

# Profielanalyse van testcores van leerlingen uit verschillende sociale milieugroepen

J. F. M. CLAESSEN en B. W. G. M. SMITS

*Instituut voor Onderwijskunde, Katholieke Universiteit Nijmegen.*

## *Samenvatting.*

*In het onderzoek worden schoolprestaties vergeleken van leerlingen uit verschillend sociaal milieu, die deel uitmaken van een representatieve steekproef van de Nederlandse basisschoolpopulatie.*

*Met behulp van lineaire discriminantanalyse kon worden aangetoond dat de testprestaties van de diverse sociale milieugroepen significant van elkaar afweken. De verschillen kwamen zowel voor bij verbale als bij numerieke tests. Tussen de milieugroepen bleken voornamelijk niveauverschillen te bestaan.*

*De binnengroepsvariantie was echter van dien aard, dat het voorspellen van sociaal milieu, op grond van overeenkomst tussen individueel profiel en de gemiddelde testprofielen van de milieugroepen, niet afweek van hetgeen bij een randomisering verwacht mocht worden.*

## 1. Inleiding.

De samenhang tussen sociaal milieu en onderwijssucces is op velerlei manieren onderzocht. Daarbij is duidelijk komen vast te staan, dat leerlingen van lagere sociale milieugroepen binnen ons huidige onderwijsstelsel nog steeds een aanzienlijke achterstand bezitten. In dit artikel zullen de verschillen tussen milieugroepen worden geanalyseerd op grond van schoolvorderingen. Aan de hand van de resultaten op een aantal schoolvorderingentests en gegevens over doubleren, afkomstig van leerlingen van een landelijke steekproef, zal worden nagegaan op welke wijze en in welke mate leerlingen van deze

groepen zich in het basisonderwijs van elkaar onderscheiden.

Aangenomen dat er tussen milieugroepen verschillen worden aangetroffen, interesseerde ons vooral de vraag of deze verschillen sterker naar voren zouden komen bij tests met een overwegend verbaal karakter dan bij tests die meer appelleren op numerieke en ruimtelijke factoren, zoals o.a. door Bernstein (1960) wordt gesuggereerd. Opgemerkt moet worden dat in dit onderzoek werd uitgegaan van gegevens, verzameld in het kader van een ijkingsproject. Een keuze van tests waarin het verbale, respectievelijk ruimtelijke en numerieke aspect nog duidelijker naar voren zou komen dan in dit onderzoek, zou met betrekking tot bovengenoemde vraagstelling meer voor de hand hebben gelegen. Een tweede vraag betreft de afgenomen leestests. Deze tests beslaan een groot deel van het verzameld testmateriaal. Onderzocht zal worden of tests waarin taalvaardigheid en begrijpend lezen een belangrijke rol spelen, meer bijdragen aan de verschillen tussen de milieugroepen dan tests bestemd voor het meten van de technische leesvaardigheid.

Uitgebreid aandacht zal voorts worden geschonken aan de aard van de verschillen tussen de milieugroepen. Deze verschillen zullen worden bestudeerd tegen de achtergrond van andere tests die tegelijkertijd aan dezelfde groepen zijn voorgelegd. Mocht blijken dat ten aanzien van een bepaalde test afwijkingen tussen milieugroepen voorkomen die significant verschillen van de afwijkingen die bij andere tests worden aangetroffen, dan duidt dit op milieugevoeligheid van de betreffende test. Milieugevoeligheid kan daarbij verwijzen naar verschillen tussen milieugroepen die daadwerkelijk bestaan, maar van een

andere orde zijn dan bij de overige tests. Anderzijds kan milieugevoeligheid ook duiden op een vertekende weergave van eventuele verschillen, doordat bepaalde groepen door de tests worden bevoordeeld c.q. benadeeld. In deze laatste betekenis kunnen tegen een dergelijke test soortgelijke bezwaren worden aangevoerd als die welke sinds de tweede wereldoorlog met name in de Verenigde Staten werden uitgebracht tegen het cultuurgebonden karakter van veel intelligentietests. Nadrukkelijk werd door tal van auteurs dan ook aangedrongen (zie Eels, Davis e.a. 1951) op het ontwikkelen van 'cultuur-faire' tests, dat wil zeggen tests die zijn opgebouwd uit items die aan de verschillende cultuurgroepen evenveel recht doen.

Tot slot zal in dit artikel worden nagegaan hoe karakteristiek in feite de profielen van de test-gemiddelden voor de onderscheiden milieugroepen zijn.

## 2. Situering van het onderzoek.

De tests\* waarop de sociale milieugroepen zullen worden vergeleken, werden afgenomen in het kader van een ijkingsonderzoek\*\*, waarin verschillende toetsen voor het basisonderwijs gelijktijdig werden geijkt. Voor dit onderzoek werd door middel van clustersampling een landelijke steekproef getrokken, bestaande uit leerlingen van het tweede tot en met het zesde leerjaar van 25 scholen voor gewoon basisonderwijs. Bij een controle achteraf bleek de steekproef ten aanzien van een aantal relevant geachte kenmerken een bevredigende representativiteit te vertonen (Mommers en Smits, 1973). Tegelijkertijd met de afname van de tests in het schooljaar 1970/1971, werden gegevens verzameld over het sociaal milieu van de leerlingen. Hiertoe werd gebruik gemaakt van een beroepenklapper, ont-

\* Een overzicht van deze tests wordt gegeven aan het eind van dit artikel.

\*\* Dit ijkingsonderzoek werd mede mogelijk gemaakt door een subsidie van de Stichting voor Onderzoek van het Onderwijs te 's-Gravenhage, Project 0180.

wikkeld door het Instituut voor Toegepaste Sociologie te Nijmegen (versie 1970). In dit klassifikatiesysteem wordt als criterium voor sociaal milieu genomen het beroep – eventueel het laatst uitgeoefende beroep – van de vader. Elk beroep wordt daarbij beoordeeld op een aantal aspecten, zoals: vereiste aanleg, formeel vereiste opleiding, het beroepsniveau etc. In dit onderzoek werd het laatst genoemde aspect, het beroepsniveau – het prestige dat een beroep geniet – aangegrepen voor de indeling van de leerlingen naar sociaal milieu. De milieugroepen die worden onderscheiden zijn:

- niveaugroep 1 ongeschoolde arbeid
- niveaugroep 2 geschoolde arbeid
- niveaugroep 3 lagere employees
- niveaugroep 4 kleine zelfstandigen
- niveaugroep 5 middelbare employees
- niveaugroep 6 hogere beroepen

Wellicht ten overvloede zij opgemerkt, dat de resultaten van de hierna te bespreken vergelijkingen bezien moeten worden in het licht van deze sociale milieu-indeling.

## 3. Analyse van de schoolvorderingen.

### 3.1. Sociaal milieu en doorstroming.

Als eerste werd nagegaan of er tussen leerlingen van de sociale milieugroepen verschillen zouden voorkomen in de frekwentie van doubleren. De gegevens die hierover bij het begin van het onderzoek (september 1970) via de schooladministraties werden verzameld, staan weergegeven in tabel 1. De tabel behoeft een korte toelichting. De percentages doublures hebben betrekking op het aantal leerlingen per milieugroep dat in het betreffende leerjaar of in één van de voorgaande leerjaren is blijven zitten. Zo betekent het percentage dat vermeld staat onder leerjaar zes bij de groep ongeschoolde arbeid, dat van de 84 leerlingen afkomstig van scholen verspreid over het gehele land, bijna 30% tenminste éénmaal heeft gedoubleerd. Geen onderscheid is gemaakt tussen leerlingen die één of meerdere malen zijn blijven zitten.

Meervoudige doublures – ongeveer 2% van het

Tabel 1. Aantal leerlingen en het percentage doublures per sociale milieugroep, gespecificeerd naar leerjaar. Gegevens afkomstig van een landelijke steekproef in het schooljaar 1970-1971 (N = 3428).

Sociaal milieu	Leerjaar twee		Leerjaar drie		Leerjaar vier		Leerjaar vijf		Leerjaar zes	
	N (686)	% doubl.	N (703)	% doubl.	N (675)	% doubl.	N (682)	% doubl.	N (682)	% doubl.
Ongeschoolde arbeid	97	18.6	113	20.4	80	21.2	68	33.8	84	29.8
Geschoolde arbeid	266	11.3	233	12.1	231	18.6	219	21.9	205	21.5
Lagere employees	130	3.8	150	6.0	154	9.1	155	19.4	158	17.7
Kleine zelfstandigen	117	10.3	129	10.8	111	16.2	128	14.1	141	19.1
Middelbare employees	36	8.3	40	15.0	48	4.2	65	18.5	47	14.9
Hogere beroepen	40	5.0	38	2.6	51	5.9	47	8.5	47	6.4
Totaal % leerlingen dat tenminste eenmaal heeft gedoubleerd	10.2%		11.5%		14.4%		19.8%		19.7%	

totaal aantal leerlingen – kwamen overigens vrijwel uitsluitend voor bij leerlingen van de twee laagste milieugroepen. De schommelingen in de percentages, die vooral bij de hoogste twee milieugroepen worden aangetroffen, mogen niet los gezien worden van het verhoudingsgewijs geringe aantal leerlingen dat van deze beide groepen deel uitmaakt. Enigszins opmerkelijk is de positie van de groep kleine zelfstandigen. Het relatief hoog aantal doublures in deze groep is te verklaren vanuit de zeer heterogene samenstelling van deze groep.

Vergelijking van de milieugroepen naar de frekwentie van doubleren toont duidelijk aan, dat de doorstroming in de lagere milieugroepen aanzienlijk minder vlot verloopt. Een eenvoudige berekening leert dat het percentage zittenblijvers in beide laagste milieugroepen ruim tweemaal hoger is dan bij de twee hoogste groepen. Een vergelijking tussen de laagste en de hoogste milieugroep laat zelfs zien dat in de laagste groep vier maal zoveel zittenblijvers voorkomen als in de hoogste groep.

Van de leerlingen die het vijfde leerjaar hebben bereikt blijkt bijna 20% een klas te hebben moeten overdoen. Een percentage dat nagenoeg gelijk is aan dat voor het zesde leerjaar.

### 3.2. Profielanalyse van de testcores van de milieugroepen.

De variatie in de frekwentie van doubleren toont

slechts aan dat er tussen milieugroepen belangrijke verschillen bestaan in het succesvol doorlopen van de school. Geen informatie wordt verkregen over de aard van de verschillen tussen de milieugroepen. De aanwezigheid van een groot aantal testcores stelde ons in staat dit nader te onderzoeken. De testresultaten van de leerlingen werden per sociale milieugroep verzameld. Dit gebeurde bovendien per leerjaar afzonderlijk omdat er in de diverse leerjaren ten dele verschillende tests waren afgenomen. De aldus verkregen verzamelingen van testgemiddelden per milieugroep zullen we in het vervolg aanduiden met de term *profiel*. Met behulp van een multivariate analyse\* werd getoetst of de profielen van de milieugroepen significant van elkaar afweken. Een analyseprocedure die geschikt is voor het vergelijken van profielen en die bovendien specificeert welke bijdrage de afzonderlijke tests aan het onderscheid tussen de profielen leveren, is lineaire discriminantanalyse. Achtereenvolgens werden in het onderzoek de volgende bewerkingen uitgevoerd. Voor elke milieugroep werden per leerjaar de gemiddelden op de ver-

\* Bij een dergelijke benadering kunnen de profielen van testgemiddelden gezien worden als coördinaten die per groep een punt (centroïde) fixeren in een multidimensionele ruimte. Een ruimte die wordt opgespannen door de tests van het profiel. Groepen met verschillende vectoren van coördinaten bevinden zich op verschillende punten in de ruimte.

Tabel 2. Per leerjaar de F-ratio's en de correlaties van de tests met de discriminantvariabele.

TESTS	leerjaar 2 (N=582)		leerjaar 3 (N=572)		leerjaar 4 (N=550)		leerjaar 5 (N=550)		leerjaar 6 (N=550)	
	F-ratio	discr.var.	correlaties F-ratio	discr.var.	correlaties F-ratio	discr.var.	correlaties F-ratio	discr.var.	correlaties F-ratio	discr.var.
Een-Minuut Test, vorm A (sept. '70)	4.93**	.64	2.62**	.43	4.03*	.46	5.84**	.48	2.27*	.30
Een-Minuut Test, vorm B (sept. '70)	4.47**	.60	2.52**	.38	2.51*	.35	4.56**	.41	1.39	.20
Een-Minuut Test, vorm A (mrt. '71)	3.54**	.54	2.89**	.43	4.03*	.45	5.30**	.51	2.36*	.33
Een-Minuut Test, vorm B (mrt. '71)	3.87**	.57	3.77**	.51	4.75*	.50	7.44**	.60	2.79*	.34
Schriftelijke Opdrachten II	8.25**	.85	8.85**	.82						
Schriftelijke Opdrachten III	9.84**	.94	8.72**	.81	6.71**	.67				
Schriftelijke Opdrachten IV			6.56**	.68						
Stillestest 3										
Stillestest 4					6.77**	.68				
Stillestest 6										
Woordenschattest										
Toets Inzichtelijk Rekenen 4-5			11.04**	.91	9.63**	.82	7.17**	.59	3.21**	.37
Toets Inzichtelijk Rekenen 5-6					3.63*	.45	15.16**	.90	5.41**	.55
Test bij schoolbegin K2							8.58**	.61	5.12**	.54
Lagere Schooltest M1 (plaatjes-inhoud)	3.07**	.44								
Lagere Schooltest M2 (leesinhoud)					7.47**	.73			4.67**	.45
Test Mech. leesv. (Wiegersma)	5.45**	.67			5.75**	.64				
I.S.I.					5.86**	.64				
Rekenfactor									10.48**	.81
Taalfactor									6.51**	.57
Verbale Intelligentiefactor									8.34**	.71
Ruimtelijke Intelligentiefactor									6.15**	.60
P.M.T.-K										
P schaal									1.94	.14
F-schaal									2.20	-.38
F+-schaal									1.07	.14
SW-schaal									1.68	-.20

\* = p < .05    \*\* = p < .01

schillende tests berekend. Hierna werd met behulp van variantieanalyse getoetst\* of de profielen van gemiddelden significant van elkaar verschilden. Dit bleek voor alle leerjaren het geval te zijn ( $p < 0.001$ ). Omdat met uitzondering van leerjaar vier de binnengroep-variantie-covariantiematrices niet voldoende homogeen waren, is een mogelijke overschatting van deze verschillen in de andere leerjaren niet uitgesloten. Vervolgens werden per leerjaar de discriminantvariabelen\*\* bepaald. Deze variabelen komen tot stand door lineaire transformatie van de testcores en wel zo dat een onderliggende dimensie wordt gevormd waarop milieugroepen maximaal van elkaar verschillen. De correlaties van de tests met de gevonden discriminantvariabele staan per leerjaar weergegeven in tabel 2. Ter vergelijking zijn hierin ook de univariate F-ratio's opgenomen. Tussen de correlaties en de hoogte van de F-ratio's bestaat een hoge mate van overeenkomst. Dit is te verklaren doordat de tests steeds op vrijwel dezelfde wijze tussen de milieugroepen discrimineren.

Zoals reeds gesteld, leefden er ten aanzien van de afgenomen tests een tweetal specifieke vragen. Richtten we ons allereerst op de vraag of tests met een sterk verbaal karakter meer bijdragen aan de verschillen tussen milieugroepen, dan tests die meer beroep doen op ruimtelijke en numerieke factoren. Aanwijzingen hiervoor zijn bij meerdere auteurs te vinden. Zo analyseerden Eels, Davis e.a. (1951) een negental intelligentietests op milieugevoeligheid. Verbale items uit deze tests bleken scherper te discrimineren tussen de milieus dan ruimtelijke en figurale items. Bernstein (1969) konkludeerde op grond van verschillen tussen sociale milieu's in prestaties op een woordenschattest en Raven's matrices-test, dat de spraakcode waarop in veel tests een

beroep wordt gedaan, onvoldoende is afgestemd op de lagere sociale milieu's. Arvey (1972) noemde de 'verbal nature of many tests' als een van de factoren die verantwoordelijk zijn voor de grote discrepanties tussen de milieu's. Ofschoon de tests in dit onderzoek van een wat andere aard zijn dan die welke in bovengenoemde studies werden afgenomen, is een vergelijking van milieugroepen op tests waarin verbale factoren een meer en minder belangrijke rol spelen, mogelijk. Deze vergelijking is, zoals tabel 2 laat zien, alleen zinvol voor de drie hoogste leerjaren. Zoals blijkt komt uit de correlaties geen duidelijk beeld naar voren. Weliswaar discrimineert de Woordenschattest in de leerjaren vier en vijf het sterkst, anderzijds blijken de rekentests in leerjaar vijf en de M1 (plaatjesinhoud) in leerjaar vier, meer bij te dragen aan de verschillen tussen de milieugroepen dan de tests voor begrijpend lezen (Stilleestests en Schriftelijke Opdrachten). In leerjaar zes is de hoogste bijdrage zelfs afkomstig van de rekenfactor uit de I.S.I., gevolgd door de verbale intelligentiefactor, terwijl de taalfactor en de ruimtelijke intelligentiefactor nagenoeg even sterk discrimineren. De hier gevonden resultaten geven geen duidelijke steun aan de bewering dat de verschillen tussen milieugroepen sterker naar voren komen bij meer verbale- dan numerieke en ruimtelijke tests. Opmerkelijk in dit verband is dat ook in een recent Westduits onderzoek (Oevermann, 1972), opgezet met het expliciete doel de resultaten van Bernstein te toetsen, tussen milieugroepen geen duidelijke discrepanties in de verschillen bij verbale en figurale intelligentietests werden aangetroffen.

De tweede vraag heeft betrekking op de leestests. Komen de verschillen tussen de milieugroepen meer naar voren bij tests voor begrijpend lezen en taalvaardigheid dan bij tests die bedoeld zijn voor het meten van de technische leesvaardigheid? Ofschoon de correlaties op de discriminantvariabele niet tussen leerjaren mogen worden vergeleken, zijn de onderlinge verschillen binnen de leerjaren voldoende groot en consistent om deze vraag positief te beantwoorden. Schriftelijke Opdrachten, de Woordenschattest en de Stilleestest 3 en 4 vertonen in alle leerjaren een hogere

\* Gebruik werd gemaakt van een drietal toetsingsgrootheden: de 'maximum likelihood ratio' van Wilks, het 'trace criterion' van Hotelling en het 'largest root criterion' van Roy.

\*\* Hiervoor werd gebruik gemaakt van het programma 'Multan'. E. Roskam, Psychologisch Laboratorium, Nijmegen, 1971.

bijdrage aan het verschil tussen milieugroepen dan de Een-Minuut-Test en de Test voor Mechanische Leesvaardigheid. Enigszins een uitzondering vormt de Stilleestest 6. Zowel in leerjaar vijf als in leerjaar zes zien we dat deze test nog slechts in geringe mate afwijkt van de bijdrage van de technische leesvaardigheidstests. Vergelijking van de correlaties van deze laatstgenoemde tests leert, dat de Mechanische Leesvaardigheidstest meer onderscheid aanbrengt tussen de milieugroepen dan de Een-Minuut-Test.

Tot slot een opmerking over de Een-Minuut-Test. Het in de profielen opnemen van de resultaten van twee afnames op beide vormen, in plaats van bijvoorbeeld het gemiddelde resultaat, gebeurde in het kader van een onderzoek naar de equivalentie van beide vormen. Deze viervoudige presentatie, waardoor de test een dominerende positie inneemt, werd in dit onderzoek gehandhaafd, omdat het niet waarschijnlijk is dat de onderlinge verhoudingen tussen de milieu's hierdoor zullen worden gewijzigd.

### 3.3. Aard van de verschillen in de testprofielen.

Na te hebben vastgesteld, dat de profielen van de milieugroepen van elkaar verschilden, zullen we ons bezig houden met de vraag van welke aard deze verschillen zijn. Er bestaan drie mogelijkheden:

1. Tussen de milieugroepen bestaan alleen niveauverschillen.
2. De verschillen worden veroorzaakt door een interactie van milieu en test(s).
3. Ofwel zij zijn te wijten aan een combinatie van niveauverschil en interactie-effect.

Beide laatste gevallen zouden duiden op een bijzondere relatie tussen test en milieu.

Zoals in de inleiding naar voren werd gebracht, kunnen met betrekking tot hierboven genoemde mogelijkheden, slechts uitspraken worden gedaan in het licht van de overige tests in het profiel. Om na te gaan of er van interactie-effect sprake is, werd per leerjaar een multivariate variantieanalyse uitgevoerd op de verschillen van de tests. Deze procedure, die aan de hand van figuur 1 wordt toegelicht, staat beschreven in

Morrison (1967). Figuur 1 toont een fictief voorbeeld van de testprofielen van twee groepen. Elk profiel bestaat uit vier gemiddelde testcores: A, B, C en D. De scores op deze tests zijn door standaardisatie in vergelijkbare schaaleneenheden uitgedrukt. Door binnen elk profiel de testcores paarsgewijs van elkaar af te trekken ontstaan verschillen. Zo komt verschilscore 1 (1') tot stand door van de scores op test B die van test A af te trekken. Wanneer nu mocht blijken, dat de vectoren van verschillen van de diverse profielen aan elkaar gelijk zijn, zoals in figuur 1 wordt geïllustreerd, dan kunnen de gevonden afwijkingen tussen de groepen alleen worden toegeschreven aan niveauverschillen.

Het bepalen of er van interactie sprake is, komt dus in wezen neer op het toetsen of de testprofielen van de groepen evenwijdig aan elkaar verlopen. Mocht dit niet het geval zijn, dan kan uit de bijdrage van de afzonderlijke verschillen worden vastgesteld welke test(s) hiervan de oorzaak is. Na verwijdering van deze test uit de testprofielen kunnen in het gereduceerde profiel opnieuw de verschillen worden bepaald. Hierna kan andermaal worden getoetst of dit tot resultaat heeft dat de profielen nu wel evenwijdig aan elkaar verlopen. Het zal duidelijk zijn dat het reduceren van de testprofielen aan grenzen gebonden is. Door het uitsluiten van tests vermindert niet alleen de variantie tussen de groepen, maar ook het aantal vrijheidsgraden.

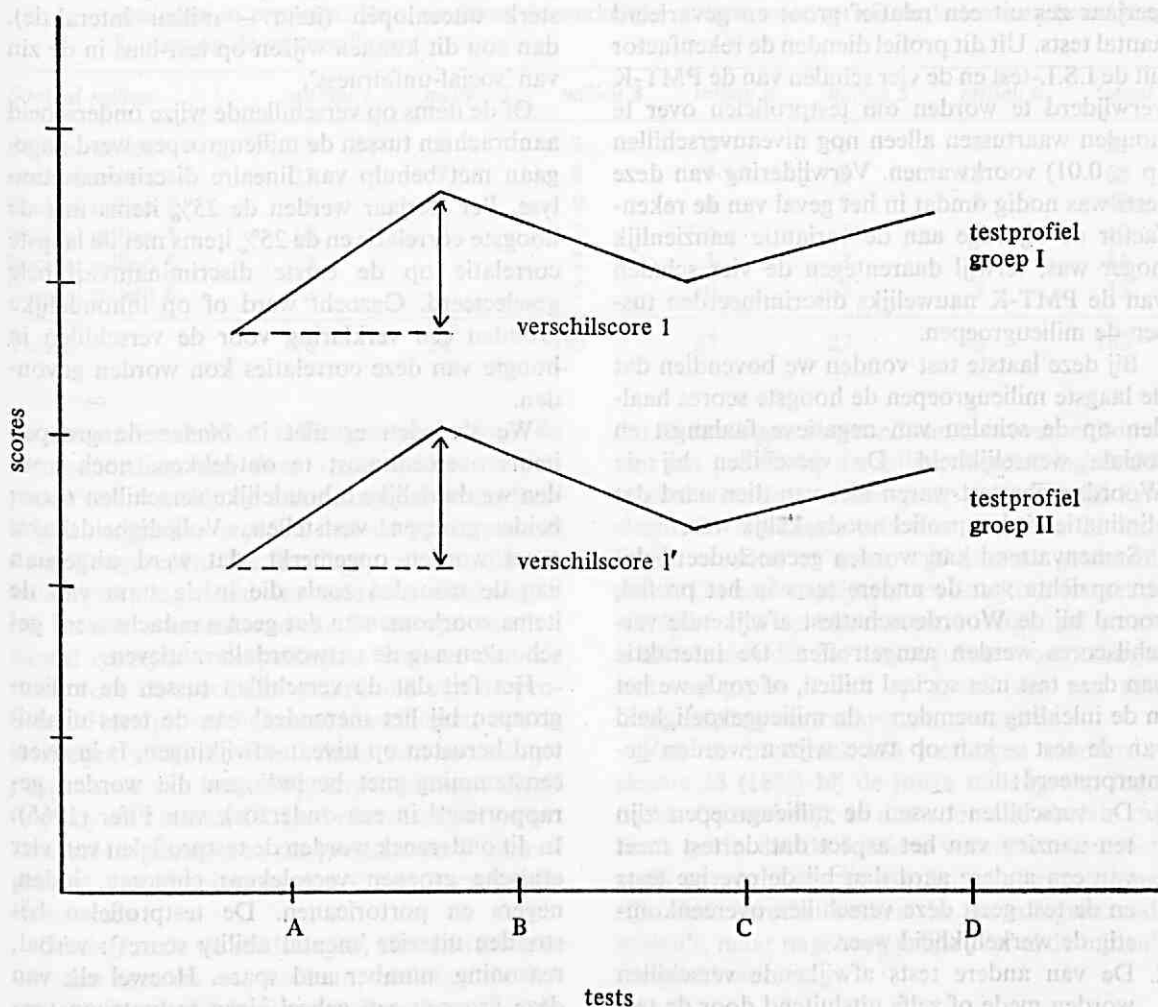
Tevens kan hieraan de vraag worden verbonden of het resterend profiel niet te zeer aan waarde heeft ingeboet, om nog als basis voor vergelijking te kunnen dienen.

### Resultaten.

Omdat de profielen in de diverse leerjaren niet gelijk van samenstelling zijn, werd de analyse op de verschillen per leerjaar uitgevoerd. De meest saillante bevindingen staan hieronder samengevat.

In leerjaar twee bleken de verschillen tussen de testprofielen van de zes milieugroepen uitsluitend te berusten op niveau-afwijkingen. Een andere

Figuur 1. Bepaling van verschillcores voor het toetsen van het interactie-effect.



situatie troffen we aan in de leerjaren drie en vier. Hier vertoonden de testprofielen geen parallel verloop. Uit de variantie van de afzonderlijke verschillcores was duidelijk dat in beide leerjaren de hoogste bijdrage afkomstig was van de Woordenshattest. Na verwijdering van deze test uit de profielen wees de opnieuw uitgevoerde analyse uit dat de gereduceerde profielen wel evenwijdig verliepen. Ook zonder de Woordenshattest bleek met behulp van multivariate variantieanalyse dat de profielen van de milieugroepen nog significant van elkaar afweken ( $p < 0.01$ ). In leerjaar vijf kon in tegenstelling tot beide

voorgaande leerjaren geen parallel verloop van de profielen worden verkregen. Ook niet na verwijdering van meerdere tests uit het profiel. Een mogelijke oorzaak hiervoor kan gelegen zijn in een vrijwel volstrekt gebrek aan homogeniteit van de varianties, hetgeen met name in dit leerjaar sterk naar voren kwam. Met uitzondering van eventueel de Woordenshattest, die wederom de hoogste bijdrage leverde aan de variantie van de verschillcores, zouden wij het niet evenwijdig verlopen van de testprofielen dan ook niet willen toeschrijven aan een daadwerkelijk interactie-effect.

Zoals tabel 2 laat zien, bestaat het profiel van leerjaar zes uit een relatief groot en gevarieerd aantal tests. Uit dit profiel dienden de rekenfactor uit de I.S.I.-test en de vier schalen van de PMT-K verwijderd te worden om testprofielen over te houden waartussen alleen nog niveauverschillen ( $p < 0.01$ ) voorkwamen. Verwijdering van deze tests was nodig omdat in het geval van de rekenfactor de bijdrage aan de variantie aanzienlijk hoger was, terwijl daarentegen de vier schalen van de PMT-K nauwelijks discrimineerden tussen de milieugroepen.

Bij deze laatste test vonden we bovendien dat de laagste milieugroepen de hoogste scores haalden op de schalen van negatieve faalangst en sociale wenselijkheid. De verschillen bij de Woordenschattest waren niet van dien aard dat eliminatie uit het profiel noodzakelijk was.

Samenvattend kan worden geconcludeerd dat ten opzichte van de andere tests in het profiel, vooral bij de Woordenschattest afwijkende verschillen werden aangetroffen. De interactie van deze test met sociaal milieu, of zoals we het in de inleiding noemden – de milieugevoeligheid van de test – kan op twee wijzen worden geïnterpreteerd:

1. De verschillen tussen de milieugroepen zijn ten aanzien van het aspect dat de test meet van een andere aard dan bij de overige tests en de test geeft deze verschillen overeenkomstig de werkelijkheid weer.
2. De van andere tests afwijkende verschillen worden mede of zelfs uitsluitend door de test zelf veroorzaakt, doordat bepaalde groepen door de test worden bevoordeeld. De verschillen berusten in feite (ten dele) op de test-bias. De test is 'social-unfair'.

Op grond van de ter beschikking staande gegevens is het niet goed mogelijk aan te geven welke van beide interpretaties het meest waarschijnlijk van toepassing is als verklaring voor het afwijkend gedrag van de Woordenschattest. Om hierover wat meer zekerheid te krijgen werd teruggegrepen op de items van deze test. De gedachte die hierbij werd gevolgd, is dat binnen de test de items op homogene wijze dezelfde factor(en) meten. Indien nu mocht blijken dat de

prestaties van milieugroepen en bepaalde items sterk uiteenlopen (item – milieu interactie), dan zou dit kunnen wijzen op test-bias in de zin van 'social-unfairness'.

Of de items op verschillende wijze onderscheid aanbrachten tussen de milieugroepen werd nagegaan met behulp van lineaire discriminantanalyse. Per leerjaar werden de 25% items met de hoogste correlatie en de 25% items met de laagste correlatie op de eerste discriminantvariabele geselecteerd. Gezocht werd of op inhoudelijke gronden een verklaring voor de verschillen in hoogte van deze correlaties kon worden gevonden.

We slaagden er niet in *binnen* de groepen items overeenkomst te ontdekken, noch konden we duidelijke inhoudelijke verschillen *tussen* beide groepen vaststellen. Volledigheidshalve moet worden opgemerkt, dat werd uitgegaan van de woorden zoals die in de stam van de items voorkomen en dat geen aandacht werd geschonken aan de antwoordalternatieven.

Het feit dat de verschillen tussen de milieugroepen bij het merendeel van de tests uitsluitend berusten op niveau-afwijkingen, is in overeenstemming met bevindingen die worden gerapporteerd in een onderzoek van Fifer (1966). In dit onderzoek worden de testprofielen van vier etnische groepen vergeleken: chinezen, joden, negers en portoricanen. De testprofielen bestonden uit vier 'mental ability scores': verbal, reasoning, number and space. Hoewel elk van deze groepen een geheel eigen testpatroon vertoonde, bleken binnen elke groep de testprofielen van lower en middle class parallel te lopen.

#### 4. De voorspelbaarheid van het sociaal milieu.

In het voorafgaande stelden we vast dat de profielen van de milieugroepen zeer duidelijk van elkaar afweken en dat deze afwijkingen voor het merendeel toe te schrijven waren aan niveauverschillen. De vraag kan tot slot gesteld worden, welke betekenis we aan deze statistische bevindingen mogen hechten? Of anders geformuleerd, hoe karakteristiek zijn in wezen de profielen van



*Tabel 3.* Klassifikatie van leerlingen (leerjaar 4) naar milieugroepen op grond van overeenkomst tussen testuitslagen en profielen. De rijen geven de werkelijke situatie weer, terwijl de kolommen betrekking hebben op de voorspellingen.

<i>Sociaal milieu</i>	<i>milieu 1</i>	<i>milieu 2</i>	<i>milieu 3</i>	<i>milieu 4</i>	<i>milieu 5</i>	<i>milieu 6</i>	<i>totaal</i>
Sociaal milieu 1	5	4	3	2	3	4	21
Sociaal milieu 2	23	7	8	9	8	10	65
Sociaal milieu 3	5	6	9	4	8	8	40
Sociaal milieu 4	10	5	3	6	3	3	30
Sociaal milieu 5	3	1	2	1	3	4	14
Sociaal milieu 6	2	1	5	0	2	3	13
Totaal	48	24	30	22	27	32	183

de zes onderscheiden milieugroepen? Voor de beantwoording van deze vraag kozen we een procedure waarbij de testuitslagen van de leerlingen vergeleken worden met de profielen van de milieugroepen. De leerlingen worden bij die milieugroep ingedeeld waarmee hun testuitslagen het meeste overeenkwamen. De hoogte van het aantal correcte voorspellingen fungeert dan als maat voor de specifieke geaardheid van de profielen. Het op overeenkomst klassificeren van de leerlingen gebeurde met behulp van een aangepaste versie van het computerprogramma Clasif (Cooley en Lohnes, 1971). Voorwaarde voor het kunnen toepassen van dit programma is homogeniteit van de binnengroep-variantie-covariantiematrices.

Omdat, zoals we reeds zagen, alleen in leerjaar vier aan deze eis was voldaan, besloten we het onderzoek tot dit leerjaar te beperken. Om de vergelijking tussen de testuitslagen en de profielen zo onafhankelijk mogelijk te laten verlopen, deelden we de leerlingen van dit leerjaar op basis van toeval in twee groepen in. Van de grootste groep, 2/3 van het aantal leerlingen, werden de profielen voor de milieugroepen bepaald. De testuitslagen van de overige leerlingen werden tegen deze profielen afgezet. De resultaten van dit predictie-onderzoek staan vermeld in tabel 3. De rijtotalen geven de aantallen leerlingen weer, die in feite tot een bepaalde sociale milieugroep behoren, terwijl de kolomtotalen betrekking hebben op de voorspelde aantallen.

Beschouwen we sociaal milieu 1, dan zien we

dat 48 leerlingen op grond van hun overeenkomst met het profiel bij dit milieu worden ingedeeld, hoewel tot deze milieugroep in werkelijkheid slechts 21 leerlingen behoren. Van deze 21 leerlingen worden er maar vijf correct geklassificeerd. Uit het vergelijken van de rijtotalen met de aantallen die op bijbehorende diagonaal staan, is af te leiden dat het aantal juiste voorspellingen niet veel groter is dan op grond van een random indeling verwacht mocht worden.

Van de in totaal 183 leerlingen worden er slechts 33 (18%) bij de juiste milieugroep ingedeeld. Opmerkelijk is dat voor alle sociale milieu's geldt, dat de foutieve voorspellingen niet, zoals verwacht mocht worden, voor een belangrijk deel terecht komen in de aangrenzende milieu's, maar nagenoeg gelijk over al de milieu's verspreid worden.

*Samenvattend* mogen we concluderen, dat hoewel de profielen van de milieugroepen zeer significant van elkaar verschillen, de testuitslagen van de leerlingen van de diverse sociale milieu's elkaar in hoge mate overlappen. Van karakteristieke profielen voor bepaalde milieugroepen is geen sprake. Met andere woorden het voorspellen van het sociaal milieu is op grond van testuitslagen (gelukkig) niet mogelijk. Deze conclusie wordt nog op een andere wijze bevestigd.

Bij het onderzoek naar de verschillen tussen de profielen van de milieugroepen bleek slechts een fractie van de proportie variantie in de eerste

discriminantvariabele voor rekening te komen van de verschillen tussen de milieugroepen. De percentages varieerden van 8% in leerjaar twee tot 14% in leerjaar vijf.

#### Overzicht van de afgenomen tests

##### A. Tests die geijkt werden:

- *Een-MinuuT Test, vorm A en B. Schoolvorderingentest voor de technische leesvaardigheid, bestemd voor het tweede tot en met het zesde leerjaar van het basisonderwijs.* Tweede, geheel opnieuw bewerkte uitgave, B. Th. Brus en M. J. M. Voeten, Nijmegen, 1972.
- *Schriftelijke Opdrachten, Schoolvorderingentests voor het begrijpend lezen, bestemd voor het tweede, derde en vierde leerjaar van het basisonderwijs.* Tweede, geheel bewerkte uitgave, B. Th. Brus en J. B. A. M. van Bergen, Nijmegen, 1972.
- *Stilleestests, bestemd voor het derde tot en met zesde leerjaar van het basisonderwijs.* J. Bakker, Nijmegen, 1966, 1972.
- *Woordenschattest, bestemd voor het derde tot en met het zesde leerjaar van de basisschool.* P. J. J. Stijnen, Nijmegen, 1972.
- *Toetsen voor Inzichtelijk Rekenen. Schoolvorderingentoetsen voor het inzichtelijk rekenen, bestemd voor het vierde tot en met het zesde leerjaar van het basisonderwijs.* M. J. C. Mommers en B. W. G. M. Smits, Tilburg, 1972.

##### B. Tests die in het kader van het validiteitsonderzoek in dit ijkingsproject werden afgenomen:

- *Test bij het Schoolbegin.* Model K2 Pintner, R., Cunningham, B. en Durost, W. N. Voor Nederland bewerkt door J. Luning Prak, Groningen, 1970.
- *Pintner-Durost Lagere Schooltest.* Twee vormen: Model M1, Schaal 1 (Plaatjesinhoud) en Model M2, Schaal 2 (leesinhoud). Nederlandse bewerking van J. Luning Prak, Groningen, 1968.
- *De Test voor Mechanische Leesvaardigheid.* S. Wiegiersma, Groningen, 1970.
- *Het Schoolvorderingen- en Intelligentiegedeelte van de I.S.I.-test, vorm II.* J. Th. Snijders en V. J. Welten, Groningen, 1968.
- *De Prestatie Motivatie Test voor Kinderen.* H. J. M. Hermans, Amsterdam, 1971.

#### Literatuur:

- Arvey, R. D., Some comments on culture fair tests. *Personnel Psychology*, 1972 (25) 433-448.
- Bernstein, B., Language and social class. *British Journal of sociology*, 1960 (11) 271-276.
- Cooley, W. and P. Lohnes, *Multivariate Data analysis*. New York, 1971.
- Cooley, W., Techniques for considering multiple measurements. In: L. Thorndike (ed.), *Educational measurement* (second edition), Washington, 1971.
- Cleary, T. A. and T. L. Hilton, An investigation of item-bias. *Educational and Psychological Measurement*, 1968 (28) 61-75.
- Echternacht, G., A quick method for determining test-bias. *ETS-research bulletin*. New Jersey, 1972.
- Eels, K. and A. D. Davis e.a., *Intelligence and Cultural differences*. Chicago, 1951<sup>2</sup>.
- Fifer, G., Social Class and cultural differences in diverse mental abilities. In: A. Anastasi (ed.), *Testing problems in perspective*. Washington 1966.
- I.T.S., *Beroepenklapper*. Nijmegen, oktober 1970.
- I.T.S., *Beroepskenmerken en sociale milieu's*. Nijmegen, 1971.
- Mommers, M. J. C. en B. W. G. M. Smits, *Toetsen voor inzichtelijk rekenen. Verantwoording en handleiding*. Tilburg, 1973.
- Morrison, D. F., *Multivariate statistical methods*. New York 1971.
- Oevermann, U., *Sprache und soziale Herkunft*. Frankfurt am Main, 1972.
- Tiedeman, D., A geometric model for the profile problem. In: A. Anastasi (ed.), *Testing problems in perspective*. Washington, 1966.

#### Curricula vitae

J. F. M. Claessen (geb. 1946) studeerde theologie aan de Hogeschool voor Theologie en Pastoraal te Heerlen; aansluitend pedagogiek aan de Katholieke Universiteit te Nijmegen, met als specialisatie onderzoek van het onderwijs; thans verbonden als onderwijskundig medewerker aan het Instituut voor Toegepaste Sociologie te Nijmegen en belast met onderzoek op het terrein van vreemde talenonderwijs.

Adres: Verlengde Groenestraat 55, Nijmegen.

B. W. G. M. Smits (geb. 1943) behaalde in 1968 zijn doctoraal examen psychologie; vanaf 1969 werkzaam bij de afdeling Research van het Pedagogisch Instituut van de K.U. te Nijmegen; sinds 1971 verbonden aan het Instituut voor Onderwijskunde; enkele publicaties op het gebied van testkonstruktie.

Adres: Erasmuslaan 40, Nijmegen.