

juist geschetste ontwikkeling. Wij verwachten wel dat hiervan een stimulans op zijn werk zal uitgaan, waardoor een verhoging van het peil van de huisartsgeneeskunde zal optreden.

De taak van de huisarts in Nederland is een zeer specifieke, gezien de continuïteit welke inherent is aan zijn werk. Wij hopen, dat het Nederlands Huisartsen Genootschap, het Nederlands Huisartsen-Instituut en de docenten in de huisartsgeneeskunde er met elkaar in zullen slagen de unieke positie van de huisarts te verstevigen.

Het bestuur heeft zich erover beraden of aan het lidmaatschap van het Genootschap eisen mogen worden gesteld. Hoewel het bestuur van ieder lid een zekere mate van activiteit verwacht ten aanzien van nascholing, wetenschappelijk onderzoek en praktijkvoering, heeft het gemeend dat aan het lidmaatschap geen eisen mogen worden verbonden. Wel is het bestuur van mening, dat het perifere lid moet worden aangemoedigd tot eigen onderzoek, deelname aan studiegroepen of andere vormen van nascholing.

Om de leden een stimulans te geven heeft het bestuur besloten een prijsvraag uit te schrijven voor alle Nederlandse huisartsen of groepen van huis-

artsen waarbij zij op de volledige medewerking van de commissies mogen rekenen. Deze prijsvraag zal eens in de drie jaren worden uitgeschreven. De inzender van het beste antwoord zal worden beloond met een prijs van vijfduizend gulden. Het onderwerp van de prijsvraag, die dit jaar wordt uitgeschreven luidt als volgt: verlangd wordt een beschrijving van één dag praktijkvoering van de huisarts. Antwoorden kunnen worden ingezonden tot 1 januari 1971; de uitslag zal bekend worden gemaakt op het N.H.G.-congres 1971. Ik hoop dat het uitschrijven van deze prijsvraag ertoe kan leiden, dat velen onder u zullen trachten tot beantwoording van de vraag te komen.

Dames en Heren, het Genootschap staat aan de vooravond van velerlei ontwikkelingen en mogelijk ook veranderingen in de huisartsen-wereld. Wij zullen de invloeden daarvan ondervinden, maar ook trachten daaraan voor zover dat in ons vermogen ligt richting te geven, zoals dat in het verleden is geschied. Tezamen met de commissies, het Nederlands Huisartsen-Instituut en de centrumbesturen hopen wij het juiste antwoord te kunnen vinden op de vragen, die het huidige tijdsgewricht ons stelt.

## *De F-scale als instrument om de autoritaire instelling te meten als individueel-psychologische variabele in een onderzoek bij huisartsen \**

DOOR DR. P. H. J. M. HEYDENDAEL, PSYCHOLOOG

*Inleiding.* De onderhavige studie vertegenwoordigt niet de eerste poging om de Amerikaanse „F-scale” \*\* bruikbaar te maken voor toegepast onderzoek in Nederland. Zowel *Albinski* als *Weima* hebben in hun academische proefschriften de autoritaire persoonlijkheid tot een centraal thema gemaakt. Beide auteurs hanteren een verkorte en deels gewijzigde versie van de oorspronkelijke Amerikaanse lijst.

Het doel van dit onderzoek is een vorm van de F-scale, die voldoende betrouwbaar is om individuen met elkander te vergelijken naar de mate waarin zij autoritair zijn ingesteld, toe te passen in een onderzoek bij een aantal Nederlandse huisartsen naar de wijze van hun praktijkvoering. In deze doelstelling is nogal de nadruk gelegd op de betrouwbaarheid. Zolang aan de eisen van betrouwbaarheid niet is voldaan, is het niet verant-

woord noch doelmatig om te beginnen aan een toegepast onderzoek. Voorlopig nemen wij hierbij aan, dat ons instrument de autoritaire instelling meet („face-validity”).

Na een korte, beschrijvende samenvatting van het begrip autoritaire persoonlijkheid, volgen in vrijwel chronologische volgorde de verschillende procedures die hebben geleid tot een redelijk betrouwbaar instrument. Daarna volgt de beschrijving van de toepassing van deze F-scale in een veldonderzoek bij huisartsen naar de wijze van hun praktijkvoering. Wij zullen ons hierbij beperken tot een kleine exploratie ter verduidelijking van de hypothesen.

De constructie van de hier beschreven en gehanteerde F-scale heeft in 1962 plaatsgevonden onder leiding van de psycholoog *Grosfeld* met medewerking van schrijver *dezes* (*Grosfeld* en *Heydendael*), beiden toentertijd verbonden aan de afdeling Research van het Gemeenschappelijk Instituut voor Toegepaste Psychologie (G.I.T.P.) te Nijmegen. In dit artikel zal een samenvatting worden gegeven van deze constructie omdat wij dit nodig achten voor een goed begrip van de resultaten van ge-

\* Uit het Instituut voor Sociale Geneeskunde van de Katholieke Universiteit te Nijmegen.

\*\* In tegenstelling tot de tot nu toe gevolgde methode om niet-Nederlandse termen bij herhaling van aanhalingstekens te voorzien, zal zulks in dit artikel alleen bij de eerste keer plaatsvinden.

noemd veldonderzoek. Voor de volledige beschrijving van de constructie wordt verwezen naar een intern rapport van het G.I.T.P. (*Grosfeld*).

*Samenvatting van de voornaamste psychologische achtergronden van de F-scale.* De F-scale is ontstaan in een milieu waarin men de psychologische kwaliteiten van het vooroordeel wilde opsporen. *Adorno en medewerkers* wilden het mechanisme van de discriminatie blootleggen en vooral het karakterologische type van de man bestuderen die discrimineert. Zij gingen uit van de hypothese dat er inderdaad een bepaald persoonlijkheidstype van deze aard zou bestaan: „the authoritarian type of man”. *Adorno* trachtte een methode te ontwikkelen met behulp waarvan bevooroordeelde mensen zouden kunnen worden onderscheiden van niet- of minder bevooroordeelde. Naarmate zijn onderzoek vorderde, wijzigde hij zijn methode en inzichten, waarbij ook het object van onderzoek steeds breder werd.

Aanvankelijk ontwikkelde men een schaal die een maatstaf moest zijn voor de graad van „antisemitisme”: de AS-scale. Toen echter bleek dat vele antisemieten op hun beurt waren bevooroordeeld ten aanzien van andere bevolkingsgroepen, ontwikkelde men een maatstaf voor „ethno-centrisme”: de E-scale. Ook dit ethno-centrisme bleek een deel te zijn van een groter geheel, namelijk van het „politiek-economisch conservatisme”, waarvoor men toen de PEC-scale ontwierp. Tenslotte bleek dat het bevooroordeeld zijn slechts een uitwendige manifestatie is van dieper liggende, deels onbewuste factoren. Daarom was het nodig een meettechniek te ontwikkelen, waarvan de proefpersoon de bedoeling niet zou kunnen onderkennen: de F-scale. De F-scale moest dus niet zozeer appelleren aan oppervlakkige opvattingen en opinies, als wel aan meer fundamentele persoonlijkheidsfactoren. De nieuwe meettechniek zou, aldus *Adorno*, moeten leiden tot „valid estimates of antidemocratic tendencies at the personality level”.

De vraag, welke diepere persoonlijkheidsfactoren deze autoritaire persoonlijkheidsstructuur constitueren, meenden *Adorno en medewerkers* te kunnen beantwoorden op grond van de vroegere fasen van hun onderzoek. De persoonlijkheidsvariabelen die hieronder volgen, hebben als uitgangspunt gediend bij de formulering van de „items” (vragen) van de F-scale. Tezamen vormen zij een min of meer duurzaam gestructureerd geheel van eigenschappen: het „autoritair syndroom” of de „autoritaire instelling”. Deze persoonlijkheidsvariabelen zijn de volgende.

Conventionalisme: de autoritaire persoonlijkheid houdt star vast aan burgerlijke waarden.

Autoritaire onderworpenheid: de autoritaire persoonlijkheid gedraagt zich onderworpen tegenover de leiders van de „in-groep”.

Autoritaire agressie: hij veracht mensen die conventionele waarden schenden.

Anti-intraceptie: hij verzet zich tegen en hij is afkerig van het subjectieve.

Bijgeloof en stereotypie: hij gelooft dat het lot van de mens op mystieke wijze is gedetermineerd, hij denkt in rigide categorieën.

Kracht en hardheid: hij interesseert zich overmatig voor tegenstellingen zoals hard en zacht, leider en volgeling, overheersing en onderworpenheid.

Afbrekende en cynische houding: hij breekt waarden af die anderen hebben opgebouwd.

Projectie-tendens: hij projecteert zijn onbewuste emotionele intenties.

Seksualiteit: hij is voortdurend bezig en bezorgd over alles wat hieraan appelleert.

*Adorno* stelde zichzelf bij de formulering van de items voor de F-scale de volgende eisen. Elk item moest samenhangen met een of meer van de bovengenoemde persoonlijkheidsfactoren; het moest zo indirect mogelijk worden gesteld; het mocht geen aanwijzing geven over de bedoelingen van het onderzoek, noch mocht het vooroordelen openlijk tot uitdrukking brengen; het moest een zeker evenwicht vertonen tussen objectieve waarheid en irrationaliteit. Een appel op de ratio mocht in geen enkel item ontbreken. Dit mocht echter niet de belangrijkste determinant zijn bij de keuze tussen „eens” en „oneens”.

Op grond van deze principes stelde *Adorno* zijn eerste ontwerp op. Elk item hield een bewering in waarover de proefpersoon zijn oordeel moest geven. Daarbij had hij de keuze uit zeven antwoorden, waarvan de uitersten luiden: „geheel eens” en „geheel oneens”. Een experiment met 295 proefpersonen leidde tot een betrouwbaarheidscoëfficiënt van 0,74. Na analyse van de eerste versie bleven negentien items over die aan *Adorno's* eisen voldeden. Tezamen met vijftien nieuwe items vormden zij de tweede experimentele versie van de F-scale. Deze versie werd beproefd op 286 proefpersonen. Dit leverde een stabiliteitscoëfficiënt op van 0,87. Een nieuw onderzoek met een schaal van 42 items bij 1 518 proefpersonen leverde een gemiddelde stabiliteitscoëfficiënt op van 0,90 (range 0,81 - 0,97).

*De constructie van de F-scale in dit onderzoek.*

Voor de experimentele versie van onze F-scale werden 126 zogenaamde stellingen verzameld. Deze items zijn afkomstig uit de volgende bronnen. Een aantal werd overgenomen van *Adorno* en van *Albinski* en een iets groter aantal werd zelf geconstrueerd. Daarbij werd uitgegaan van twee principes. Elk item moest duidelijk samenhangen met één van de negen door *Adorno* genoemde persoonlijkheidsvariabelen die tezamen het autoritaire syndroom constitueren. Verder moesten de aantallen items behorende bij elk van deze negen persoonlijkheidsvariabelen ongeveer gelijk zijn, namelijk veertien. In dit aantal zijn de overgenomen items inbegrepen. In het totaal bevatte de experimentele versie van de F-scale dus  $9 \times 14 = 126$  items.

Genoemde experimentele versie van de F-scale is afgenomen aan 525 eerstejaarsstudenten van beiderlei kunne, in leeftijd variërend van zestien tot dertig jaar. De meeste items waren dusdanig geformuleerd dat het antwoord „volkomen eens” werd verondersteld te duiden op een autoritaire instelling. Tien items waren echter zo geformuleerd dat het antwoord „volkomen oneens” duidde op een autoritaire instelling. Bij deze items moest bij de „scoring” op de puntenschaal dus worden omgekeerd. Nadat voor een proefpersoon elk item was gewaardeerd, werd de totale „ruwe” score berekend door optelling van de waarden van alle items.

Bij de item-analyse – een methode ter controle van de bruikbaarheid van de vragen of stellingen waaruit een schaal bestaat – werden de volgende twee maatstaven aangelegd. Het „discriminerend vermogen” van elk item moest voldoende zijn. Dat wil zeggen: met behulp van elk item afzonderlijk moest men in staat zijn mensen met een sterke en mensen met een zwakke autoritaire instelling van elkander te onderscheiden. Verder moest de „voorkeursindex” van elk item aan bepaalde voorwaarden voldoen. De voorkeursgraad geeft aan, welk percentage personen uit de onderzochte groep een bepaald item in de verwachte richting beantwoordt.

Voor de bepaling van het discriminerende vermogen van de items kan men verschillende maatstaven aanleggen. In dit verslag zullen wij vanwege de technische aard van dit probleem de gevolgde werkwijze niet verder vermelden. Onder voorkeursindex wordt verstaan: het percentage personen dat een bewering in autoritaire zin beantwoordt. Voor een schaal als de gebruikte is het zinvoller wanneer elk item „een gelijke kans” heeft om de totale score in positieve of negatieve richting te beïnvloeden. Het gaat immers om de onderscheiding van mensen met alle graden van autoritaire instelling en niet alleen om bijvoorbeeld de onderscheiding tussen zeer autoritair en uiterst autoritair.

In statistische termen betekent dit, dat de gemiddelde voorkeursgraad van alle items – dus van de schaal als geheel – gelijk zou moeten zijn aan  $p = 0,50$  en dat de voorkeursindices van de afzonderlijke items eveneens zo dicht mogelijk bij deze waarde zouden moeten liggen. De grenzen werden in dit onderzoek gesteld op  $p = 0,20$  en  $p = 0,80$ .

Een derde voorwaarde voor de itemselectie luidde als volgt: de t-toets voor het verschil tussen de gemiddelde score van de „hoge” groep en de gemiddelde score van de „lage” groep onderzochten moet significant zijn op het vijf procent niveau. Bij deze derde voorwaarde gaat het eveneens om een maatstaf voor het discriminerende vermogen van het item. Deze maatstaf diende als een extra controle. De resultaten van de drie genoemde grootheden per item zullen vanwege hun technische karakter niet worden vermeld. Na de selectie bleven 63 items over.

Nadat de items van deze experimentele F-scale waren geselecteerd met betrekking tot hun discriminerende vermogen en hun voorkeursindex, moest

nog de betrouwbaarheid van de test als geheel – dus de stabiliteit of de consistentie van het meetinstrument – worden bepaald. Dit kan op verschillende manieren worden gedaan. Aangezien het niet goed mogelijk is de F-scale bij dezelfde proefpersonen binnen korte tijd nogmaals af te nemen, kon de „test-retest”-methode niet worden gebruikt. Daarom werd de voorkeur gegeven aan de „split-half”-methode.

Volgens deze methode werden de als geschikt geselecteerde items verdeeld in twee gelijkwaardige groepen. Voor iedere persoon werden twee scores berekend. De grootste moeilijkheid bij deze methode is de zorg voor de gelijkwaardigheid van de beide „testhelften”. Aan deze voorwaarde kon worden voldaan door de keuze van 24 paren items. De beide testhelften, die dus elk 24 items bevatten, konden als gelijkwaardig worden beschouwd. Van de 63 items waren dus slechts 48 gebruikt; de overige vijftien items werden voorlopig buiten beschouwing gelaten. Voor 105 proefpersonen, die een aselechte steekproef vormden uit de groep die voor dit onderzoek was ingezet, werden vervolgens de scores berekend voor elk van de beide testhelften. Met behulp van deze gegevens werd de betrouwbaarheid van de F-scale op verschillende manieren vastgesteld.

Men kan twee gelijkwaardige testhelften beschouwen als twee paralleltests. Door correlatie van de scores van de twee testhelften verkrijgt men dan een maat voor de betrouwbaarheid. In dit geval bedroeg deze correlatie:  $r_{tt} = 0,63$  (volgens Pearson). Dit werd te laag gevonden. Omdat de beide testhelften inderdaad volledig gelijkwaardig waren, kon worden uitgegaan van de paralleltestbetrouwbaarheid. Deze kan dan worden gecorrigeerd voor de volledige testlengte (dus 48 items) met behulp van een andere formule. De zo gevonden stabiliteitscoëfficiënt bedroeg  $r_{tt} = 0,77$ . Men mag veilig aannemen dat de betrouwbaarheid van deze F-scale-versie van 48 items 0,77 bedraagt. De standaardfout bedroeg 0,04. Met een waarschijnlijkheid van 95 procent mag men dus aannemen dat de stabiliteitscoëfficiënt van deze F-scale ligt tussen 0,73 en 0,81.

Tot op dit punt van het onderzoek is gewerkt met de ruwe scores op de experimentele versie van 126 items en met de ruwe scores op de voorlopige Nederlandse versie van 48 items. Deze ruwe score werd steeds berekend door optelling van de score op alle items die tot deze versies behoorden. Zoals gezegd kan de score van een afzonderlijk item variëren van een tot zeven punten. Toen een voorlopige versie was samengesteld waarvan de betrouwbaarheid bekend was, rees de vraag of nu reeds een voorlopige normering kon worden gegeven. Onder „norm” moet men verstaan een vergelijkingswaarde, waaraan men zich bij de beoordeling van een testscore oriënteert. In principe zijn normen op de voorgestelde F-scale inderdaad nu mogelijk. Men moet zich bij de interpretatie van de testscore op grond van de te geven normen echter

steeds realiseren dat zij zijn ontleend aan een proefgroep van slechts 525 studenten; zij zijn dus voorlopig.

De eerste vraag die bij voorlopige normering aan de orde komt is, welke soort normen wenselijk zijn. Omdat het bij gebruik van de schaal ook zal gaan om vergelijking tussen individuen, zijn variabiliteitsnormen gewenst, in dit geval standaardcores. Wij hebben een aantal test-technische procedures gevolgd; hierop zal echter in het kader van dit artikel niet verder worden ingegaan.

Voordat de schaal in verder onderzoek kon worden gebruikt, leek een onderzoek naar de „response-set”, – de mate van instem-tendentie van een persoon – zinvol.

Het begrip response-set duikt bij de constructie en validatie van bijna alle psychologische vragenlijsten steeds weer op. Een van de belangrijkste kritieken op de Amerikaanse F-scale's was de mogelijkheid van response-set bij de beantwoording van de items.

Uit hier niet nader te noemen onderzoekingen is gebleken, dat autoritaire instelling en response-set alleen dan geen enkel verband met elkander hebben, wanneer men de items zowel positief als negatief formuleert. Bestaat de lijst uit alleen positief geformuleerde vragen of stellingen, dan blijkt tot op een bepaalde hoogte invloed uit te gaan van de response-set. Instemmen met alles wat autoriteit is (hier dus instemmen met een bewering afgezien van de inhoud van die bewering) kan een aspect zijn van de autoritaire instelling.

Als hypothese van ons onderzoek naar de response-set stelden wij dan ook, dat personen met een hoge score op de F-scale meer instem-tendentie vertonen, terwijl de invloed van deze instem-tendentie bij de laag-scorenden veel minder zal zijn. Om dit na te gaan is een onderzoek gebeurd, dat wij echter vanwege de technische aard hier niet zullen weergeven. Uit de resultaten ervan mogen wij concluderen dat bij personen met een hoge totaalscore, dus met een hoge mate van autoritaire instelling, meer sprake is van een instem-tendentie dan bij personen met een lage score op de F-scale.

*Een onderzoek met de beschreven F-scale bij een aantal huisartsen.* De summier beschreven geconstrueerde F-scale werd in het kader van een uitgebreid onderzoek naar de wijze van praktijkvoering aan 73 huisartsen afgenomen. Het afnemen geschiedde uiteraard vrijwillig en individueel. Bij de interpretatie van de verkregen uitkomsten dient men wel ermede rekening te houden, dat een aantal huisartsen niet voelde voor het invullen van deze F-scale. Gezien deze omstandigheid en het feit dat de voorlopige normering van de geconstrueerde schaal berustte op een onderzoek bij eerstejaarsstudenten, werden de scores van de huisartsen opnieuw genormaliseerd in een zogenaamde stamine-schaal. Het zag er overigens naar uit, dat de huisartsen gemiddeld lager scoorden op de schaal dan de studenten. De score-range bij de huisartsen

liep van 105 tot en met 211. De spreiding is groot. In *tabel 1* zijn de genormaliseerde scores voor de 73 huisartsen weergegeven.

*Tabel 1. Frequentieverdeling van F-scale-scores van 73 huisartsen volgens een genormaliseerde verdeling.*

Interval ruwe score	f	Klasse
105 - 121	3	I
122 - 127	5	II
128 - 145	9	III
146 - 159	12	IV
160 - 175	15	V
176 - 184	12	VI
185 - 196	9	VII
197 - 205	5	VIII
206 - 211	3	IX

Deze verdeling op de F-scale werd vervolgens afgezet tegen een aantal variabelen: de leeftijd, het beroep van de vader van de arts, de grootte van de woonplaats waar hij zijn jeugd had doorgebracht, de grootte van de plaats waar hij zijn praktijk uitoefent en het totale aantal patiënten uit de praktijk. De verwachting was, dat tussen de genoemde variabelen en de scores op de F-scale geen significant verband zou bestaan, dit op grond van vroegere dergelijke onderzoekingen bij studenten met dezelfde en een verkorte F-scale (*Grosfeld*).

*Tabel 2. Verdeling van de F-scale-klassen van 73 huisartsen naar leeftijdsklassen.*

F-scale-klassen	Leeftijdsklassen			
	30-39	40-49	50-59	60-69
I	2	1		
II	2	3		
III	3	4	2	
IV	2	8	1	1
V	6	5	4	
VI	3	4	5	
VII	4	3	2	
VIII	1	3	1	
IX	1	1	1	
Totaal	24	32	16	1

Na samentrekking van de gegevens van *tabel 2* in negen cellen werd met behulp van de Chi<sup>2</sup>-test nagegaan of de verdeling afweek van de te verwachten verdeling. Het resultaat was verre van significant: Chi<sup>2</sup> = 0,96 bij df = 4.

Na samentrekking van de gegevens van *tabel 3* in zes cellen werd via de Chi<sup>2</sup>-test nagegaan of de verdeling afweek van de verwachte verdeling. Het resultaat was significant: Chi<sup>2</sup> = 6,89 bij df 2 (0,025 < p < 0,05). Huisartsen wier vader vanuit het gezichtspunt van beroep tot de hogere sociale lagen kon worden gerekend, scoorden relatief lager op de F-scale. De scores van deze categorie concen-

Tabel 3. Verdeling van de F-scale-klassen van 73 huisartsen naar beroep van de vader.

F-scale-klassen	Beroep van de vader		
	Academicus Grote onderneming Hoger personeel	Middelbaar personeel Middenstand	Agrariër Lager personeel Arbeider
I .....	2	1	
II .....	3	1	1
III .....	4	5	
IV .....	6	6	
V .....	8	7	
VI .....	10	2	
VII .....	1	6	2
VIII .....	2	3	
IX .....	1	2	
Totaal .....	37	33	3

treren zich relatief meer op de middenklassen van de F-scale, terwijl de scores van de huisartsen uit de overige beroepsklassen relatief meer voorkomen in de hogere F-scale-klassen.

Ten aanzien van de verdelingen van de F-scale-klassen van de huisartsen naar respectievelijk de plaats waar zij hun jeugd hebben doorgebracht en de plaats waar zij momenteel hun praktijk uitoefenen – beide variabelen geoperationaliseerd naar inwoneraantal – constateerden wij geen significante verschillen:  $\chi^2$  bedroeg respectievelijk 0,09 bij  $df = 4$  en 0,64 bij  $df = 4$ .

Tabel 4. Verdeling van de F-scale-klassen van 73 huisartsen naar het totale aantal patiënten van ieders praktijk.

F-scale-klassen	Aantal patiënten		
	1 300 - 2 900	2 901 - 4 500	4 501 of meer
I .....	3		
II .....	3	1	1
III .....	2	4	3
IV .....	5	5	2
V .....	6	5	4
VI .....	6	3	3
VII .....	3	5	1
VIII .....	2	2	1
IX .....	–	1	2
Totaal .....	30	26	17

Na samentrekken van de gegevens van tabel 4 in negen cellen werd via  $\chi^2$  getoetst of de verdeling afweek van de verwachte verdeling. Het resultaat was niet significant:  $\chi^2 = 1,31$  bij  $df = 4$ .

De resultaten van deze eerste exploratie doen vermoeden dat alleen het sociale milieu voor zover weergegeven door het beroep van de vader, duide-

lijk heeft te maken met de sterkte van het autoritaire syndroom. De meest voor de hand liggende verklaring van dit verband is wellicht de vaker vermelde opvatting dat in de middenklassen bij de opvoeding meer de nadruk ligt op het normerende aspect ervan. Een dergelijke samenhang tussen sociaal milieu en individueel-psychologische gegevens is ook geconstateerd in het onderzoek naar kleurenvoorkeur bij kinderen (*Grosfeld en Heydendael*). Ook hier werd dezelfde interpretatie gegeven.

Waren er aanwijzingen geweest, dat bovenstaande uitkomsten duidelijk een bepaalde tendens te zien gaven, dan zouden verdere statistische bewerkingen, zoals correlatieberekeningen, zinvol zijn geweest. De verschillen zijn echter te klein om hiervan veel te verwachten.

*Hypothesen voor verder onderzoek.* De F-scale is, zoals reeds gezegd, aan de huisartsen afgenomen in het kader van een uitgebreid onderzoek naar de wijze van hun praktijkvoering. Dit onderzoek is verricht door een arts verbonden aan het Instituut voor Sociale Geneeskunde van de Universiteit te Nijmegen; het zal op een later tijdstip in extenso worden gepubliceerd.

Ten einde na te gaan of het gebruik van de F-scale in dit en in een ander onderzoek zinvol is en of de onderzoeker kan komen tot gerichte hypothesen, hebben wij hoewel slechts summier, enkele vragen wat verder – bijna casuïstisch – geëxploreerd. Daartoe hebben wij de vijf laagst-scorende en de vijf hoogst-scorende huisartsen tegenover elkander gezet. Zij werden op de volgende punten onderling vergeleken: het verwijzingspercentage, het aantal patiënten per uur spreekuur, het aantal patiënten per uur visiterijden, het aantal opgegeven maatschappelijk onaangepaste gezinnen per 1 000 patiënten en het aantal irregulaire patiënten per week. Er zijn natuurlijk ook andere maten mogelijk.

1 *Het verwijzingspercentage.* Omdat zich onder de vijf laagst-scorenden drie apotheekhoudende huisartsen bevonden, was het nodig eerst een correctiefactor in te voeren. Op grond van de verwijzingspercentages van 1966 uit de ziekenfondsgegevens, werden de afzonderlijke gemiddelden berekend van de apotheekhoudende en niet-apotheekhoudende artsen.

De eerste categorie verwees gemiddeld 33,8; de tweede gemiddeld 41,5 per 100 verzekerden. De correctiefactor bedraagt derhalve  $41,5/33,8 = 1,40$ . Na deze correctie te hebben toegepast werden de verwijzingsindices van de vijf hoogst-scorenden en de vijf laagst-scorenden opgeteld. De som bedraagt respectievelijk 218,5 en 218,6. Dit verschil is zo minimaal dat verder toetsen geen zin heeft. Op grond van deze summier exploratie kan dus voorlopig niet worden aangenomen dat de score op de F-scale als operationalisering van de autoritaire instelling heeft te maken met het aantal verwijzingen van de huisarts.

2 *Het aantal patiënten per uur spreekuur.* Het ge-

middelste aantal patiënten per uur spreekuur van de vijf laagst-scorende huisartsen bedraagt 10,0; van de vijf hoogst-scorende huisartsen 11,0. Dit verschil is klein. Uit de spreiding van de cijfers lijkt er echter toch sprake te zijn van de mogelijkheid van een groter verschil, getuige de gegevens van *tabel 5*.

*Tabel 5. Voor verklaring zie tekst.*

	Aantal patiënten per uur spreekuur					
	8	9	10	11	12	13
Laagst-scorenden	1	1	1	1	1	—
Hoogst-scorenden	—	—	3	—	1	1

Het lijkt ons, dat bij verder onderzoek een samenhang kan zijn te verwachten tussen de score op de F-scale en het aantal patiënten per uur spreekuur in die zin, dat de hoog-scorenden relatief meer patiënten per uur spreekuur zullen afhandelen. Deze verwachting is meer gebaseerd op het verschil in spreidingen van de aantallen dan op de gemiddelden.

3 *Het aantal patiënten per uur visiterijden.* Het gemiddelde aantal patiënten per uur visiterijden van de laagst-scorende groep huisartsen bedraagt 3,4; dat van de hoogst-scorende groep 4,8. Ook de spreiding van deze aantallen wijst in de richting van de hypothese dat hoog-scorende huisartsen gemiddeld meer patiënten per uur visiterijden afhandelen dan laag-scorende huisartsen (*tabel 6*).

*Tabel 6. Voor verklaring zie tekst.*

	Aantal patiënten per uur visiterijden				
	2	3	4	5	6
Laagst-scorenden	1	2	1	1	—
Hoogst-scorenden	—	1	1	1	2

Deze voorlopige bevinding zou overeenkomen met de onder punt 2 geopperde veronderstelling.

4 *Het aantal door de huisarts zelf opgegeven maatschappelijk onaangepaste gezinnen per 1 000 patiënten.* Bij dit gegeven werd uitgegaan van de veronderstelling dat hoog-scorende artsen relatief meer zogenaamde onmaatschappelijke gezinnen zullen opgeven. Ondanks het kleine aantal huis-

artsen dat wij voor deze exploratie hanteerden, lijkt het erop dat de veronderstelde samenhang is te verwachten. Het gemiddelde aantal onmaatschappelijke gezinnen per 1 000 patiënten van de vijf laagst-scorende huisartsen bedraagt 4,6, terwijl dit aantal bij de vijf hoogst-scorende huisartsen veel hoger ligt, namelijk 8,6.

5 *Het totale aantal irregulaire patiënten per week.* De verwachting ten aanzien van dit punt is, dat hoog-scorenden relatief minder irregulaire patiënten per week zullen hebben. Deze veronderstelling vindt enigermate steun in de uitkomsten van onze beperkte vergelijking: het gemiddelde van de vijf hoogst-scorenden bedraagt 17,6 tegen dat van de vijf laagst-scorenden 20,4. Hoewel het verschil niet significant is, zou het de moeite waard zijn genoemde veronderstelling in een verder onderzoek uit te diepen.

Hoewel bovenstaande exploraties een uiterst voorlopig karakter hebben — de invloed van de praktijkgrootte hebben wij vanwege de kleine aantallen niet in bovenstaande beschouwingen betrokken — lijkt het ons samenvattend de moeite waard in een onderzoek naar de praktijkvoering van huisartsen meer gebruik te maken van psychologische variabelen. Per slot van rekening is het de huisarts zelf die voor een deel zijn praktijk en alles wat daarmee samenhangt in de hand heeft. Verwonderlijk is het daarom, dat tot nu toe de huisarts in een dergelijk onderzoek of in beschouwingen daarover, slechts als een variabele werd beschouwd in een sociale context en nauwelijks als een individu met historisch gegroeide en actuele eigenaardigheden.

- Adorno, T. W., E. Frenckel-Brunswick, D. J. Levinson en R. N. Sanford, *The Authoritarian Personality*, New York, 1950.
- Albinski, M. *De onderwijzer en de cultuuroverdracht*. Assen, 1959.
- Grosfeld, J. A. M. *Op zoek naar een Nederlandse versie van de F-scale*. Intern G.I.T.P.-rapport, Nijmegen, 1962.
- Grosfeld, J. A. M. en P. H. J. M. Heydendaal (1965) *Begripsvaliditeit als middel tot controle op dubieuze psychodiagnostische hulpmiddelen*, geïllustreerd aan de Lüscher-sneltest. *Gawein*, 14, 20.
- Weima, J. *Authoritaire persoonlijkheid en anti-papisme*. Hilversum, 1963.