

Van der Wolfs onderzoek naar schooluitval

Verslag van een secundaire analyse*

M. KALMIJN EN
P.G. SWANBORN

*Sociologisch Instituut, Rijksuniversiteit
Utrecht*

Samenvatting

De belangrijkste conclusie uit de dissertatie van Van der Wolf (1984) is dat leerlingen op scholen met een meer traditioneel onderwijsleerklimaat meer schoolsucces hebben dan leerlingen op progressieve scholen. De methodologische kritiek op zijn onderzoek heeft twijfel gezaaid over de geldigheid van die conclusie. In een heranalyse van Van der Wolfs data laten we zien dat de clusters met behulp waarvan Van der Wolf zijn belangrijkste hypothese bevestigde nauwelijks verschillen in onderwijsstijlkenmerken. Vervolgens analyseren we op schoolniveau. De partiële correlaties tussen de onderwijsstijl en sommige indicatoren van schoolsucces blijven significant. Dit resultaat kan worden gezien als een gedeeltelijke bevestiging van Van der Wolfs conclusie. Het is zeker geen sterke bevestiging gezien de ernstige tekortkomingen in de data.

1 Inleiding

In 1984 is de orthopedagoog Van der Wolf gepromoveerd op het proefschrift 'Schooluitval'. Hierin wordt verslag gedaan van een onderzoek naar de invloed van schoolinterne factoren op schooluitval. Als schooluitvalindicatoren beschouwt de auteur het percentage verwijzingen naar het buitengewoon on-

* Met dank aan J.C. van der Wolf voor het ter beschikking stellen van zijn data en aan J. Hox en B. Schijf voor de technische assistentie daarbij. Tevens danken wij P. de Graaf en H. Ganzeboom voor hun commentaar.

derwijs, het percentage doorstromers naar het LBO, LEAO en LAVO, het percentage doorstromers naar het MAVO, en het percentage zittenblijvers. Voor wat betreft de schoolinterne factoren legt de auteur de nadruk op de onderwijsstijl. Centraal in het onderzoek en in de publiciteit die het heeft gekregen, staat het ook door Van der Wolf niet verwachte resultaat dat er op scholen met een meer traditionele onderwijsstijl minder 'schooluitval' voorkomt dan op scholen met een meer progressieve of vernieuwende onderwijsstijl. Het onderzoek van Van der Wolf sluit aan op een halverwege de jaren zestig in de Verenigde Staten begonnen onderzoekstraditie waarin naast de invloed van milieufactoren de invloed van het onderwijsleerklimaat op leerlingprestaties wordt onderzocht (zie voor een recent overzichtartikel Van de Grift, 1985).

Op dit proefschrift is methodologische kritiek gekomen (De Graaf & Swanborn, 1985; Kreft, 1985; Dronkers, 1986). De suggestie wordt gewekt dat aan de geldigheid van de conclusies met recht getwijfeld kan worden. In dit artikel rapporteren we over een heranalyse van Van der Wolfs data. Voor een goed begrip is het nodig vooraf Van der Wolfs werkwijze, zijn belangrijkste resultaten en de kritiek die er op het onderzoek gekomen is samen te vatten.

2 De werkwijze van Van der Wolf

Dataverzameling

a. Gedurende enige jaren vindt er op een aantal scholen en een Pedagogische Academie in Amsterdam een kwalitatief onderzoek plaats. Deze onderzoeksfase levert een aantal schoolkenmerken op die relevant zouden zijn voor de probleemstelling.

b. Op 13 openbare lagere scholen in Amsterdam worden gegevens verzameld: 1411 leerlingen worden klassikaal schriftelijk geënquêteerd en 122 leerkrachten vullen een vragenlijst in over hun achtergrond, hun manier van lesgeven en de wijze waarop zij omgaan

met problemen op school.

c. Op schoolniveau worden onder meer gegevens verzameld over 'schooluitval'. De afhankelijke variabelen zijn dus op schoolniveau gemeten.

Analyse van leerkrachtgegevens (N = 122)

a. Uit de gegevens worden 29 samengestelde leerkrachtvariabelen geconstrueerd.

b. Alle leerkrachtvariabelen worden gezamenlijk geanalyseerd met een principale componenten analyse. Van der Wolf vindt drie componenten die inhoudelijk verre van eenduidig zijn. Toch weet Van der Wolf deze onder drie namen te vangen: controlerend (o.a. leerstofgericht en gebruiken van straf en beloningen), steunzoekend (o.a. aversie tegen regels en dwang, minder gericht op schoolse kennis) en coöperatief (o.a. tevredenheid over schoolhoofd en collega's, tevredenheid over didactische werkwijze van de school en weinig waargenomen geplaagde kinderen). De componenten bevatten niet alleen onderwijsstijlkenmerken, maar ook zaken als mening over het schoolgebouw, percentage waargenomen lastige kinderen, klassegrootte en taakbelasting.

c. Op basis van de drie componenten maakt Van der Wolf drie schalen. Hoe de schalen precies zijn geconstrueerd, wordt niet vermeld.

d. De leerkrachtschalen worden geaggregeerd tot variabelen op schoolniveau. Van der Wolf vermeldt niet of de variantie vooral tussen de scholen optreedt – in dat geval kan aggregatie zinvol zijn – of dat juist ook tussen de leerkrachten binnen elke school verschillen bestaan.

Analyse op schoolniveau (N = 13)

a. Onderzocht wordt in welke mate de scholen verschillen in milieusamenstelling en het percentage mediterrane leerlingen. Milieusamenstelling wordt alleen indirect gemeten via de zogenaamde 'onderwijsstimuleringspunten'. Deze gegevens zijn echter incompleet (vier van de dertien scholen hebben geen score) en worden door Van der Wolf ook niet betrouwbaar geacht. Er bestaan verschillen tussen de scholen op genoemde kenmerken, maar er wordt niet getoetst op significantie. Krefte heeft dit wel gedaan voor het percentage mediterrane leerlingen en deze verschillen bleken significant (Krefte, 1985). Dit wijst er

op dat er verschillen in aanvangsniveau kunnen zijn.

b. Op basis van de drie afhankelijke variabelen (zittenblijven, doorstroomnaar het LBO enz., doorstroom naar het MAVO en verwijzing naar het BUO) worden de scholen via een clusteranalyse tot drie clusters gegroepeerd (nadat één school uit de analyse is verwijderd). Zo is er een cluster met veel schooluitval, een met weinig 'schooluitval' en een er tussen in. Het eerste cluster noemt hij 'prestatiegericht', het tweede 'affectief welzijnsgericht' en het derde noemt hij 'cognitief welzijnsgericht'. Daarbij moet opgemerkt worden dat de afzonderlijke outputvariabelen niet tot dezelfde rangorde van scholen leiden. De namen wijzen op een indeling op basis van schoolkenmerken. De clustering is echter gebaseerd op schoolsuccesindicatoren. De uitspraak 'op "prestatiegerichte" scholen hebben de leerlingen meer schoolsucces dan op "welzijnsgerichte" scholen' is dus per definitie waar.

Analyse op clusterniveau

a. Er wordt nagegaan of de clusters verschillen in beginniveau. Het beginniveau wordt geïndiceerd door Cito-leestoetsen die zijn afgenomen in de eerste klas van elke school. De verschillen tussen de clusters zijn niet significant. De clusters hebben dus een gelijk gemiddeld beginniveau.

b. Om de probleemstelling (dragen schoolkenmerken bij tot schooluitval?) direct te beantwoorden worden de gemiddelden op de schalen 'steunzoekend', 'controlerend' en 'coöperatief' tussen de clusters vergeleken. De verschillen zijn significant. Het cluster met de meeste 'schooluitval' scoort het laagst op de controlerende dimensie en het laagst op de coöperatieve dimensie. Het cluster met de minste schooluitval scoort het laagst op de steunzoekende dimensie (blz. 93). Dit is het belangrijkste resultaat in het kader van de probleemstelling.

De analyse van de gegevens uit de leerling-enquête (N = 1411) zullen wij hier niet uitgebreid bespreken. Op een soortgelijke wijze als bij de leerkrachtgegevens worden er vijf leerlingenschalen geconstrueerd: agressiviteit, eenzaamheid, negatieve mening over de gang van zaken in de klas, slechte ervaring met leerkrachten en negatieve mening over mede-

leerlingen. De schalen worden geaggregeerd tot variabelen op klasniveau. Deze variabelen worden gecorreleerd met leerkrachtvariabelen. Hierbij worden de leerlingsschalen als afhankelijke variabelen opgevat. Daarnaast worden de verschillen tussen de clusters op deze schalen onderzocht.

3 De kritiek

Ten aanzien van de *onafhankelijke variabelen* blijven nogal wat procedures onduidelijk. Daarnaast wordt bij de aggregatie van leerkrachtvariabelen niet vermeld hoe groot de verschillen per school tussen de leerkrachten zijn. Als leerkrachten op een school zeer uiteenlopende oordelen hebben over wat er op school gebeurt, kunnen we niet over een betrouwbare schoolvariabele spreken.

Bovendien bestaan de drie schalen uit een merkwaardig samenraapsel van kenmerken. De waardering van het schoolgebouw rekent de auteur bij dezelfde dimensie als de beoordeling van de samenwerking met collega's ('coöperatieve dimensie') en het positief waarderen van de contacten met de ouders rekent hij bij dezelfde dimensie als het niet gericht zijn op schoolse kennis ('steunzoekende dimensie'). Deze voorbeelden laten zien dat de auteur zich bij de toewijzing van schoolkenmerken aan dimensies niet door voorafgaande inhoudelijke ideeën heeft willen laten leiden. Gevolg is dat de gevonden resultaten ten aanzien van de schoolinterne factoren niet eenduidig iets over de invloed van onderwijsstijlkenmerken zeggen. Het idool van de 'in de empirie gefundeerde theorie' waardoor Van der Wolf zich in het kwalitatieve vooronderzoek laat leiden, lijkt hem ook hier parten te spelen: de empirie als 'rockbottom' van onze kennis. Dat procedure artefacten teweeg kunnen brengen, en vooral dat zonder theoretische inzichten in de aard van de variabelen en hun relaties data weinig voorstellen, komt de lezer herhaaldelijk in gedachten.

Ten aanzien van de *afhankelijke variabelen* zijn eveneens kritische kanttekeningen te plaatsen. Van der Wolf brengt de drie afhankelijke variabelen samen onder de titel 'schooluitval'. Dit is een wat ongelukkig gekozen naam, omdat doorstroming naar het voortgezet onderwijs niet als 'schooluitval'

beschouwd kan worden. Omdat het de auteur wel degelijk gaat om het falen van kinderen op school (blz. 9), zou de titel 'schoolsucces' een betere keuze zijn. BUO-verwijzingen en doorstroom naar het voortgezet onderwijs kunnen beide onder deze noemer worden gebracht. Het zittenblijven is in dit kader geen geschikte variabele. In welke mate scholen hun leerlingen laten doubleren is behalve van hun prestaties ook afhankelijk van de onderwijsstijl van de school. Op sommige scholen probeert men de leerlingen die achter zijn met extra individuele begeleiding bij te werken zonder dat men ze laat doubleren. Daardoor is het best mogelijk dat twee scholen met gemiddeld even goede prestaties leverende leerlingen sterk verschillen in het percentage zittenblijvers. Wij zullen in de analyse deze variabele niet opnemen¹.

Het construeren van clusters op basis van de afhankelijke variabelen is niet verstandig omdat daardoor invloeden van verschillende onafhankelijke variabelen niet meer gescheiden kunnen worden. Bovendien is het geven van namen aan clusters die een indeling op basis van onafhankelijke variabelen suggereren onjuist en misleidend. Een belangrijker argument tegen clusters van scholen is echter dat het onnodige verspilling van informatie is. Alle relevante variabelen zijn minstens op schoolniveau gemeten en de analyse kan dus op dat niveau plaatsvinden.

Het is interessant zich te verdiepen in dit punt. Clustering van scholen kan bijna alleen begrepen worden vanuit de veronderstelling van de onderzoeker dat er 'typen van scholen' bestaan die systematische verschillen vertonen op zowel de onafhankelijke als de afhankelijke variabelen. Als we op schoolniveau analyseren dreigen we deze systematische verschillen niet meer te kunnen waarmaken door de aanwezigheid van ideosyncratische variantie. Aggregatie tot clusters zou tot wegmiddelen van deze ideosyncratische variantie leiden. Gezien het kleine aantal scholen is deze laatste veronderstelling echter weinig realistisch. Zeker als we bedenken dat temidden van de veelheid van kenmerken waarop scholen van elkaar verschillen er niet één is die qua samenhang met schoolsucces er duidelijk uitspringt.

Ten aanzien van de *mogelijk storende invloeden* merken we op dat het onduidelijk is

wat Van der Wolf doet met het percentage mediterrane leerlingen. Hij vermeldt slechts dat de clusters niet significant verschillen in deze percentages (blz. 94). Sommige tabellen staan tweemaal in het proefschrift, maar een tabel hiervan ontbreekt. Er wordt ook niet getoetst of de scholen significant verschillen in deze percentages. Zoals Kreft heeft laten zien is dat duidelijk wel het geval (Kreft, 1985).

Daarnaast heeft de auteur het Cito-leesniveau in de eerste klas opgenomen. Op zichzelf is dat een goede controlevariabele. De gegevens hebben echter betrekking op meerdere generaties zesde-klassers. Als het beginniveau van de leerlingen in de loop der tijd voor een aantal scholen verschillend is veranderd, dan is deze variabele niet geschikt. De afhankelijke variabelen zijn immers op één moment, namelijk in 1982, gemeten.

Met betrekking tot de derde controlevariabele, sociaal milieu, merkt Van der Wolf zelf op dat deze onbetrouwbaar is gemeten. Het weglaten van een dergelijke variabele is een ernstige omissie. Het constant houden op de invloed van het Cito-niveau in de eerste klas is niet voldoende, omdat de invloed van het sociaal milieu gedurende de gehele lagere schoolloopbaan aanwezig is. De Graaf vond bijvoorbeeld bij een heranalyse van de gegevens van Meijnen effecten van het sociaal economisch milieu op het advies van de onderwijzer en het prestatieniveau in de zesde klas, gegeven de invloed van het prestatieniveau in de eerste en derde klas (De Graaf, 1985, blz. 359). Deze effecten bleken van dezelfde omvang als de effecten van milieu op de eerder gemeten prestatiescores.

We sluiten deze kritiek af met de stelling dat de analyses nodeloos ingewikkeld zijn. De probleemstelling laat zich op een veel eenvoudiger wijze analyseren. In de rest van dit artikel wordt van zo'n analyse verslag gedaan.

4 *Secundaire analyses*

Bij de secundaire analyses hebben we ons beperkt tot de vraag in hoeverre een traditionele/progressieve manier van lesgeven samenhangt met schoolsucces. Eerst gaan we na of Van der Wolfs clusters verschillen in ken-

merken van de onderwijsstijl. Daarna analyseren we op schoolniveau de verbanden tussen schoolsucces en onderwijsstijl. Tot slot toetsen we een paar extra hypothesen die gaan over schijnbaar causale verbanden en interpretaties van verbanden.

4.1 *De clusters van Van der Wolf nader bekeken*

We beperken ons in deze analyse tot leerkrachtvariabelen die inhoudelijk betrekking hebben op de onderwijsleersituatie (zie Tabel 1). Met deze selectie trachten we, voor zover de data het toelaten, zoveel mogelijk aan te sluiten bij eerder verricht onderzoek naar schoolkenmerken (Bennett, 1976; Meijnen, 1982, blz. 135-155). De door ons gekozen kenmerken vallen deels onder 'organisatie en management', 'controle en sanctionering' en 'inhoud van het curriculum' (Meijnen, 1982, p. 145-147).

Met variantieanalyse gingen we na in welke mate de variantie in 20 leerkrachtvariabelen veroorzaakt wordt door verschillen tussen leerkrachten en in welke mate ze veroorzaakt wordt door verschillen tussen de scholen. Voor slechts zeven variabelen was de F-toets uit de variantieanalyse significant². Voor de overige variabelen zit de spreiding meer binnen de scholen (tussen de leerkrachten) dan tussen de scholen. Dat betekent dat het niet erg veel zin heeft deze variabelen als schoolkenmerken op te vatten. Het betekent echter niet dat de clusters niet significant kunnen verschillen op deze variabelen. Ook kan een schaal bestaande uit deze variabelen wel significante verschillen vertonen. We houden deze variabelen daarom in de analyse.

Om de betrouwbaarheid van de meting van de onderwijsstijl te vergroten hebben we een schaal gemaakt op de genoemde leerkrachtvariabelen. Een factoranalyse leverde na rotatie geen eenvoudige structuur in twee of meer duidelijk onderscheiden dimensies op. Bij de schaalconstructie bleek dat de betrouwbaarheid substantieel verhoogd kon worden door het verwijderen van zeven items. De geconstrueerde schaal noemen we ONDERWIJSSTIJL (Cronbach's alpha is .68; 13 items, N = 122)³.

We gaan vervolgens na in hoeverre de clusters significant verschillen op de 20 enkelvoudige onderwijsstijlvariabelen en op de

schaal ONDERWIJSSTIJL (zie Tabel 1). (Alle onderwijsstijlkenmerken zijn zodanig gecodeerd dat een hoge score een traditionele onderwijsstijl betekent. Cluster 1 heeft minder 'schooluitval' dan cluster 2, en cluster 2 heeft minder 'schooluitval' dan cluster 3.)

We zien dat de clusters significant verschillen op slechts vier van de 20 kenmerken. Clusters van scholen die meer 'schooluitval' vertonen dan andere clusters geven minder vaak huiswerk op, laten hun leerlingen vaker zelf hun plaats kiezen, eisen minder vaak van hun leerlingen dat ze stil moeten zijn en belonen hun leerlingen minder vaak. De overige kenmerken laten geen of niet significante verschillen tussen de clusters zien. Voor 9 kenmerken zijn de niet-significante verschillen wel in de verwachte richting. Voor 7 kenmerken is dat niet het geval. De clusters blijken voorts (net) niet significante verschillen te vertonen op de schaal ONDERWIJSSTIJL. Deze waarde ligt op de rand van significantie ($p = .06$). De verschillen zijn ook in de richting die overeenkomt met Van der Wolfs conclusie (cluster 1 heeft een meer

traditionele onderwijsstijl dan cluster 2 en cluster 2 heeft een meer traditionele onderwijsstijl dan cluster 3).

De conclusie is dat de clusters van scholen niet overtuigend verschillen in onderwijsstijl. De hypothese van Van der Wolf kan niet goed worden bevestigd als we hem volgen in zijn weinig voor de hand liggende clusterprocedure en daarbij een eigen en eenvoudiger te interpreteren operationalisering van de onderwijsstijl gebruiken. Het ligt echter meer voor de hand hypothesen te toetsen op het niveau met het grootst mogelijk aantal eenheden, dat wil zeggen op schoolniveau.

4.2 Verbanden op schoolniveau

Om de hypothese te toetsen kunnen we volstaan met het berekenen van de correlaties tussen de schoolsuccesindicatoren en de schaal ONDERWIJSSTIJL. Het zittenblijven beschouwen we zoals reeds gezegd niet als indicator voor schoolsucces.

We gebruiken twee controlevariabelen in de analyse: percentage buitenlandse leerlingen en Cito-leestoets. Daarmee is maar zeer

Tabel 1 Variantie-analyses op schoolkenmerken (12 scholen* zijn eenheden en 3 clusters zijn groepen); $p < .05$

Schoolkenmerk	Gemiddelde van cluster			F-ratio
	1	2	3	
Tafeltjes twee aan twee	1.3	1.0	1.1	0.6 ns
Vrij rondlopen toegestaan (-)	1.7	1.7	1.6	0.3 ns
Vrij praten toegestaan (-)	1.8	1.7	1.5	1.2 ns
Toestemming vragen bij klas verlaten	1.5	1.5	1.5	0.1 ns
Verplicht stil zijn	1.5	1.3	1.2	9.2
Huiswerk krijgen	1.4	1.1	1.3	5.2
Samenwerken aan zelf te kiezen taak (-)	3.8	3.7	3.5	1.8 ns
Samenwerken aan opdrachten (-)	3.4	3.6	3.4	0.4 ns
Zelfstandig werken aan zelf te kiezen taak (-)	3.8	3.5	3.2	2.6 ns
Cijfers op werk schrijven	1.6	1.6	1.4	0.3 ns
Beloningen geven	1.7	1.3	1.2	5.1
Voorrechten intrekken	1.4	1.3	1.3	0.1 ns
Leerlingen bepalen zelf waar ze zitten (-)	3.2	2.9	2.7	7.3
Leerlingen op grond van prestaties op aparte plaats	0.9	0.8	0.8	0.6 ns
Tafels uit het hoofd leren	1.9	1.8	1.9	0.8 ns
Frontaal tot klas spreken	2.0	2.0	2.2	0.5 ns
Cijfers hardop voorlezen	1.2	1.1	1.2	0.6 ns
Strafwerk geven	1.4	1.3	1.3	0.6 ns
Leerlingen klas uit sturen	1.0	1.1	1.1	1.5 ns
Leerlingen naar hoofd sturen	1.4	1.7	1.4	2.1 ns
ONDERWIJSSTIJL (Z-scores)	0.199	-0.072	-0.242	4.0 ns

* één school is niet in een van de clusters opgenomen (zie Van der Wolf, 1984, p. 94)

ten dele rekening gehouden met alle variabelen van een adequaat model (zie paragraaf 3).

Bij de inspectie van de strooiingsdiagrammen bleek de school die Van der Wolf uit de analyses had gelaten steeds een duidelijke uitbijter te zijn. Verwijdering van deze school heeft tot gevolg dat correlaties van de onderwijsstijl met de doorstroomindicatoren met gemiddeld .05 omhoog gaan. De correlatie tussen het percentage verwijzingen naar het BUO en de onderwijsstijl verandert zelfs spectaculair: van $-.13$ naar $-.43!$

Nets als in de vorige paragraaf laten we deze school hier uit de analyse. Dit verkleint het aantal gegevens met 8%, we houden nog 12 scholen over. Zo'n klein aantal eenheden is niet gebruikelijk in statistische analyse van data, maar dat wil niet zeggen dat er dan geen correlaties kunnen worden berekend. Een klein verband zal echter minder snel significant zijn. De power is derhalve gering.

Ondanks het feit dat het aantal eenheden zeer klein is zijn sommige correlaties zo hoog dat ze significant zijn.

(De correlatiematrix is opgenomen in de Bijlage.)

De (partiële) correlaties met ONDERWIJSSTIJL zijn deels wel significant en in de voorspelde richting. Naarmate een school een traditionelere manier van lesgeven heeft, is de doorstroom naar het MAVO hoger (net niet significant) en de doorstroom naar het LBO lager, gegeven de invloed van het aanvangsniveau zoals gemeten door de Cito-leestoets in de eerste klas en het percentage mediterrane leerlingen. De (partiële) correla-

Tabel 2 *Partiële en gewone correlaties tussen schoolsuccesindicatoren en ONDERWIJSSTIJL constant gehouden op percentage mediterrane leerlingen en aanvangsniveau (N=12 en $p < .05$)*

	ONDERWIJSSTIJL	
	corr.	part.corr.
Perc. doorstroom naar het LBO	$-.56$	$-.56$
Perc. doorstroom naar het MAVO	$.57$	$.54$ ns
Perc. naar de brugklas (VWO/HAVO)	$-.02$ ns	$-.05$ ns
Percentage BUO-verwijzingen	$-.43$ ns	$-.48$ ns

ties met de doorstroom naar de brugklas zijn vrijwel nul. De onderwijsstijl zorgt dus hoofdzakelijk voor een differentiatie tussen de twee laagste niveaus van doorstroming. Het percentage verwijzingen naar het BUO is lager naarmate de onderwijsstijl traditioneler is, gegeven de invloed van de controlevariabelen. Dit partiële verband is kleiner dan de verbanden met de doorstroom en is niet significant. Opgemerkt moet worden dat door constant houden op de invloed van het aanvangsniveau de partiële correlaties hoger worden en dat door constant houden op de invloed van het percentage mediterrane leerlingen de partiële correlaties weer afnemen.

De conclusie is dat in enkele gevallen de resultaten wijzen in de richting van een bevestiging van de hypothese van Van der Wolf. Veel verbanden zijn echter niet significant, hetgeen met een zo klein aantal eenheden ook niet verwacht kon worden. Door het feit dat de verbanden niet significant zijn en het feit dat niet alle relevante controlevariabelen zijn gemeten, kunnen we niet spreken van een deugdelijke bevestiging. Een falsifiëring is het echter zeker niet.

Open blijft de interpretatie van de verbanden. De vraag is ten eerste of het verband als causaal verband is op te vatten (onderwijskenmerken leiden tot 'schooluitval'). Stel dat er van causaliteit sprake is, dan kunnen we ons voorts afvragen hoe de werking van de schoolkenmerken opgevat moet worden. Deze vragen komen hierna aan de orde. Interpretatie van gegevens op geaggregeerd niveau blijft overigens een hachelijke zaak, omdat men nooit weet bij welke mensen een hoge score op een variabele samen gaat met een hoge score op een andere variabele. Zo kan bijvoorbeeld de vraag of het juist kinderen uit arbeidersmilieus zijn die van een strakkere onderwijsleersituatie profiteren niet worden beantwoord⁴.

4.3 De verbanden nader bekeken

De hypothese luidt dat een traditionele onderwijsstijl tot minder 'schooluitval' leidt omdat kinderen meer en beter leren op zulke scholen. De onderwijsstijlkenmerken worden verondersteld een causale invloed te hebben. Voorwaarde voor causaliteit is onder andere dat een verband niet het gevolg mag zijn van een gemeenschappelijke achterliggende variabele. In het materiaal konden we

een mogelijk achterliggende variabele vinden: de ervaring van de leerkracht. Omdat jonge onderwijzer(e)s(en) niet lang geleden hun opleiding tot onderwijzer(es) hebben afgerond is het mogelijk dat ze optimistisch zijn over de mogelijkheid van het in praktijk brengen van onderwijskundige idealen. Jonge leerkrachten zouden dus een meer progressieve onderwijsstijl kunnen hebben. Als jonge leerkrachten door hun onervarenheid er tevens toe bijdragen dat de prestaties van de leerlingen minder zijn dan bij ervaren leerkrachten, kan er een schijnbaar causaal verband bestaan tussen onderwijsstijl en leerprestaties.

Op leerkrachtniveau hebben we het verband berekend tussen het aantal jaren onderwijservaring en onderwijsstijlkenmerken. Het blijkt echter dat onervaren leerkrachten nauwelijks een meer progressieve onderwijsstijl hebben dan ervaren leerkrachten. Het verband met ONDERWIJSSTIJL is slechts .03 en niet significant. Er kan dus geen sprake zijn van een schijnbaar causaal verband veroorzaakt door de ervaring van de leerkracht. De hypothese moet worden verworpen.

Als we van een causale invloed uitgaan kunnen we ons verder afvragen hoe deze invloed loopt. In het materiaal hebben we een mogelijk intermedieërende variabele kunnen vinden: de orde in de klas. We kunnen nagaan of de leerlingen op scholen met een traditionele (strengere) onderwijsstijl betere prestaties leveren omdat deze scholen er in slagen orde te houden, waardoor de leerlin-

gen zich beter op het werk kunnen concentreren en daardoor tot betere prestaties kunnen komen. We hebben deze hypothese als volgt getoetst. In de enquête voor de leerlingen is op twee manieren gevraagd naar de orde in de klas. Ten eerste is gevraagd naar wat leerlingen waarnemen dat er in de les gebeurt ('mijn klasgenoten zijn erg rustig', 'kinderen uit mijn klas plagen elkaar gemeen', 'in de klas zijn de kinderen meestal lekker bezig' enz.). Ten tweede is gevraagd naar de activiteiten van de leerlingen zelf (andere kinderen van het werk houden, vechten met andere kinderen, de juf voor de gek houden, in de bank krassen enz.). De scores op de twee soorten kenmerken hebben we gesommeerd tot twee schalen, respectievelijk ORDE 1 en ORDE 2⁵. Om verbanden met de onderwijsstijl en 'schooluitval' te berekenen hebben we deze schalen geaggregeerd tot variabelen op schoolniveau⁶. Volgens de hypothese moet het verband tussen onderwijsstijl en schoolsucces verdwijnen als we constant houden op de invloed van de orde in de klas. De partiële correlaties staan in Tabel 3⁷. De gewone correlaties staan in de Bijlage.

Uit de correlaties (zie Bijlage) zien we dat traditionele scholen er volgens de leerlingen zelf beter in slagen orde te houden dan niet-traditionele scholen, en dat naarmate er meer orde in de klassen is het percentage dat doorstroomt naar het LBO veel lager is en het percentage dat doorstroomt naar het MAVO veel hoger is. De mate van orde in de klas heeft geen significante invloed op de doorstroom naar de brugklas. De hypothese ondervindt geen volledige ondersteuning omdat bij constant houden op de invloed van de tweede orde-variabele het verband tussen de onderwijsstijl en de doorstroom (LBO en MAVO) nagenoeg verdwijnt, terwijl dit niet het geval is als we constant houden op de invloed van de eerste orde-variabele (Tabel 3). Alleen als we aan de leerlingen vragen wat ze zelf op school voor kattenkwaad uitvoeren wordt het verband dus veel kleiner. Dit ligt aan het feit dat de correlatie tussen schoolsucces en de tweede orde-variabele veel hoger is dan tussen schoolsucces en de eerste orde-variabele waarin wordt gevraagd naar de waarneming van de orde in de klas (.85/-.74 tegen .26/-.51). Het is niet duidelijk waarom het verband tussen schoolsuc-

Tabel 3 *Partiële correlaties tussen schoolsuccesindicatoren en ONDERWIJSSTIJL constant gehouden op: (1) aanvangsniveau sec; (2) aanvangsniveau en de orde in de klassen op school volgens twee metingen (N=12 en p<.05)*

	(1)	(2)
	ORDE 1 ORDE 2	
Perc. doorstroom naar het LBO	-.56	-.53 ns
Perc. doorstroom naar de MAVO	.54 ns	.43 ns
Perc. doorstroom naar HAVO/VWO	-.05 ns	.36 ns
Perc. BUO-verwijzingen	-.43 ns	-.37 ns

ces en orde hoger is als we aan de leerlingen vragen wat ze zelf uithalen.

Een onverwacht resultaat is dat, terwijl de gewone correlatie niet significant was, bij constant houden op de invloed van de orde in de klas, traditionele scholen significant minder vaak hun leerlingen naar het BUO verwijzen dan progressieve scholen ($-.66$). De partiële correlatie wordt hier dus niet zwakker maar juist sterker (negatief).

5 Conclusie

Het proefschrift van Van der Wolf verdient waardering voor de ijver waarmee omvangrijk en rijk materiaal is verzameld. Ook de globale werkwijze – een ‘interpretatief’ onderzoek voorafgaand aan en funderend voor een kwantitatieve benadering – kan in principe een goede aanpak zijn.

Des te meer valt te betreuren dat, terwijl er zoveel werk is gemaakt van leerkracht- en leerlingvariabelen, de data niet op het niveau van leerkrachten ($N = 122$) of klassen geanalyseerd kunnen worden. Met betrekkelijk weinig moeite hadden outputvariabelen per klas verzameld kunnen worden. Nu is het zo dat variantieanalyses laten zien dat de variantie tussen de leerkrachten binnen sommige scholen aanzienlijk is, en dat aggregatie tot een ‘schoolkenmerk’ in veel gevallen een bedenkelijke manoeuvre is.

Voorts is onbegrijpelijk dat voor wat betreft belangrijke controlevariabelen zoals de opleiding en het beroep van de ouders, zo gemakkelijk wordt geargumenteed. Immers de invloed van milieuvariabelen op schoolsucces is zoals bekend aanzienlijk en bovendien veel sterker dan de invloed van (school)klasvariabelen. In een onderzoek gericht op deze laatste dient daarom de invloed van milieuvariabelen zeer zorgvuldig gecontroleerd te worden. Controleren op de invloed van het Cito-leesniveaue in de eerste klas is zeker niet voldoende.

Daarnaast zijn er variabelen op schoolniveau waarop Van der Wolf grote verschillen rapporteert. Voor wat betreft het percentage leerlingen uit mediterrane landen is het verschil tussen de clusters niet significant; tussen de scholen loopt het echter op van 15 tot 31%. Aangezien er op clusterniveau geen significante verschillen in leesniveau in de eerste

klas bestaan acht Van der Wolf het gerechtvaardigd om te stellen dat verschillen in beginniveau geen storende invloed zullen hebben op het verband tussen schoolsuccesindicatoren en leerkrachtvariabelen. Het is echter onbegrijpelijk waarom er op dit niveau wordt geanalyseerd.

Een factor die geheel onbelicht blijft betreft de stabiliteit van zowel onafhankelijke als afhankelijke variabelen door de jaren heen. Wat is de frequentie van, en de invloed van docentwisselingen, methodeveranderingen enz.?

Ook is het construeren van de schalen van de onafhankelijke variabelen erg onbevredigend. Zonder inhoudelijke argumenten laat Van der Wolf een factoranalyse bepalen hoe de schalen geconstrueerd moeten worden. Het gevolg hiervan is dat effecten van de samengestelde variabelen zeker niet eenduidig kunnen worden geïnterpreteerd.

In onze heranalyse hebben we voor zover mogelijk met bovengenoemde bezwaren rekening gehouden. Het kleine aantal casussen en het ontbreken van relevante controlevariabelen zijn echter bezwaren waaraan we ook in onze heranalyse niet ontkomen.

Eerst hebben we laten zien dat als we de door Van der Wolf geconstrueerde clusters op een andere manier analyseren op onderwijsstijlkenmerken, zijn hypothese niet bevestigd wordt. Zelf zag Van der Wolf in het resultaat van de analyse van de clusters een bevestiging. Vervolgens hebben we op schoolniveau de hypothesen getoetst met partiële correlaties. Deze veel eenvoudiger manier van analyseren laat zien dat zijn conclusie zeker niet weerlegd kan worden. In enkele gevallen zijn er zelfs aanwijzingen voor een ondersteuning van zijn conclusie. Naarmate een school een traditionelere onderwijsstijl heeft is het percentage leerlingen dat naar het MAVO gaat hoger en het percentage leerlingen dat naar het LBO gaat lager, gegeven de invloed van twee controlevariabelen (Cito leesniveau en buitenlandse leerlingen). Opvallend is dat de onderwijsstijl een negatief effect heeft op het laagste doorstroomniveau, terwijl het een positief effect heeft op het een na laagste doorstroomniveau. De onderwijsstijl heeft alleen een differentiërende werking voor de laagste twee niveaus. We kunnen echter niet spreken van een deugde-

lijke bevestiging gezien het beperkte aantal cases waardoor veel verbanden niet significant zijn en gezien de onvolledigheid van de gegevens.

Al met al is het te betreuren dat Van der Wolfs veelomvattende onderzoek zo weinig is ingebouwd in bestaande Nederlandse onderzoekstradities. Zowel ten aanzien van schoolvariabelen en multiniveau-analyse als ten aanzien van multivariate analyse en causale modellenbouw kent ons land voldoende deskundigen die met enkele suggesties ertoe hadden kunnen bijdragen dat Van der Wolfs werk aanzienlijk vruchtbaarder had kunnen zijn dan thans gerealiseerd is.

Noten

1. Meijnen beschouwt doubleerbeleid ook als schoolkenmerk (Meijnen, 1982, blz. 144).
2. De variabelen waarvan de F-waarde significant is ($p < .05$) zijn: plaatsen rangschikken twee aan twee, leerlingen lopen vrij rond, leerlingen mogen praten als ze dat willen, huiswerk krijgen, individueel werken aan vrije opdracht, cijfers hardop voorlezen, beloningen geven.
3. De schaal ONDERWIJSSTIJL bestaat uit de volgende variabelen: leerlingen bepalen zelf waar ze zitten, plaatsen twee aan twee gerangschikt, vrij rondlopen, hardop praten toegestaan, toestemming vragen om klas te verlaten, leerkracht verwacht dat leerlingen stil zijn, huiswerk geven, samenwerken aan vrije opdrachten, individueel werken aan vrije opdracht, cijfers op werk zetten, beloningen geven aan de besten, strafwerk geven, leerlingen het lokaal uitsluiten. De items zijn eerst omgecodeerd zodat een hoge score een traditionele onderwijsstijl betekent, en vervolgens omgezet in standaardcores. De Z-scores zijn daarna opgeteld en het resultaat is gedeeld door het aantal items.
4. Deze vraag is onder andere door Meijnen behandeld in een onderzoek naar 184 scholen (Meijnen, 1980).
5. De orde-schalen zijn net als de schaal ONDERWIJSSTIJL geconstrueerd. ORDE 1 bestaat uit de volgende items: rustige klasgenoten, er wordt goed orde gehouden, klasgenoten pesten elkaar gemeen, klasgenoten worden gauw kwaad, weinig ruzie in de klas, juf heeft ruzie met leerling, klasgenoten maken grappen over de juf (Cronbachs alpha .54, 7 items, $N = 1132$). ORDE 2 bestaat uit de volgende items: toch doorgaans als meester vraagt op te

houden, niet doen wat meester zegt, brutale antwoorden, juf pesten, anderen van werk houden, vechten, pesten, anderen voor de gek houden, op bank krassen, schreeuwen door de les, tekenen tijdens les, laat van pauze terugkomen, eerder van school gaan, klas uitlopen, juf voor de gek houden, lawaai maken, zingen, zoemen en fluiten, stiekem spelletjes doen (Cronbachs alpha .89, 18 items, $N = 1132$). Een factoranalyse op leerling-niveau liet twee duidelijk te onderscheiden dimensies zien.

6. Een variantie-analyse laat zien dat de variantie binnen de scholen (tussen de leerlingen) kleiner is dan tussen de scholen (F-waarden voor de schalen 5.5 respectievelijk 3.5, beide significant). Aggregatie is dus zinvol.
7. We hebben hier niet constant gehouden op de invloed van het percentage mediterrane leerlingen omdat het aantal onafhankelijke variabelen in verhouding tot het aantal cases daardoor te groot zou worden.

Literatuur

- Bennett, N., *Teaching styles and pupil progress*. London: Open Books, 1976.
- Dronkers, J., Boekbespreking. *Pedagogische Studiën*, 1986, 63, 41-43.
- Graaf, P. de, & P. G. Swanborn, Boekbespreking. *Mens en Maatschappij*, 1985, 60, 311-312.
- Graaf, P. de, Culturele hulpbronnen en schoolloopbanen in het lager onderwijs. *Mens en Maatschappij*, 1985, 60, 345-366.
- Grift, W. van de, Onderwijsleerklimaat en leerlingprestaties. *Pedagogische Studiën*, 1985, 62, 401-414.
- Kreft, I., Enige aantekeningen bij een empirisch onderzoek in het lager onderwijs te Amsterdam. *Tijdschrift voor Onderwijsresearch*, 1985, 10, 189-194.
- Kreft, I., Bespreking van G. W. Meijnen, Van zes tot twaalf. *Comenius*, 1985, 5, 95-101.
- Meijnen, G. W., Schooltypen in het lager onderwijs en milieuspecifieke leerprestaties. *Mens en Maatschappij*, 1980, 55, 385-410.
- Meijnen, G. W., *Schoolloopbanen in het lager onderwijs*. Sociologisch Instituut Groningen, 1982.
- Meijnen, G. W., *Van zes tot twaalf* (SVO-reeks 79). Harlingen: Flevodruk, 1984.
- Wolf, J. C. van der, *Schooluitval; een empirisch onderzoek naar de samenhang tussen schoolinterne factoren en schooluitval in het regulier onderwijs*. Dissertatie U. v. A., Lisse: Swets & Zeitlinger, 1984.
- Wolf, J. C. van der, Schooluitval in de l'art pour l'art kritiek. *Tijdschrift voor Onderwijsresearch*, 1985, 10, 195-199.

Curricula vitae

M. Kalmijn is doctoraalstudent in de sociologie aan de Rijksuniversiteit Utrecht (vakgroep Theoretische Sociologie en Methodenleer) en aan de Rijksuniversiteit Groningen (vakgroep Algemene Theoretische Sociologie).

P. G. Swanborn is hoogleraar in de methoden en technieken van sociaal-wetenschappelijk onder-

zoek aan de Rijksuniversiteit Utrecht, vakgroep Theoretische Sociologie en Methodenleer.

Adres: Sociologisch Instituut, Rijksuniversiteit Utrecht, Postbus 80.140, 3508 TC Utrecht

Manuscript aanvaard 3-4-'86

Summary

Kalmijn, M. & P. G. Swanborn. 'School drop out revisited: a secondary analysis of Van der Wolf's data.' *Pedagogische Studiën*, 1986, 63, 435-444.

The main conclusion drawn in Van der Wolf's PhD thesis is that pupils from schools with a traditional teaching style are more successful in their schoolcareer than pupils from schools with a progressive teaching style. The abundance of methodological critical remarks on Van der Wolf's methods casts doubt on the validity of this conclusion. We reanalysed Van der Wolf's data. Van der Wolf analysed clusters of schools. We show that these clusters hardly differ in teaching style. In an analysis on the school-level (N=12) we showed that for some outputvariables traditional schools indeed tend to have pupils with more success in school than progressive schools. This result partially confirms Van der Wolf's conclusion. It is certainly not a strong confirmation, because of serious shortcomings in the data.

Bijlage

Correlatiematrix (N = 12 scholen)

(2)	.08								
(3)	-.42	.31							
(4)	.22	-.32	-.90						
(5)	.48	.01	-.25	-.20					
(6)	-.40	-.06	.36	-.40	.09				
(7)	.03	-.25	-.56	.57	-.02	-.43			
(8)	.26	.05	-.44	.56	-.30	-.26	.28		
(9)	.06	.13	.26	-.51	.57	.27	-.47	-.73	
(10)	-.34	.30	.85	-.74	-.26	.08	-.62	-.55	.27
(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	

- (1) Citoscore op leestoets in eerste klas
- (2) Percentage mediterrane leerlingen
- (3) Doorverwijzing naar LBO
- (4) Doorverwijzing naar MAVO
- (5) Doorverwijzing naar brugklas
- (6) Verwijzing naar BUO
- (7) Onderwijsstijl (hoe hoger de score, hoe traditioneler)
- (8) Ervaring leerkracht
- (9) Gepercipieerde wanorde in de klas (hoge score, veel wanorde)
- (10) Zelf veroorzaakte wanorde in de klas