Vaak, één of enkele malen per maand
--	----------------------------------------------	----------------------------------------	-------------------------------------

Oefeningen in het structureren van alinea's en gehele stelproducten

Oefeningen in het opbouwen van een betoog; hoe een onderwerp systematisch te behandelen

Oefeningen in het schrijven van onderdelen van een betoog; alleen de inleiding; de hoofdvraag; het slot

Oefeningen in het verzamelen en selecteren van materiaal uit boeken, handboeken e.d.

Oefeningen in het denken over een onderwerp en het ontwikkelen van een stelproduct (bv. het in grote lijnen opzetten)

	Zelden of nooit, hoogstens één keer per jaar	Soms, één of enkele malen per kwartaal	Vaak, één of enkele malen per maand
Oefeningen in beeldspraak, stijlfiguren en het concreet of juist beeldend gebruik van taal			
Oefeningen in het analyseren van het publiek van het stel- produkt en oefeningen in het schrijven voor een bepaald pu- blik			
Oefeningen in de verhalende tekstsoort			
Oefeningen in de beschrijvende tekstsoort			
Oefeningen in de uiteenzettende, beschouwende tekstsoort			
Oefeningen in de argumente- rende tekstsoort			

## Feedback

Verschillende personen kunnen commentaar geven op de stelprodukten die leerlingen schrijven. Geef aan hoe vaak de onderstaande mogelijkheden zich in de klas voordoen.

	Zelden of nooit, hoogstens één keer per jaar	Soms, één of enkele malen per kwartaal	Vaak, één of enkele malen per maand
Leerlingen krijgen commentaar, suggesties en/of verbeteringen van de eigen docent			
Leerlingen krijgen commentaar suggesties en/of verbetering van een of meer andere docenten of hospitanten			
Leerlingen krijgen commentaar, suggesties en/of verbeteringen van hun klasgenoten			
Leerlingen geven commentaar e.d. op hun eigen stelprodukten			

Op welke wijze geven verschillende personen commentaar op de stelprodukten die uw leerlingen schrijven? Geef aan hoe vaak de onderstaande mogelijkheden zich in deze klas voordoen.

	Zelden of nooit, hoogstens één keer per jaar	Soms, één of enkele malen per kwartaal	Vaak, één of enkele malen per maand
Commentaar van de docent (of een andere volwassene) wordt in geschreven vorm aan de leerling aangeboden: elke leerling krijgt beoordeelde/gecorrigeerde stelopdracht terug om voor zichzelf door te lezen			
Commentaar van de docent (of van een andere volwassene) wordt in een persoonlijk gesprek tijdens of na de les aangeboden			
Commentaar van de docent (of een andere volwassene) wordt aangeboden aan de hele klas: de belangrijkste fouten uit een aantal stelprodukten worden bediscussieerd			
Commentaar van klasgenoten wordt gegeven in kleine groepjes, waarbij leerlingen elkaars stelprodukten bespreken			
Commentaar van klasgenoten wordt in schriftelijke vorm uitgewisseld			

Commentaar op stelprodukten kan op verschillende momenten gegeven worden. Geef aan hoe vaak de onderstaande mogelijkheden zich in de klas voordoen.

	Zelden of nooit, hoogstens één keer per jaar	Soms, één of enkele malen per kwartaal	Vaak, één of enkele malen per maand
Ik geef commentaar terwijl de leerlingen bezig zijn de stelopdracht te maken			
Ik geef commentaar e.d. meteen na het maken van de stelopdracht (tot 3 dagen erna)			
Ik geef commentaar kort na het maken van de stelopdracht (4 tot 7 dagen erna)			



Docenten kunnen hun beoordeling van stelprodukten op verschillende manieren uitdrukken. Geef over elk van de onderstaande manieren aan hoe vaak u die gebruikt bij uw stelonderwijs in deze klas.

	Zelden of nooit, hoogstens één keer per jaar	Soms, één of enkele malen per kwartaal	Vaak, één of enkele malen per maand
Ik ken aparte cijfers, letters of woorden toe aan een aantal aspecten van het stelprodukt (bv. inhoud, organisatie)			
Ik geef een algemeen commentaar in geschreven vorm aan het eind van het stelprodukt in 2 of 3 regels (globale opmerkingen over de kwaliteit)			
Ik geef gedetailleerd commentaar (inclusief suggesties voor verbeteringen in de vorm van geschreven opmerkingen in de kantlijn en/of tussen de regels			
Ik geef gedetailleerd commentaar (inclusief suggesties voor verbetering) op slechts één bepaald aspect van het stelprodukt (bv. inhoud, stijl)			
Ik geef commentaar op sterke en zwakke punten en markeer verschillende soorten fouten en zwakheden en de van toepassing zijnde correcties met symbolen die bij de leerlingen bekend zijn			
Ik geef commentaar op sterke en zwakke punten en geef verschillende soorten fouten en zwakheden aan in de tekst en op een aparte lijst waarop de belangrijkste soorten fouten en moeilijkheden beschreven zijn			
Ik geef een sterk punt aan in elk stelprodukt			
Ik geef commentaar en stel verbeteringen voor; vervolgens herzien de leerlingen hun stelprodukt geheel: ze schrijven een 'herziene versie'			
Ik geef commentaar en stel verbeteringen voor; vervolgens verbeteren de leerlingen alleen de gebrekkige passages (gedetailleerde revisie)			
Ik geef commentaar en stel verbeteringen voor, maar ik vraag geen revisie van de leerlingen			