1. Statistische computerprogramma's als foutenbron: uitzondering of regel?

G.P. VAN DER VORST

KORTE SAMENVATTING

Bij sociaal-wetenschappelijk en marktonderzoek wordt tegenwoordig veel gebruik gemaakt van computers voor de uitvoering van rekenwerk van statistische aard. Weinig wordt onderkend dat de daarbij benodigde 'statistische' computerprogramma's belangrijke foutenbronnen kunnen zijn. In dit artikel wordt ruim aandacht geschonken aan deze gevaarlijke voetangel op het toch al moeilijk begaanbare gebied van kwantitatief sociaal-wetenschappelijk en marktonderzoek.

1. INLEIDING

Het meten van menselijke eigenschappen en gedragingen moet zonder meer beschouwd worden als een gecompliceerde en omstreden activiteit, zeker in onze huidige samenleving die knap ingewikkeld is geworden. Sociaal-wetenschappelijke onderzoekers en marktonderzoekers die zich er mee bezighouden, en dus hun observaties reduceren tot getallen en in die vorm analyseren, bewegen zich onmiskenbaar op een moeilijk begaanbaar gebied vol voetangels en klemmen. Bij dit laatste moet vooral gedacht worden aan de vele potentiële foutenbronnen in steekproefonderzoek. Bovendien weten dergelijke onderzoekers zich binnen hun eigen wetenschapscomplex bekritiseerd door een toenemend aantal sociale wetenschappers die andere, kwalitatieve methoden voorstaan om inzicht te verkrijgen in menselijk gedrag. Waarbij volgens velen ook nog eens de sociale wetenschappen in een crisis verkeren!

Ondanks dit alles wordt er nog steeds op grote schaal gemeten door sociale wetenschappers en marktonderzoekers. Zij moeten daartoe van veel markten thuis zijn en beschikken over specifieke kennis en vaardigheden. Hier past enige relativering. Het is namelijk wel zo dat de meeste (alle?) voetangels en klemmen onderhand gemarkeerd zijn. Gekwalificeerde onderzoekers zouden dus in staat moeten zijn een betrekkelijk veilige onderzoeksroute uit te stippelen. En met voldoende kennis en steun van anderen, zoals eventueel van methodologen, statistici, enquêteurs en respondenten, zou zo'n route vervolgens zonder al te veel problemen afgelegd moeten kunnen worden.

Aan het eind van de onderzoeksroute beschikt men dan over het begeerde cijfermateriaal. Eind goed, al goed? Dat is te zeggen. Als het verworven databestand omvangrijk is of als men ingewikkelde statistische berekeningen wil uitvoeren, is er nog een verwerkingsprobleem. In de zestiger jaren zijn sociaal-wetenschappelijke onderzoekers begonnen dergelijke problemen op te lossen

met behulp van computers. Deze computertoepassing heeft sedertdien een hoge vlucht genomen die tot op heden voortduurt en waarbij ongetwijfeld ook steeds meer gebruik zal worden gemaakt van microcomputers.

Het gebruik van computers bij omvangrijk rekenwerk is alleszins begrijpelijk. Met de elektronische rekenhulpmiddelen is rekencapaciteit beschikbaar gekomen die de menselijke rekenvermogens verre te boven gaat. Snelheid en precisie: de twee ontbrekende elementen! Weinigen onderkennen echter dat het verwerkingsprobleem vandaag de dag nog steeds bestaat en nu zelfs gecompliceerd wordt door het toenemende gebruik van microcomputers. Bij gebruik van een computer moet dit apparaat namelijk eerst geprogrammeerd worden voor de specifieke computertoepassing. En dat is weer heel lastig en dus duur. Daarmee zou het verwerkingsprobleem getransformeerd zijn in een programmeerprobleem, ware het niet dat er inmiddels, vooral voor de grotere computers (mainframes en minicomputers), een groot aanbod is aan programmatuur (ook wel: software) voor rekenwerk van statistische aard. Het gebruik van statistische programmatuur van derden, vooral van standaardprogrammatuur, is dan ook groot bij universiteiten en andere instellingen die over grotere computers beschikken. Ook of vooral uit kostenoverwegingen wordt dergelijke programmatuur gekocht en grootscheeps gebruikt.

Het probleem is hiermee echter in werkelijkheid niet opgelost; in dit artikel zal ik aangeven dat een computer in veel gevallen beschouwd kan worden als een boobytrap, zij het dat er geen opzet in het spel is. De statistische programmatuur moet gezien worden als de explosieve lading die onmerkbaar en met onbekende kracht kan ontploffen. De bewerking van het databestand kan zodoende leiden tot waardeloze eindresultaten. En dat is een ernstige zaak die dringend aandacht behoeft van alle betrokkenen.

2. HET PROGRAMMEERPROBLEEM

Ik kan me goed voorstellen dat veel lezers vreemd zullen hebben opgekeken van de wending aan het slot van de inleiding. Immers, een computer wordt toch juist als hulpmiddel aangewend om tot een hoge verwerkingssnelheid en -precisie te komen? *In principe* is dat ook mogelijk maar in de praktijk doen zich op beide punten vaak belangrijke problemen voor. Dat heeft alles te maken met het feit dat een computer niet zonder meer voor elke reken- of andere activiteit te gebruiken is. Zoals in de inleiding al is aangegeven, is er een extra hulpmiddel nodig om gegevens met een computer te be- en verwerken: computerprogrammatuur. Dit artikel heeft voornamelijk betrekking op het gebruik van programmatuur voor statistisch rekenwerk, hierna verder aan te duiden met de gebruikelijke maar wat ongelukkig gekozen term 'statistische programmatuur'.

De beperking tot statistische programmatuur is uitsluitend bepaald door de doelgroep van dit jaarboek en dus van dit artikel. Het gaat hier echter om een algemeen probleem dat in feite van groot maatschappelijk belang is maar door slechts weinigen als zodanig wordt onderkend: het programmeerprobleem.

Om het ontstaan van dit probleem te kunnen begrijpen, is een terugblik nodig. De eerste inspanningen op computergebied waren gericht op de constructie van apparatuur. Het programmeren van de eerste computers was een inge-

wikkelde, eigenlijk ondoenlijke opgave. Nadien zijn de inspanningen vele jaren vooral gericht gebleven op de perfectionering van de apparatuur. Het programmeren werd daarbij wel gemakkelijker maar nooit gemakkelijk. Het omgaan met een grote maar eindige machineprecisie vereist bij bepaalde toepassingen bijvoorbeeld specifieke mathematische kennis die relatief weinigen gegeven is. Ook weinig programmeurs!

Het 'gat' tussen de basisfaciliteiten, computerapparatuur en de doorgaans bijgeleverde (bedrijfs)systeemprogrammatuur, en de specifieke toepassing ervan is in feite bijzonder groot. En dat is gerust een maatschappelijk probleem te noemen, gelet op het grootscheepse gebruik van computers in onze samenleving. Dit kan bijvoorbeeld geïllustreerd worden met de kritiek van een groep verontruste computerprogrammeurs op de nieuwe programmeertaal ADA, waarvan de ontwikkeling is geïnitieerd door het Pentagon. ADA wordt zelfs één van de belangrijkste mogelijke oorzaken genoemd van een kernoorlog! Daarbij wordt gedoeld op de onvolkomenheden die ook aan deze (hogere) programmeertaal schijnen te kleven.

Het gebruik van computerapparatuur waarborgt in ieder geval niet de beoogde verwerkingssnelheid en -precisie. En dat strookt bepaald niet met de beeldvorming bij het grote publiek hetgeen misschien wel het feitelijke probleem is. Het gebrek aan adequate programmeercapaciteit en dus programmatuur wordt ook wel aangeduid met 'de software-crisis', een term die dateert van een Nato-conferentie die zo'n vijftien jaar geleden werd gehouden over de toen al gesignaleerde problemen bij het praktische gebruik van computers. Nu verliest de term 'crisis', zeker bij de huidige economische malaise, door het frequente gebruik aan betekenis dan wel kan het gebruik ervan irritatie oproepen. Ik zou daarom liever van een programmeerprobleem spreken. De signalering van een crisissituatie op computergebied heeft overigens nog niet geleid tot een stucturele verbetering. Niet voor wat betreft de beeldvorming bij het grote publiek en (dus?) ook niet voor wat betreft het programmeerprobleem zelf.

Maar, zullen misschien veel lezers denken, hoe zit het dan met de ruimtevaart en andere spectaculaire zaken die alleen met behulp van computers konden worden gerealiseerd? Het transporteren van mensen naar de maan vereist toch een ongelofelijke precisie? Dat is waar maar daar kan direct tegenover worden gesteld dat voor de bedoelde projecten vaak grote fondsen beschikbaar waren, hetgeen een hogere graad van perfectie mogelijk maakt, en dan nóg kwamen er fouten voor in de gebruikte programma's (waarvoor uiteraard weinig 'reclame' is gemaakt). Bij maanreizen wordt daarom constant een grote groep van programmeurs beschikbaar gehouden om tijdens de vlucht geconstateerde programmafouten op afstand te verbeteren.

De komst van 'de' microcomputer wordt door sommigen gezien als een oplossing maar is hooguit een complicerende factor met betrekking tot de situatie op programmeergebied. Er is namelijk misschien al wel de nodige (statistische) standaardprogrammatuur voor 'micro's' voorhanden maar het aanbod is volmaakt ondoorzichtig. De toch al minimale aandacht voor het kwaliteitsaspect raakt nu nog verder versnipperd, althans voorlopig. Voorts zullen fouten in micro-standaardprogrammatuur niet zo gauw door individuele microcomputergebruikers gemeld worden bij de leverancier, op wiens ondersteu-

ning ook minder controle en pressie mogelijk is. De grote mate van beunhazerij die het terrein van de statistische programmatuur kenmerkt, zowel voor wat betreft de ontwikkeling als het gebruik ervan, neemt alleen maar toe door de decentralisering van computergebruik met microcomputers. Er kan dus gerust gesproken worden van een complicatie van het programmeerprobleem.

Voor lezers die langzamerhand het idee hebben gekregen dat hier revolutionaire taal wordt uitgeslagen, zij snel opgemerkt dat het hier beschreven programmeerprobleem algemeen bekend is in de computerwereld waar men niet zo gauw opkijkt van fouten. Het optreden van fouten in programmatuur wordt ervaren als onlosmakelijk verbonden met programmeren.

Tot zover over het perspectief waarin men het verdere betoog moet zien.

3. STATISTISCHE PROGRAMMATUUR

Maar snel terug naar het beperkte onderwerp van dit artikel: het gebruik van statistische programmatuur door sociaal-wetenschappelijke en marktonderzoekers. Het zal nu al wel duidelijk zijn geworden dat er geen florissant beeld zal worden geschetst. Dat zou natuurlijk ook ten enemale onmogelijk zijn waar het bedoelde programmagebruik zich immers afspeelt in de overlapping van twee crisissen, de crisis in de sociale wetenschappen en de software-crisis.

Het is misschien wel zinvol om vooraf duidelijk te stellen dat er niet wordt getwijfeld aan de goede trouw van de gemiddelde onderzoeker/computergebruiker. Wel en vooral aan de kritische zin!

Alleen een gebrek aan kritische zin heeft namelijk kunnen leiden tot het huidige, intensieve gebruik van statistische standaardprogrammatuur binnen het sociaal-wetenschappelijk en marktonderzoek. De voordelen (kostenbesparing e.d.) van het gebruik van dergelijke 'voorgebakken' programma's lijken weliswaar op het eerste gezicht evident maar kunnen vrijwel nooit opwegen tegen de bezwaren die er aan verbonden zijn. Dit zou voor eenieder meteen duidelijk moeten zijn. Wat te denken van, bijvoorbeeld, de volgende twee essentiële vragen:

1. hoe is de nauwkeurigheid van de eindresultaten, gerelateerd aan de nauwkeurigheid van de invoergegevens?
2. wie is verantwoordelijk en wie aansprakelijk voor fouten en onnauwkeurigheden die veroorzaakt zijn door programmafouten?

Zonder een duidelijk antwoord op dergelijke vragen kun je toch maar beter niet een computerprogramma gebruiken bij vaak ingewikkeld statistisch rekenwerk, zou je denken. Kostenoverwegingen moeten immers gebaseerd zijn op kwaliteitsmeting en -garanties? Toch wordt vooral de bekende statistische standaardprogrammatuur, zoals SPSS, BMDP, SAS en P-STAT, ook in Nederland honderdduizenden keren per jaar gebruikt zonder dat de leverancier enig uitsluitsel heeft geboden inzake vraag 1. en de in vraag 2. bedoelde aansprakelijkheid geheel bij de gebruiker ligt die hiermee een zware last op de schouders heeft gekregen!

De voorgaande beweringen behoeven nu onderhand enige onderbouwing. Hierna zullen vanuit de praktijk kanttekeningen worden geplaatst bij zaken die volgens het gros van de onderzoekers tot evidente voordelen van compu-

tergebruik gerekend moeten worden: een hoge verwerkingssnelheid en een grote precisie. Daarbij zal de meeste aandacht besteed worden aan de precisie.

3.1. *De tijdwinst: betrekkelijk*

Voor relativering van het tijdvoordeel dat onlosmakelijk verbonden lijkt met het gebruik van (statistische) programmatuur zal best nog wel begrip te vinden zijn. Bijvoorbeeld bij onderzoekers die al eens een computer hebben gebruikt voor hun statistische berekeningen. Die weten dat het vooral voor beginners lang kan duren voor de eerste resultaten geoogst kunnen worden van computerverwerking via statistische methoden van hun onderzoeksgegevens. Dat wordt veelal teruggebracht tot onnozelheid van de doorsnee-beginner, 'alle begin is moeilijk', maar heeft ook alles te maken met de op computergebied algemeen gangbare opvattingen over 'gebruikersvriendelijkheid'. Wat bijvoorbeeld te denken van het op zijn minst merkwaardige feit dat de leverancier van het grote centrale computersysteem van de Rijksuniversiteit te Utrecht tot op heden gebruikers verplicht de geschatte maximale duur van de verwerking, het geschatte maximale geheugenbeslag e.d. te specificeren in het achttallige stelsel! Dit soort zaken schrikt natuurlijk af en dat geldt zeker ook voor de wijze waarop de verwerking van gegevens met behulp van grote IBM-computers gestuurd moet worden; een voorbeeld:

```
//jobkaart
//TIME=(0,30), REGION=512 k
//SPSSSTAP EXEC SPSS
//FT04F001 DD DSN=&&SAVE,DISP=(NEW,PASS),
//DCB=(RECFM=VB),UNIT=SYSDA,
//SPACE=(TRK,(5,5))
(SPSS-stuurgegevens)
//SASSTAP EXEC SASPLOT
//FILEIN DD DSN=&&SAVE, DISP=(OLD,PASS)
(SAS-stuurgegevens)
```

Daarbij komt dat een verantwoorde keuze van een statistisch programma een uiterst lastige opgave is, zodat een bewuste keuze een tijdrovend karwei zou zijn. Het programma zelf en de bijbehorende documentatie lenen zich doorgaans niet goed voor beoordeling. En dat geldt ook voor de meer ervaren computergebruiker. Dit alles wordt niet als een belangrijk probleem ervaren, programma's worden, ook omwille van de tijd, gekozen zonder een gedegen beoordeling door wie dan ook. Leveranciers maken hier misbruik van door eenzijdig aandacht te schenken aan de mogelijkheden van hun produkt en voortdurend te sleutelen aan de programmatuur, voornamelijk of louter ter vergroting van de toegankelijkheid.

Voorts kan het invoeren van gegevens in een computer ook nog een heel karwei zijn, zeker als dat via ponskaarten gaat, en dat geldt in wat mindere mate voor het controleren en interpreteren van computeruitvoer.

Tenslotte is bij incidentele testen gebleken dat er bij computergebruik relatief forse verschillen kunnen optreden in pure computertijd bij de oplossing van hetzelfde rekenprobleem met verschillende programma's en/of verschillende computers. Dit kan geïllustreerd worden aan de hand van recente testge-

gevens met betrekking tot rekenwerk met de IBM- en de CDC (Control Data Corp.)-versie van SPSS. Daarbij werd een bepaalde statistische toetsingsgrootheid uitgerekend voor twee steekproeven van verschillende grootte. Voor een steekproefgrootte van 144 stuks was de rekentijd op de IBM-computer 0.18 seconde en op de CDC-computer 0.36 seconde, en bij een steekproefgrootte van 1440 bedroeg de rekentijd resp. 0.39 en 15.2 seconden. Helaas is de belangstelling voor dit soort testresultaten nu nog vrijwel nihil.

Dit laatste is, gegeven de totale tijdsinvestering, wel begrijpelijk.

Een belangrijke tegenwerping zou op dit moment kunnen zijn dat er gewoon geen alternatief is. Doe het zelf maar eens met pen en papier. En dat is natuurlijk zo, tenminste als de precisie van de eindresultaten naar behoren moet zijn. En wat dat betreft, er valt toch niet tegen een computer op te boksen?

3.2. *De precisie van eindresultaten: onzeker en dus onbetrouwbaar!*

Waar de tijdwinst door computerverwerking vaak betrekkelijk kan worden genoemd, zijn echter zeker kanttekeningen te plaatsen bij de precisie. Een beschouwing over de cijfermatige nauwkeurigheid die in feite in acht wordt genomen bij de verwerking van gegevens met behulp van een computersysteem plus (statistische) applicatieprogrammatuur, kan prima worden opgehangen aan een beantwoording van de twee belangrijke vragen die aan het begin van 3. zijn opgeworpen met betrekking tot de nauwkeurigheid van de eindresultaten en de verantwoordelijkheid voor programmafouten.

Inzake de verantwoordelijkheid voor fouten en onnauwkeurigheden die zijn veroorzaakt door programmafouten (vraag 2.) verschaffen de leveranciers van statistische pakketten meestal wel de vereiste duidelijkheid. Maar dan alleen in de koop-, huur- of onderhoudsovereenkomst waarin altijd elke verantwoordelijkheid wordt afgewezen. En die contracten krijgt niet iedere gebruiker onder ogen terwijl de 'tussenhandel' (rekencentra e.d.) te weinig aandacht schenkt aan en vraagt voor dergelijke belangrijke clausules.

Voorbeelden van clausules:

'While LICENSOR has carefully developed... and... has been tested for accuracy and proper functioning, LICENSOR disclaims, except as to reasonable operability, any responsibility for the accuracy or correctness of...' of

'It is further understood that neither the University of..., the authors of..., nor the sponsoring agencies will be held responsible for any computational or other error that may arise from the use of...' of

'Although the program has been working satisfactorily for all data analysed so far, no claim is made that it is free of error and no warranty is given to the accuracy and functioning of the program.'

Het gaat in deze gevallen om bekende en veel gebruikte statistische programmatuur. De namen van de bewuste programma's zijn niet vermeld omdat in vrijwel ieder contract soortgelijke clausules voorkomen.

Het is dus duidelijk dat de verantwoordelijkheid geheel afgeschoven wordt. Op wie? Uiteindelijk op de computergebruiker die echter meestal totaal niet in

staat is om de verwerkingsprecisie te volgen en te controleren.

Het is dus van vitaal belang voor de gebruikers om zelf de kans op fouten en problemen goed in te (kunnen) schatten. Gelet op het probleem om een gefundeerde keuze uit het aanbod van programmatuur te kunnen maken, als geschetst onder 3.1. hiervoor, is dat echter onmogelijk.

Die laatste constatering wint aan belang als we nu nog eens vraag 1. proberen te beantwoorden. Het zal al duidelijk zijn geworden dat de leverancier niets garandeert inzake de te verkrijgen nauwkeurigheid. Bij studie van de handboeken (manuals) van statistische pakketten blijkt bovendien al gauw dat er meestal zelfs niets over wordt vermeld in de documentatie. Dat laat dus ruimte voor belangrijke en lastige vragen, gelet op de praktische onmogelijkheid van controle en het programmeerprobleem dat in 2. is besproken.

Het lijkt daarom op dit moment zinnig om eens een blik op de praktijk te werpen. Zitten er fouten in (statistische) programma's, hoe vaak en wat gebeurt daarmee? En dié vragen kunnen weer wél beantwoord worden. Er worden doorgaans veel fouten in programma's gevonden en voor de correctie daarvan is men meestal aangewezen op de leverancier die geen of nauwelijks contractuele verplichtingen hiertoe kent. In de praktijk wordt dus frequent aangetoond dat de nauwkeurigheid van computerresultaten op zijn best onzeker is.

Het vertrouwen van computergebruikers in de precisie van computerresultaten is echter onvoorstelbaar. In 1967 (Longley, 1967) is al eens onderzoek verricht naar de resultaten van destijds veel gebruikte programma's voor regressie-analyse met betrekking tot een 'normaal' probleem. De nauwkeurigheid van de resultaten bleek enorm te variëren (voor kenners: van het teken van de regressiecoëfficiënten tot een nauwkeurigheid van 6 à 7 decimalen). Er zijn daarna nog wel eens onderzoeken uitgevoerd naar de precisie van resultaten verkregen met verschillende (statistische) programma's. Daarbij werden telkens wel miskleunen gesignaleerd in ook bekende programmatuur (zie bv.: Wampler, 1972; Francis, 1973; Golhar & Skillings, 1976; Stemerdink, 1982).

Hoe dan ook, de slotsom van de voorgaande opmerkingen lijkt overduidelijk. Computerresultaten die niet nagerekend of 'bewezen' kunnen worden, moeten worden beschouwd als onnauwkeurig dan wel als onzeker. En dus in ieder geval als onbetrouwbaar!

In de volgende paragraaf zal nader worden ingegaan op veel voorkomende fouten in programma's en documentatie en zullen voorbeelden worden aangereikt.

4. FOUTEN IN STATISTISCHE PROGRAMMA'S EN DE BIJBEHORENDE GEBRUIKERSDOCUMENTATIE

Met computers kunnen in principe uiterst precieze berekeningen worden uitgevoerd. De precisie – ook van de afzonderlijke basisoperaties in de hardware – is echter wel eindig en kan resulteren in een relatief grote afwijking van de juiste uitkomst. De totale afwijking wordt mede bepaald door het gekozen rekenschema (algoritme). Bij de berekening van gemiddelden en standaarddeviaties kan door de keuze van een bepaald rekenschema al van alles mis gaan, zij

het dat het bij voorbeelden vaak wel om tamelijk exotische gegevens gaat. In handboeken wordt hier wel aandacht aan geschonken, maar de daar gegeven waarschuwingen zouden algemener moeten klinken. In dit verband moet ook nog worden gezegd dat de mate van precisie kan verschillen voor computers van verschillend fabrikaat, een gegeven waarmee bij omzetting (conversie) van programma's vaak weinig rekening wordt gehouden.

Er zijn echter nog meer faktoren die de precisie van de eindresultaten van computerverwerking kunnen beïnvloeden:
- programmeerfouten of tekortkomingen van programmatuur;
- fouten en tekortkomingen in de programmadocumentatie en -uitvoer.

De doorsnee gebruiker is zich van deze foutenbronnen nauwelijks of niet bewust.

4.1. *Programmeerfouten*

Bij het programmeren kunnen fouten van verschillende aard worden gemaakt en er worden in de praktijk ook inderdaad veel fouten gemaakt; programmeren moet worden gezien als een ingewikkelde menselijke activiteit. Een extra complicatie kan gelegen zijn in de kwaliteit van de gebruikte programmeertaal en het bijbehorende vertaalprogramma (compiler), waarmee programmatekst omgezet kan worden in machine-instructies.

Ook in veel gebruikte (standaard)programmatuur komen allerlei soorten fouten voor. Bij verbetering van fouten worden vaak nieuwe geïntroduceerd. Het is in de computerwereld een volkomen aanvaard gegeven dat grote programma's legio fouten bevatten. Van dit laatste getuigen ook de citaten uit mededelingen van de zijde van de leverancier van de CDC-versie van SPSS: '...This error may cause a mode error, or the results for all lists but the first may be incorrect', '...the results are strange... if not, the results are unpredictable and bad', '...the results are generally too large' enzovoorts. In weer een ander geval wordt meegedeeld dat in plaats van het gewenste resultaat de wortel daaruit wordt afgedrukt (Overigens wel netjes dat deze leverancier ontdekte fouten kenbaar maakt aan in principe alle gebruikers).

Hoewel ontdekte fouten soms wel worden gepubliceerd en meestal wel na enige (vaak: geruime) tijd worden gepubliceerd, is het maar de vraag hoeveel gebruikers (tijdig) op de hoogte komen van die fouten en, misschien nog belangrijker, wat er resteert aan onontdekte fouten. Voortdurend worden namelijk nieuwe fouten gevonden in versies van SPSS en andere (standaard)programma's. Zo bleek een bepaalde berekening met de CDC-versie van SPSS te resulteren in een kanswaarde ter grootte van 1.18. Een kanswaarde kan echter nooit groter zijn dan één. In andere gevallen werd de betreffende kanswaarde, hoewel binnen de theoretische grenzen, stelselmatig te hoog berekend; dit werd pas bij toevallig narekenen ontdekt. Ook bestaat vaak onduidelijkheid over de wijze waarop getoetst is: één- of tweezijdig. Bij dit alles dient bedacht te worden dat een kanswaarde vaak als een belangrijk sluitstuk wordt beschouwd van een statistische analyse.

Voorts blijken vaak te weinig controles ingebouwd te zijn in (standaard)programma's. Men dient hierbij te bedenken dat standaardprogrammatuur

geschreven is voor verwerking van 'normale' problemen. Dit kan bij verwerking van onjuiste, onbedoelde of exotische gegevens leiden tot ongedefinieerde akties als delen door nul (Het inbouwen van veel controles kan overigens ook het gevaar vergroten dat gebruikers zich nog meer op computers gaan verlaten. Deze signaleren dan immers toch ook de vergissingen?).

Veel programmafouten betreffen onnauwkeurigheden in de rekenprocedures. We hebben hier bijvoorbeeld te maken met afrondfouten. Tussentijdse afrondingen kunnen meer ingewikkelde rekeningen geheel verstoren.

Verder moeten we hier ook rekening houden met het ontbreken van adequate controles waardoor bijvoorbeeld doorgerekend kan worden met 'negatieve' kwadraten of waarbij aftrekkingen en delingen van bijna gelijke getallen tot ernstige rekenfouten kunnen leiden.

Ook speelt hier de rekenprocedure een rol. De nauwkeurigheid van bij voorbeeld een iteratieproces (proces van stapsgewijze benadering) wordt mede bepaald door het gekozen stopcriterium en daarmee de precisie van de eindresultaten.

4.2. *Dokumentatie*

De tweede categorie fouten die de precisie van de eindresultaten beïnvloeden, bestaat uit fouten en tekortkomingen in de programmadocumentatie. Hierin wordt vaak geen aandacht besteed aan de mate van precisie van resultaten, ook niet los van eventuele onnauwkeurigheden in de data. De gebruikte algoritmen worden soms niet gepubliceerd. Algoritmen worden ook nogal eens 'overgenomen' door de ene leverancier van de andere, hetgeen de controle van resultaten door verwerking van gegevens met twee verschillende programma's bij voorbaat tot een dubieuze tijdsinvestering maakt. De documentatie bevat bovendien meestal geen waarschuwingen voor fouten en onnauwkeurigheden.

Het voorgaande geldt ook voor de verklarende tekst (of het ontbreken ervan) in de uitvoer.

5. GEBRUIKERSFOUTEN

Een andere categorie fouten betreft die van de gebruiker zelf. Men kan fouten maken bij de keuze van een programma, door de probleemstelling aan te passen en/of door niet te toetsen op vooronderstellingen. Dergelijke fouten spruiten vaak voort uit een gebrek aan kennis van de gebruikte methoden. Ook kan men fouten maken in de programmasturingen (vertaling van de gebruikerswensen in de invoervoorschriften voor het programma) of door deze weg te laten en gebruik te maken van de zogenaamde 'defaults'.

Door gebruik van programmadefaults kunnen met veel standaardprogrammatuur zonder veel moeite ingewikkelde berekeningen uitgevoerd worden. Door middel van enkele programmasturingen kan men uitvoer bewerkstelligen (resultaten voor het uitzoeken) zonder voor lastige keuzen en vragen te worden gesteld: 'grand fishing expeditions'. In feite volgt men bij het gebruik van defaults wegen die door de ontwerper van het pakket zijn aangegeven. Dat gebeurt veel, zeker bij het gebruik van ingewikkelde technieken. Ook hier-

bij is vaak sprake van aanpassing van de probleemstelling (uit tijdsoverwegingen, onbedoeld of uit onmacht?).

Gebruikersfouten kunnen onopgemerkt blijven en desastreuze gevolgen hebben voor de eindresultaten. Het betreft hier veel voorkomende fouten.

Ook zijn er fouten en onnauwkeurigheden in de invoergegevens mogelijk. Analyses waarbij 'struktureel' sprake is van onvolkomenheden in de invoergegevens, worden in de literatuur ook wel aangeduid met 'gigo-research' (garbage in, garbage out). Waar bij veel sociaal-wetenschappelijk onderzoek en marktonderzoek data relatief onnauwkeurig zijn, wordt over de gevolgen daarvan voor de eindprecisie meestal niets in de documentatie vermeld. Er wordt in de praktijk gewoon geen rekening gehouden met dit gegeven.

6. MAATREGELEN

6.1. *Inleiding*

Tot zover kritiek op de gang van zaken. Veel onderzoekers zullen daar moeite mee hebben. Daarom wordt hier nogmaals benadrukt dat de goede trouw van de betrokken computergebruikers niet in twijfel wordt getrokken.

Het zal hopelijk duidelijk zijn geworden dat het hier gaat om een poging (potentiële) rekenaars te bewegen tot een kritische benadering van het gebruik van het hulpmiddel computer en de daarbij benodigde programmatuur. En tot een kritische benadering van de overige betrokkenen bij samenstelling, beschikbaarstelling en praktisch gebruik van programmatuur. Dat zou namelijk kunnen leiden tot een belangrijke verbetering van de huidige situatie, ook al zou een dergelijke benadering in eerste instantie gebaseerd zijn op wantrouwen ten aanzien van de kwaliteit van programmatuur. Zulk wantrouwen is alleszins gerechtvaardigd, eigenlijk onvermijdelijk en noodzakelijk.

Hoog tijd voor wat meer opbouwende taal, saneringsmaatregelen of oplossingen, als een soort handreiking aan de 'slachtoffers', de betrokken computergebruikers? Of juist niet? Immers, zolang binnen de groep van (potentiële) rekenaars met de computer onvoldoende inzicht bestaat in de stand van zaken zijn saneringsmaatregelen en oplossingen waarschijnlijk niet aan de orde.

Het suggereren ervan kan zelfs een hachelijke onderneming zijn. Het onderwerp is namelijk niet nieuw en bij eerdere discussies over de kwaliteit van statistische programmatuur verwekten kritische geluiden slechts een kortdurend schokeffect. Opgeroepen frustraties richtten zich doorgaans op de critici, althans op de door de critici aangedragen alternatieven, maatregelen e.d. die aldus de ongewenste functie kregen van bliksemafleider (te algemeen, te weinig pragmatisch of iets van dien aard).

Nu moeten de bedoelde discussies dan ook vaak gezien worden als woordenwisselingen bij frontale botsingen van twee tegenovergestelde meningen. Enerzijds het 'gevestigde' standpunt: leve toegankelijke en veelomvattende programmatuur! Anderzijds ideeën, zoals vervat in dit artikel, die er op neer komen dat het gebruik van programmatuur voor rekenwerk maar 'linke soep' is en weinig ruimte bieden voor populaire oplossingen.

Er is eigenlijk maar één echt goede oplossing denkbaar voor de korte termijn opnieuw en overdacht beginnen met de aanmaak van programmatuur en de beschikbare programmatuur laten voor wat ze is (verbeteren van de bestaande programmatuur zou een lapmiddel zijn). En dat is onmiskenbaar een impopulaire oplossing voor de problemen, gelet op de vlucht die het rekenen met computers heeft genomen en de afhankelijkheid die daardoor ontstaan is. Het zou wel een afdoende oplossing zijn!

Natuurlijk kan dit wat genuanceerd worden en het gebruik van bepaalde, nu reeds in omloop zijnde (standaard)programmatuur, bij voorbeeld voor de produktie van frequentie- en kruistabellen, zou misschien nog door de vingers gezien kunnen worden. Maar 'daarna' houdt het al gauw op en zullen (potentiële) rekenaars hun hand niet voor de correcte werking van programmatuur in het vuur (kunnen) steken.

Geluiden voor wat betreft de langere termijn kunnen positiever zijn. Immers, bij een nieuwe start kan gebruik gemaakt worden van de inmiddels opgedane ervaringen. Die moeten dan wel geanalyseerd worden. Ook zonder diepgaande analyse kunnen echter nu al punten geduid worden waarop verbetering nodig en mogelijk zal zijn.

6.2. *Voorlichting*

Zo wijzen de voorgaande kanttekeningen in de richting van een andere aanpak van de voorlichting omtrent de (on)mogelijkheden van computers. Waar elk lid van onze samenleving wel op de een of andere wijze 'automatiseringsconsument' is of zal worden, dient er gesleuteld te worden aan de beeldvorming rond computers. Een inhaaloperatie waarbij ook met spoed en bekwaam gewerkt moet worden aan de totstandkoming en verdere ontwikkeling van een consumentenbewustzijn. Dit zal betrokkenheid en actie eisen van rekencentra, informatici, numerici en statistici alsook van leveranciers van computerprogrammatuur.

Voor wat betreft de inbreng hierbij van de zijde van leveranciers, lijkt in Nederland echter nog geen basis voorhanden. Zo schijnen consumentenorganisaties in Nederland al jaren aan te dringen op een betere informatieverschaffing door leveranciers, vooral betreffende de niet-levensmiddelen. Hiervoor zou de overheid het wettelijk kader moeten scheppen. Het belang van wettelijke regelingen kan in dit verband niet genoeg benadrukt worden, gelet op de verwachting dat het gebruik van computerelektronica en programmatuur in de komende jaren een ingrijpende invloed zal hebben op onze samenleving.

6.3. *Ontwikkeling consumentenbewustzijn*

Het is de vraag of een verbetering van de voorlichting een belangrijke voorwaarde is voor het ontstaan van consumentenbewustzijn. Het moet toch mogelijk zijn dat de 'automatiseringsconsument' zijn rechten herkent. Waar blijven consumentenorganisaties zoals de 'Consumentenbond'? Misschien is een aparte organisatie wenselijk.

Computerprogrammatuur is een produkt waaraan eisen gesteld kunnen

worden, zoals kwaliteitseisen. Dat gebeurt echter nauwelijks; het marktgebeuren wordt vrijwel geheel beheerst door de leveranciers. Met alle gevolgen van dien!

Eén van die gevolgen is dat de (potentiële) gebruiker van computerprogrammatuur voor rekenwerk geconfronteerd wordt met een vrijwel onoplosbaar informatieprobleem: hoe valt overzicht van en doorzicht in de 'markt', het aanbod, te verkrijgen? Vooral doorzicht valt nauwelijks te verwerven. Hiertoe moet namelijk de prijs/prestatie-verhouding voor de in aanmerking komende programmatuur bepaald kunnen worden ter vergelijking en zouden in ieder geval de (on)mogelijkheden van het produkt grondig bekend moeten zijn. Dat zou specifieke kennis, ervaring, bruikbare kwaliteitscriteria en veel tijd vereisen. Aan deze combinatie van eisen wordt door individuen zelden voldaan. Bundeling van kennis is dan ook een eerste vereiste. Het probleem in deze wordt in de praktijk echter door weinigen onderkend en gemis aan kennis wordt niet als een hinderlijke beperking voor computergebruik ervaren. Men laat zich bij de keuze van programmatuur doorgaans door derden leiden of door louter beschikbaarheid van bepaalde programmatuur. Hiervoor is genoegzaam aangetoond hoe gevaarlijk dat wel niet is.

6.4. *Erkenning van programmeren als hoog gekwalificeerde arbeid*

Programmatuur is niet zomaar een produkt. De produktie ervan vereist vakmanschap.

Het is momenteel echter een volkomen geaccepteerde zaak dat Jan & Alleman allerlei programma's samenstelt en zijn produkten, vaak voor geld, slijt aan anderen; de snelle opkomst van de microcomputer is, zoals al gezegd, een stimulans gebleken voor deze beunhazerij. Het ontwerpen van programmatuur moet echter beschouwd worden als hoog gekwalificeerde arbeid. Spoedige erkenning van dit gegeven zou ook bereikt kunnen worden bij een goede voorlichting en vervolgens door het stellen van kwaliteitseisen door de consument.

Voor programmeren voor derden zou men feitelijk over een soort licensie moeten beschikken, terwijl het vak zonder meer verder nog een gedegen (beroeps)code vereist. Dergelijke zaken kunnen misschien alleen van overheidswege op gang gebracht worden, waar de bestaande automatiseringsclubs nadrukkelijk in gebreke zijn gebleven tot op heden. Dit alles zolang de ontwikkelingen op computergebied het samenstellen van programmatuur niet ingrijpend vereenvoudigd zullen hebben. In het algemeen zou meer aandacht besteed moeten worden aan het bestaan en de verdere ontwikkelingen van mathematische hulpmiddelen om de correctheid van programma's aan te tonen.

6.5. *Verduidelijking en sanering van vakgebieden en relaties*

Met een beschouwing van de programmeur komt men er echter niet. De taken en relaties in 'software-land' moeten goed afgepaald worden.

Daarbij lijken allereerst de takenpakketten van instellingen als rekencentra in het geding. Dergelijke instellingen besteden soms 60 à 70% van de tijd aan 'onderhoud' van programmaprodukten van derden alsook aan daarmee gere-

lateerde taken. Vaak zonder benul van de kwaliteit van de beheerde programmaprodukten! Tijd voor de ontwikkeling van (bijvoorbeeld onderhouds)normen, een echt vakgebied waardig. De verschillende taken van instellingen als rekencentra moeten bovendien meer afgeleid zijn van de wensen en behoeften bij de (potentiële en eigenlijk alleen de kritische) computergebruikers. Daar mankeert het nogal aan. Dat is wel begrijpelijk omdat de *huidige* computergebruikers zich door grote verschillen in opvattingen, behoeften en computerfaciliteiten nauwelijks (laten) organiseren. Dat is echter geen reden voor de rekencentra om het heft niet zelf in handen te nemen, buiten de broodnoodzakelijke activiteiten als voorlichting en evaluatie.

De bijzondere positie van rekencentra als intermediair tussen software-consumenten en -producenten dient gestalte te krijgen. De rekencentra dienen dan wel meer nadrukkelijk dan voorheen voor de belangen van hun gebruikers op te komen. Ter verontschuldiging voor de nalatigheid in deze van rekencentra moet wel gesteld worden dat de wettelijke bescherming van computerprogrammatuur (ook) in Nederland niet geregeld is. Onder druk van afzonderlijke gebruikers(groepen) moet programmatuur beschikbaar gemaakt worden van leveranciers die dan contracten voorleggen onder het motto 'slikken of stikken'.

Tenslotte dringt zich in dit verband de vraag op wat geëist mag worden van informatici, numerici, statistici enz.

6.6. *Evaluatie*

Daarmee zijn we toegekomen aan de mogelijkheid van samenwerkingsprojecten. Voor de verdere gang van zaken van groot belang, omdat hoe dan ook aan een echte verbetering van de situatie grote *gezamenlijke* inspanningen ten grondslag moeten liggen.

Dat geldt niet alleen de aanmaak van programmatuur maar ook de formulering en toepassing van kwaliteits- en evaluatiecriteria. Evaluatie is dringend noodzakelijk, de prioriteit moet zo snel mogelijk verlegd worden van meer naar GOED.

Er is overigens op het gebied van statistische computerprogrammatuur hier en daar wel schuchter wat aan evaluatie gedaan. De resultaten zijn vaak nauwelijks bruikbaar en worden in ieder geval nauwelijks gebruikt. Met een lijst met criteria en overwegingen voor evaluatie, enkele jaren geleden in Amerika opgesteld door Francis c.s. (Francis, Heiberger & Velleman, 1975; Velleman & Welsch, 1976) is het net zo vergaan als het bekende boek 'A discipline of programming' (Dijkstra, 1976): er is bijzonder veel aan gerefereerd, het lijkt aanzienlijk minder gelezen en wordt zo te zien nauwelijks in praktijk gebracht. Hierbij wordt altijd tijdgebrek als voornaamste reden voor nalatigheid aangevoerd. Een kwestie van prioriteitenstelling. Hiermee is tegelijk aangegeven waarom inventarisatie en evaluatie niet als eerst noodzakelijke lange termijnactiviteiten zijn genoemd. Dat zijn ze natuurlijk wel – hoe is de situatie nu precies? wat moet er veranderen? – maar de opstelling tot nu toe van de betrokken 'computerdeskundigen' (leveranciers van programmatuur, samenstellers, rekencentra e.d.) lijkt te wijzen op de noodzaak van gebruikersdwang. In

dat licht is dit artikel in ieder geval tot stand gekomen. En dus ook de opsomming van gewenste/noodzakelijke activiteiten in deze paragraaf.

7. EPILOOG

Wat te denken van de computergebruiker die meende dat met de introductie van BMDP* een Brede Maatschappelijke Discussie inzake Programmatuur zou zijn ingezet? Van een brede maatschappelijke discussie inzake programmatuur is echter geen enkele sprake. En dat zou misschien ook wel overdreven zijn, alhoewel het gebruik van computers in onze samenleving kritisch bezien zou moeten worden.

Toch is in de afgelopen jaren in Nederland wel eens kritisch gediscussieerd over het gebruik van statistische (standaard)programmatuur. Bijvoorbeeld binnen de Vereniging Voor Statistiek, bij de oprichting van een Contactgroep Statistische Programmatuur (thans Sectie Statistische Programmatuur) en bij een latere jaarvergadering van die contactgroep; in beide gevallen was het discussiethema: 'Statistische pakketten, een zegen of een ramp?'. Een en ander heeft echter niet geleid tot een kritische beweging rond het gebruik van statistische programmatuur.

Incidentele critici krijgen nu soms een applaus maar hebben het meestal zwaar te verduren. Misschien is dat wel afhankelijk van de presentatie en eigenlijk is het niet zo belangrijk. Belangrijker is het effect van kritiek op wat langere termijn. En dat is tot nu toe – zowel na lof als na hoon – nog niet veel geweest. In ieder geval niet een spoortje van opwaartse druk op de activiteiten van welk samenwerkingsverband dan ook op dit gebied. Ook geen druk van gebruikers op het doen en laten van rekencentra en van rekencentra (samen met 'hun' computergebruikers) op leveranciers. Niets daarvan!

De vraag is hoe spijtig dat allemaal is. Immers, iedereen vermaakt zich of verdient zijn brood? De vraag moet ieder maar voor zichzelf beantwoorden. Mogelijk kunnen de overwegingen in het voorgaande daarbij van nut zijn.

BIBLIOGRAFIE

Dijkstra, E.W., 1976, *A discipline of programming,* Prentice-Hall Inc.

Francis, I., 1973, A comparison of Several Analysis of Variance Programs, *Journal of the American Statistical Association,* Vol. 68, No. 344 (December) blz. 860.

Francis, I., R.M. Heiberger en P.F. Velleman, 1975, Criteria and Considerations in the Evaluation of Statistical Program Packages, *The American Statistician,* Vol. 29, No. 1 (Februari) blz. 52.

Golhar, M.B. en J.H. Skillings, 1976, A comparison of Several Analysis of Variance Programs with Unequal Cell Size, *Communications in Statistics.*

Longley, J.W., 1967, An Appraisal of Least-Squares Programs for the Electronic Computer from the point of View of the User, *Journal of the American Statistical Association,* Vol. 62, blz. 819.

Stemerdink, G.J., 1982, Vergelijking van rekennauwkeurigheid bij verschillende statistische pro-

* BioMeDical Programs, een pakket van statistische standaardprogramma's.

grammapakketten, *Kwantitatieve Methoden,* Vol. 3, No. 5, blz. 116.
Velleman, P.F. en R.E. Welsch, 1976, On Evaluating Interactive Statistical Program Packages, *Communications in Statistics,* blz. 197.
Wampler, R.H., 1972, Some Recent Developments in Linear Least-Squares Computations, *Proceedings of the Interface: Computer Science and Statistics,* 6th Annual Symposium, Berkely.

2. Survey-onderzoek onder buitenlandse werknemers, kan dat wel?

H. WENTHOLT

SAMENVATTING

Aan de hand van een beschrijving van de opzet en uitvoering van een survey-onderzoek onder 8 nationaliteiten buitenlandse werknemers in Nederland, maakt de auteur duidelijk dat dergelijk onderzoek bijzonder gecompliceerd is. Het vergt veel creativiteit, improvisatie en dus juist ook een meer dan gebruikelijke, discipline van de onderzoeker. Een routine-kwestie zal onderzoek onder buitenlandse werknemers voorlopig nog niet kunnen zijn. De winst van het in 1980 en '81 uitgevoerde NOS-onderzoek 'Massamedia en buitenlandse werknemers in Nederland'* bestaat, behalve uit belangwekkende resultaten en conclusies, uit het 'bewijs' dat het in ieder geval mogelijk is om de vele moeilijkheden, voetangels en klemmen van representatief onderzoek onder deze andere culturen binnen onze samenleving, te overwinnen. En daaraan werd en wordt nogal eens getwijfeld.

INLEIDING

Het aantal uitgevoerde en in uitvoering zijnde onderzoeken onder zgn. etnisch-culturele minderheidsgroepen in ons land neemt de laatste tijd sterk toe. Toch zijn de onderzoeken, die uitgaan van de survey-methode en die althans pogen een voor een bepaalde populatie representatieve beschrijving van gedragingen, meningen en houdingen te verwezenlijken, nog steeds zeer schaars.

De redenen daarvoor kunnen er vele zijn. Zowel m.b.t. de survey als methode, als m.b.t. het trekken van representatieve steekproeven.

1.1. *Het survey aspect*

De methode van een vraaggesprek (als communicatie-proces en als sociale situatie) is een methode die méér thuishoort in de cultuur van geïndustrialiseerde landen met relatief ver ontwikkelde onderwijs-systemen, dan in zgn. niet westerse culturen. Het hanteren in zo'n vraaggesprek van een systematisch opgezette vragenlijst, maakt die 'vreemdheid' tot de cultuur van de ondervraagde alleen maar groter.

Het is hier in deze inleidende zinnen niet de plaats om fundamenteel in te gaan op alle verschillende factoren, die in een vraaggesprek-situatie bij een surveybenadering, van invloed zijn op het uiteindelijk 'produkt' van een dergelijk

* Wentholt, Hans – 'Massamedia en buitenlandse werknemers in Nederland', NOS-Programmadiensten Publiek en Programma Onderzoek, Hilversum, februari 1982. [Volledig rapport 391 pagina's; samenvatting 31 pagina's.]

vraaggesprek. D.w.z. op de geregistreerde antwoorden op gestelde vragen, die verondersteld worden de werkelijkheid van de betreffende ondervraagde te representeren. Velen zullen de werkelijkheidswaarde van dergelijke resultaten met bijv. Nederlandse respondenten al relativeren; des te meer reden voor scepsis is er m.b.t. survey-onderzoek onder bijv. de voornamelijk mediterrane groepen buitenlandse werknemers in ons land. Een survey-onderzoek onder buitenlandse werknemers over een onderwerp als de cultuur- en communicatie-kloof, die er bestaat tussen hen en de hen omringende Nederlandse samenleving, bedient zich ironisch genoeg van een methode waarin, zowel inhoudelijk als 'procedureel', de onderzochte culturele kloof zèlf tot uitdrukking komt.

Zoals gezegd, dit vormt een intensivering en verruiming van een probleem, dat zich bij survey-onderzoek in onze eigen autochtone samenleving ook reeds voordoet. Toch wordt dergelijk onderzoek sinds enige decennia in de westerse wereld uitgevoerd en zelfs op steeds grotere schaal en t.a.v. steeds meer aandachtsgebieden. Kennelijk kan de survey-methode bij bepaalde probleem- en vraagstellingen resultaten opleveren, die bijv. beleidsmakers op velerlei gebieden en niveaus, als ook wetenschapsmensen, valide en daarom soms ook bruikbare inzichten verschaffen. Het feit, dat de survey-methode nog zo weinig voor onderzoek onder etnisch-culturele minderheden in ons land is toegepast, zou er op kunnen duiden, òf dat de voor survey-onderzoek in aanmerking komende probleem- en vraagstellingen zich voor deze groepen in de Nederlandse samenleving nog nauwelijks hebben voorgedaan, òf dat de methode altijd op grond van de zojuist aangestipte problemen bewust werd verworpen.

Als dat zo zou zijn, is dat in het licht van het ontbreken van een duidelijk beleid t.a.v. minderheden zowel als in het licht van de ongelooflijk vele en moeilijke problemen die bij het interviewen van buitenlandse werknemers moeten worden opgelost, begrijpelijk. Maar zeker ook te betreuren. Juist omdat er een duidelijk en onderbouwd beleid hoogst noodzakelijk is.

Een aantal problemen bij het afnemen van systematische vraaggesprekken aan buitenlandse werknemers in het kader van survey-onderzoek, komen hierna onder par. 2 aan de orde. Ook de beperkingen die daardoor ontstaan. Eén, aan de survey-opzet gerelateerd probleem, verdient in deze inleiding nog aparte aandacht:

1.2. *Het representativiteits-aspect*

Survey-onderzoek is meestal steekproef-onderzoek. Het trekken van representatieve steekproeven uit de populatie buitenlandse werknemers in een land als Nederland is een ongehoord moeilijke, tijdrovende en dus kostbare zaak. En, t.o.v. het 'representativiteits-ideaal' zelfs een riskante zaak.

Sceptici t.a.v. de survey-methode zijn veelal tegelijkertijd ook zeer relativerend over deze methodische representativiteits-doelstelling. Het woord 'ideaal' zou men in dit verband al helemaal niet in de mond willen nemen. Toch kan representativiteit voor bepaalde beleids- en/of onderzoekdoelstellingen zonder meer essentieel zijn.

Bij het onderzoek 'Massamedia en buitenlandse werknemers in Nederland', dat ik recentelijk voor de NOS-Programmadiensten televisie en radio heb begeleid en gerapporteerd, was dit bijvoorbeeld het geval. Zowel ter evaluatie van de eigen programma's voor buitenlandse werknemers, als voor het in kaart brengen van het meer algemene media-gedrag en het verkrijgen van een beeld van de functie daarvan t.b.v. de bestaans-ontwikkeling van groepen buitenlanders als deel van de Nederlandse samenleving, alsmede van de ideeën, wensen en verwachtingen daarover van de diverse nationaliteiten buitenlanders zèlf, ging het de NOS essentieel om de Turken, Marokkanen, Grieken etc. als *totale groep*. Als 'massa' zo men wil. En niét om geëmancipeerde, nederlands sprekende, voorhoede-personen uit de verschillende nationaliteiten, onder wie sociaal-wetenschappelijk onderzoek naar heersende maatstaven inderdaad best mogelijk is. Het effectueren van een emanciperende doelstelling d.m.v. een massa-medium als tv, dient volgens de NOS (en ook volgens ondergetekende) wezenlijk uit te gaan van een gerichtheid op 'een ieder' uit de groep waar men zich op richt. En daarmee wordt de representativiteit van een onderzoek, dat een zodanige beleidsdoelstelling moet onderbouwen, zulk een principiële zaak, dat woorden als 'representativiteits-ideaal' niet eens meer als een farce hoeven te klinken. Als de programma-inspanningen zich – misschien wel juist – ook richten op de analfabeten of de meest geïsoleerden onder de doelgroepen, óók op de illegaal hier verblijvende buitenlanders; als zij juist óók door alle politieke en levensbeschouwelijke scheidingslijnen binnen de verschillende populaties héén willen breken, dan zal het beleidsonderbouwende onderzoek dat zèker moeten doen.

In de navolgende paragrafen geef ik vanuit de in deze inleiding neergelegde optiek, een feitelijk en methodisch verslag van het 'hoe' van het uitgevoerde NOS-onderzoek. Het survey-onderdeel daarvan was, gegeven de doel- en vraagstelling, essentieel. Maar, uitgaande van de beperkingen van zo'n survey-aanpak, zeker voor de onderzochte mediterrane nationaliteiten: even essentieel waren de *ter uitdieping* en *aanvulling* hierop uitgevoerde talrijke *groepsgesprekken*. Enerzijds met 'doorsnee'-personen uit de onderzochte populaties, anderzijds – toch (!) – met allerlei voorhoede-groepen. De bedoeling van deze verslaggeving is aan te geven, dat uitvoering van representatief survey-onderzoek onder buitenlandse werknemers in ons land, moeilijk maar niet onmogelijk is. We hebben bij de NOS, wegens te verwachten problemen, lang geaarzeld om het te doen (zo'n jaar of vijf), maar zijn achteraf uiterst tevreden, uiteindelijk gewoon in het diepe te zijn gesprongen.

2. DE OPZET VAN HET UITGEVOERDE ONDERZOEK

2.1. *Steekproeftrekking en selectie*

2.1.1. *Survey-onderzoek onder volwassenen*

2.1.1.a. *Selectie en representativiteit*

De wenselijkheid van het trekken van representatieve steekproeven uit 8 verschillende populaties buitenlandse werknemers in Nederland is hiervoor uitvoerig beargumenteerd. In de praktijk is op beperkte wijze aan die wenselijkheid tegemoet gekomen kunnen worden, waarbij *prioriteit* is gegeven aan

representativiteit t.a.v. de verdeling van individuen over de verschillende socio-economische strata boven representativiteit t.a.v. de regionale verdeling van elk der populaties in Nederland.

Daarom is gekozen voor het trekken van a-selecte steekproeven uit *volledige* adressen-bestanden (= de *totale* populatie) van elk van de te onderzoeken nationaliteiten, maar dan in een beperkt aantal gemeenten.

D.w.z. dat het grootste gewicht van de inspanningen van het onderzoekteam om een steekproef te trekken is gelegd bij het werkelijk zo compleet en volledig maken van het adressen-bestand in een bepaalde gemeente om niet een bias in te bouwen, die bijvoorbeeld bestaat uit het missen in je steekproef van individuen/gezinnen die illegaal in ons land verblijven, etc. Bij een onderzoekvraagstelling vanwege de overheid zou dat misschien nog aanvaardbaar kunnen zijn; bij de onderhavige onderzoekvraagstelling vanwege de NOS (met haar taakstelling) uiteraard juist niet. Dàt er beperkingen t.a.v. het representativiteits-'ideaal' gesteld moesten worden, en derhalve prioriteitskeuzen moesten worden gemaakt, kwam door redenen van financiële en organisatorische (maar uiteindelijk natuurlijk in wezen óók van financiële) aard. De daadwerkelijke samenstelling van volledige adressen-bestanden van buitenlandse werknemers – en vooral van sommige nationaliteiten –, zowel als de trekking daaruit, bleek zo'n gecompliceerde en arbeidsintensieve taak, dat het tegen aanvaardbare kosten en in een aanvaardbaar tijdsbestek nauwelijks op landelijke schaal te organiseren viel. Vooral omdat de *kwaliteit* en de *toegankelijkheid* van de adressenbestanden van buitenlandse werknemers zo volstrekt verschillend zijn in de diverse Nederlandse gemeenten. Bij de uiteindelijke keuze van Rotterdam en Utrecht als de twee steden waar het survey-onderzoek is uitgevoerd, heeft – behalve dat sommige van de te onderzoeken nationaliteiten buitenlanders daar (vrijwel) geheel geconcentreerd zijn – zeker ook meegespeeld, dat daar in 1980/81 een behoorlijk ontwikkeld stelsel van Nederlandse ondersteuningsorganisaties en van buitenlandse zelf-organisaties aanwezig was. De steekproeven zijn namelijk zowel in Rotterdam als in Utrecht – na vergelijkend onderzoek – uiteindelijk getrokken uit adressenbestanden van diverse plaatselijke stichtingen voor buitenlandse werknemers, omdat zij a) vollediger bleken te zijn dan de gemeentelijke adressen-bestanden en b) met name in Rotterdam *méér in overeenstemming* zijn met de door de gemeente zèlf gepubliceerde demografische gegevens! Het trekken van de steekproeven is meestal met de hand en door leden van het onderzoekteam zèlf uitgevoerd. Een voorbeeld van de grote arbeidsintensiviteit van een onderzoek als het onderhavige, indien men de representativiteits-eis – althans binnen een bepaalde stad – zover mogelijk tracht door te voeren, omdat dit vanuit de aard van de vraagstelling uiterst relevant wordt geacht.

Zoals reeds gezegd, de keuze van Rotterdam en Utrecht werd ook bepaald door de concentratie of de zelfs vrijwel exclusieve aanwezigheid aldaar van de te onderzoeken nationaliteiten in ons land. De Kaapverdianen in Nederland bijvoorbeeld wonen vrijwel uitsluitend in Rotterdam. Deze eis van 'voldoende aanwezigheid' was niet alleen van belang, om daar statistisch zowel als economisch verantwoord het survey-onderzoek te kunnen uitvoeren, maar ook omdat in de onderzoek-opzet tevens besloten lag, dat de groepsgesprekken met

'doorsnee'-volwassenen en -jongeren in dezelfde lokale (deel-)populaties gevoerd moesten worden. Immers door een 'voldoende aanwezigheid' is het ook waarschijnlijker dat er voldoende mogelijkheden bij instellingen en op scholen zijn, om de opzet van groepsgesprekken uitvoerbaar te maken.

2.1.1.b. *Steekproefdesign*

Uitgangspunt was dat pèr nationaliteit een netto-steekproef van n = 200 personen van 16 jaar en ouder zou worden bereikt. De totale non-respons door adressen-mutaties (een aanzienlijke faktor, zoals ons bekend was!), weigering en onbereikbaarheid van bewoners van adressen werd geschat op gemiddeld 50% over de verschillende nationaliteiten. Daarom diende het design in principe een a-select getrokken bruto steekproef van n = 300 personen per nationaliteit te omvatten. Ofwel, bij ondervraging van 1 persoon per adres/huishouden: een bruto steekproef van n = 300 adressen.

Tabel 1. Steekproef design

		n personen van 15 jaar en ouder in de populatie*	verlangde bruto n van personen = adressensteekproef
R'dam	Spanjaarden	2508	300
	Italianen	678	x
	Joegoslaven	3083	300
	Portugezen	2424	150
	Kaapverdianen		150
Utrecht	Turken	2770	150 x + y = 300
	Marokkanen	2471	150
	Grieken	750	150
	Italianen**		y

* volgens demografische gegevens gemeenten per 1-1-1979.

** uiteindelijk is, lopende het onderzoek, nog een extra steekproef Italianen in Utrecht getrokken moeten worden. 'Er bij' trekken in Rotterdam was na de eerste selectie uit het bestand onaanvaardbaar, i.v.m. een daaruit voortvloeiende te dichte bezetting van bepaalde straten waar Italianen bijeen wonen (buren-beïnvloeding).

Het design in tabel 1 wijkt af van de direct er aan voorafgaand neergelegde uitgangspunten. Waarom?

Methodisch: besloten werd Turkse en Marokkaanse vrouwen *niet* via individuele vraaggesprekken op het huisadres te ondervragen. Het moet zowel onmogelijk als onwenselijk worden beschouwd, om te trachten in de gezinssituatie een door de echtgenoot onbeïnvloed gesprek te voeren, zoals vanuit onze (westerse) onderzoek-methodologie noodzakelijk zou zijn. Daarom is gekozen voor een benadering van vrouwen in deze Islamitische bevolkingsgroepen in het kader van les- en naaigroepen. Daar konden zij onder begeleiding en in het bijzijn van de vertrouwde begeleidster in een – kennelijk door de echtgenoot geaccepteerde situatie (maar buiten zijn aanwezigheid) de vragenlijsten invullen. In het steekproefdesign zijn daarom voor de Turken en Marokkanen alleen bruto steekproeven van N = 150 (mannelijke) personen opgenomen. Natuurlijk levert deze benaderingswijze een inbreuk op de nagestreefde represen-

tativiteit van het onderzoek op. Zij het een onvermijdelijke. Er is sprake van een afwijkende benaderingswijze als ook van een afwijkende selectie: je krijgt toch vaker met – relatief gesproken – nog wat mindere geïsoleerde vrouwen te doen, dan in een doorsnee groep vrouwen het geval zou zijn. Dat moet in de beschouwing der resultaten worden verdisconteerd.

Financieel: Uit de adressenbestanden van de Portugezen, Kaapverdianen en Grieken konden slechts steekproeven getrokken worden, die de helft van de bedoelde omvang hadden. Reden: de NOS-Dienst Radio Programma, voor wie met name het onderzoek onder deze nationaliteiten werd uitgevoerd, had niet méér financiële middelen beschikbaar.

De keuze tussen Rotterdam en Utrecht voor die nationaliteiten, die zowel in de ene als de andere stad aanwezig zijn, is gemaakt op grond van de omstandigheid in welke stad een grotere bevolkingsconcentratie aanwezig was (betreft bijv. de Grieken en – aanvankelijk – de Italianen), of in welke stad het adressen-bestand beter of completer was, dan wel in welke stad de steun die de organisaties in het kader van het *totale* onderzoek konden geven, functioneler was (betreft bijv. de Turken en Marokkanen)*.

Quotering naar geslacht: Zoals al eerder gesteld: het uitgangspunt van de a-selecte steekproeftrekking was voor dit onderzoek een belangrijke en principiële zaak. Vanwege een te verwachten volstrekt *niet* a-selecte non-respons bij deze mediterrane groepen, in de richting van een veel grotere weigering door vrouwen, is besloten toch op één kenmerk te quoteren, n.l. geslacht. Leidraad daarbij waren de demografische gegevens voor de verschillende nationaliteiten van de betreffende gemeenten. (In praktijk bleek maar al te zeer, hoe die quotering nodig was!)

2.1.2. *Survey-onderzoek èn groepsgesprekken onder jongeren*

Om praktische zowel als methodologische overwegingen is de selectie van jongeren niet ook via a-selecte steekproeftrekking uit de adressenbestanden geschied.

1. de adressenbestanden hielden géén gezinssamenstelling in, onbekend was derhalve of er jongeren op een adres aanwezig waren. Om de noodzakelijk geachte netto 100 jongeren per nationaliteit te bereiken zouden grotere bruto-steekproeven dan n = 150 getrokken moeten worden en ook arbeidsintensiever adressen moeten worden bezocht.
2. individuele vraaggesprekken met kinderen in de thuis-situatie zijn moeilijk te realiseren. Zeker met meisjes en zeker in de Islamitische groepen. Ouders zullen zich er vaak mee bemoeien of in ieder geval erbij aanwezig willen zijn, en dus het vraaggesprek beïnvloeden.

Daarom is gekozen voor een benadering van jongeren via scholen in de gemeenten waar ook de volwassenen van de betreffende nationaliteiten zijn ondervraagd. In de schoolsituatie zijn de vragenlijsten individueel, zij het in een klassikale situatie, onder begeleiding van een lid van het onderzoekteam ingevuld. *Met dezelfde groepen zijn in principe ook de groepsgesprekken gevoerd.*

* Het benaderen van Turkse en Marokkaanse vrouwen in les- en naaigroepen bleek bijvoorbeeld in Utrecht beter te organiseren.

2.1.3. *Groepsgesprekken met 'doorsnee'-personen en met 'voorhoede'-groepen*

Hoewel de uitgevoerde groepsgesprekken uiteraard juist *niet* in het kader van representativiteits-aanpak thuishoren, maar ter kwalitatieve uitdieping resp. ter aanvulling daarvan dienen, worden zij hier volledigheidshalve genoemd omdat zij in de totale onderzoek-opzet zo essentieel geacht werden en ook bleken te zijn. Zowel gezien de aard van de vraagstelling als gezien de beperktheid van een survey-onderzoek in het licht van die vraagstelling. De groepsgesprekken met doorsnee-personen, *ter uitdieping* van de survey-resultaten, zijn gevoerd met personen uit de zelfde stadspopulaties waaruit ook de personen-steekproeven voor het survey-onderzoek zijn getrokken. De gespreksgroepen werden gevormd door en/of in samenwerking met 'culturele' zelf-organisaties van de verschillende nationaliteiten in Rotterdam en Utrecht, dan wel met organisaties, die zich met vorming, onderwijs of andersoortige sociaal-culturele activiteiten bezighouden. Zeer bewust is daarbij getracht partners te vinden, die niet politiek of maatschappelijk eenzijdig waren gericht of samengesteld. Vandaar de beperking tot organisatie-partners in het sociaal-culturele veld, al moet daarbinnen ook nog scherp voor eventuele politiek-maatschappelijke eenzijdigheden gewaakt worden. Voor de gesprekken werden deelnemers aan de activiteiten van deze instellingen, zonder aanzien des persoons, uitgenodigd. Alle groepsgesprekken dienden door de leden van het onderzoek-team zelf gevoerd te worden, zonodig met tussenkomst van een tolk.

Zoals gezegd, bij het voeren van deze groepsgesprekken was representativiteit niet relevant. Wèl de noodzakelijke kwalitatieve uitdieping van de representatieve, maar beperkte gegevens. En intensiever, dáár, waar die uitdieping meer noodzakelijk was. Daarom diende er ter uitdieping van de vraaggesprekken met vrouwen in het survey-onderzoek bij elke nationaliteit tenminste ook één gesprek met *uitsluitend* vrouwen gevoerd te worden. In gemengde groepen, zo was de verwachting, zouden mannen toch zeer licht de boventoon voeren. Daardoor zouden zowel de specifieke problemen van vrouwen als het benodigde extra accent op vrouwen in de uitdieping van het survey-onderzoek onvoldoende tot hun recht komen. In de uitdieping van het survey-onderzoek moest met name het aspect van de 'wensen en verwachtingen' ten aan zien van radio en tv en programma's voor en over buitenlandse werknemers extra aandacht krijgen. Dergelijke thema's laten zich immers met name moeilijk, en slechts ten dele, in de gestandaardiseerde vraagstelling van een vragenlijst vangen.

De gesprekken met leden – vertegenwoordigers van de kaders van politieke, maatschappelijke en religieuze *eigen* organisaties in de verschillende nationaliteiten, werden vooral belangrijk geacht om het geluid – in verscheidenheid – te laten horen van organisaties die bepaalde doelstellingen m.b.t. het leven en werken van de buitenlanders in Nederland trachten te verwezenlijken. En, die op die manier ook verantwoordelijkheid dragen voor de levenssituatie van buitenlanders. Het ging niet om de uitdieping van het survey-onderzoek, maar om *aanvulling* ervan: het op een systematische wijze bij elkaar brengen van visies en wensen m.b.t. de inspanningen van radio en tv om te helpen de problemen van de buitenlanders op te lossen. Visies en wensen, die wel vaker

– weten we – maar via andere kanalen en zonder de systematiserende en objectiverende tussenkomst van onderzoekers op de bestuurs- of redactietafels van de omroepen of op het bordje van overheid en politici gelegd worden.

Aan de selectie van groepen werden de volgende eisen gesteld:

a) Gesprekken dienden te worden gevoerd met vertegenwoordigers van het kader van de grotere landelijke of koepel-organisaties, dan wel – bij het ontbreken daarvan – met een groep van voor het onderzoek bijeengebrachte vertegenwoordigers van lokale organisaties.
b) De selectie diende in principe het volledige scala van organisaties te omvatten. Onder 'maatschappelijke' organisaties werden behalve vakbonden, organisaties met sociaal-culturele en educatieve doelen en doeleinden van maatschappelijke zorg, ook organisanisaties gericht op het (school-)onderwijs begrepen.

Natuurlijk begeven onderzoekers zich met zo'n opzet op glad ijs: De onbekendheid van Nederlanders met de situatie van buitenlanders kan ook hen parten spelen. Hoe krijg je de zekerheid dat het bedoelde scala binnen de gestelde randvoorwaarden ook volledig is? De enige mogelijkheid is, je zo goed mogelijk bij vertrouwd te achten en deskundige instanties te laten inlichten.

Besloten werd hiertoe alle kanalen en contacten, die in de loop van het onderzoek her en der in Nederland waren aangeboord te gebruiken en de aldus verkregen opgave van eigen organisaties tot slot nog eens te checken tegen een zo volledig en zo betrouwbaar mogelijke administratie, als die van het Nederlands Centrum voor Buitenlanders (NCB).

Ook deze gesprekken dienden door de leden van het team te worden gevoerd, zonodig met begeleiding van een tolk. De samenstelling van de groep gesprekspartners was uiteraard voor de verantwoordelijkheid van de uitgenodigde organisatie. De uitnodiging gold het (leidinggevend) kader van de betreffende groeperingen. Bij de uitnodiging werd de organisatie in kennis gesteld van de gespreksonderwerpen, opdat er van een voorbereide gesprekssituatie sprake kon zijn.

2.2. *De vragenlijst/de leidraad voor de groepsgesprekken*

2.2.1. *Opzet en indeling*

Na een zeer uitvoerige voorbereiding, niet alleen om in overleg met betrokkenen de vraagstelling te kunnen formuleren, maar juist ook om via tientallen gesprekken meer kennis te verkrijgen van culturele eigenheden van de te onderzoeken groepen en de mogelijke consequenties daarvan voor de te verwachten interviewsituatie bij de verschillende categorieën respondenten (mannen, vrouwen, jongeren, politieke of regionale sub-groepen, taal-groepen, analfabeten, etc.), is uiteindelijk één – meest uitgebreide, althans meest gedetailleerde – vragenlijst ontworpen. Daarop konden voor specifieke groepen en situaties zonodig aanpassingen worden gemaakt. Bijvoorbeeld:

– voor de Turkse en Marokkaanse vrouwen zijn bepaalde onderdelen weggelaten. Dit is besloten omdat de vragenlijst erg uitgebreid was en hier en daar toch te gecompliceerd voor deze beide categorieën.
– voor de vragen omtrent de persoonlijke en de gezinssituatie, het verblijf in

het thuisland en in Nederland, werd voor mannen en vrouwen een aparte lijst ontwikkeld,
- aan de jongeren is een lijst voorgelegd, die hier en daar minder gedetailleerd was dan de 'moeder'-lijst, maar bijvoorbeeld uitgebreider t.a.v. een specifiek onderwerp als het radioprogramma voor buitenlandse jongeren etc.

De totale voorbereiding en opstelling van de vragenlijst heeft het onderzoekteam ca. 4 à 5 maanden gekost.

Uiteindelijk omvatte de vragenlijst 11 onderdelen ('hoofdstukken') t.w. 1. Algemene personalia, 2. Kijkgedrag, meningen etc. Nederlands tv-aanbod, 3. Kijkgedrag, meningen etc. 'eigen' Paspoort-programma's, andere Paspoort-programma's, 4. Luistergedrag, meningen etc. Nederlands radio-aanbod, thuislandradio, 5. Luistergedrag, meningen etc. 'eigen' NOS-programma's voor buitenlandse werknemers, 6. Interesses en wensen m.b.t. mogelijke onderwerpen in radio- of tv-programma's voor de eigen nationaliteit, 7. Sociale contacten n.a.v. radio- en tv-programma's voor buitenlanders, 8. Gedrag, meningen etc. gedrukte en overige media, 9. Meer specifieke persoonlijke gegevens, 10. Ruimte voor spontane opmerkingen/'boodschappen' aan de programmamakers, 11. (in te vullen door de enquêteur) Gegevens m.b.t. het verloop van het vraaggesprek, de interview-situatie, waarneembare sociaalmaatschappelijke achtergrond van de ondervraagde.

De hoofdstuk-indeling van de vragenlijst lag ook ten grondslag aan de leidraad voor de groepsgesprekken. Deze leidraad omvatte daarom, voor zowel de 'doorsnee'- als de 'voorhoede'-groepsgesprekken, de volgende thema's:
– Kijk- en luistergedrag en meningen m.b.t. de Nederlandse tv en radio, – Kijkgedrag en meningen m.b.t de eigen 'Paspoort'-programma's, – Luistergedrag en meningen m.b.t. de eigen radiprogramma's, – Wensen en verwachtingen m.b.t. de Nederlandse tv en radio, in het bijzonder m.b.t. programma's voor en over buitenlandse werknemers, – Beschrijving van de eigen levenssituatie, c.q. van de eigen groep.

2.2.2. *De voertaal*

De vragenlijst is uiteraard in het Nederlands ontworpen en daarna in de 7 talen, die in de 8 te ondervragen nationaliteiten gesproken, c.q. verstaan worden, vertaald. (Voor de Portugezen en Kaapverdianen werd van één, in het Portugees gestelde, vragenlijst uitgegaan.) Met de Nederlandse concept-lijst zijn een 20-tal proefgesprekken gevoerd door de onderzoekteamleden met buitenlanders in de verschillende nationaliteiten, die de Nederlandse taal machtig waren. Op grond van deze ervaringen zijn vele bijstellingen gepleegd, alvorens tot vertaling werd overgegaan.

De vragenlijst omvatte zowel open als gesloten (geprecodeerde) vragen. In totaal 200. De door de enquêteur te stellen vraag was in de betreffende taal gesteld; de eventuele geprecodeerde antwoorden in het Nederlands, i.v.m. de latere scoringen, (her)-codering en verponsing door Nederlanders. De antwoorden op de open vragen dienden daartoe eveneens in het Nederlands genoteerd te worden. Dit betekende dus de noodzaak van 2-talige enquêteurs.

2.3. *Werving en training van enquêteurs*

2.3.1. *De uitgangpunten en de praktijk van de werving*

In een verslag van een onderzoek onder een bevolking(-'s groep) uit onze Westerse samenleving zal aan dit onderwerp niet vaak een volwassen paragraaf gewijd hoeven worden. Hier is dat wel het geval, omdat de identiteit van de enquêteur, gegeven de taal-problemen, de culturele verschillen tussen de onderzochte nationaliteiten onderling en de verschillen van de mediterrane cultuur met de 'cultuur van de onderzoeker', zo bijzonder belangrijk is in het licht van een interviewsituatie, die gestandaardiseerd moet zijn èn betrouwbare resultaten moet opleveren. Daarom werden vooraf een aantal uitgangspunten gesteld:

a) De voertaal in de vraaggesprekken is de eigen taal van de respondent: het gesprek wordt zonder tussenkomst van een tolk gevoerd. De culturele en intellectuele kloof tussen onderzoeker en respondent wordt daarmee nog zo klein mogelijk gehouden.
b) De enquêteurs dienen op zijn minst 2-talig te zijn, d.w.z. de betreffende taal zowel als Nederlands te spreken. T.a.v. de keuze om enquêteurs uit de 8 verschillende nationaliteiten aan te trekken, die Nederlands spreken, dan wel Nederlanders, die de betreffende talen spreken, werd uiteindelijk voor de laatste mogelijkheid de voorkeur uitgesproken. Doorslag gaf, ondanks een wat grotere afstand tot de respondent, dat interviewer en verantwoordelijk onderzoeker dan in ieder geval de zelfde cultuur delen, waardoor het beter mogelijk is om een methodisch verantwoorde en gestandaardiseerde interview-uitvoering te bewerkstelligen. Zoals reeds gezegd: mediterrane culturen verdragen zich minder goed met de 'cultuur' van het objectiverende vraag- èn antwoord-gedrag, dat voor een vraaggesprek in het kader van sociaal-wetenschappelijk onderzoek noodzakelijk wordt geacht. Bovendien is bij het interviewen door een landgenoot-enquêteur de kans niet denkbeeldig, dat de politieke, maatschappelijke of religieuze identiteit van de interviewer een rol gaat spelen. Hetgeen voor de Nederlandse onderzoek-leiding nauwelijks te controleren is. Het selecteren van zoveel mogelijk werkelijk neutrale buitenlandse interviewers is moeilijk en evenzeer voor de onderzoeker niet altijd controleerbaar.
c) Vrouwelijke respondenten dienen door vrouwelijke enquêteurs te worden ondervraagd. Voor sommige van de te onderzoeken nationaliteiten moet dit als een absolute voorwaarde worden gesteld. Om reden van een noodzakelijk gestandaardiseerde respondent-benadering diende deze voorwaarde derhalve algemeen gesteld te worden. Praktisch betekende dit dat per nationaliteit minstens 50% van de aan te trekken enquêteurs, vrouwen dienden te zijn.

Op grond van de ervaring met de proefgesprekken (gemiddelde gespreksduur 1½ à 2 uur), werd uitgegaan van 10 uit te voeren gesprekken per enquêteur. In totaal dienden derhalve voor de geplande aantallen gesprekken 110 enquêteurs, verspreid over 8 taalgroepen te worden aangetrokken. Aanvankelijk werd gekozen voor een werving van Nederlandse enquêteurs via universitaire instituten voor de diverse relevante talen, via particuliere studiecentra en migrantencentra met de betreffende talen sprekende Nederlanders in dienst.

Tabel 2. Samenstelling enquêteurs-korps per nationaliteit van de doelgroep, geslacht en landsaard van de enquêteur.

Nationaliteit doelgroep	geslacht enquêteurs		landsaard enquêteurs		Totaal	
	vrouw	man	nederl.	buitenl.	reëel	gewenst
Joegoslaven	12	4	5	11	16	20
Italianen	6	3	8	1	9	20
Turken	–	8	–	8	8	10
Marokkanen	–	6	2	4	6	10
Portugezen	3	6	2	7	9	10
Spanjaarden	14	6	14	6	20	20
Kaapverdianen	5	5	1	9	10	10
Grieken	6	4	8	2	10	10
Totaal	46	42	40	48	88	110

Toen dat na zeer intensief zoeken niet volledig lukte, moest ook de werving van personen uit de diverse verlangde buitenlandse nationaliteiten aanvaardbaar worden geacht. Daarbij is overigens getracht toch vooral via universitaire instituten te werven, omdat geacht werd dat de noodzaak van een (sociaal-wetenschappelijke) methodische aanpak van de vraaggesprekken bij deze categorie personen het best duidelijk zou zijn te maken en te verzekeren.

Tabel 2 laat zien hoe het resultaat van de enige maanden vergende werving – waarbij tientallen instanties en tussenpersonen moesten worden ingeschakeld – zich verhoudt tot de geschetste uitgangspunten.

De verlangde aantallen enquêteurs werden, ondanks alle inspanningen, niet altijd bereikt. De niet gestandaardiseerde samenstelling van de enquêteursploeg naar 'Nederlanderschap' of 'buitenlanderschap' van de betreffende personen moest – zoals gezegd – worden geaccepteerd. De verlangde 2-taligheid werd in alle gevallen wèl bereikt.

De samenstellen naar sekse van de enquêteurs voldeed, behalve bij de Portugezen, eveneens en zelfs ruimschoots aan de gestelde eis.

De Nederlandse enquêteurs waren in het gros van de gevallen doctoraalstudenten aan de in de arm genomen universitaire instituten. Ook de buitenlandse enquêteurs waren vaak studenten aan de R.U. Utrecht en de Erasmusuniversiteit (bijv. via de bureaus voor buitenlandse studenten en via persoonlijke relaties van reeds aangeworven nederlandse enquêteurs, die vaak hun taalkennis al in activiteiten met buitenlanders in de praktijk brengen). Ook via de stichtingen voor buitenlandse werknemers, via de eigen culturele verenigingen van buitenlanders, zijn tenslotte enquêteurs aangeworven. De vereiste 2-taligheid was daarbij eis (en voldeed ook als norm voor de overdraagbaarheid van onderzoekprincipes, etc. tijdens de instructie en de training). Veel moeite werd door het onderzoekteam gegeven aan het zich vergewissen van een zo groot mogelijke 'neutraliteit' van de buitenlandse enquêteurs.

2.3.2. *Instructie, training en begeleiding*

De 8 aangeworven ploegen enquêteurs hebben ieder apart (dus pèr nationaliteit) in Hilversum een instructie ontvangen. Daarbij werden achtergronden en

inhoud van de vragenlijst doorgenomen, werd een vraaggesprek geoefend en werden instructies m.b.t. benadering en selectie van personen op de geselecteerde adressen gegeven. Elke enquêteur kreeg een vast aantal te bezoeken adressen mee. Elke ploeg kwam, na het uitvoeren van 2 interviews door iedere enquêteur, opnieuw met een lid van het team bijeen. Daarbij werden ervaringen uitgewisseld, instructies herzien of aangescherpt en werden alle uitgevoerde interviews volledig met degene die ze afgenomen had, doorgenomen. Vrijwel 100% van de ingevulde vragenlijsten kon – eventueel na verduidelijking en aanvulling door de enquêteur – worden geaccepteerd als verwerkbare resultaten. Als leerproces voor enquêteurs en begeleiders is deze tweede instructie bijzonder nuttig en zinvol geweest. Ook in de periode daarna zijn de teamleden vooral bij sommige nationaliteiten (Turken, Marokkanen, Joegoslaven en Kaapverdianen), nog zeer intensief met een persoonlijke begeleiding van enquêteurs bezig geweest. Menig vragenlijst is de dag na invulling nog persoonlijk door teamleden met de betreffende enquêteur doorgesproken. De sociale contacten, die annex hiermee opbloeiden, bleken soms dringend nodig om de 'moed er bij de enquêteurs (in enkele probleem-nationaliteiten) in te houden'. Dit gold met name de Turken, Marokkanen en Kaapverdianen. Ook hebben de enquêteursploegen elkaar onderling vaak nog ondersteund en geholpen.

2.3.3. *Onderlinge afstemming van enquêteurs en steekproeven: invloed op de representativiteit*

Op zichzelf is het verstrekken van bijvoorbeeld 15 adressen aan iedere enquêteur met de opdracht 10 gesprekken te realiseren, door op 1 adres 1 random te kiezen persoon van 16 jaar en ouder te ondervragen, waarbij vrouwen alleen door vrouwelijke enquêteurs worden ondervraagd, al ingewikkeld. Bovendien was, zoals aangegeven in par. 2.1.1.b, in de steekproefopzet gequoteerd op het kenmerk geslacht, i.v.m. de te verwachten grote non-respons onder vrouwen.

In de praktijk is daarom aan elke enquêteur en enquêtrice in elke proeg opgegeven hoeveel mannen en hoeveel vrouwen hij of zij diende te interviewen (mannen interviewden uiteraard alleen mannen). Door het onmiddellijk instu-

Tabel 3. De bereikte streekproeven naar sekse, absoluut en procentueel

		Joegosl.	Ital.	Turk.	Marok.	Port.	Spanj.	Kaapv.	Griek.	Totaal
		n	n	n	n	n	n	n	n	n
man		66	41	86	38	64	104	23	14	436
vrouw		82	27	14*	26*	33	67	9	32	290
Totaal	n	148	68	110	64	97	171	32	46	726
		%	%	%	%	%		%	%	%
man		45	60	86	59	66	61	72	30	60
vrouw		55	40	14	41	34	39	28	70	40
Totaal	%	100	100	100	100	100	100	100	100	100

* ondervraagd in een groepssituatie.

ren naar de 'centraalpost' van *elke* ingevulde vragenlijst, hetgeen vanuit controle en correctie-oogpunt toch al noodzakelijk was, kon worden bijgehouden in hoeverre de quota gehaald werden. Er trad inderdaad een veel grotere non-respons onder vrouwen op, hetgeen werd gecompenseerd door op een bepaald moment vrouwelijke enquêteurs op te dragen, uitsluitend nog (random te selecteren) vrouwen op de te bezoeken adressen te ondervragen. Het moment van ingrijpen lag voor de verschillende nationaliteiten niet gelijk, omdat ook de mate van non-respons onder de vrouwen verschillend was. Doordat de verdeling van adressen per enquêteurs-ploeg zodanig was geweest, dat iedere enquêteur een 'eigen wijk' had, leverde dit geen belangrijke vertekening van de man-vrouw-verdeling naar de spreiding over de betreffende stad op. Toch is de quotering naar geslacht, alsmede de wijze waarop die getracht is te realiseren, uiteraard van invloed op het principe van de a-selecte trekking en dus op de representativiteit van de steekproef voor de populatie.

3. DE REALISATIE VAN HET SURVEY-ONDERZOEK ONDER VOLWASSENEN

3.1. *De bereikte steekproeven en de non-respons*

1. De gewenste aantallen ondervraagden per nationaliteit van n = 200 zijn slechts benáderd. En dan nog in zeer wisselende mate: variërend van 20% 'succes' bij de Kaapverdianen tot 84% bij de Spanjaarden. Omdat de bruto steekproeven soms tot boven n = 300/n = 150 moesten worden uitgebreid, hangt dit mede samen met de geringe aantallen aanwezige adressen (= gezinnen/huishoudens) in een bepaalde populatie in de gekozen stad. Deze beperking heeft zich vooral bij de Italianen, Kaapverdianen en Grieken doen voelen. De Italianen zijn uiteindelijk zelfs in 2 steden ondervraagd. De kwantitatieve omvang van de non-respons is niet automatisch 100% minus het percentage 'succes', i.v.m. een wisselende mate van aanspreken en zelfs uitbreiden van de bruto steekproefomvang om aan de gewenste nettoomvang te komen. Daarom wordt hierna in tabel 4 apart verslag gedaan van de opgetreden non-respons.
2. De verdeling man-vrouw binnen elke nationaliteit is geen, op basis van a-selectiviteit tot stand gekomen, representatieve verdeling der seksen. Nochthans kan in zijn totaliteit van een redelijke mate van a-selectiviteit van de

Tabel 4. Non-respons bij 5 onderzochte nationaliteiten

Nationaliteit	% gesl. interviews	% niet kloppend adres/ verhuisd met onbekende bestemming/nooit thuis	% weigering
Joegoslaven	53%	39%	8%
Italianen	51%	37%	12%
Marokkanen	38%	33%	29%
Spanjaarden	57%	33%	10%
Kaapverdianen	32%	43%	25%

Bovengenoemde percentages zijn betrokken op de per nationaliteit in de praktijk zeer wisselende aantallen uitgezette én bezochte adressen.

tot stand gekomen steekproeven gesproken worden. Er is derhalve sprake van een onder de gegeven omstandigheden 'best haalbare' representativiteit van de onderzoek resultaten.

3. De kolom Totaal staat uitsluitend voor het geheel van alle ondervraagden in dit onderzoek. Deze totaal-groep is 'nergens' representatief voor. Zeker niet voor het totaal van de buitenlandse werknemers in Nederland, of zelfs maar in Utrecht of in Rotterdam.

Door onvolkomenheden in de administratie kan alleen een verantwoorde weergave van de opgetreden non-respons worden geleverd voor vijf van de acht nationaliteiten. Van de gepleegde contacten met Turkse, Griekse en Portugese adressen is door de enquêteurs helaas tè gebrekkig verslag gedaan, om daar een betrouwbaar inzicht in te krijgen.

Oorzaken voor non-respons, die te maken hebben met fouten of achterhaaldheden in het adressen-bestand waren kwantitatief gesproken de belangrijksten. Een adres werd bijvoorbeeld na 3× bezocht te zijn, en niemand die voor een interview in aanmerking kwam, te hebben aangetroffen, als 'niet thuis' aangemerkt. De percentages weigering tot medewerking bevinden zich op een redelijk acceptabel niveau. Het wat hogere percentage onder de Marokkanen en Kaapverdianen is, gegeven de in paragraaf 2.3.2. gememoreerde ervaringen van de enquêteurs, zeker verklaarbaar. Vanuit die zelfde ervaringen van enquêteurs kan worden gesteld, dat de respons-problemen onder de Marokkanen ook het grootst van allemaal geweest zijn. Ook groter dan bijv. onder de Turken, waarvan hier helaas geen kwantitatieve verantwoording is te geven. Het hoge non-respons percentage onder de Kaapverdianen heeft in werkelijkheid wat lager gelegen omdat op de aangetroffen adressen nog al eens geen Kaapverdianen, maar Portugezen bleken te wonen. Dezelfde fouten, maar dan andersom, zaten in het Portugese adressenbestand. De aldus onjuist gebleken adressen zijn toch weer gebruikt, maar dan in de andere – juiste – steekproef. Per saldo is derhalve het non-respons percentage in tabel 4 in negatieve zin 'geflatteerd'.

3.2. *De uitgevoerde individuele vraaggesprekken*

3.2.1. *Arbeidsproduktiviteit der enquêteurs*

Op grond van tabel 2 kan men berekenen dat 52% van de enquêteurs vrouwen waren en 48% mannen. Evenzo dat 55% buitenlander was en 45% Nederlander. Van de uitgevoerde 726 vraaggesprekken blijken maar weinig daar van afwijkende percentages interviews door vrouwen, respectievelijk mannen te zijn uitgevoerd als hun aandeel in het enquêteurs-korps bedroeg. Idem wat betreft de percentages door Nederlanders en buitenlanders gerealiseerde gesprekken. Met andere woorden: alle onderscheiden categorieën enquêteurs hadden bij benadering een gelijke effectieve arbeidsproduktiviteit.

Dit resultaat betekent dat de opdracht aan de mannelijke en vrouwelijke enquêteurs, ondanks alle complicaties van andere instructies etc. vrijwel volledig is geslaagd.

3.2.2. *De gerealiseerde afstemming van de sekse van de interviewer op die van de ondervraagde*

Tabel 5. Gerealiseerde afstemming sekse enquêteur en sekse van de ondervraagde (totale aantal ondervraagden)

Geïnterviewd door:	mannelijke resp.	vrouwelijke resp.	Totaal
enquêteurs	63%	12%	43%
enquêtrices	37%	88%	57%
Totaal	100	100	100
(n=)	(440)	(286)	(726)

3.2.3. *De interviewsituatie*

Deze paragraaf is geheel gebaseerd op de door de enquêteurs aan het eind van de vragenlijst gegeven beoordelingen van het zojuist door hen met een respondent gevoerde gesprek.

Het blijkt dat in alle nationaliteiten, behalve de Italianen, in 9 of meer van de 10 gevallen het interview metterdaad in de taal van de ondervraagde is gevoerd. In 1 of minder van de 10 gevallen was de voertaal Nederlands. Dat bij de Italianen de voertaal in 4 op de 10 keren Nederlands was, vindt zijn verklaring in het grote percentage Nederlandse echtgenotes van Italianen, die tot de steekproef 'Italianen' zijn gerekend.

De opzet om de taal van de ondervraagde te gebruiken mag derhalve geslaagd genoemd worden. Naar het oordeel van de enquêteurs konden – en dat hangt daar uiteraard sterk mee samen – ondervraagde en interviewer elkaar ook in vrijwel alle gevallen goed begrijpen en verstaan:

Tabel 6. Kon U elkaar goed verstaan en begrijpen?

	Joegosl.	Ital.	Turk.	Marok.	Port.	Spanj.	Kaapv.	Grie.	Totaal
Goed	98	98	88	80	98	96	91	95	95
niet zo goed	1	2	11	20	2	4	9	5	5
slecht	1	–	1	–	–	–	–	–	0
Totaal	100%*	100%	100%	100%	100%	100%	100%	100%	100%

* 100% = 100% van alle gevoerde vraaggesprekken.

Alleen onder de Marokkanen en Turken is het percentage wat lager. Uit de toelichtingen van de Marokkaanse enquêteurs, blijkt dit vooral te zijn opgetreden als de ondervraagden Berbers waren (een taalbarrière binnen één nationaliteit dus!). Ook wantrouwen ten opzichte van de bedoelingen van de enquêteur speelden er een rol in.

Het interview blijkt vaak in aanwezigheid van nog andere personen dan de respondent te zijn geweest. De percentages van de gevallen waarin dat het geval was, variëren in de verschillende nationaliteiten van minimaal 67% (Kaapverdianen) tot maximaal 84% (Grieken).

Het aantal personen, dat erbij aanwezig was, varieerde zeer sterk per natio-

naliteit. Meestal betrof het 1 of 2 personen, maar bij de Turken en Marokkanen waren respectievelijk in 80% en 60% van de gevallen 3, 4 of zelfs méér personen aanwezig. Dat laatste hangt natuurlijk sterk samen met de woon- en leefsituatie van de ondervraagden in de betreffende nationaliteit. Bij de Turken betreft het overigens, net als bij de andere nationaliteiten, overwegend gezinsleden. Bij de Marokkanen waren 'de meerdere anderen' het vaakst medepension-bewoners!

Belangrijk is vooral echter hun invloed op het gesprek:

Tabel 7. In welke mate bepaalden 'anderen' mede de inhoud van de antwoorden? (Allleen de gevallen, dat er anderen aanwezig waren)

	Joegosl.	Ital.	Turken	Marok.	Port.	Spanj.	Kaapv.	Griek.	Totaal
niet	57%	60%	25%	34%	52%	55%	53%	57%	50%
gering	34%	29%	71%	44%	38%	37%	40%	34%	41%
sterk	9%	10%	5%	22%	9%	8%	7%	9%	9%
Totaal	100%	100%	100%	100%	100%	100%	100%	100%	100%

Aangenomen mag worden dat invloed van de omstanders, in ieder geval bij de interviews met de Islamitische respondenten, in nogal wat gevallen aanwezig is geweest. Zo'n groeps-invloed kan uiteraard zijn weerslag hebben op de representativiteit van resultaten, omdat er een soort 'cluster'-effect van uitgaat. Anderzijds staat deze groepsbeïnvloeding wel voor een stukje werkelijkheid in de mediterrane culturen, die bij de respondent misschien een natuurlijker antwoord-houding toelaat, dan het geval zou zijn, indien hij in zijn eentje door een 'buitenstaander' over 'van alles en nog wat' ondervraagd wordt. Het zijn bovendien onvermijdelijke condities waaronder zo'n onderzoek plaatsvindt. Het is echter goed ze te (onder-)kennen. (Anders lag het bij de te verwachten aanwezigheid van ouders bij thuis te houden interviews met jongeren. Dié was aanleiding een andere onderzoekmethodiek te kiezen. De beïnvloeding zou de bedoeling van een *apart* (!) jongeren-onderzoek dan immers in de kern ondergraven hebben.)

De medewerking van de ondervraagden werd over het algemeen door de enquêteurs positief beoordeeld. Onder de Kaapverdianen, Turken en Marokkanen ligt het percentage in dit opzicht positief beoordeelde vraaggesprekken echter slechts rond de 60%. De in par. 2.3.2. gememoreerde noodzakelijke persoonlijke begeleiding van de enquêteurs door de leden van het onderzoekteam, hangt hier duidelijk mee samen.

De meest gehoorde, door de enquêteurs doorgegeven kritiek van de ondervraagden op het onderzoek, betrof de lengte van het vraaggesprek en het 'nog eens het radio-deel van de lijst moeten doen' als het televisie-gedeelte was afgewerkt. De kritiek op de lengte van de vragenlijst is niet zo onverklaarbaar als we de volgende tabel in ogenschouw nemen.

Behalve de reële lengte van de vragenlijst zullen ook de verschillende levensstijlen een rol gespeeld hebben; de kortere vragenlijsten voor de Portugezen en Kaapverdianen leidden ook tot kortere vraaggesprekken, voor de Grie-

Tabel 8. Duur van de gevoerde vraaggesprekken.

	Joeg.	Ital.	Turk.	Marok.	Port.	Spanj.	Kaapv.	Griek.	Totaal
1 uur of minder	28%	38%	31%	10%	68%	55%	89%	22%	42%
1 – 1½ uur	48%	32%	36%	31%	28%	26%	11%	41%	34%
1½ – 2 uur	12%	12%	20%	21%	3%	14%	–	24%	13%
2 uur of meer	10%	16%	13%	38%	1%	5%	–	13%	10%
Totaal	100%	100%	100%	100%	100%	100%	100%	100%	100%

ken heeft dat zeker niet zo uitgewerkt. De kroon wordt overigens gespannen door de Marokkanen: in bijna 4 op de 10 gevallen duurde het gesprek 2 uur of meer. Het zal niet toevallig zijn dat de Marokkaanse enquêteurs ook verhoudingsgewijs het meest over moeilijkheden tijdens de gesprekken hebben gerapporteerd. Anderzijds heeft de specifiek bij de Marokkanen vaak voorkomende pension-situatie er duidelijk ook mee te maken.

4. EEN SAMENVATTEND OORDEEL OVER DE RESULTATEN-VERZAMELING

Op grond van het in de paragrafen 1 en 2 gestelde, lijkt het juist een al hiervoor gemaakte kwalificatie te herhalen: het is de 'best haalbare' resultatenverzameling, die gegeven de beschikbare tijd, de beschikbare financiën (*totale* onderzoek-budget, incl. verwerking: ca. f 135.000,—*) en de samengestelde vragenlijst, voor het onderzoek-team mogelijk was.

Belangrijk voor het ondanks de noodzaak van voortdurende bijstellingen en improvisatie tòch merendeels geslaagde 'vasthouden' aan het gekozen onderzoek-design, waren zeker ook de omvang, de samenstelling èn de 'pioniersgeest' van het onderzoek-team. Het team was klein (4 personen), waardoor er korte communicatielijnen bestonden en dus de discipline, die zo nodig was om trouw te blijven aan de gekozen opzet, gemeenschappelijk kon worden aangehouden. De spreiding van disciplines in het team, n.l. een cultureel antropologe (kennis van culturen en levenswijzen), een agoge en een maatschappelijk werkster (kennis van groepsprocessen, ervaring met mensen en instanties benaderen) en een sociaal-psycholoog (kennis van onderzoekmethoden en -technieken), waarvan de meesten ook werkelijk met buitenlanders hadden gewerkt, bleek eigenlijk zonder meer noodzakelijk. De bereidheid van de teamleden om in de begeleiding van enquêteurs, het verkrijgen van medewerking van onnoemelijk veel personen en instanties, de eigen tijd en persoonlijkheid zodanig in te zetten, dat er *vertrouwen* ontstond, was van onschatbare waarde voor het slagen van zowel selectieprocedures als de individuele en groeps-interviews zèlf. En als zodanig cruciaal voor de *kwaliteit* van het onderzoek.

* Voor dit bedrag is tevens een survey-onderzoek onder 325 jongeren in de 8 nationaliteiten uitgevoerd en zijn 17 discussies met 'doorsnee-groepen' volwassenen, 40 discussies met groepen jongeren en 27 discussies met de kaders van eigen organisaties gehouden en verwerkt.

Het verstrekken van een routine-opdracht aan – geeft niet welk – onderzoek-bureau, is iets dat toekomstige opdrachtgevers van onderzoek onder buitenlandse werknemers, zich voorlopig nog niet kunnen permitteren. Zoals het eventueel gekozen onderzoek-bureau zich voorlopig nog niet moet voorstellen dat een routine-uitvoering van zo'n onderzoek mogelijk zou zijn.

3. Het gebruik van im-/exportstatistieken

A.J. OLIVIER EN M.J.C. ZUURHOUT

1. INLEIDING

Binnen marketing-organisaties worden jaarplannen gemaakt. Die jaarplannen zijn dan gebaseerd op de meest recente ontwikkelingen en de lange termijn ondernemingsdoelstelling en resulteren in een zeer concrete tonnage- of eenhedendoelstelling per produkt. Deze doelstelling wordt vanzelfsprekend onderbouwd met een concreet marketing-plan, een geformuleerde marketing-mix en een geakkoordeerd marketing-budget. Voert men dit plan uit, dan zal er behoefte zijn aan een mechanisme dat controleert of de aan het begin van de planningsperiode geformuleerde doelen bereikt gaan worden.

Daar waar het gaat om internationale marketing-operaties zal men bij voorkeur gebruik moeten maken van een im-/exportmonitor. Wij verstaan hieronder een gestandaardiseerde onderzoekmethode waarin zorgvuldig gedefinieerde kengetallen volgens een strak tijdschema gemeten worden.[1]

In dit artikel wordt nader ingegaan op de specifieke toepassingen van monitoren bij de formulering van het internationale marketing-beleid.

Er wordt vooral ingegaan op de systematische vertekeningen die er onderkend kunnen worden in het door de overheid ter beschikking gestelde cijfermateriaal in de vorm van invoer- en uitvoerstatistieken.[2]

2. DEFINIËRING VAN TOTAALMARKTEN

Het in beeld brengen van de eigen omzetten en het relateren van de eigen omzetten aan de totaalmarkt is een eerste stap om te komen tot een oordeel over de mate waarin men in staat geweest is de in het marketing-plan geformuleerde doelstellingen te bereiken. Het vaststellen van de eigen leveringen aan derden is een eenvoudige zaak. Dit in tegenstelling tot het definiëren en kwantificeren van een totaalmarkt.

In de wijze waarop totaalmarkten gekwantificeerd kunnen worden is een drietal niveaus te onderscheiden.[3]

1. aanbodsniveau
2. handelsniveau
3. vraagniveau.

2.1. *De totaalmarkt op aanbodsniveau*

Een schatting van de totaalmarkt dient men – waar mogelijk – te baseren op

de goederen die door de gezamenlijke fabrikanten met succes op de markt aangeboden worden. Er zijn hierbij twee situaties denkbaar:
a. de fabrikanten wisselen gegevens uit
b. de totaalmarkt kan gereconstrueerd worden

Ad a. - De uitwisseling van gegevens

De concurrerende fabrikanten hebben overeenstemming bereikt over de uitwisseling van afzetcijfers. Een onafhankelijke instelling neemt de afzetcijfers in ontvangst en levert de deelnemers de optelsom van alle afzetten. Deze onafhankelijke instelling is veelal een accountantskantoor, een studiecentrum of een bedrijfsvereniging. Situaties waarin concurrenten toch samenwerken zijn het resultaat van verstandig management.

Ook in situaties waarin niet alle concurrenten deelnemen kunnen soms acceptabele schattingen gemaakt worden.

Deze methode brengt de werkelijke afzet op directe wijze in beeld, hetgeen zal resulteren in relatief nauwkeurige schattingen. Desalniettemin dient de gebruiker zich terdege te realiseren dat ook in deze opgaven fouten kunnen sluipen. Wij noemen een aantal in de praktijk geconstateerde tekortkomingen:
- de fabrikant 'vergeet' een gedeelte van zijn afleveringen op te geven
- de fabrikant definieert de markt anders en geeft bepaalde leveringen systematisch niet op
- de fabrikant geeft de omzet van bepaalde dochterondernemingen niet op
- de fabrikant hanteert een andere periode.

Ad b. - De reconstructie van de totaalmarkt

De totaalmarkt op basis van het vraagniveau kan ook geconstateerd worden door produktiestatistieken te combineren met im- en exportcijfers.

De totaalmarkt ('apparant consumption')[3] wordt gedefinieerd als

Verkochte produktie + Invoer – Uitvoer

2.2. *De totaalmarkt op handelsniveau*

Door gebruik te maken van bijvoorbeeld een Nielsen detaillistenpanel kan men totaalmarkt-gegevens verkrijgen. Omdat bij dit type meetinstrument het steekproefkader de doelgroeppopulatie niet volledig afdekt, ontstaat er een structureel dekkingsverschil (sample bias).

Zolang men het bestaan van dit verschijnsel maar onderkent kunnen de uitkomsten van dit type meetinstrument desondanks zonder al te veel schroom gebruikt worden, temeer daar het er toch voornamelijk om gaat te bezien hoe het eigen marktaandeel of het marktaandeel van een bepaalde concurrent zich ontwikkelt. Aanname hierbij is dan, dat de ontwikkelingen in de afgedekte delen van de doelgroeppopulatie gelijk zijn aan de ontwikkelingen in het niet afgedekte deel van de doelgroeppopulatie. Dit neemt niet weg dat over- en onderdekking ook kan ontstaan op het niveau van het marktaandeel. Dit laatste moet veel ernstiger opgevat worden, omdat dit wel een ernstige bedreiging kan zijn van de betrouwbaarheid van de gegevens.[4, 5]

Is deze verhouding aanwezig, dan kan de totaalmarkt als volgt ingeschat worden.

Wij gaan uit van het volgende:

$$\frac{S_{wt}}{S_{nt}} = \frac{T_{wt}}{T_{nt}} \tag{1}$$

dan is:

$$T_{wt} = \frac{S_{wt}}{S_{nt}} \times T_{nt} \tag{2}$$

S_{wt} = werkelijke verkopen van de fabrikant in periode t,
S_{nt} = verkopen zoals dit door Nielsen gerapporteerd wordt in periode t,
T_{nt} = de werkelijke totaalmarkt in periode t.
T_{wt} = de werkelijke totaalmarkt in periode t.

2.3. *De totaalmarkt op vraagniveau*

Indien de totaalmarkt niet op de hierboven beschreven wijze in beeld gebracht wordt, dan resteren in principe drie alternatieven:
1. gebruik maken van de gegevens uit een consumentenpanel
2. gebruik maken van enquêtes
3. gebruik maken van indicatoren waarmee de totaalmarkt bepaald kan worden

Ad 1.
De praktijk leert dat men op markten van 'fast moving consumer goods' die met grote regelmaat door consumenten gekocht worden, een nauwkeurige totaalmarkt-schatting kan baseren op een consumentenpanel.

Er is echter een zekere risico verbonden aan het feit dat de nauwkeurigheid terugloopt wanneer een produkt minder frequent gekocht wordt. Een consumentenpanel is dan ook niet geschikt voor het schatten van totaalmarkten van niet-duurzame consumptiegoederen.

Ad 2.
In dat geval resteert geen andere totaalmarkt-schattingsmethode dan de consumentenenquête. De consumentenenquête zal moeten vaststellen hoeveel procent van de ondervraagden een produkt gekocht heeft en hoe lang het geleden is dat deze aankoop werd verricht. Omdat het hier gaat om subjectieve oordelen zal na empirische validatie subjectief door middel van correctiefactoren bijgestuurd moeten worden.

Ad 3.
Op industriële markten werkt ook dit minder goed en kan men beter terugvallen op strak gedefinieerde formele relaties, waarbij men de totaalmarkt schat aan de hand van een aantal exogene variabelen.

In dit artikel wordt ingegaan op de problemen en de mogelijkheden die bedrijven hebben die zich voor hun totaalmarkt-schattingen moeten baseren op de 'apparant consumption' benadering.

Deze bedrijven stuiten op een tweetal betrouwbaarheidsproblemen:

a. de betrouwbaarheid van de binnenlandse produktiestatistieken
b. de betrouwbaarheid van de im- en exportstatistieken

Ad a.

Produktiestatistieken in Nederland 'dekken' bijna nooit de totale branche. In vele branches bevindt zich het merendeel der bedrijven in de categorie met minder dan 10 werknemers, terwijl het CBS rapporteert over bedrijven met 10 of meer werknemers. Veelal wordt de meest gedetailleerde informatie verstrekt voor bedrijven met 50 of meer werknemers. Daarnaast worden de produktiecijfers binnen een bepaalde branche soms onvoldoende gedetailleerd weergegeven. Er zijn branches waarvoor alleen Benelux-produktiecijfers bekend zijn (rubberverwerkende industrie) waaruit de Nederlandse produktie al helemaal niet kan worden gedestilleerd. Produktiestatistieken vormen dan ook een geheel afzonderlijk probleem, deels overigens ook door wettelijke (beschermings)maatregelen veroorzaakt. De betrouwbaarheid van de binnenlandse produktiestatistieken is dan ook op z'n minst discutabel te noemen.

Ad b.

Ook de betrouwbaarheid van de im- en exportstatistieken laat in een belangrijk aantal gevallen te wensen over. Er is een aantal systematische vertekeningen in dit cijfermateriaal aanwezig en in dit artikel wordt in de volgende hoofdstukken ingegaan op de soorten bias die deze vertekeningen veroorzaken.

Een goede kennis van de wijze waarop men deze statistieken kan gebruiken stelt de manager in staat de totaalmarkt beter te schatten. Daarnaast verkrijgt hij de mogelijkheid om de exportmarkten beter in beeld te brengen en beter vast te stellen welke exportmarkt voor hem aantrekkelijk is en welke niet. Voor een nadere verhandeling over de totaalmarktschattingen die gebaseerd zijn op consumentenpanels of retail audits wordt verwezen naar eerdere publikaties.[2]

3. BRONNEN VOOR BIAS IN IM-/EXPORTGEGEVENS

3.1. *Registratie bias*

De overheid heeft zich belast met het registreren van de internationale goederenbeweging.

De registratie van de in- en uitvoergegevens vindt in principe plaats aan de grens en is in wezen een bijprodukt van de in- en uitklaringsprocedures.

Deze procedures waren vroeger noodzakelijk in verband met de betaling van invoerrechten en vinden nu plaats in het kader van de vrijstelling voor de op het produkt rustende verplichting tot BTW-betaling (het zogenaamde 0-tarief).

De exporterende onderneming moet aan de hand van boeken en bescheiden kunnen aantonen dat de goederen het land verlaten hebben.

Dit kan gebeuren aan de hand van een uitvoerdocument of aan de hand van een betalingsbewijs van de buitenlandse importeur via een buitenlandse bank.

Door de voortschrijdende liberalisatie van het handelsverkeer blijven grensregistraties steeds meer achterwege en moet steeds meer een beroep gedaan worden op de vrijwillige medewerking van de im- en exporteurs. Dit heeft de volgende consequenties:

1. Sommige importeurs doen geen of een gedeeltelijke opgave.
2. Sommige importeurs hebben de neiging de posten minder verfijnd te detailleren; dit ter beperking van de hoeveelheid werk.
3. Door gebrek aan ervaring worden er fouten gemaakt in de nomenclatuur.
4. De gegevens komen daarnaast later beschikbaar en de kans op fouten wordt groter.
5. Fouten worden soms door het CBS achterhaald en zeer regelmatig vinden achteraf rectificaties plaats.
6. Doordat een fabrikant/exporteur gebaat is bij een zo klein mogelijke administratieve belasting, wordt er veelal weinig acht geslagen op een juiste detaillering van de nomenclatuur.

3.2. *Bestemmingsbias*

Het is dan ook onterecht om te veronderstellen dat de invoerstatistieken van het ene land overeenkomen met de uitvoerstatistieken van het andere land.

Dat kan geïllustreerd worden aan de hand van onderstaande tabel:

Tabel 1. De vergelijkbaarheid van importen en exporten.

Nomenclatuur 1.702.600	Karamel sucres et melasses, caramelises. Zucker und melassen karamelisiert.	
	West-Duitsland importeert uit Frankrijk	Frankrijk exporteert naar West-Duitsland
1977	160.000 kg	125.000 kg
1978	151.000 kg	135.000 kg
1979	96.000 kg	80.000 kg
1980	103.000 kg	78.000 kg

Bron: Officiële invoerstatistiek Duitsland en uitvoerstatistiek Frankrijk.

De opgaven verschillen indien beide landen met elkaar vergeleken worden. Daarnaast wijken de trendontwikkelingen van elkaar af.

Voor deze opmerkelijke verschijnselen is in zijn algemeenheid een drietal oorzaken aan te wijzen:

1. Verschillen in registratie-tijdstippen ten gevolge van opslag in douane-entrepôts.
2. Afwijkingen tussen het land van opgegeven bestemming en land van feitelijke bestemming.
   Tijdens de opslag van goederen in het douane-entrepôt of tijdens het transport kunnen goederen worden doorverkocht en daarmee een andere bestemming in een ander land krijgen.
   Deze transacties worden meestal niet gecorrigeerd in de exportstatistieken

en vertekenen de werkelijkheid derhalve.
3. Verschillen in registratietijdstip.

3.3. *Nomenclatuur bias*

Nomenclaturen zorgen voor een overzichtelijke weergave van de statistische gegevens die de overheid verzamelt. Men is bij het inschatten van totaalmarkten afhankelijk van de nomenclaturen die de overheid gebruikt. Door uit te gaan van die nomenclaturen van de statistieken van de buitenlandse handel werkt men met een zorgvuldig gedefinieerd kengetal omdat precies omschreven is wat onder een bepaald nummer valt.

Uit de statistiek blijkt:
- de hoeveelheid waarin gemeten wordt
- de waarde waarin gemeten wordt
- de periode waarover gemeten wordt.

In theorie zouden er dus geen fouten gemaakt kunnen worden bij de bepaling onder welk nomenclatuurnummer het produkt valt. Men dient er echter rekening mee te houden dat in de praktijk verschillende interpretaties ten aanzien van onder welk nomenclatuurnummer het produkt hoort mogelijk zijn. Dit geldt met name voor produkten waarbij het niet absoluut duidelijk is dat ze onder een bepaald nomenclatuurnummer vallen zoals bij nieuwe produkten bijvoorbeeld kunsthars. De ene douanebeambte die een formulier invult kan namelijk een andere categorie kiezen dan zijn collega.

In Europa zijn de statistieken gestandaardiseerd volgens de CCCN (Customs Cooperation Council Nomenclature). Voor deze afkorting schrijft men ook weleens BTN (Brusselse Tarief Nomenclatuur) omdat de BTN bestaat uit de circa 1.100 viercijferige zogenaamde Brusselse tariefposten. In de EEG werkt men met de Nimexe-goederennomenclatuur. Deze is ontwikkeld uit de CCCN. De Nimexe-goederennomenclatuur is opgesteld door het Bureau voor de Statistiek der Europese Gemeenschappen in samenwerking met de nationale statistische diensten van de lidstaten, ten behoeve van de statistieken van de buitenlandse handel. Het aantal Nimexe-posten bedraagt ongeveer 7.300 en bestaat uit zes cijfers. De lidstaten van de EEG zijn sinds 1 januari 1978 verplicht de Nimexe-goederennomenclatuur toe te passen waardoor een vergelijking van de buitenlandse handel van de lidstaten goed mogelijk is. Voor nationale doeleinden kan iedere post van de Nimexe verder worden onderverdeeld. Zo heeft Nederland zijn statistiekposten met één cijfer uitgebreid waardoor het totaal op zeven komt. De Verenigde Naties hebben de goederen ingedeeld volgens de SITC (Standard International Trade Classification). Deze is te vergelijken met de BTN.

Hoewel de internationale standaardisatie steeds verder doorgevoerd wordt, is het nog steeds mogelijk dat de definities van produkten en produktvariëteiten van land tot land verschillen. Naarmate de verfijning groter wordt is de kans op misverstand groter.

3.4. *Verwerkingsbias*

Voor veel produkten zijn im- en exportcijfers verkrijgbaar. Toch zijn niet alle

goederen die de grens overgaan terug te vinden in de statistieken van de buitenlandse handel. Dit komt zowel vanwege geheimhouding als vanwege het niet bereiken van de publikatielimiet. Hieronder wordt dit nader toegelicht.

3.4.1. *Geheimhouding*

Als het aantal exporteurs gering is of als er sprake is van één exporteur, is het mogelijk dat de omvang van de individuele exporteur af te leiden is uit de officiële uitvoerstatistiek. Door de bekendmaking van die cijfers zou de concurrentie te weten kunnen komen welke transacties met het buitenland door een bepaalde onderneming zijn gedaan. De belangen van een industriële onderneming zouden dan in ernstige mate kunnen worden geschaad. In die gevallen kan men vragen om niet te worden opgenomen in de uitvoerstatistiek.

3.4.2. *Publikatielimiet*

Alleen de douane-aangiften die een bepaalde waarde en hoeveelheid te boven gaan worden gesommeerd. Men hanteert dus een verwerkingsdrempel. Voor het werken met een im-/exportmonitor is dit gegeven niet zo van belang om daar rekening mee te houden omdat het om zeer kleine hoeveelheden gaat van geringe waarden.

In 1981 vielen 126 invoerposten en 366 uitvoerposten (2% respectievelijk 7% van de totale waarde) onder de geheimhoudingsregeling. Het ligt voor de hand dat het CBS er naar streeft het totaal van de posten die onder de geheimhoudingsregeling vallen tot een minimum te beperken.

De hierboven beschreven vormen van bias behoren alle tot de registratie/verwerkingsbias.

Alleen door zorgvuldige vergelijking in de tijd en cross-sectie-analyse zal men de tekortkomingen van de im- en exportgegevens ontdekken en deze waar nodig kunnen corrigeren. Het belang van monitoren blijft eveneens duidelijk. Immers, alleen door de data gedurende lange tijd te verzamelen en een aantal bronnen met elkaar te vergelijken leert men de tekortkomingen kennen van het cijfermateriaal.

4. DE PRIJZENREGISTRATIE

In de importstatistieken worden niet alleen hoeveelheden gemeten, maar stelt men ook de waarde vast. Dit laatste kengetal kan de manager, bij een juist gebruik, waardevolle informatie geven over de exportmarkten. Alvorens in te gaan op het nut van deze informatie is het goed allereerst stil te staan bij de wijze waarop deze gegevens geregistreerd worden.

4.1. *Exportprijsstellingen*

Bij uitvoer van produkten gaat men uit van de wereldmarktprijs van de goederen op het moment dat deze goederen de grens passeren.

Exportprijs:

'De waarde bij uitvoer is gelijk aan de waarde, die de goederen hebben op

de plaats en het tijdstip waarop ze *de grens passeren* (f.o.b.-waarde). Daarbij gaat men uit van de normale prijs, dat wil zeggen de prijs, die op het moment van uitvoer kan worden bedongen ingevolge een onder voorwaarden van vrije mededinging tussen twee onafhankelijke partijen tot stand gekomen koop en verkoop.'

Definitie: Bron statistiek van de buitenlandse handel.

In een belangrijk aantal gevallen vindt er geen levering f.o.b.-grens plaats en zal de prijs ingeschat moeten worden.

Primair gebeurt dit door de verlader van de goederen en wordt zijn opgave gecontroleerd door de douanebeambte. In laatste instantie wordt de opgave opnieuw gecontroleerd door de computerprogramma's van het CBS door vergelijking van de opgave met cijfers uit het verleden. Het is duidelijk dat hier een aantal subjectieve waardeoordelen aanleiding kan zijn tot een vertekening van de werkelijkheid.

Worden op een uitgevoerd produkt export-subsidies verstrekt, dan worden deze subsidies in mindering gebracht op de geëxporteerde waarde en hierdoor wordt de prijs per stuk of per kg omlaag gebracht.

4.2. *Importprijsstellingen*

Indien een goed geïmporteerd wordt, wordt niet de f.o.b.-grens prijs gehanteerd, maar hanteert men de c.i.f.-prijs.

Deze prijs hoeft in de meeste gevallen niet ingeschat te worden, maar kan vanuit de papieren die de goederen begeleiden vrij objectief vastgesteld worden.

Er wordt immers òf op c.i.f.-condities geleverd òf de verlader heeft een aanvullende nota waarop zijn specifieke transportkosten vermeld staan.

Importprijs:

'De normale prijs van de ingevoerde goederen bevat alle op de goederen drukkende kosten tot aan de plaats van binnenkomst in het douanegebied, met uitzondering van de in het binnenland verschuldigde belastingen en heffingen (c.i.f.). Voor de goederen herkomstig uit niet-EEG-landen is dit de waarde c.i.f. buitengrens van de Gemeenschap, voor overige goederen de waarde c.i.f.-grens van *Nederland*.'

Definitie: Bron statistiek van de buitenlandse handel.

Indien men een schatting wenst te maken van het prijspeil waarop in een exportmarkt zaken gedaan worden, kan men zich beter baseren op de importstatistieken van het importerende land en niet op de eigen exportstatistieken. Dat men zich bij voorkeur dient te verzekeren van een inzicht in beide statistieken teneinde een zo volledig mogelijk totaalbeeld te verkrijgen en op deze wijze tekortkomingen sneller te onderkennen behoeft in het licht van het voorafgaande geen nadere toelichting.

Wanneer het gaat om de betaalde prijs per kg zijn er dus systematische ver-

schillen tussen importstatistieken en exportstatistieken. Ter illustratie volgen hieronder de prijzen per kg van de in West-Duitsland geïmporteerde Karamel.

Tabel 2. Vergelijking tussen exportprijzen en importprijzen. Jaar 1979.

Ingevoerd vanuit:	Prijsniveau van geïmporteerd produkt	Prijsniveau van de corresponderende exportstatistiek van het betreffende land	Prijsverschil
	prijs per kg in hfl.	prijs per kg in hfl.	
Frankrijk	2,66	2,55	– 11 ct
België + Luxemburg	1,42	1,46	+ 4 ct
Nederland	3,50	3,36	– 14 ct
Italië	1,28	1,10	– 18 ct
United Kingdom	2,17	2,09	– 8 ct

Bron: Officiële invoer- en uitvoerstatistieken.

Het feit dat bij de importcijfers c.i.f.-prijzen en bij de exportcijfers f.o.b.-prijzen geregistreerd worden kan de vastgestelde prijsverschillen onmogelijk volledig verklaren.

De systematiek in de prijsverschillen ontbreekt immers. De meeste importprijzen zijn weliswaar hoger dan de exportprijzen, maar niet allemaal (bijvoorbeeld België en Luxemburg).

Een tweede bron van afwijking is de valuta waarin gehandeld wordt. Veelal is dit de valuta van het exporterende land. Soms ook de valuta van het importerende land of de valuta van een derde land (bijvoorbeeld Zwitserse franken of U.S.-dollars).

De statistieken van de diverse landen kennen alleen maar de eigen valuta. En er vindt derhalve te allen tijde op zijn minst aan één van beide zijden van de grens een valuta-omrekening plaats.

Deze omrekening wordt uitgevoerd door de douanebeambte die op het importformulier de waarde in de valuta van het importerende land aangeeft.

In situaties waarin hevige koersschommelingen plaatsvinden of de-/revaluaties kan het niet tijdig aanpassen van de koersen aanleiding geven tot afwijkingen tussen de gemiddeld betaalde prijs zoals het importerende land dat vaststelt en de wijze waarop het exporterende land dat vaststelt.

De prijsverschillen tussen im- en exportstatistiek blijken bij volgtijdelijke analyse niet stabiel te zijn maar ook van jaar tot jaar sterk te kunnen verschillen. Dit betekent dat men niet alleen bij cross-sectie-analyse maar ook bij longitudinale analyses rekening moet houden met een vertekening van de gebruikte cijfers.

Bij tijdreeks-analyses van individuele in geld gewaardeerde markten ontstaan twee additionele problemen. De eerste is het fluctueren van wisselkoersen, vooral in die situaties waarin twee met elkaar vergeleken valuta's ten opzichte van het goud ieder voor zich wezenlijk andere posities ingenomen hebben.

Het tweede probleem ontstaat wanneer twee landen in de tijd met elkaar

vergeleken worden en in deze twee landen doet zich een wezenlijk andere inflatie voor.
Weliswaar zullen de wisselkoersen tussen beide valuta aangepast worden, maar hierbij kunnen zich aanmerkelijke time-lags voordoen die volgtijdelijke analyse bemoeilijken.

Het herleiden van geldwaarden naar Europese rekeneenheden lost het hierboven geschetste probleem van de koersverschillen bij een cross-sectie-analyse op en wanneer wij er daarnaast nog toe overgaan de Europese rekeneenheden terug te herleiden tot constante prijzen met peiljaar A winnen ook de vergelijkingen in de tijd aanmerkelijk in waarde.

Hier staat dan tegenover dat de prijzen die uitgedrukt zijn in de Europese rekeneenheid voor veel exporterende en importerende managers slechts een abstracte waarde hebben.

5. HET GEBRUIK VAN IM-/EXPORTGEGEVENS

5.1. *Algemeen*

Indien wij er, door systematische analyse, in slagen om de mogelijke vertekeningen van het cijfermateriaal in beeld te brengen, ontstaat een belangrijk meetinstrument, dat ons in de allereerste plaats in staat stelt om de apparant consumption vast te stellen.

Daarnaast kan dit materiaal gebruikt worden om te bepalen welke exportmarkten interessant zijn en tegen welke prijsniveaus de produkten aangeboden moeten worden om kans van slagen te hebben op een exportmarkt.

Wij zullen deze laatste onderwerpen in de navolgende subparagrafen bespreken.

5.2. *De potentiële exportmarkten*

In veel situaties is de omvang van de exportmarkt onbekend. Dit is meestal een gevolg van de omstandigheid dat men niet over panelgegevens beschikt, of dat de produktiestatistieken niet voldoende verfijnd uitgesplitst zijn of soms helemaal niet beschikbaar zijn.

Men kan in deze situaties gebruik maken van gespecialiseerde adviesbureaus die via hun relaties in het desbetreffende land kennis hebben van de omvang van een totaalmarkt of een 'best guess' kunnen afgeven.

Een tweede methode om na te gaan welke exportmarkten interessant kunnen zijn is het vaststellen van importoverschotten. Men gaat er dan vanuit dat het niet zozeer gaat om de vraag hoe groot de vraag in een bepaald land is, maar om de vraag in welke mate een land in staat is in de eigen behoefte te voorzien met de in het land aanwezige produktiecapaciteit.

Hoe groter het importoverschot, des te groter de mogelijkheid voor een buitenlandse firma om toe te treden tot een lokale markt.

In een land met een exportoverschot is zelfs een dermate grote produktiecapaciteit aanwezig, dat voor het buitenland geproduceerd kan worden en toetreden tot een dergelijke markt is heel wat minder aantrekkelijk.

Ter illustratie geven wij u het voorbeeld van de Duitse en de Italiaanse karamel im- en exporten.

Uit de im-/exportcijfers van Italië voor het produkt karamel blijkt het volgende:

Tabel 3. Karamel in Italië.

× 1.000 kg	1977	1978	1979	1980
export	295	710	990	1160
import	141	112	77	26
exportprijs	f 1,19	f 1,11	f 1,15	f 1,24
importprijs	f 1,88	f 1,79	f 2,23	f 2,62
exportoverschot	154	598	913	1134

Gezien deze cijfers zou men kunnen concluderen dat er in Italië geproduceerd wordt tegen lage prijzen en dat een toetredende exporteur/fabrikant weinig kans zal krijgen om een deel van de markt te veroveren.

Tabel 4. Karamel in Duitsland.

× 1.000 kg	1977	1978	1979	1980
export	487	344	223	226
import	1.246	1.867	1.941	2.557
exportprijs	f 1,82	f 2,19	f 2,63	f 2,84
importprijs	f 1,87	f 1,56	f 1,66	f 1,71
importoverschot	759	1.523	1.718	2.331

Indien men de im- en exportcijfers van West-Duitsland met die van Italië gaat vergelijken, blijkt dat men beter vanuit Nederland naar West-Duitsland kan gaan exporteren dan naar Italië.

In West-Duitsland is namelijk sprake van een importoverschot van 2.331.100 kilogram in 1980, terwijl dit importoverschot eveneens aanzienlijke stijging vertoont. Daarentegen is in Italië sprake van een exportoverschot dat bovendien aanzienlijk groeit.

Men zou dus, uitgaande van deze cijfers, kunnen concluderen dat export vanuit Nederland naar West-Duitsland meer kans op succes biedt dan naar Italië. Ook kan uit deze cijfers geconcludeerd worden dat het makkelijker zal zijn om de Italiaanse prijzen in West-Duitsland te onderbieden dan de Italiaanse prijzen in Italië zelf.

Meestal is het wenselijk naast de im- en exportcijfers tevens de produktiecijfers te hebben, zoals reeds vermeld in paragraaf 5.2. In het geval van de karamel in West-Duitsland weet men dat er een toename is van het importoverschot. Daaruit mag echter niet geconcludeerd worden, dat de relatieve concurrentiepositie van de Duitse karamel producerende industrie duidelijk verminderd is. Je weet in deze situatie namelijk niet of het groeiende importoverschot in West-Duitsland een gevolg is van een toename in het verbruik van karamel of een gevolg is van de vermindering van de Duitse produktie. Door het verzame-

len van produktiecijfers kan men een antwoord op deze vraag vinden. Toch mag men uit deze cijfers wel het vermoeden uitspreken dat er goede kansen liggen voor export naar West-Duitsland.

5.3. *Prijsanalyse*

Is eenmaal vastgesteld dat de Duitse markt aantrekkelijk is, dan kan op basis van de beschikbaar zijnde importstatistieken een prijsfrequentieverdeling per land opgesteld worden.

Tabel 5. 0.04 Bondsrepubliek Duitsland
1.702.600 Zucker u. Melassen karamelisiert
Import 1980

Land	Prijs per kg in f	× 1.000 kg
Ierland	1,14	216
Italië	1,24	818
België/Luxemburg	1,46	557
U.S.A.	1,74	328
Engeland	2,55	439
Frankrijk	3,02	103
Nederland	3,29	73

De Italianen domineren de importmarkt met scherpe prijzen. Ierland is in 1980 voor het eerst op de Duitse markt aanwezig.

Desalniettemin blijft er boven dit prijsniveau ook nog voldoende volume over waarbij met name het geringe aanbod in de prijsklasse ± 1,75 - ± 2,55 opvalt. Wel is duidelijk dat het Nederlandse prijsniveau dermate hoog is dat men zichzelf grotendeels uit de markt geprijsd heeft.

Een confrontatie van deze prijsfrequentieverdeling met de kostprijs en gewenste winstbijdrage geeft een inzicht in de haalbaarheid van de exportplannen van een karamelproducerende fabrikant naar Duitsland.

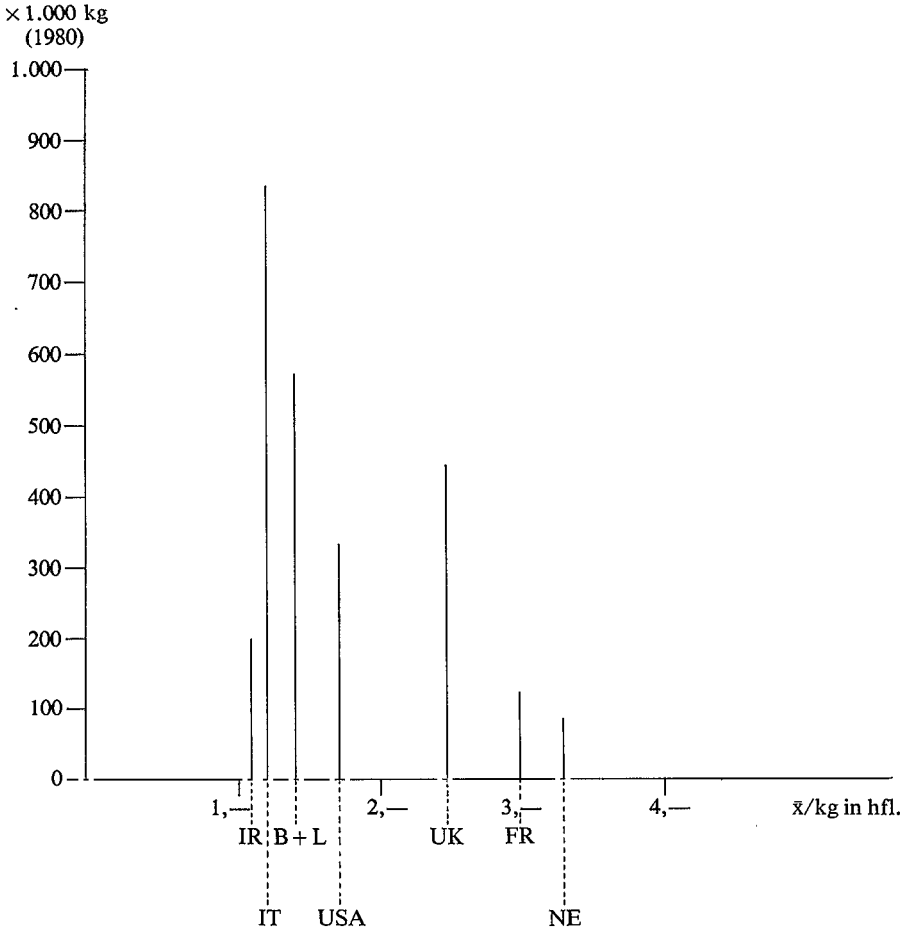
6. CONCLUSIE

Het bepalen van de totaalmarkt is een belangrijk onderdeel van een marketingplan. Hierbij gaat het niet alleen om het bepalen van de thuismarkt, maar ook om het bepalen van de omvang van de exportmarkten.

In dit artikel werd ingegaan op de vertekeningen die zich kunnen voordoen bij importcijfers en exportcijfers. Hoewel de vertekeningen op zichzelf genomen aanzienlijk zijn, blijkt het dat men nuttige informatie uit im-/exportstatistieken kan ontlenen.

Men dient dan wel meerdere cijferbronnen met elkaar te vergelijken omdat men niet kan volstaan met een eenmalige meting.

De gegevens moeten permanent in beeld gebracht worden teneinde onregelmatigheden zo snel mogelijk op te sporen.



BIBLIOGRAFIE

1. Leeflang, P.S.H. en Olivier, A.J. 1981. 'Market Intelligence Monitoren', *Tijdschrift voor marketing* (Juni), pp. 10-15.
2. Kersten, H.M.P. en Bethlehem, J.G. 1982. 'Foutenbronnen in Steekproefonderzoek', *Jaarboek van de Nederlandse Vereniging van Marktonderzoekers,* pp. 44-65.
3. Zwan, A. v.d. en Verhulp, J. 1980. *Grondslagen en techniek van de marktanalyse,* H.E. Stenfert Kroese b.v., Leiden, pp. 98-127.
4. Leeflang, P.S.H. en Olivier, A.J. 1980. *'What is wrong with the Audit Data we use for Decision Making in Marketing',* Proceedings 33rd Esomar Congress, pp. 219-239.
5. Leeflang, P.S.H. en Olivier, A.J. 1980. *'Facing Panel non-response, consequences and solutions',* Proceedings 35th Esomar Congress, pp. 417-439.

4. EIM-bestedingenindex

een nationaal onderzoek naar uitgaven van personen en huishoudens

J.B. AMIABEL EN H.H.F. GALAMA

SAMENVATTING
Na een stukje ontstaansgeschiedenis en een beschrijving van de doelstellingen, de aanwendingsmogelijkheden en de organisatorische constructie van de EIM-Bestedingenindex, worden in dit artikel enkele technische aspecten van het project en van het gehouden proefonderzoek belicht. Het artikel wordt afgesloten met een blik op de uitgangspunten voor een verdere ontwikkeling van het project.
Uitgangspunt van dit artikel is het proefonderzoek (n = 200 huishoudens), dat in juni 1983 voltooid werd en waarvan op korte termijn wordt besloten of het onderzoek als EIM-project zal worden gecontinueerd en uitgebreid.

1. WAAROM EEN EIM-BESTEDINGENINDEX?

1.1. *Signalen omtrent de behoefte aan een consumentenpanel*

Reeds in november 1977 heeft de Commissie Research Midden- en Kleinbedrijf [1] in haar aan het Ministerie van Economische Zaken uitgebrachte rapport, zowel voor het meso-onderzoek als ten behoeve van het macro-onderzoek, aangedrongen op een snelle start van een continu consumentenonderzoek. Dit met name ter verhoging van het inzicht in de veranderingen van de vraag van de consument en veranderingen in de bestedingen naar distributiekanalen en distributievormen. Dit ten behoeve van de wens om door middel van extrapolatie te komen tot prognoses.

Begin 1981 werd door het Ministerie van Economische Zaken (E.Z.) een werkgroep ingesteld, met als doel het verzamelen van gegevens, die gebruikt kunnen worden voor het doen van voorspellingen omtrent de ontwikkeling van diverse categorieën van consumptieve bestedingen. Behalve E.Z. waren in deze werkgroep vertegenwoordigd, het Centraal Planbureau (C.P.B.), het Centraal Bureau voor de Statistiek (C.B.S.) en het Economisch Instituut voor het Midden- en Kleinbedrijf (E.I.M.).

Als activiteiten van deze werkgroep kunnen worden genoemd een analyse van tijdreeksen 1970 tot en met 1978 van Eurostat-statistieken betreffende de EG-landen en het toevoegen van enkele vraagstellingen aan de consumenten conjunctuuronderzoeken (C.C.O.) van het C.B.S.

Goed bruikbare, met name gedifferentieerde en actuele gegevens om aan de doelstelling van de werkgroep te voldoen, bleken niet voorhanden.

1.2. *De informatiebehoefte bij het EIM*

Afgezien van het voorgaande, leeft bij het EIM als onderzoekinstituut en bij

haar zusterinstelling, het Centraal Instituut voor het Midden- en Kleinbedrijf (CIMK), ter realisering van hun doelstellingen – economisch en sociaal onderzoek respectievelijk voorlichting en advies ten behoeve van het midden- en kleinbedrijf – voor tal van onderzoekterreinen sinds jaar en dag de behoefte aan actuele cijfers inzake de relatieve verdeling van uitgaven over bestedingsdoelen naar distributiekanalen en distributievormen. Dus primair is er de behoefte aan een goed inzicht in de ontwikkeling van de marktaandelen. Daarnaast bestaat de behoefte aan goede informatie voor het verrichten van studies naar huidig en voor de toekomst te verwachten koopgedrag en winkelkeuze en naar bestedingen in relatie tot socio-economische kenmerken en geografische criteria.

Bij de behandeling van de doelstellingen van dit project zal nader op de aangegeven informatiebehoefte worden ingegaan.

1.3. *Reeds bestaande informatiebronnen*

Alvorens werd overgegaan tot het opzetten van een nieuw instrumentarium werd uiteraard nagegaan in hoeverre bestaande informatiebronnen aan de allerwegen gesignaleerde behoefte aan informatie en onderzoekmogelijkheden zouden kunnen voldoen. Het doorlopend budgetonderzoek (D.B.O.) van het C.B.S. komt nog het meest tegemoet aan deze informatiebehoefte. Geconstateerd werd echter, dat dit onderzoek onvoldoende aansluit bij de behoefte aan actuele gegevens en bovendien onvoldoende mogelijkheden biedt voor het volgen en analyseren van ontwikkelingen binnen sectoren en deelpopulaties. Tenslotte werd vastgesteld, dat het met name ten behoeve van het midden- en kleinbedrijf (M.K.B.) gewenste inzicht in distributiepatronen en consumentengedrag – in relatie tot de bestedingen – moeilijk te realiseren is met het C.B.S.-instrumentarium.

Door het CBS en het EIM is gemeenschappelijk geconstateerd, dat er ten aanzien van beide onderzoeken kan worden gesproken van complementariteit. Besloten is tot een taakverdeling met afspraken over samenwerking en onderlinge afstemming.

De opgevangen signalen, de specifieke informatiebehoefte ten behoeve van het M.K.B. en de bereikte overeenstemming omtrent taakverdeling met het C.B.S. hebben geleid tot de 'geboorte' van het Proefprojekt Bestedingenindex bij het EIM.

2. DOELSTELLINGEN EN AANWENDINGSMOGELIJKHEDEN

2.1. *Doelstellingen bezien vanuit het EIM*

Centrale doelstellingen, die aan de opzet en inrichting van de EIM-bestedingenindex ten grondslag liggen zijn:

A. het kunnen vaststellen van de relatieve verdeling van uitgaven over doelen en distributiekanalen/-vormen en de ontwikkelingen daarin.

Dit voor het vaststellen van marktaandelen:

a. voor te onderscheiden sectoren als Midden- en Kleinbedrijf (M.K.B.), klein en middelgroot, grootfiliaalondernemingen, warenhuizen, postorderbedrijven en verbruikscoöperaties in het kader van de studie Korte en Middel-

lange Termijnontwikkeling (consequenties van de economische ontwikkelingen op de gang van zaken in het midden- en kleinbedrijf);

b. voor te onderscheiden bedrijfsgroepen (zoals: kleding/textielzaken, kruidenierszaken en bijvoorbeeld bedrijfstypen in de horeca) in het kader van bedrijfstakstudies en branche-onderzoeken;

c. voor te onderscheiden distributievormen (geparallelliseerd, discount/niet-discount, warenhuis/speciaalzaak) ten behoeve van het volgen van ontwikkelingen.

Het kunnen vaststellen van marktaandelen voor genoemde en andere categorieën is essentieel voor:

- herleiding van omzetontwikkelingen uit de ontwikkeling van de totale consumptieve bestedingen;
- inzicht in concurrentieverhoudingen, die van invloed zijn op de ontwikkeling van winstmarges die, samen met de ontwikkeling van omzet en kosten, bepalend zijn voor de ontwikkeling van het winstinkomen, dat uiteindelijk centraal staat.

B. het kunnen verrichten van studies:

- naar huidig en toekomstig koopgedrag en naar winkelkeuze
  * analyse naar de relaties tussen prijs, kwaliteit en hoeveelheid
  * effecten van inkomensontwikkeling op het koopgedrag in relatie tot aspecten van houding en gedrag
  * winkelkeuze in relatie tot kwaliteit en afstand
- naar de (ontwikkeling van) bestedingen in relatie tot kenmerken van huishoudens en personen (onder meer huishoudensgrootte, inkomen en vrij besteedbaar inkomen van huishoudens en personen, leeftijdsgroep) en geografische criteria.

2.2. *Aanwendingsmogelijkheden voor het MKB-veld en in ruimer verband*

Uit de bovenstaande, overigens nog globale, aanduiding van doelstellingen en daarvoor te verrichten studies, moge blijken dat de EIM-Bestedingenindex grote mogelijkheden biedt voor de werkterreinen – en derhalve voor de taakvervulling – van het EIM en het CIMK. Niet alleen kan worden voorzien in actuele en specifiek op probleemstellingen toegesneden informatie en kunnen basisgegevens worden verstrekt voor onder meer studies betreffende margeontwikkelingen, inkomens- en prijselasticiteiten, voor sector en branche structuuronderzoeken en voor ruimtelijk- en vestigingsplaatsonderzoek. Ook voor het beleidsonderbouwend onderzoek ten behoeve van het MKB-veld, alsmede voor ruimere beleidstoepassing door overheid, instellingen en overig bedrijfsleven biedt het instrument grote aanwendingsmogelijkheden.

Het zijn immers niet alleen de consumptieve bestedingen naar artikelgroepen en naar detailhandelskanalen, die worden geregistreerd. Het gaat bij dit onderzoek om alle uitgaven van personen en huishoudens, dus ook overdrachten/bestedingen in de sfeer van sparen, beleggen, giften, abonnementen, vakanties e.d. alsmede de min of meer vaste financiële verplichtingen als belastingen, verzekeringen, huren, hypotheek, leningen en gebruikskosten van telefoon, gas, water en elektra.

3. DE ORGANISATORISCHE EN FINANCIËLE CONSTRUCTIE

In feite wordt de bestedingenindex ontwikkeld door twee partijen, te weten het EIM en Intomart B.V. Het EIM is in de huidige constructie te typeren als opdrachtgever/coördinator, eigenaar en gebruiker van gegevensbestanden, belast met de commerciële exploitatie en de financiering. Intomart kan in dit verband worden omschreven als panelhouder/uitvoerder, belast met de ontwikkeling van het instrumentarium en het steekproefbeheerssysteem alsmede het verrichten van veldwerk, panelonderhoud, gegevensverwerking en de technische analyse en evaluatie.

Het EIM heeft dit project primair opgezet om het sociaal en economisch onderzoek en de voorlichting en advisering ten behoeve van het midden- en kleinbedrijf met een goed toegesneden informatiestroom te kunnen voeden. Ter financiering van de voortzetting en verdere uitbouw van het instrument is het noodzakelijk dat informatie wordt geleverd aan derden tegen betaling.

4. TECHNISCHE ONTWIKKELING VAN HET INSTRUMENT

4.1. *Afbakening van de doelgroep, registratie-eenheid en registratie-ingangen*

De doelstelling van het bestedingenonderzoek voor wat betreft het registratiesysteem zou technisch als volgt geformuleerd kunnen worden: het voor Nederland in kaart brengen van de uitgaven van personen en huishoudens naar bestedingsdoel, bestedingsplaats en betaalwijze. Gegeven deze doelstelling zou iedere persoon of ieder huishouden dat in Nederland uitgaven doet in de doelgroep opgenomen moeten worden. Daar van de zeer jeugdigen niet verwacht mag worden dat zij in staat zijn de beoogde registraties adequaat te kunnen verrichten, zijn personen jonger dan 12 jaar buiten de doelgroep gelaten. Verondersteld mag worden dat daarmee echter slechts een marginaal onderdeel van het totaal der bestedingen buiten de registratie zal vallen.

Daar vele bestedingen niet een persoonlijk karakter hebben, maar gedaan worden voor een geheel huishouden, is het huishouden gekozen als uitgangspunt van de doelgroep. Het onderzoek richt zich daarmee op huishoudens, maar niet in die zin, dat uitsluitend één persoon van het huishouden de registratie van bestedingen verricht. Ieder lid van het huishouden van 12 jaar of ouder registreert zelf de door hem of haar verrichte bestedingen. Voorwaarde van deelname aan het onderzoek is dat *alle* leden van het huishouden bereid zijn hun bestedingen te registreren.

Ook het begrip 'besteding' diende afgebakend en geoperationaliseerd te worden. Er is sprake van een besteding telkens wanneer een betaling plaatsvindt. Ruil in natura valt hiermede dus buiten de registratie. Daarnaast is ter voorkoming van overbelasting van het panel ook voor de te registreren bestedingen een ondergrens getrokken. Continu worden uitsluitend bestedingen van 10 gulden of meer geregistreerd. Omdat het hiermede buiten de registratie vallende deel van bestedingen met name voor specifieke markten niet onbelangrijk is, werd in de proefperiode onderzocht of het mogelijk is dat het panel voor een korte periode óók bestedingen beneden 10 gulden registreert. In paragraaf 6.2 komen wij op dit experiment terug.

Eerder in dit artikel werden reeds enkele kernbegrippen genoemd. Deze worden hier nader toegelicht. Telkenmale wanneer een panellid een betaling verricht registreert hij of zij een aantal zaken die de bestedingen karakteriseren. Geregistreerd wordt:

1. Datum van besteding
2. Bedrag van besteding
3. Betaalwijze; de panelleden beschikken over een geordend overzicht (codelijst) van een voor het onderzoek relevante indeling naar de wijze waarop men een besteding betaald kan hebben. In totaal worden 32 verschillende wijzen van betaling onderscheiden. Het panellid registreert aan de hand van een twee-cijferige code de op de besteding van toepassing zijnde betaalwijze.
4. Bestedingsplaats; de panelleden beschikken tevens over een geordend overzicht (codelijst) van een voor het onderzoek relevante indeling naar plaatsen waar (of personen en instanties aan wie) men de betaling kan hebben verricht. In totaal worden circa 450 verschillende bestedingsplaatsen onderscheiden. Dit grote aantal onderscheiden bestedingsplaatsen is niet uitsluitend ingegeven door de wens nagenoeg onbegrensde flexibiliteit te hebben in rapportagemogelijkheden door voor verschillende doeleinden juist voor die doeleinden specifiek relevante subgroepen van bestedingsplaatsen te kunnen samenstellen, het grote aantal is ook een gevolg van het feit dat aangesloten moet worden bij de terminologie waarmee panelleden bestedingsplaatsen definiëren om een betrouwbare registratie mogelijk te maken. Een met name genoemde aanduiding als Albert Heijn, Miro, Gamma, etc. sluit nauwer aan bij de belevingswereld van de panelleden dan onderzoeksbegrippen zoals discounter, warenhuis, zelfbedieningszaak of bouwmarkt.
5. Bestedingsdoel; antwoord op de vraag wat men heeft gekocht of waarvoor men heeft betaald wordt gegeven door de codering van het bestedingsdoel. Er worden 17 verschillende hoofddoelen onderscheiden, zoals bijvoorbeeld levensmiddelen, kleding, apparaten, etc. Elk van de hoofddoelen is weer onderverdeeld naar bestedingscategorieën. Zo is het hoofddoel 'kleding' weer onderverdeeld naar sportkleding, onder-/nachtkleding, herenbovenkleding, dames-bovenkleding, kinder- en babykleding, stoffen en materialen, accessoires, herstel van kleding en andere betalingen in verband met kleding. Op het niveau van bestedingscategorie worden circa 130 verschillende bestedingsdoelen onderscheiden. Ook hier geldt voor het grote aantal onderscheiden bestedingsdoelen eenzelfde argumentatie als bij punt 4 weergegeven.

4.2. *Split-run*

Hiervoor werden de algemene opzet en ingangen van het continue registratiesysteem beschreven. Deze opzet komt optimaal tegemoet aan die situaties waarbij een besteding voor één concreet doel wordt verricht. In de praktijk zijn echter situaties denkbaar waarbij met één besteding – één betaling – meerdere bestedingsdoelen worden gediend.

Met name bestedingen in supermarkten kunnen dit naar bestedingsdoel he-

terogene karakter hebben. Om te onderzoeken in welke mate aan dit probleem in het ontwerp van het registratiesysteem tegemoet gekomen kan worden, is bij de helft van het panel met een alternatief registratiesysteem geëxperimenteerd waarbij per besteding het hoofddoel van die besteding wordt aangegeven en tevens de besteding verbijzonderd wordt naar de specifieke bestedingscategorieën. Deze meer uitgebreide vorm van registratie wordt in het vervolg als 'versie-B' aangeduid, in tegenstelling tot de algemene registratievorm die wij 'versie-A' zullen noemen.

4.3. *Basislijst*

Naast de continue registratie van bestedingen werd een vragenlijst opgesteld aan de hand waarvan op ad-hoc basis van de aan het onderzoek meewerkende huishoudens een aantal gegevens vastgelegd werd. In de vragenlijst werd met name ingegaan op achtergrondgegevens, die het huishouden karakteriseren, periodieke uitgaven, woonlasten en woonsituatie, attitudes tot zaken als werkgelegenheid, sparen en eigen financiële situatie terwijl daarnaast het inkomen van alle leden van het huishouden werd vastgelegd.

Op grond van deze informatie kunnen binnen het panel afzonderlijke subgroepen onderscheiden worden, over welke groepen afzonderlijk gerapporteerd kan worden.

5. WERVING EN REPRESENTATIVITEIT

5.1. *Twee fasen procedure*

De opbouw en continuering van een representatief panel stelt bijzondere eisen aan de opzet van werving. De ervaring heeft geleerd dat bepaalde segmenten uit de bevolking sterk van elkaar verschillen ten aanzien van de mate waarin men bereid is mee te werken aan panelstudies. Een procedure waarbij men alle soorten huishoudens – segmenten – in gelijke mate benadert in de bruto steekproef, resulteert in de praktijk wegens grote responsverschillen in een panel waarbij bepaalde segmenten sterk zijn over- of ondervertegenwoordigd. Vandaar dat gekozen werd voor een werving in twee fasen.

In de eerste fase werd een basisbestand opgebouwd van een duizendtal huishoudens. Deze huishoudens zegden na een algemene uitleg van het doel van de studie in principe medewerking toe. Tevens werden bij die gelegenheid enkele belangrijke – het huishouden karakteriserende – gegevens vastgelegd.

In de tweede fase werd uit dit basisbestand een proefpanel van 200 huishoudens gerecruteerd. Werd voor de opbouw van het basisbestand de gebruikelijke procedure van a-selecte benadering gekozen, voor de werving in de tweede fase werd juist zeer select gewerkt. Voor 'moeilijke' segmenten werden meer contacten gelegd en voor 'gemakkelijke' segmenten juist minder dan gemiddeld, opdat het uiteindelijke panel ondanks responsverschillen een zo goed mogelijke afspiegeling zou vormen van de populatie.

Het spreekt voor zich dat bij een beperkte omvang – 200 huishoudens – het panel onmogelijk representatief kan zijn naar een veelheid van variabelen. Variabelen die in de proeffase speciale aandacht kregen zijn gezinsgrootte,

opleiding en inkomen van de hoofdkostwinner.

5.2. *Steekproefprocedure*

Voor de opbouw van het basisbestand werd een steekproefprocedure gevolgd die binnen de grenzen van praktische beperkingen elk huishouden een gelijke kans geeft in dit bestand opgenomen te worden. In concreto werd de volgende procedure gevolgd. Voor alle Nederlandse gemeenten werd het aantal afgiftepunten bepaald. Dit is een uitermate goede afspiegeling van het aantal huishoudens per gemeente daar in tegenstelling tot het aantal woningen per gemeente in afgiftepuntenbestanden nog onderscheid wordt gemaakt naar individuele (meerdere) woningen per adres, terwijl tevens de mogelijkheid is geboden om afgiftepunten die geen woonadres weerspiegelen (winkels, kantoren e.d.) buiten de telling te laten. Zo werd voor alle gemeenten voor heel Nederland een cumulatieve verdeling gemaakt van het aantal afgiftepunten van woonadressen. Op basis van het totale aantal afgiftepunten, de beoogde steekproefomvang en de omwille van efficiency gewenste clustering van minimaal 4 huishoudens per in de steekproef vallende gemeente werd een trekkingsfactor bepaald. Langs deze weg werd bepaald welke gemeenten in de steekproef opgenomen werden en met welk aantal te benaderen huishoudens. Aan de hand van deze informatie – bewerkt tot postcodetrajecten – trok de PTT de adressen uit haar afgiftepuntenbestand.

De getrokken adressen fungeerden als startadres. Indien op het betreffende adres geen medewerking aan de ondervraging verkregen kon worden (3× niet thuis, weigering van één of meerdere leden van het huishouden van 12 jaar of ouder) werd volgens een in de enquêteursinstructie omschreven procedure, uitgaande van dit startadres, een vervangend adres bepaald. Deze procedure herhaalde zich eventueel.

5.3. *Respons en representativiteit in de eerste fase*

In tabel 1 wordt een overzicht gegeven van de respons-ervaring bij de opbouw van het basisbestand.

Tabel 1: Responsoverzicht werving 1e fase

Resultaat van de contactpoging	abs	%
gesprek en in principe bereid tot medewerking	986	49
3× niet thuis	255	13
weigering van één of meerdere leden van het huishouden	693	34
andere reden voor non-respons, zoals ziekte e.d.	80	4
Totaal	2014	100

De helft van de contactpogingen heeft dus geresulteerd in de beoogde situatie van een kort vraaggesprek waarbij de leden van het huishouden in principe bereidheid tot medewerking aan het bestedingenonderzoek toezegden.

De verdeling van de in het basisbestand opgenomen huishoudens naar de bij

de ondervraging vastgelegde kenmerken, werd naast de bekende landelijke verdelingen naar deze kenmerken geplaatst. Zelfs voor in de praktijk relatief moeilijk tot respons te bewegen groepen, zoals alleenstaanden, ouderen, inwoners van de drie grote steden, bleek het basisbestand bijzonder goed aan te sluiten bij de landelijke verdelingen. De grote zorg besteed aan de steekproeftrekking en de introductie van het onderzoek bij de huishoudens, zal hier debet aan zijn.

5.4. *Respons en representativiteit in de 2e fase*

In tabel 2 wordt een overzicht gegeven van de responservaring bij de opbouw van het proefpanel (200 huishoudens) uit het basisbestand.

Tabel 2: Responsoverzicht werving 2e fase

	A-versie		B-versie		Totaal	
Resultaat van de contactpoging	abs	%	abs	%	abs	%
medewerking	130	64	91	42	221	53
opzegging	5	2	10	5	15	4
uitgestelde weigering	3	1	10	5	13	3
weigering	64	32	105	49	169	40
Totaal	202	100	216	100	418	100

Ruim de helft van de uit het basisbestand benaderde personen werkt mee aan het onderzoek. Uit tabel 2 blijken duidelijke verschillen in responsresultaten tussen de A-versie en de B-versie. Het laatste – meer gecompliceerde – registratiesysteem stuit op meer problemen bij de werving. Ook ten aanzien van paneluitval lijkt de B-versie grotere problemen te geven.

De voornaamste redenen voor het weigeren van medewerking in de 2e fase zijn de verwachtingen dat het registreren teveel tijd kost, te bewerkelijk of zelfs te moeilijk is. Daarnaast speelt het strakke selectiecriterium dat alle leden van het huishouden van 12 jaar en ouder medewerking moeten verlenen ook een belangrijke rol.

Ook in deze fase werd de verdeling van het proefpanel naar een aantal kenmerken naast de landelijke verdelingen gelegd. Het proefpanel blijkt ondanks de beperkte omvang een redelijke afspiegeling van Nederlandse huishoudens. Alleenstaanden zijn iets oververtegenwoordigd, terwijl ouderen, lager opgeleiden en de lagere sociale klassen iets ondervertegenwoordigd zijn. Dit hangt samen met de responservaring. Doordat vereist is dat *alle* leden van het huishouden van 12 jaar en ouder meewerken, is de kans op non-respons bij grotere huishoudens relatief groter. De lagere respons bij ouderen, lager opgeleiden en de lagere sociale klassen zal samenhangen met de mate waarin men 'onderzoek-minded' is en de mate waarin men de toch niet eenvoudige registratie denkt te kunnen uitvoeren.

6. BEVINDINGEN TEN AANZIEN VAN DE REGISTRATIE

6.1. *Visuele en computermatige controle*

De registratieformulieren worden maandelijks door de panelleden ingezonden. Voordat met de computermatige verwerking van de gegevens van de maanden januari en februari werd aangevangen, werden al deze bestedingen stuk voor stuk op juistheid van invulling gecontroleerd. Nadat de gegevens op de computer zijn ingevoerd, wordt een zeer groot aantal systematische controles uitgevoerd op waarschijnlijkheid van combinaties van de registratieingangen; brood pleegt men niet bij een benzinepomp te kopen, bij de bakker wordt niet met een acceptgirokaart afgerekend, etc.

De bevindingen van deze controles hebben ons geleerd dat registratie via de A-versie slechts bij uitzondering problemen geeft. De geconstateerde problemen werden grotendeels veroorzaakt door onvolkomenheden of onduidelijkheden in de gemaakte indelingen naar bestedingsdoel en bestedingsplaats. Deze codelijsten zijn inmiddels aangepast. Bij de registratie via de B-versie bleek daarnaast dat met name het verbijzonderen van de besteding naar bestedingscategorie voor meerdere panelleden een moeilijke zaak is.

6.2. *Bestedingen beneden tien gulden*

In paragraaf 4.1 beschreven wij de benedengrens van registratie en de reden voor deze grens. In maart werd deze grens voor de laatste twee weken losgelaten en vroegen wij de panelleden ook de bestedingen beneden 10 gulden te registreren. De bereidheid tot medewerking aan deze proef was enorm; 97% van de huishoudens was hiertoe direct bereid, terwijl het enkele huishouden dat twijfelde niet zozeer aarzelde vanwege de extra belasting, maar reeds een totale beëindiging van de registratie-activiteiten overwoog.

Het aantal geregistreerde bestedingen verdubbelde nagenoeg voor de laatste twee weken, terwijl de totale waarde van de geregistreerde bestedingen slechts met circa 7 procent toenam. Voor specifieke bestedingsdoelen zijn de bestedingen beneden 10 gulden echter van grotere betekenis; zo constateerden wij voor levensmiddelen een waardetoename van circa 18 procent.

Deze bevindingen onderschrijven de veronderstelling dat een continue registratie van óók bestedingen beneden 10 gulden tot een overbelasting van het panel zal leiden. Deze bestedingen vertegenwoordigen echter een waarde die relatief van beperkte omvang is. Een registratie van deze bestedingen voor een beperkte periode hoeft daarmee de waarde van het onderzoek niet wezenlijk te schaden. De meest ideale procedure zou die procedure zijn waarbij volgens een roulerend systeem steeds een beperkt deel van het panel ook bestedingen beneden 10 gulden registreert. Op die wijze worden ook voor deze bestedingen de seizoenspatronen en trends adequaat en snel in beeld gebracht. Gegeven het relatief grote aantal bestedingen op het waarderingstraject beneden 10 gulden kan zulks betrouwbaar geschieden bij een relatief kleinere steekproef, c.q. een deel van de totale steekproef.

6.3 *Verhoudingen tussen bestedingsdoelen*

Ter evaluatie van de validiteit van de uitkomsten van de continue registratie der bestedingen is als eerste ingang de verhouding tussen bestedingen aan verschillende hoofddoelen onderzocht.

De verhoudingscijfers, zoals wij die in andere bronnen aantreffen, betreffen steeds jaarcijfers. De verhoudingscijfers van het onderzoek zoals wij die in tabel 3 presenteren, zijn uitsluitend gebaseerd op de maanden januari tot en met maart, zodat een marginale afwijking ten gevolge van een niet constant seizoenspatroon niet uitgesloten is.

Voor een optimale vergelijking van cijfers uit verschillende bronnen is het essentieel dat de cijfers betrekking hebben op exact hetzelfde fenomeen, c.q. dat de gehanteerde definities overeenkomen. Wij hebben getracht zo nauw mogelijk aan te sluiten bij definities van andere bronnen, dat er desalniettemin discrepanties tussen definities blijven bestaan, zal evident zijn.

Tabel 3: Aandeel in totale consumptieve bestedingen in procenten

Bron	levens-middelen	kleding & schoeisel	huur en huurwaarde	gas, water & elektriciteit
Indexpanel (83)	20,2	7,2	13,5	6,6
EIM (82) [2]	19,4	7,9	–	–
D.B.O. (80) [3]	21,4	8,5	14,3	5,0
Nat. Rek. (81) [4]	19,8	8,9	10,3	5,7

In het Indexpanel werd voor de eerste drie maanden van 1983 een relatief aandeel van 20,2 procent van de totale waarde der bestedingen gemeten voor het hoofddoel 'levensmiddelen'. Verschillen met andere bronnen kunnen volledig aan verschillen in definities worden toegeschreven, zo goed als onderlinge verschillen tussen de drie aangehaalde bronnen ook daarmee samenhangen.

Voor kleding en schoeisel lijkt het in onze studie gevonden aandeel aan de lage kant te zijn. Wij brengen in deze echter de opmerking in herinnering dat de cijfers niet gecorrigeerd zijn voor seizoenspatronen en dat voor de textielbestedingen doorgaans voor de eerste drie maanden een index van 86,2 wordt gebruikt waarbij de jaarlijkse uitgaven geïndexeerd zijn op 100. Daarnaast moet opgemerkt worden dat de cijfers van D.B.O. en de Nationale Rekeningen respectievelijk gebaseerd zijn op 1980 en 1981 en dat blijkens EIM-cijfers [2] het relatieve volume van bestedingen in de sector 'kleding en schoeisel' de laatste jaren afneemt.

Voor huur en huurwaarde zijn de cijfers van DBO en Nationale Rekeningen in geval van 'eigen huis' gebaseerd op toegerekende huurwaarden, terwijl het indexcijfer strikt gebaseerd is op feitelijke lasten, waarvan hypotheeklasten in geval van eigen woning een bijzonder belangrijke component is. Wij zien dat het aandeel volgens onze definitie – die essentieel verschilt van de aangehaalde bronnen – toch redelijk in de buurt komt van studies die een andere weg volgen ter bepaling van het aandeel van dit bestedingsdoel.

Voor de post 'gas, water en elektriciteit' moet allereerst opgemerkt worden dat het D.B.O.-cijfer exclusief 'water' is. In de Nationale Rekeningen [4] zien

wij dat de post water op een aandeel van circa 0,5% geschat moet worden. Daarnaast brengen wij in herinnering dat zowel het D.B.O.-cijfer als het cijfer van de Nationale Rekeningen van enkele jaren terug is en dat het in de huidige economische situatie evident is dat uitgaven voor gas, water en elektra een relatief groter aandeel van het gezinsbudget zijn gaan opeisen, welke verschijnsel in het index-cijfer teruggevonden wordt.

6.4. *Niveau van bestedingen*

Naast de verhouding tussen bestede bedragen aan verschillende bestedingsdoeleinden is gekeken naar het niveau van de bestedingen voor een aantal bestedingsgroepen.

Een volledige dekking van 100% is het doel waarnaar bij panelonderzoek gestreefd wordt, maar wordt in de praktijk nimmer gerealiseerd. Een volledige dekking behoeft ook geen onverbiddelijke eis te zijn, daar op zeer verantwoorde en betrouwbare wijze voor een systematische onderdekking gecorrigeerd kan worden. Uiteraard stijgt de betrouwbaarheid van de resultaten en wint daarmee het onderzoek aan waarde naarmate een hogere dekking gerealiseerd wordt.

In tabel 4 geven wij een overzicht van het niveau van bestedingen – na seizoenscorrectie – in guldens geprojecteerd naar jaarbasis van zowel de Indexstudie als het D.B.O. [3]. Het is een illustratie van de mate waarin een proefpanel van beperkte omvang in een periode van opbouw en experimenten uitkomsten kan geven die op dit aggregatieniveau overeenstemmen met andere bronnen.

Tabel 4: Besteding per huishouden op jaarbasis in guldens.

	Index-panel	D.B.O. [3]
totaal consumptieve bestedingen	31.860,–	32.314,–
voeding	7.369,–	6.911,–
kleding en schoeisel	2.320,–	2.738,–

7. DE UITGANGSPUNTEN VOOR EEN VERDERE ONTWIKKELING

7.1. *Conclusies van de onderzoekgroep (EIM-Intomart)*

De onderzoekgroep concludeert:

- dat de uitkomsten van het proefonderzoek qua niveau en verdelingen goed tot zeer goed overeenstemmen met geraadpleegde bronnen;
- dat, gelet op bovenstaande conclusie en gezien de geringe problemen van respondenten bij het invullen van aankoopplaatsen, de verwachting gerechtvaardigd is dat ook de verdelingen naar distributiekanalen goed voor het voetlicht gebracht kunnen worden. Toetsing van uitkomsten betreffende distributie vindt thans plaats;
- dat de twee toegepaste registratiemethoden beide goed toepasbaar zijn gebleken, maar dat de B-versie vanwege de complexiteit meer inzet van de respondent vergt, hetgeen merkbaar is bij werving en respons. De opbouw

en continuering van een via het B-systeem registrerend representatief panel is daardoor wezenlijk kostbaarder dan een even omvangrijk via de A-versie registrerend panel. Gelet op deze overweging verdient het registratiesysteem-A de voorkeur. De gewenste nadere detaillering van de besteding ingeval van verzamelaankopen – met name relevant voor levensmiddelen – vereist echter registratie via de B-versie. Aangezien deze aankopen met hoge frequentie geschieden, kan worden volstaan met een kleine subgroep in de B-versie registrerende panelleden ter bepaling van verdeelsleutels waarmee in de A-versie geregistreerde verzamelaankopen verbijzonderd kunnen worden;
- dat de bestedingen beneden de 10 gulden goed in beeld gebracht kunnen worden. Ter vermijding van overbelasting van het panel en gelet op de hoge frequentie van bestedingen onder de 10 gulden is het verantwoord en aan te bevelen volgens een roulerend systeem steeds een beperkt deel van het panel bestedingen onder de 10 gulden te laten registreren.

7.2. *Beoordeling en advies van de adviescommissie*

Door het Bestuur van het EIM is ter begeleiding van het proefonderzoek een adviescommissie benoemd, waarin zitting hebben vertegenwoordigers van universitaire instellingen, overheid en bedrijfsleven.

De Adviescommissie Bestedingenindex heeft begin mei 1983 het proefonderzoek op diverse aspecten beoordeeld. De commissie concludeert samenvattend, dat hetgeen is voorgelegd aan technische opzet, aan ervaringen met de informatievergaring en aan onderzoekuitkomsten niet tegenspreekt, dat er sprake is van een haalbaar en goed meetinstrument. De commissie adviseert het bestuur van het EIM het panel te continueren voor de rest van het lopend jaar, met nadere besluitvorming omtrent continuering/uitbreiding per ultimo november 1983.

Een langere periode met beperkte steekproef wordt wenselijk geacht om:
- het instrument in technisch opzicht te kunnen verfijnen en bij te stellen;
- op basis van onderzoekresultaten de interesse op de markt nader te kunnen mobiliseren en in te vullen.

7.3. *Voortgang*

Door de onderzoekgroep is thans voorgesteld het panel te continueren voor de rest van 1983 met een beperkte steekproefuitbreiding naar 300 huishoudens (150 A-versie en 150 B-versie). De voorgestelde uitbreiding dient om bepaalde methodologische zaken (B-versie en registratie onder de 10 gulden) in te passen en nader te analyseren. Voorts om in het algemeen de analysemogelijkheden enigszins te vergroten, hetgeen ook de mogelijkheid biedt de onderzoekuitkomsten commercieel aantrekkelijker te kunnen presenteren.

Het voornemen is het panel gefaseerd uit te breiden. Uitgangspunt was en is een steekproefbasis van 1000 huishoudens (±2400 personen van 12 jaar en ouder) en in een latere fase, afhankelijk van de belangstelling op de markt, uit te breiden naar 2000 huishoudens. Bij de uitbreiding van 1000 naar 2000 bestaat het voornemen om de steekproef niet zonder meer representatief voor

de Nederlandse bevolking te vergroten, maar ten behoeve van studies bij deelpopulaties bepaalde categorieën te oversamplen.

Zomer 1983 wordt door het bestuur van het EIM een besluit genomen omtrent het al dan niet continueren van de bestedingenindex als EIM-project. De financiering voor de verdere ontwikkeling van het project zal hierbij een zwaarwegende factor zijn.

Samenvattend kan ter afsluiting worden gesteld:

- dat het ontwikkelde instrument naar haalbaarheid van informatievergaring en in technisch opzicht wordt beoordeeld als een goed meetinstrument;
- dat het project steunt op een breed draagvlak:
  * een groot aantal toepassingsvelden;
  * talrijke analysemogelijkheden;
  * een geconstateerde grote behoefte aan de – op grotere steekproefbasis – te bieden informatie;
- dat de veronderstelde teruggang in bestedingen door het instrument goed in kaart gebracht kan worden als niet tezelfder tijd de teruggang in onderzoekgerichte bestedingen leidt tot gebrek aan financiële middelen voor de verdere ontwikkeling van het project.

7.4. *Laatste ontwikkelingen*

Ten tijde van het gereed maken van de drukproef van dit artikel, heeft het bestuur van het E.I.M. besloten de E.I.M.-Bestedingenindex voor de rest van 1983 te continueren met enige uitbreiding van de steekproef. Aan de aan dit besluit gekoppelde voorwaarde, dat de hiervoor vereiste financiering op korte termijn zou worden verkregen, is inmiddels voldaan. Het project kan thans verder worden ontwikkeld, terwijl per ultimo november 1983 wederom een positief besluit mag worden verwacht, indien op dat tijdstip voldoende participatie is verkregen voor de begin 1984 na te streven steekproefbasis van 1.000 huishoudens.

[1] Deze commissie, onder voorzitterschap van Prof. dr. A. Bosman, werd ingesteld door de Staatssecretaris van Economische Zaken met de opdracht te adviseren omtrent:
- een programma voor research voor het midden- en kleinbedrijf;
- coördinatie van activiteiten van verschillende instellingen;
- aansluiting van de beleidsresearch van de algemeen sociaal-economische politiek van de overheid en de beleidsresearch ten behoeve van het midden- en kleinbedrijf.

[2] De EIM-Mededelingen, 1982, no. 6.

[3] Statistisch Bulletin, CBS, 38e jaargang no. 100, 19 nov. 1982

[4] Nationale Rekeningen, CBS, 1981.

5. Ontwikkelingen in het bestedingsklimaat en in de vraag naar elektrotechnische consumentenartikelen

J.W. VAN DER BEN

SAMENVATTING

Sinds 1980 worden wij geconfronteerd met een daling van de reële inkomens. De hiermee samenhangende teruggang van de bestedingen is gepaard gegaan met een meer dan proportionele vermindering van de vraag naar duurzame consumptiegoederen. Bij nadere beschouwing blijkt dat de negatieve inkomensontwikkeling zeer verschillend uitwerkt op de verschillende bestedingscategorieën binnen het gebied van de duurzame artikelen.

Doordat de gunstige invloeden, die inwerken op de vraag naar elektrische apparaten, de ongunstige invloeden overtreffen is de vraag naar deze produkten tot dusver weinig gevoelig geweest voor het ongunstige bestedingsklimaat. Ook de komende jaren zal de elektrotechnische industrie in staat zijn om een relatief sterk aanbod in de markt te leggen.

1. INLEIDING

Na enkele decennia van sterke groei van de particuliere bestedingen worden wij, reeds enkele jaren achtereen, geconfronteerd met een teruggang van de consumptieve uitgaven.

In Nederland is deze afname van de consumptie geprononceerder dan in de meeste andere westerse industrielanden.

Deze teruggang van de bestedingen is gepaard gegaan met een meer dan proportionele vermindering van de vraag naar duurzame artikelen. Bekijken wij de bestedingen aan duurzame artikelen echter voor een aantal componenten, dan blijken er grote verschillen te zijn in de gevoeligheid voor de teruggang van de consumptieve uitgaven.

In de context van de teruglopende vraag naar duurzame artikelen worden als voorbeeld regelmatig produkten van de elektrotechnische industrie genoemd zoals televisie-apparaten en wasmachines.

Bij een wat meer genuanceerde beschouwing blijkt dat de vraagontwikkeling voor deze produkten echter niet representatief is voor de bedoelde consumptieve reacties.

Ingegaan zal worden op een aantal zowel positieve- als negatieve invloeden op de vraag naar elektrotechnische consumentenartikelen.

Tegen deze achtergrond zal, nadat enkele opmerkingen gemaakt zijn over de potentiële economische ontwikkeling, een globale schets gegeven worden van

het perspectief van het aanbod van – en van de vraag naar – produkten van de elektrotechnische industrie.

2. BESCHIKBAAR – EN 'VRIJ BESTEEDBAAR' INKOMEN

Het niveau van de consumptieve bestedingen en de verdeling daarvan over de verschillende bestedingsmogelijkheden worden door een reeks van factoren beïnvloed.

Wij noemen in dit verband het niveau – en de ontwikkeling – van het reële inkomen, de verhouding tussen het niet vrij besteedbare – en het vrij besteedbare inkomen, de koopbereidheid (Katona 'willingness to buy'), de levensstijl, de levensfase waarin men verkeert, etc.

In het kader van deze notitie zullen wij de ontwikkeling van het reële inkomen, en van het vrij besteedbare deel ervan, nader bekijken.

In de periode 1967 – 1977 nam het reëel vrij beschikbare inkomen per werknemer gemiddeld met 3,2% per jaar toe. Sindsdien worden wij geconfronteerd met de volgende ontwikkeling.

Tabel 1. Reëel inkomen per werknemer (procentuele mutaties t.o.v. het voorafgaande jaar).

	1978	1979	1980	1981	1982
Beschikbaar reëel inkomen	+3,2	+1,8	−1,4	−3,7	−2,0

(Bron: Centraal economisch plan 1983).

In de afgelopen jaren is een zeer duidelijke breuk in de ontwikkeling ontstaan. In 1980 tekent zich de omslag af naar een dalend reëel inkomen. Sindsdien duurt de daling voort en ook voor 1983 wordt een teruggang van 3% voorzien.

Over 4 jaar gecumuleerd levert dit een vermindering van het reële inkomen op van ca. 10%. Ten opzichte van de trend uit de jaren 1967 – 1977 (+3,2% per jaar) ontstaat zodoende in 4 jaar tijd een afwijking van ca. 25%!

Hoe ontwikkelde zich nu in dezelfde periode het vrij besteedbare deel van het inkomen?

Het C.B.S. heeft zich nooit gewaagd aan het meten en registreren van het discretionaire-inkomen. Het NIPO heeft echter, op basis van metingen, een lange-termijn-reeks opgebouwd van het mediaan-inkomen van het nederlandse gezin en van het mediaan-bedrag dat een gezin nodig heeft om 'net aan rond te komen'. Tussen deze twee grootheden bevindt zich de 'vrij besteedbare ruimte'. In het hiernavolgende bekijken wij de ontwikkeling van deze 'vrij besteedbare ruimte' voor dezelfde periode (1967 – 1982) als gehanteerd werd bij de gegevens over het reëel beschikbare inkomen. Om tot een vergelijkbare basis te komen werden de nominale cijfers van het NIPO uitgedrukt in constante prijzen.

Tabel 2. Ontwikkeling van de 'vrij besteedbare ruimte' (in guldens).

	Nominaal	In prijzen van 1970	Mutatie reëel vrij beschikbaar inkomen in %	Vrij besteedbaar in % van het inkomen
1967	2.700	3.100 }	+ 3,6	28
1977	8.000	4.400 }	(gem. per jaar)	33
1978	8.800	4.650	+ 5,5	31
1979	9.250	4.700	+ 1,0	34
1980	8.300	3.900	− 17,0	29
1981	8.800	3.900	0	30
1982	10.100	4.350	+ 11,0	33

(Bron: Combinatie van NIPO en C.B.S.).

Voor de periode 1967 – 1977 zien wij een toename van de vrij besteedbare ruimte van 3,6% per jaar, tegenover in dezelfde periode een toename van het reële inkomen van 3,2% per jaar*.

Het is plausibel dat bij groeiende welvaart de vrij beschikbare ruimte iets sneller groeit dan het inkomen. We zien dan ook dat het aandeel van de vrij besteedbare ruimte in het inkomen toeneemt van 28% in 1967 tot 33% in 1977.

Over wat langere termijn gezien, en zonder grote discontinuïteiten in de ontwikkeling, geven de NIPO-gegevens een alleszins waarschijnlijke ontwikkeling aan van de beschikbare ruimte in de gezinsinkomens.

Anders wordt dit wanneer wij het verloop van de NIPO-cijfers met betrekking tot de beschikbare ruimte van jaar tot jaar bezien in de periode 1978 – 1982. Deze jaarmutaties vertonen een uiterst grillig verloop.

We constateren een scherpe terugval in de reëel vrij besteedbare ruimte in 1980 van 17% en een even opvallende stabilisatie in 1981 en een toename in 1982 met 11%.

Deze mutaties wijken sterk af van wat er in de betrokken jaren met de reële inkomens is gebeurd. Bij de mutaties op jaarbasis, met name in perioden van discontinuïteit van ontwikkelingen, zijn psychologische reacties kennelijk een rol gaan spelen.

Het ziet er naar uit dat, na een langdurige ervaring van krachtig en gestaag toenemende reële inkomens, de omslag in 1980 een 'schokeffect' heeft veroorzaakt.

De 'schrikreactie' heeft men dan ondermeer vertaald in het gevoel dat men belangrijk meer nodig had om 'net aan rond te komen'.

Daarna zien we een tegengestelde beweging. Terwijl het reële inkomen terugloopt, stijgt de door middel van het onderzoek te becijferen beschikbare ruimte weer. Na het verwerken van het schokeffect is vermoedelijk weer een meer realistische kijk op de situatie ontstaan.

Ook de door het C.B.S. gepubliceerde reeks betreffende de consumentenin-

* Aangetekend moet worden dat de meetbasis van beide reeksen niet geheel dezelfde is. Het reëel beschikbare inkomen van werknemers betreft gehuwde werknemers met 2 kinderen en met een inkomen juist beneden de laagste premiegrens der sociale verzekering. Het NIPO werkt met de mediaanwaarde van een steekproef uit alle huishoudens die in woningen gevestigd zijn.

dex, een samengestelde index op basis van attitude onderzoek, laat in de periode van het 4e kwartaal 1979 tot het 4e kwartaal 1980 een sterke afname van het vertrouwen zien.

We zullen in de volgende paragraaf zien dat de omslag in 1980, en de gevolgen daarvan voor het consumentenvertrouwen en de beleving van de vrij besteedbare ruimte, zich op specifieke wijze vertaald hebben in het bestedingsgedrag.

3. CONSUMPTIEVE BESTEDINGEN EN DE VRAAG NAAR DUURZAME CONSUMPTIEGOEDEREN

Hoe is nu in de periode 1978 – 1982, waarin de omslag zich heeft afgetekend, de ontwikkeling van de reële bestedingen geweest?

Tabel 3. Consumptieve bestedingen van gezinshuishoudingen in constante prijzen (procentuele mutaties t.o.v. het voorafgaande jaar)

	1978	1979	1980	1981	1982
Totaal	+4,4	+2,9	−0,4	−2,6	−0,4
Waarvan:					
Voedings- en genotmiddelen	+2,6	+3,0	− 0,7	−0,9	−2,5*
Overige goederen en diensten	+4,6	+4,6	+ 2,2	−0,4	−0,5*
Duurzame consumptiegoederen	+3,8	−0,3	− 4,7	−6,9	−4,5*
Waarvan:					
Personenauto's	+6,9	+2,2	−27,4	−4,9	
Textiel en kleding	+2,7	−0,7	− 2,0	−6,4	
Huish. Art. en woninginrichting	+5,1	−2,2	+ 2,0	−9,1**	
Elektro-technische apparaten	+7,0	−4,0	+ 7,0	−0,5	+5,5

(Bron: C.B.S. en Philips Nederland).

* Voorlopige indicaties.

** In deze reeks is ook elektro-technische huishoudelijke apparatuur begrepen. Aangezien de markten voor deze produkten een vrijwel constant niveau vertonen zijn de aangegeven mutaties vooral aan het niet elektro-technische pakket in deze reeks toe te schrijven.

We zien dat de omslag in het verloop van het reële inkomen zich ook in de bestedingen aftekent.

Wanneer we de 3 hoofdgroepen van de bestedingscategorieën bekijken, t.w. voedings- en genotmiddelen, overige goederen en diensten en duurzame consumptiegoederen, dan is het duidelijk dat de vraag naar duurzame consumptiegoederen het meest onder druk is komen te staan.

Wanneer we binnen de bestedingen aan duurzame artikelen naar enkele belangrijke subcategorieën kijken (waarvan de CBS-cijfers nog niet verder gaan dan 1981) dan blijken de reacties op de inkomensontwikkeling zich op verschillende wijze te vertalen naar de onderscheiden bestedingscategorieën.

De consumptieve reactie in het omslag-jaar 1980, dat wil zeggen het jaar van de 'schokreactie' is, voor wat betreft de duurzame artikelen, vrijwel ge-

heel tot uitdrukking gekomen in de vraag naar personenauto's. De totale bestedingen aan duurzame consumptiegoederen vertoonden, in constante prijzen, een daling in 1980 van f 2,1 mld. De bestedingen aan personenauto's (voor niet-zakelijk gebruik) liepen, eveneens in constante prijzen, terug met f 1,9 mld.

Het ziet er zodoende naar uit dat de consument (d.w.z. een beperkt aantal individuele consumenten) voor een eerste bijsturing van zijn bestedingen een grote en uitstelbare uitgave gehanteerd heeft.

In 1981 zien we dat, bij een voortdurende daling van het reële inkomen, het bestedingeneffect zich begint te spreiden over meer bestedingscategorieën.

We zien een versterkte daling van de vraag naar textiel en kleding en een krachtige reactie bij de woninginrichting. De omzet in de detailhandel voor woninginrichting en huisraad daalt nominaal van 1980 tot 1982 met 15% (Bron: C.B.S. Maandstatistiek van de binnenlandse handel en dienstverlening maart 1983), hetgeen een reële daling is van ca. 20%.

In 1982 zien we dan bovendien dat ook de vraag naar voedings- en genotmiddelen sterker begint te reageren op de inkomensontwikkeling.

In tabel 3 wordt tevens aangegeven hoe het verloop is geweest, in constante prijzen, van de vraag naar elektrotechnische apparaten (de gegevens hebben betrekking op de consumentenartikelen uit dit pakket).

De fluctuaties in deze reeks zijn voor een belangrijk deel terug te voeren op de invloed van grote sportevenementen op de kleurentelevisiemarkt (wereldkampioenschappen voetbal in 1978 en 1982 en olympische spelen in 1980). Na 1980 is er een toenemende invloed van de snel groeiende markt voor videorecorders.

De gecumuleerde groei wijkt belangrijk af van het verloop van de totaalcijfers van de vraag naar duurzame consumptiegoederen.

Terwijl in de jaren 1980, 1981 en 1982 de vraag naar duurzame artikelen, gecumuleerd voor deze drie jaren, in volume afneemt met 15%, stijgt in dezelfde periode het volume van de vraag naar elektrotechnische apparaten met ruim 12%.*

Voor 1983 tekent zich een verdere toename af van tenminste 5%.

In de volgende paragraaf zullen wij op de achtergronden van de vraag naar elektrische apparaten nader ingaan.

4. DE VRAAG NAAR ELEKTROTECHNISCHE APPARATEN

Wij constateerden reeds dat de vraag naar elektrotechnische apparaten zich in recente jaren relatief gunstig ontwikkelt temidden van de afnemende bestedingen aan duurzame consumptiegoederen.

Welke ongunstige- en gunstige factoren beïnvloeden de vraag naar deze apparaten?

De belangrijke ongunstige factor van het afnemende reële inkomen is in de

* Overigens moet worden opgemerkt dat in lopende prijzen het verloop van de afzet een minder sterke groei vertoont dan de toename van de volumecijfers. De prijzen van het pakket elektrotechnische apparaten dalen in de betrokken periode namelijk met ca. 2% per jaar. Dit in tegenstelling tot de prijzen van het merendeel van de andere consumptiegoederen die overwegend een stijging vertonen.

vorige paragrafen uitvoerig aan de orde gekomen. Ongetwijfeld zijn er een aantal markten voor elektrische apparaten die hiervan de nadelige invloed ervaren. Zo heeft zich bijvoorbeeld, parallel aan de teruggang in de automarkt, in de markt voor autoradio een niveau-verlaging van ca. 150.000 apparaten gemanifesteerd.

Een andere ongunstige factor is in een aantal markten de combinatie van hoge bezitcijfers met geleidelijke verbeteringen van de aangeboden produkten. Deze produktverbeteringen worden wel continu gerealiseerd maar zijn, voor wat betreft de aangeboden functies, niet opvallend genoeg om een extra impuls te geven aan de vervangingsvraag.

Een aantal produkten uit het pakket van de elektrische huishoudelijke apparaten moet tot deze categorie gerekend worden.

In dit verband moet bijvoorbeeld gedacht worden aan koelkasten, die in 98% van de gezinnen worden aangetroffen (9% van de gezinnen bezit daarbij een 2e koelkast) en aan artikelen als strijkijzers (114 strijkijzers per 100 huishoudingen) en koffiefilterapparaten (bezit 84%).

We zullen nu de gunstige factoren, die duidelijk dominant zijn ten opzichte van de ongunstige factoren, eens nader bekijken.

Als eerste gunstige factor kan genoemd worden de voortgaande toename van het aantal huishoudens.

Tabel 4. Ontwikkeling van aantal inwoners en huishoudens en van de gemiddelde gezinsgrootte.

	Aantal inwoners (×1 mln.)	Aantal huishoudens (×1 mln.)	Gemiddeld aantal personen per huishouden
1975	13,7	4,5	2,98
1980	14,1	4,9	2,82
1985	14,6	5,2	2,74

(Bron: Combinatie van gegevens).

Enerzijds door de toename van het aantal inwoners, anderzijds door de voortgaande gezinsverdunning, neemt het aantal huishoudens in de periode 1975 – 1985 toe met ruim 700.000 eenheden.

Een tweede factor die een gunstige invloed heeft op de vraag naar elektrische apparaten is de ontwikkeling naar individualisering van het gebruik en de invloed daarvan op het bezit aan apparaten. Dit effect doet zich met name voor bij de zogenaamde 'entertainment-apparatuur' (geluids- en beeldapparatuur).

Audio-apparatuur heeft destijds zijn entree in de huishoudens gemaakt als 'huiskamerradio'. Sindsdien heeft deze apparatuur vele verschijningsvormen gekregen en is zich, voor wat betreft het gebruik, gaan spreiden door de woning.

Zo treffen we thans een portable radio aan in de keuken, een klokradio in de slaapkamer, geluidsapparatuur in de kamers van de kinderen, enz.

Ook televisie apparaten beginnen thans de eerste tekenen van spreiding door het huis te vertonen.

Onderstaande gegevens illustreren dit spreidings-effect in samenhang met de individualisering van het gebruik.

Tabel 5. Aantal apparaten met vaste standplaats per 100 vertrekken in 1982.

	Woon-kamer	Keuken	Slaap-kamer ouder(s)	Slaap-kamer zoon(s)	Slaap-kamer dochter(s)	Overige vertrek-ken
Hifi-stereo-app.	154	–	–	57	37	1
Port.radio en -cassette-apparatuur	29	29	57	61	65	9
Televisie-apparaten (kleuren en zw./w.)	91	–	10	13	8	1

(Bron: Interact).

Deze spreiding van apparatuur is tot stand gekomen door een wisselwerking tussen welvaartsontwikkeling en het aanbod van de industrie. Tevens heeft de snelle penetratie van centrale verwarming, en het andere gebruik dat hierdoor van de woning mogelijk werd, tot deze individualisering bijgedragen. Daarnaast heeft natuurlijk de algemene tendens tot individualisering, als sociologische factor, een rol gespeeld.

Als derde gunstige factor voor de ontwikkeling van de vraag naar elektrische apparaten willen wij het aanbod van nieuwe produkten met nieuwe functies noemen, evenals de toevoeging van functies aan bestaande produkten.

Met name in de sector van wat genoemd wordt de consumentenelektronica zijn in de afgelopen jaren belangrijke innovaties beschikbaar gekomen.

Gebleken is dat nieuwe functies, respectievelijk wezenlijke verbeteringen, op het gebied van geluids- en beeldapparatuur een zodanig 'consumer-appeal' hebben dat de opmars van deze apparatuur niet door negatieve inkomensontwikkelingen te stuiten is. Een duidelijk voorbeeld hiervan vormt de markt voor videorecorders. De Nederlandse markt, gemeten in stuks, ontwikkelde zich als volgt:
1979: 45.000
1980: 78.000
1981: 135.000
1982: 249.000

De afzet in de 1e helft van 1983 geeft aan dat de groei in deze markt onverminderd voortduurt.

Een ander voorbeeld is de markt van televisie-spellen. Wanneer wij ons in de cijfers beperken tot de spellen die, om ze te gebruiken, worden aangesloten op het televisie-apparaat, dan was de markt hiervoor in Nederland in 1981: 25.000 stuks, in 1982: 100.000 stuks en wordt voor 1983 een niveau van 175.000 stuks verwacht.

Zo zien wij bijvoorbeeld ook dat de markt voor kleurentelevisie-apparaten zich goed weet te handhaven. Dit ondanks het feit dat deze markt reeds voor 50% bestaat uit vervangingsaankopen (vervanging van een kleuren tv-apparaat door een ander kleuren tv-apparaat). Voor een belangrijk deel is dit te danken aan de toevoeging van teletekst-faciliteiten aan de apparaten. Hier-

door is aan het kleurentelevisie-apparaat een nieuwe functie toegevoegd.

Een vierde gunstige factor voor de vraag naar elektrische apparaten is de ontwikkeling van de relatieve prijs van de produkten afkomstig van de elektrotechnische industrie. De technologische ontwikkelingen gaan in een zeer hoog tempo, met name op het gebied van de elektronica. Deze leiden tot een sterke trendmatige daling van de prijs van componenten en werkt eveneens krachtig door in de montagetechnologie.

Gerelateerd aan de volume-output per medewerker vertoont de elektrotechnische industrie een trendmatige groei van de produktiviteit van ca. 10% per jaar. Vele andere industrietakken geven een produktiviteitsgroei te zien van 1 à 3% per jaar.

Hierdoor kunnen elektronica-produkten met vergelijkbare functies steeds goedkoper worden aangeboden, respectievelijk kan voor dezelfde prijs steeds meer inhoud geboden worden.

De relatieve prijs van deze produkten laat zodoende een gunstige ontwikkeling zien. (Voor de elektrotechnische industrie betekent dit overigens dat de volume-afzet, gerelateerd aan het Bruto Nationaal Produkt, procentueel meer dan evenredig moest stijgen om met de nominale ontwikkeling 'in de pas te blijven').

Er zijn nog meer positieve invloeden te noemen. Zo zijn bepaalde artikelen, zoals wasmachines en stofzuigers, betrekkelijk ongevoelig voor omgevingsontwikkelingen door hun uitgesproken utiliteitskarakter. Andere artikelen zijn bij uitstek geschikt als cadeauartikel en hebben zich daardoor goed kunnen handhaven.

In het kader van deze notitie denken wij echter voldoende duidelijk gemaakt te hebben dat er een aantal gunstige factoren zijn, die tot dusver sterker gebleken zijn dan de ongunstige factoren, en die ook een zekere belofte voor de toekomst inhouden.

Met name de innovaties op het gebied van de consumentenelektronica en in het bijzonder van de apparatuur die tegemoet komt aan de behoefte aan 'entertainment' blijkt een gunstige tegenkracht te vormen tegen de negatieve invloeden van het verslechterde bestedingsklimaat.

De vraag naar duurzame consumptiegoederen wordt meer dan gemiddeld getroffen door de negatieve inkomensontwikkeling. Wij menen te hebben aangetoond dat dit zeer verschillend kan uitwerken op de verschillende bestedingscategorieën binnen het gebied van de duurzame artikelen.

Overigens betekent de gesignaleerde ontwikkeling in de vraag naar elektrische apparaten dat, gegeven de dalende inkomsten, door bezuinigingen elders geld is vrijgemaakt voor de bestedingen.

In het artikel van den Admirant en van Raay ('consumentenbezuinigingen in tijden van recessie', tijdschrift voor marketing, februari en april 1983) worden een tiental bezuinigingsstrategieën genoemd.

Kennelijk worden deze ook gebruikt om verschuivingen binnen de bestedingsmogelijkheden te realiseren.

Hierbij moet voor wat betreft de elektrische apparaten worden aangetekend dat deze slechts 2% van de totale particuliere consumptie vertegenwoordigen en ca. 10% van de vraag naar duurzame consumptiegoederen. Dit houdt in

dat 0,2% aanpassing in de totale consumptie reeds ruimte biedt voor een extra groei van de vraag naar elektrische apparaten van 10%.

5. HET PERSPECTIEF

5.1. *Sociaal Economisch*

De combinatie van krachtige en langdurige economische groei en inflatoire impulsen hebben tot een sterke groei van de investeringen geleid. Na de verstoring van het groei-patroon door opeenvolgende explosieve ontwikkelingen van de energieprijzen werden wij geconfronteerd met een wereldwijde structurele overcapaciteit.

Wij bevinden ons thans in een moeilijke aanpassingsperiode welke gepaard gaat met capaciteitsvernietiging.

Dit teloor gaan van capaciteit is vermoedelijk groter dan uit de cijfers blijkt. Overal waar namelijk gebrek aan winstvermogen de mogelijkheden beperkt om de installaties te vernieuwen treedt snelle veroudering van het machinepark op.

Bovendien is de overheid in de problemen gekomen doordat vele verplichtingen vanuit de groei-situatie zijn aangegaan, die onder de huidige economische omstandigheden niet meer betaalbaar blijken te zijn. De Nederlandse overheid is op dit punt nog wat verder doorgeschoten dan de overheden in de meeste andere landen van het industriële westen. Dit ondermeer als gevolg van het beleid waarbij de aardgasbaten gebruikt zijn als dekking van korte termijn uitgaven. Nu de aardgasinkomsten teruglopen schept het gevoerde beleid additionele problemen.

De heden ten dage nogal eens te beluisteren mening dat de wereldeconomie nu definitief gedwongen is in een situatie van beperkte – of nul – groei wordt niet door ons onderschreven.

Op zichzelf moet de exponentiële groei in het tempo van de afgelopen decennia een keer zijn grenzen vinden.

Wie echter enig zicht heeft op datgene wat er aan huidige en potentiële technologische mogelijkheden nog in de pijplijn zit die weet dat de genoemde grenzen voorlopig nog niet aan de orde zijn. Met name het tegen steeds lagere prijzen beschikbaar komen van intelligente apparatuur biedt zodanige mogelijkheden voor besturing en automatisering dat ons nog spectaculaire veranderingen in ondermeer de industriële voortbrenging te wachten staan.

Arbeidstijdverkorting die nu gehanteerd wordt als middel om tot arbeidsdeling te komen zal te zijnertijd uit de economische groei en de produktiviteitsstijging gefinancierd kunnen worden.

Hernieuwde economische groei en een structurele verkorting van de arbeidstijd zullen dan hand in hand gaan.

Daarbij zullen de inspanningen om de negatieve gevolgen van de economische groei (zoals de belasting van het milieu) beter onder controle te krijgen aanmerkelijk moeten worden vergroot. Deze inspanningen zullen echter tevens een sterk groeiend marktgebied vertegenwoordigen en daarmee een eigen bijdrage leveren aan de economische ontwikkeling.

Kortom hernieuwde economische groei ligt zonder twijfel in het perspectief.

Niemand zal echter precies kunnen voorspellen of de betrokken versnelling al zeer nabij is of dat voor doorbraken met een brede basis tot omstreeks 1990 gewacht moet worden.

De Nederlandse economie functioneert overigens al enige jaren slechter dan die van het merendeel van de andere westerse industrielanden.

Nederland zal in het profiteren van de versnelling, vooral voor wat betreft de aanvangsfase daarvan, vermoedelijk achter blijven bij het O.E.E.S.-gemiddelde.

Dit ondermeer doordat de vermogensverhoudingen in de Nederlandse bedrijven zeer vergaand verslechterd zijn. Dit verschijnsel draagt, in combinatie met het sociaal-politieke klimaat, bij tot een zekere mate van inflexibiliteit van de Nederlandse samenleving.

Op wat langer zicht zijn wij zeker optimistisch voor wat betreft het herstel van de economische groei.

Voor de komende jaren moet rekening gehouden worden met het voortduren van de druk op de reële inkomens. Het herstel van de trendmatige groei van de particuliere bestedingen zal nog wel even op zich laten wachten.

5.2. *Het perspectief voor de elektrische apparaten*

De technologische ontwikkelingen op het gebied van de elektronica gaan in een verbluffend tempo.

Gesteld kan worden dat de elektronische omwenteling nog maar pas is begonnen.

Belangrijke trends die zich daarbij aftekenen zijn ondermeer:

- het snel goedkoper worden van I.C.'s (integrated circuits) en van microcomputers maakt het mogelijk om voor weinig kosten geheugens en intelligentie in veel apparatuur in te bouwen
- de betrokken componenten worden bovendien alsmaar kleiner en vergen steeds minder energie waardoor het aantal toepassingsmogelijkheden wordt verruimd
- apparatuur, waarbij zeer hoge eisen worden gesteld aan de precisie, kan beschikbaar komen op een consumentenprijsniveau.
  Als voorbeeld hiervan kan genoemd worden de recent op de Nederlandse markt gebrachte LaserVision-apparatuur (beeldplatenspeler).
  De beeldplaat wordt afgetast door een laserstraal. Deze volgt het spoor op de plaat met de nauwkeurigheid van een fractie van een micron terwijl de plaat draait met een lineaire snelheid in de orde van 30 km/uur.

De consument zal in de komende jaren een stroom van nieuwe produkten en mogelijkheden op zich af zien komen. De eerste tekenen van deze nieuwe generatie apparatuur zijn inmiddels zichtbaar geworden in de vorm van de eerder genoemde LaserVision (beeldplatenspeler), de Compact-Disc (audiospeler waarbij een plaatje met een doorsnee van 12 cm wordt afgetast met een laserstraal), de draagbare videorecorder en de home-computer. Popularisering van de video-camera, mogelijkheden op het gebied van de 'elektronische fotografie', verdere uitbouw van elektronische spellen, goedkopere printers, televisie-

apparaten met geheugens, verruiming van de teletekst-faciliteiten, enz. komen in snel tempo naderbij.

Naast deze ontwikkelingen zal het ook technisch mogelijk worden om tot een veel ruimere benutting te komen van de bestaande informatie-infrastructuur (uitstaande park TV-ontvangers, kabelnetten en telefonieverbindingen). Veel van deze mogelijkheden vergen echter niet alleen technische voorzieningen maar zullen ook wettelijk geregeld moeten worden.

We denken hierbij bijvoorbeeld aan de mogelijkheden van 'elektronic-mail' (door het snel goedkoper worden van geheugens komt tekstcommunicatie via de telefoonlijn binnen ons bereik) en van elektronisch betalingsverkeer.

Met deze korte schets proberen wij duidelijk te maken dat het aantal mogelijkheden van de elektronica vrijwel onbegrensd is.

Beperkingen voor het tempo van de realisatie hiervan zijn hoogstens gelegen in de wegen die moeten worden gecreëerd om de nieuwe ontwikkelingen in de samenleving in te passen.

In de consumentenelektronica zal het aandeel van de nieuwe produkten sterk toenemen en hiermee samenhangend het aandeel van de televisie- en audio-apparatuur afnemen.

De nieuwe produkten namen op wereldbasis in 1979 ca. 5% van de markt van consumentenelektronica voor hun rekening.

In 10 jaar tijd zal het aandeel van deze nieuwe produkten (videorecorders, compact-disc, laservision, spellen, home-computers, camera's) opgelopen zijn tot ca. 40%. Daarnaast zullen ongetwijfeld nog nieuwe produkten beschikbaar komen die nu nog niet geconcretiseerd zijn.

Concluderend kunnen wij stellen dat veel van de in paragraaf 4 genoemde gunstige factoren ook in de komende jaren hun invloed zullen blijven uitoefenen.

Vergeleken met een aantal andere bestedingsmogelijkheden in het vlak van de duurzame consumptiegoederen zal de elektronische industrie een relatief sterk aanbod in de markt kunnen leggen.

Alhoewel de afzet van elektrische apparaten wel degelijk ook de negatieve invloed van de verminderde reële inkomens ondergaat mag toch verwacht worden dat de elektrische apparaten voor deze omgevingsontwikkelingen minder gevoelig zullen zijn dan gemiddeld het geval is bij de duurzame consumptiegoederen.

Dat de elektro-technische industrie, gegeven de uiterst felle internationale concurrentie, ook zijn specifieke uitdagingen kent is een ander verhaal, dat niet past in de opzet van deze notitie.

Aan de mogelijkheden om nieuwe markten te creëren zal het echter zeker niet ontbreken.

6. Milieubesef in ontwikkeling; veranderingen in meningen over en de belangstelling voor het milieuvraagstuk in Nederland sinds 1970

L.J.M. SCHREURS

SAMENVATTING

De toename van de maatschappelijke belangstelling voor de milieuproblematiek vanaf het midden der jaren zestig is sinds 1970 gepaard gegaan met onderzoek naar meningen over en aandacht van de Nederlandse bevolking voor problemen in het fysisch milieu. In deze bijdrage wordt het verloop van de belangstelling van onze samenleving voor het milieuvraagstuk in de laatste twaalf à dertien jaar aan de hand van de resultaten van deze onderzoeken en opiniepeilingen geschetst. Daarbij wordt ingegaan op een aantal methodisch-technische problemen, waarmee de onderzoeker bij het vergelijken van gegevens uit zeer uiteenlopende bronnen wordt geconfronteerd.

1. INLEIDING

Het begin van een meer dan incidentele belangstelling voor problemen in het fysisch milieu kan men zowel voor Nederland als daarbuiten in de tweede helft van de jaren zestig situeren. Na deze periode, die ook wel als de alarmfase wordt gekenschetst, krijgt de toenemende aandacht voor het milieuvraagstuk een extra-stimulans door het verschijnen van het rapport van Meadows aan de Club van Rome (Meadows, 1972).

De verhoogde belangstelling voor de milieuproblematiek begint na 1970 haar weerslag te krijgen in het beleid van de overheid en gaat tegelijkertijd gepaard met de eerste opiniepeilingen en sociaal-wetenschappelijke onderzoekingen op het betreffende terrein.

In deze bijdrage staan ontwikkelingen in de maatschappelijke belangstelling voor het milieuvraagstuk in Nederland vanaf het begin van de zeventiger jaren centraal. Omdat het door de samenleving gepercipieerde belang van de problematiek in principe op diverse, uiteenlopende wijzen kan worden achterhaald, laten wij in paragraaf 2. enkele relevante indicatoren de revue passeren.

Ons verder beperkend tot de publieke opinie over milieuproblemen als indicator voor de belangstelling van de Nederlandse samenleving voor het milieuvraagstuk, wordt ingegaan op de zeer verschillende manieren waarop dit verschijnsel in de afgelopen decennia is gemeten.

In de vierde paragraaf vindt een schetsmatig overzicht plaats van de veranderingen in de maatschappelijke aandacht voor de milieuproblematiek in Nederland.

De bijdrage wordt afgesloten met een aantal conclusies en een methodisch-technische discussie over het volgtijdelijk, diachronisch meten van de belangstelling voor het milieuvraagstuk c.q. het milieubesef van de bevolking.

2. INDICATOREN VOOR DE MAATSCHAPPELIJKE BELANGSTELLING VOOR HET MILIEUVRAAGSTUK

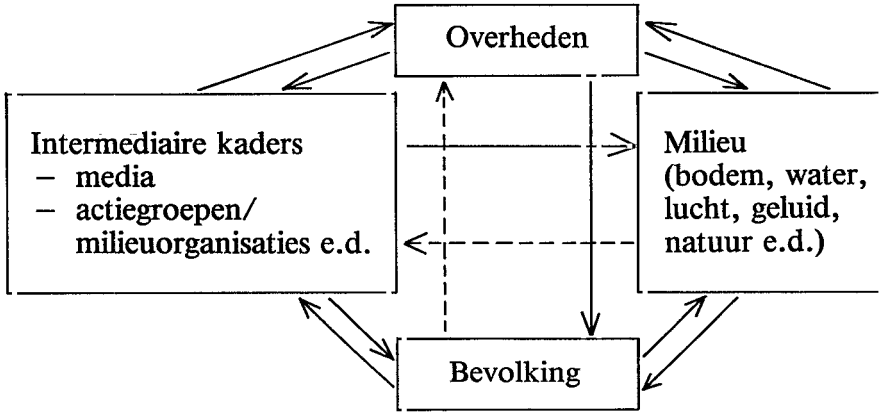
De aandacht van de samenleving voor milieuproblemen kan op een aantal manieren worden getraceerd. Zonder volledigheid na te streven noemen wij:

1. historische en politicologische analyses van het overheidsbeleid en beleidsmaatregelen op het betreffende terrein naar aard en omvang;
2. inhoudsanalyses van de aandacht voor de milieuproblematiek in publiciteitsmedia;
3. analyses van milieu-actiegroepen en milieu-organisaties naar aard, omvang en doelstelling;
4. onderzoek ten aanzien van geregistreerde klachten en andere uitingen met betrekking tot het milieu;
5. survey-onderzoek naar meningen, opvattingen, houdingen en gedragingen van de bevolking in relatie tot het fysisch milieu.

De genoemde indicatoren kunnen niet los van elkaar worden gezien, doch zijn ten nauwste aan elkaar gerelateerd. In figuur 1 wordt dit schematisch weergegeven.

De bevolking reageert op stimuli uit het milieu of percipieert de fysische omgeving en de problemen die zich daarin voordoen indirect via de media, actiegroepen, milieu-organisaties en andere intermediaire kaders. Op grond daarvan worden – beïnvloed door de eigen situatie – activiteiten ontplooid, die op hun beurt weer van invloed kunnen zijn op het milieu.

Een soortgelijk proces speelt zich af bij de overheden hetgeen resulteert in een bepaald beleid. Dit beleid komt mede tot stand op grond van signalen, die de overheid rechtstreeks of via intermediaire kaders vanuit de bevolking ontvangt (klachten, acties, berichten in media).



Figuur 1. Interrelaties tussen indicatoren.

Speelt de bevolking ten aanzien van het milieu een dubbelrol (ontvanger én instigator van stimuli), dit geldt evenzeer voor de overheid: enerzijds speelt zij in op feitelijke veranderingen in het milieu, op berichten in de pers, op activiteiten van actiegroepen en milieu-organisaties en op veranderingen in de beleving, houdingen en gedragingen van de bevolking; anderzijds zal zij via wetgeving, voorlichting en andere beleidsmaatregelen dit geheel actief (trachten te) beïnvloeden.

Mag de voorafgaande schets enigszins statisch aandoen, de ontwikkelingen in de maatschappelijke aandacht voor het milieuvraagstuk zijn dit zeker niet. Binnen deze ontwikkelingen zijn een aantal fasen te onderscheiden, die aan de hand van Downs' 'issue-attention-cycle'-theorie als volgt zijn te omschrijven (Downs, 1972; Ester en Schreurs, 1981):

fase 1: *pre-probleemfase*
Hoewel het vraagstuk als zodanig reeds bestaat, beperkt de belangstelling ervoor zich louter tot enkele experts.

fase 2: *alarmerende ontdekking en overmatig enthousiasme*
Als gevolg van een aantal dramatische gebeurtenissen of voorspellingen over ophanden zijnde catastrofen groeit de publieke belangstelling voor het probleem snel. Dit groeiend probleembesef gaat volgens Downs gepaard met een optimistisch geloof in de oplossing ervan.

fase 3: *besef van de maatschappelijke kosten*
Deze fase wordt gekenmerkt door een groeiend besef, dat de materiële én immateriële maatschappelijke kosten die aan het oplossen van het vraagstuk verbonden zijn veel hoger zijn dan aanvankelijk werd geacht.

fase 4: *geleidelijke vermindering van de publieke belangstelling*
Het idee dat de kosten van bestrijding c.q. oplossing van het vraagstuk zeer hoog zijn en daaraan gekoppeld dat men zich wellicht zelf ook bepaalde offers zal moeten getroosten, doet geleidelijk de publieke belangstelling voor het probleem verminderen. Dit leidt volgens Downs tot een van de volgende reactiewijzen: ofwel men voelt zich *bedreigd* door de omvang van het probleem, ofwel men voelt zich *ontmoedigd* doordat een snelle oplossing uitblijft, ofwel men raakt *verveeld* door het probleem.

fase 5: *post-probleemstadium*
Het maatschappelijke vraagstuk verschuift van het centrum van de publieke belangstelling naar de periferie en bereikt daarmee het laatste stadium van de cyclus.

Toegespitst op de maatschappelijke aandacht voor de milieuproblematiek in Nederland kan men stellen, dat de eerste fase de situatie in ons land tot het midden der zestiger jaren kenmerkt.

De tweede fase moeten we in de tweede helft van de jaren zestig en het begin van de jaren zeventig plaatsen. In die periode verscheen het eerste rapport van

de Club van Rome, werd een aanvang gemaakt met een milieubeleid en zagen een groot aantal actiegroepen en milieu-organisaties het eerste daglicht. De derde en de vierde fase kunnen we in de laatste tien jaar situeren: het eens zo enthousiast opgerichte Ministerie van Volksgezondheid en Milieuhygiëne komt in een relatief steeds moeilijker positie te verkeren, de 'grote schoonmaakbeurt' blijkt geen oplossing voor het vraagstuk en de louter technologische ingrepen blijken onvoldoende soelaas te bieden. Daarnaast richt de publieke belangstelling zich meer op nieuwe maatschappelijke problemen als de werkloosheidsproblematiek.

Hiermee wil echter niet gezegd zijn, dat het milieuvraagstuk volledig naar de periferie van de publieke belangstelling is verschoven. De problematiek is dermate veelomvattend en raakt zoveel burgers en maatschappelijke geledingen, dat het post-probleemstadium bij lange na niet is bereikt.

Bij het hanteren van elk van de indicatoren voor de aandacht van de samenleving voor het milieuvraagstuk komen methodisch-technische problemen om de hoek kijken. In de volgende paragraaf wordt een aantal van deze problemen beschreven.

Speciale aandacht zal hierbij worden geschonken aan het meten van de publieke opinie ten aanzien van de milieuproblematiek.

3. HET METEN VAN MILIEUBESEF; ENKELE METHODOLOGISCHE OVERWEGINGEN

In de vorige paragraaf zijn enkele indicatoren geschetst aan de hand waarvan (ontwikkelingen in) de belangstelling c.q. aandacht voor het milieuvraagstuk kan (kunnen) worden achterhaald. Genoemd werden analyses van de frequen-

Tabel 1. Geregistreerde telefoonmeldingen van particulieren over verontreiniging en hinder, per provincie.

	Klachten over					
	lucht	water	geluid	bodem	overige	Totaal
Friesland						
1975	36	90	–	18	34	178
1976	61	102	17	39	12	231
1977	55	107	7	40	26	235
1978	77	122	16	35	30	280
Drenthe						
1975	28	18	24	2	18	90
1976	32	13	16	11	20	92
1977	15	3	10	–	14	42
1978	18	16	18	2	17	71
Gelderland						
1975	–	–	–	–	–	2.111
1976	–	–	–	–	–	2.060
1977	–	–	–	–	–	1.454
1978	–	–	80	–	–	1.386
Noord-Holland[1]						
1975	(212)	(10)	(28)	(8)		258

1976	(395)	(19)	(53)	(15)		482
1977	243	69	16	18	11	424
1978/'79[2]	212	20	76	22	34	768
Zuid-Holland						
Rijnmond						
1975	6.577	37	969	13	83	7.679
1976	6.988	44	1.097	21	162	8.312
1977	7.005	38	975	11	154	8.183
1978	6.726	28	1.308	–	133	8.195
Dordrecht[3]						
1975	408	10	84	–	49	494
1976	682	16	82	–	63	817
1977	1.692	29	84	–	58	1.863
1978	607	30	107	–	82	805
Zeeland						
1976	15	–	–	–	–	15
1977	24	–	3	–	–	27
1978	20	–	9	1	–	30
Noord-Brabant[4]						
1975[5]	335	75	130	100	40	680
1976[6]	585	95	180	90	130	1.080
1977	287	57	132	75	138	689
1978	295	45	161	120	257	876
Eindhoven						
1975	360	40	337	1.161	442	2.340
1976	412	67	448	1.088	802	2.817
1977	389	17	316	797	415	2.088
1978	263	38	350	518	1.162	2.331

[1] De cijfers tussen haakjes zijn berekend uit de gemiddelde verdeling over de periode 1-10-'73 tot 1-4-'77.
[2] Meetperiode: 1 april – 31 maart.
[3] Een aantal klachten is in meerdere rubrieken opgenomen, waardoor het totaal afwijkt van de som van de rubrieken.
[4] Excl. Eindhoven.
[5] Meetperiode: september – december.
[6] Meetperiode: januari – oktober.

Bron: Algemene Milieustatistiek 1977 – 1978, CBS, 1980.

tie, de omvang en het tijdstip van georganiseerde maatschappelijke acties – hierbij kan men zowel denken aan acties van milieugroeperingen als aan beleidsmaatregelen van de overheid – en inhoudsanalyses van berichten die in de pers zijn verschenen. Op deze mogelijkheden gaan wij hier niet verder in; men raadplege hiervoor bijvoorbeeld 'Milieu van jaar tot jaar' van de Centrale Raad voor de Milieuhygiëne (een periodieke uitgave van de Staatsuitgeverij) en Tellegen en Willems (red.) 'Milieu-aktie in Nederland' (1978).

Een andere, zij het meer indirecte mogelijkheid biedt het monitoren van klachten bij officiële instanties (Schreurs, 1979: 88).

Uit tabel 1 blijkt, dat zulk een registratie niet in geheel Nederland plaatsvindt of heeft plaats gevonden, dat klachten ten aanzien van een aantal milieucomponenten niet overal worden gemonitored of dat een specificatie van de onderscheiden klachten ontbreekt. Daarnaast moeten deze gegevens ook om een andere reden met de nodige voorzichtigheid worden gehanteerd. Mensen die klagen zijn vaak niet representatief voor de groep, buurt, stad of regio

waaruit zij voortkomen en hoeven niet noodzakelijkerwijs meer hinder bij dezelfde mate van milieuverontreiniging te ervaren als de anderen (zie o.a. McKennell, 1970; Samuels, 1971; Weinstein, 1976). Ze wijken in een aantal opzichten, zoals beroep en inkomen van hun medeburgers af. Hiermee wil echter *niet* gezegd zijn, dat de klachen onterecht zijn. Uit onderzoek in de gemeente Uithoorn kwam bijvoorbeeld naar voren, dat het aantal klachten per wijk, gecorrigeerd voor het aantal huishoudens in die wijken, in grote mate overeenstemde met de omvang van luchtverontreiniging in de betreffende buurten (Van Hattum e.a., 1980).

Een mogelijkheid om een goed beeld te krijgen van de publieke aandacht voor en ervaring met (aspecten van) de milieuproblematiek vormt het ondervragen van een representatieve steekproef uit de bevolking. Doch ook bij dit zogenaamde survey-onderzoek kan men met de nodige problemen worden geconfronteerd. Deze hebben onder andere betrekking op de omschrijving en operationalisering van het begrip milieu, de formulering van de vragen en de volgorde waarin deze aan de respondent worden voorgelegd.

Uit een oriënterende studie van Van der Meer e.a. (1977) onder een aantal burgers en ambtenaren bleek dat het begrip milieu op uiteenlopende wijze wordt ingevuld. Over het algemeen werden er evenveel alle aspecten omvattende definities gegeven als opsommingen van mogelijke bedreigingen van het milieu. Bij de opsomming van de verschillende aspecten viel een aantal punten op (Van der Meer, 1981: 9):

- de onaantastbare eerste plaats van de lucht-, bodem- en watervervuiling;
- het nagenoeg geheel ontbreken van geluidshinder;
- de onduidelijke positie van zwerfvuil: door sommigen op de eerste plaats gezet, door anderen uitgesloten; bovendien aangemerkt als wel hinderlijk, maar niet bedreigend;
- de opvatting over veiligheid en kans op calamiteiten: wel bedreigend maar niet hinderlijk, en daarnaast door sommigen niet als groot probleem gezien.

Ook uit eigen onderzoek is gebleken, dat 'lawaai' nauwelijks als een vorm van milieuverontreiniging wordt beschouwd (ILBO/Veldkamp, 1981: 27).

Als belangrijke aspecten van de problematiek kwamen naar voren:

- de vervuiling van lucht, water en bodem;
- afval van fabrieken en industrie;
- afval op straat (hondepoep);
- kerncentrales en hun afvalprodukten;
- uitlaatgassen van auto's.

Voorts bleek, dat de invulling van het begrip milieuvervuiling afhankelijk is van plaatselijke omstandigheden.

Dat een en ander ook tijdsgebonden is, kan worden geïllustreerd aan het gegeven, dat in een grootschalig milieu-onderzoek in de Rijnmond (G.M.O., 1976) waarvan het veldwerk in het begin van de jaren zeventig werd gehouden in een lijst met vijftig aspecten van de omgeving het woord bodem (-verontreiniging) totaal ontbreekt.

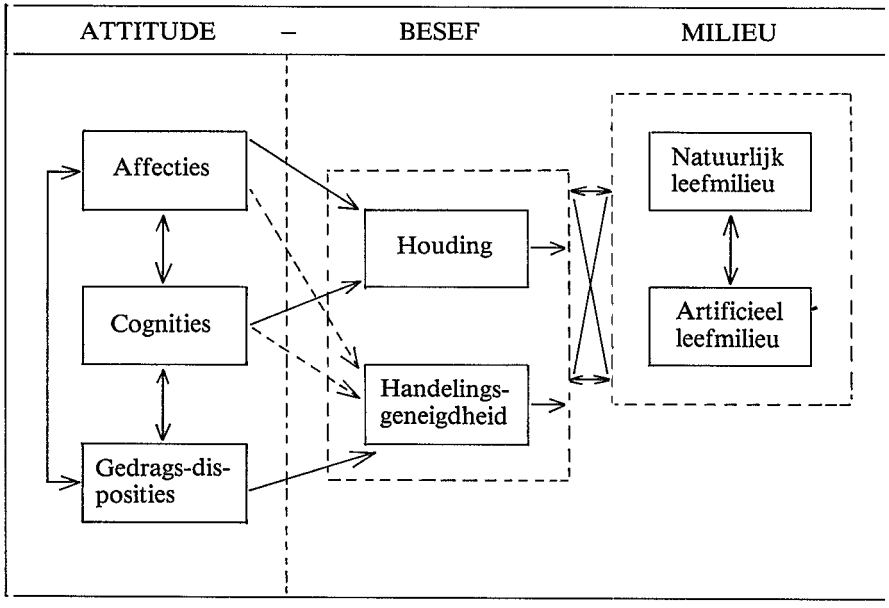
Werden in het voorafgaande de termen belangstelling en aandacht voor het

milieuvraagstuk gehanteerd, in het sociaal-wetenschappelijk en publieke opinie-onderzoek is sinds meer dan een decennium het begrip 'milieubesef' ingeburgerd. (Zie o.a.: C.R.M., 1970; Schreurs en Nelissen, 1975 en 1977; Ester, 1979; Van der Meer, 1981).

Onder milieubesef wordt het geheel van opvattingen over de bescherming, het beheer en de aantasting van het leefmilieu, alsmede de daarmee samenhangende gedragsdisposities verstaan (Schreurs en Nelissen, 1975: 10).

De inhoud van het begrip milieubesef is ruimer dan belangstelling c.q. aandacht alleen en omvat naast cognitieve en evaluatieve ook gedragsdispositionele (conatieve) elementen. Bij de operationalisering van milieubesef treedt tussen de verschillende onderzoekers de nodige divergentie op.

Schreurs en Nelissen (1975) onderscheiden binnen het begrip een houdings- en een handelingsgeneigdheidsaspect, terwijl het woord milieu in ruimere zin wordt opgevat (figuur 2).



Figuur 2. Aspecten van milieubesef.

Ester (1979) werkt het begrip in navolging van de eerder genoemde auteurs op dezelfde wijze uit, maar beperkt zich tot het natuurlijk leefmilieu.

Van der Meer (1981) heeft de mate van milieubesef bij de bevolking aan de hand van een viertal indices gemeten. Als eerste index hanteerde hij de vraag naar het belang van de milieuproblematiek in vergelijking met andere maatschappelijke problemen. De tweede index voor milieubesef wordt gevormd door een attitude-schaal, die betrekking heeft op het relativeren van milieuproblemen. Deze schaal werd reeds eerder gebruikt in het Geïntegreerd Milieu-onderzoek in de Rijnmond (G.M.O., 1976). Als derde index voor het belang dat men aan de problematiek hecht is een combinatie van vragen gehanteerd over de noodzaak, die men ervaart om bepaalde aspecten van het milieuvraagstuk op te lossen. De vierde index betreft de mate waarin men het milieu

hanteert als cognitieve coderingsdimensie bij het waarnemen van de omgeving. Deze index kan worden aangeduid als milieusaillantie of milieugevoeligheid.

Omdat de operationalisering van Van der Meer het nauwst aansluit bij de grote verscheidenheid, waarmee milieubesef vanaf 1970 in Nederland is gemeten, zullen wij zijn indeling tot uitgangspunt maken bij de beschrijving van de ontwikkeling in het milieubesef van de Nederlandse bevolking in de afgelopen twaalf à dertien jaar. De onderzoeksmatige aandacht heeft zich daarbij doorgaans op het natuurlijk milieu (water, bodem, lucht, geluid, flora en fauna) gericht.

4. MILIEUBESEF IN ONTWIKKELING

4.1. *Inleiding*

In deze paragraaf zullen aan de hand van een negental uiteenlopende bronnen veranderingen in het milieubesef van de Nederlandse bevolking vanaf 1970 worden getraceerd. In tabel 2 is weergegeven, welke bronnen het in concreto betreft.

Tabel 2, Geraadpleegde bronnen.

Titel	Jaar Veldwerk	Jaar Publicatie	Auteur/ Instantie
Milieubesef	1970	1970	Min. C.R.M.
Geïntegreerd Milieuonderzoek	1972	1976	O.L. Rijnmond
F.E.M.-enquêtes	1973 – 1978	1973 – 1978	Veldkamp/Marktond.
Zorgen voor Morgen	1974/1975	1975	S.M.O.
Sociaal-Cultureel Rapport	1975/1980	1982	S.C.P.
Milieubesef en Milieugedrag	1977	1979	Ester/IvM.
Attitude en Milieugedrag	1978	1981	Van der Meer
Longitudinaal Milieubelevingsonderzoek	1980/1981	1981	ILBO/Veldkamp
Milieu en Economie	1982	1982	Intomart

Bij het schetsen van de veranderingen zullen de door Van der Meer (1981) onderscheiden indices (a. belang; b. relativering; c. noodzaak tot oplossing; d. saillantie) als leidraad worden gehanteerd. Dat dit niet altijd mogelijk is, zal in het navolgende blijken. Tevens zal naar voren komen, dat verschillen in de vraagstelling en in de onderzochte populatie een vergelijking van de onderzoeksresultaten ernstig kunnen bemoeilijken.

4.2. *Het belang van het milieuvraagstuk*

Het belang van de milieuproblematiek wordt in de hierboven genoemde studies op zeer verschillende wijzen geregistreerd.

In het Geïntegreerd Milieuonderzoek (G.M.O.) werd de respondenten verzocht veertien op losse kaartjes omschreven maatschappelijke problemen te rangschikken. Bij de FEM-enquêtes werd telkens een identieke, open vraag gesteld naar het belangrijkste en het op één na belangrijkste probleem, waarmee Nederland te maken heeft. Ook Van der Meer confronteerde zijn respondenten met een aantal op kaarten vermelde vraagstukken, zij het dat het hier maar elf problemen betrof, vroeg hun de drie belangrijkste daaruit te selecteren om vervolgens binnen deze drie een rangorde naar belangrijkheid aan te geven. In het Longitudinaal Milieubelevingsonderzoek werd zowel naar het absolute als naar het relatieve belang van maatschappelijke vraagstukken gevraagd; het absolute belang werd via een vijfpuntsschaal gemeten, terwijl het relatieve belang via een daaraan voorafgaande open vraag naar de drie belangrijkste problemen van onze samenleving werd achterhaald. In de Intomart-enquête tenslotte dienden de ondervraagden het belang van de milieuproblematiek tegen dat van de werkloosheid af te wegen. In de overige onderzoeken werd een soortgelijke vraag niet gesteld of werden de uitkomsten ervan niet gerapporteerd.

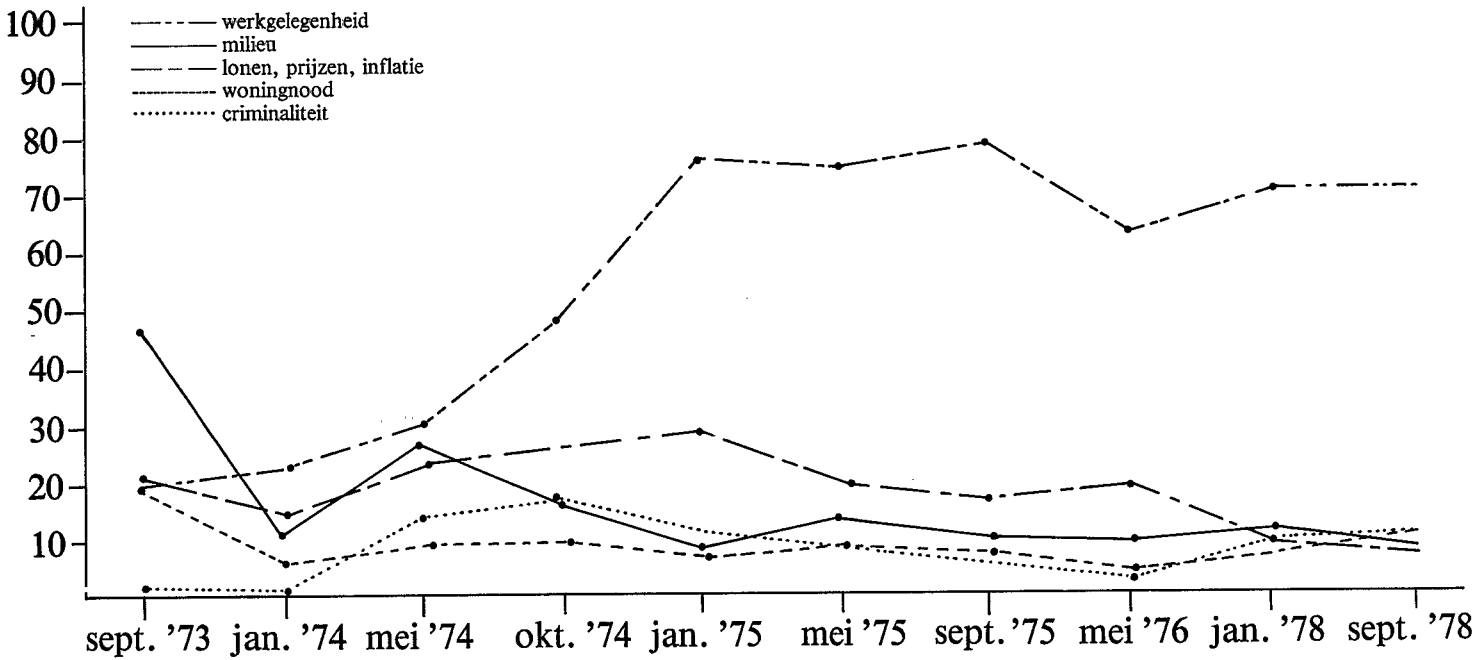
De diversiteit in vraagstelling maakt het op één lijn brengen van de onderzoeksresultaten uit de periode 1970 – 1982 nagenoeg onmogelijk. Een gunstige uitzondering betreft het tijdvak september 1973 tot september 1978. In die periode heeft Veldkamp Marktonderzoek in opdracht van het Financiëel-Economisch Magazine (FEM) een tiental metingen naar conjunctuurbeleving en -verwachtingen onder wisselende steekproeven (N = ca. 600) uit de Nederlandse bevolking verricht, waarbij een identieke vraagstelling inzake het belang van maatschappelijke vraagstukken aan de respondenten is voorgelegd.

De resultaten van deze enquêtes worden in figuur 3 weergegeven.

Met de trend in figuur 3 stroken ook andere onderzoeksresultaten:

- in het Geïntegreerd Milieu-onderzoek (veldwerk 1972) kenden zowel de respondenten uit de Rijnmond als die uit het Noordzeekanaalgebied aan milieuverontreiniging de belangrijkste plaats toe; de steekproef uit de totale Nederlandse bevolking zet weliswaar gezondheid op de eerste en milieuverontreiniging op de tweede plaats, doch deze verschijnselen zijn nauw bij elkaar betrokken waardoor nauwelijks van een afwijking kan worden gesproken. Beide problemen hebben qua belangstelling een duidelijke voorsprong op de overige (voorgeselecteerde) vraagstukken binnen de nationale samenleving;
- bij Van der Meer (veldwerk 1978) komt – na verwijdering van vraagstukken die in de andere onderzoeken niet aan de ondervraagden zijn voorgelegd – het belang van de milieuproblematiek na het werkgelegenheidsvraagstuk en problemen rond criminaliteit en misdaad op de derde plaats.

Wanneer wij ons beperken tot het als zodanig door de bevolking ervaren belangrijkste maatschappelijke probleem, dan wordt het in bepaalde mate mogelijk de trend in deze belangstelling tot 1982 door te trekken. Hierbij zij opgemerkt, dat de gegevens uit 1980/81 niet zijn ontleend aan een representatieve steekproef uit de Nederlandse bevolking, doch aan aselecte steekproeven uit een viertal contrasterende gemeenten (ILBO/Veldkamp, 1981). Voorts, dat de data uit 1982 afkomstig zijn uit een telefonische enquête (N = 501),



N.B. 1 In de figuur zijn de gesommeerde percentages van het belangrijkste en het op één na belangrijkste probleem weergegeven.
N.B. 2 In januari 1974 brak de oliecrisis uit. Dit had tot gevolg, dat 56% van de respondenten het grootste belang aan 'oliecrisis/energievoorziening' toekende, en verklaart de plotselinge terugval van andere vraagstukken.

Figuur 3. Het belang van het milieuvraagstuk in de periode 1973 – 1978.

Tabel 3. Het relatieve belang van het Milieuvraagstuk vanaf 1973 (percentages belangrijkste probleem).

		Milieu	Werkgelegenheid	Inflatie/ lonen/ prijzen	Woningnood	Criminaliteit
	sept. '73	27	8	15	11	1
	jan. '74*	3	6	6	2	–
	mei '74	14	14	14	5	7
	okt. '74	7	31	16	4	7
FEM/	jan. '75	3	58	9	1	3
Veldkamp	mei '75	3	57	7	3	3
	sept. '75	3	58	6	1	1
	mei '76	3	47	9	1	1
	jan. '78	4	53	2	2	3
	sept. '78	1	53	2	3	3
V.d. Meer	1978**	8	37	15	5	17
ILBO	1980	6	40	12	4	3
Intomart	1982***	12	25	–	–	–

* Zie NB 2 bij figuur 3.
** Zie voor aard van de vraagstelling de voorafgaande pagina's.
*** Dit is een schatting op grond van de uitkomsten van de vraag naar het relatieve belang van milieu en werkgelegenheid t.o.v. elkaar en het percentage aandeel van beide problemen t.a.v. het totaal der vraagstukken.

waarin een afweging werd gevraagd van twee maatschappelijke vraagstukken (Intomart, 1982).

Uit figuur 3 en tabel 3 kan, zij het met enige voorzichtigheid, worden geconcludeerd dat het milieuvraagstuk – na in het begin van de zeventiger jaren een toppositie te hebben ingenomen – zich qua relatief belang handhaaft bij de vier à vijf issues, die door de samenleving als belangrijk worden beschouwd. Er zijn zelfs enige – zij het vage – aanwijzingen voor een toename van het relatieve belang van milieuproblemen in de afgelopen jaren.

Een en ander zegt overigens niets over het absolute belang van de problematiek. Uit het Longitudinaal Milieubelevingsonderzoek (ILBO/Veldkamp, 1981) blijkt, dat het *absolute* belang van het milieuvraagstuk, naast dat van andere problemen, door de bevolking veel hoger wordt gesteld.

4.3. *Bezorgdheid over en relativering van milieuproblemen*

Als tweede index voor veranderingen in het milieubesef van de Nederlandse samenleving hebben wij de begrippen 'relativering' c.q. 'erkenning' van milieuproblemen geïntroduceerd.

Het begrip relativering van het milieuvraagstuk is ontleend aan het Geïntegreerd Milieuonderzoek (O.L. Rijnmond, 1976) en werd daar ook wel aangeduid met 'probleemontkenning'.

Relativering c.q. probleemontkenning vormt een tegenpool van erkenning van en bezorgdheid over milieuproblemen. Hoe deze relativering en bezorgdheid ten aanzien van de milieuproblematiek zich in de laatste elf à twaalf jaar heeft ontwikkeld wordt thans geschetst.

Tabel 4. Bezorgdheid over milieuvervuiling in Nederland, 1970 – 1980.

Maakt zich zorgen over milieuvervuiling	1970	1975	1980
– zeer veel	18,7	15,8	18,1
– veel	23,8	23,6	27,6
– tamelijk veel	28,0	33,9	34,2
– weinig	21,0	21,4	16,2
– helemaal niets	8.5	5,3	3,9
	(N = 1510)	(N = 1545)	(N = 1595)

Bron: Sociaal Cultureel Rapport 1972, S.C.P.

In het Sociaal Cultureel Rapport 1982 (S.C.P.) wordt een overzicht gegeven van de mate waarin de bevolking zich zorgen maakt over milieuvervuiling in 1970, 1975 en 1980. De gegevens van 1970 zijn ontleend aan de eerdergenoemde Milieubesefenquête, welke in opdracht van het toenmalig Ministerie van CRM werd uitgevoerd.

Tabel 4 toont de resultaten.

Vergelijkt men de gecombineerde percentages zeer veel/veel evenals die van weinig/helemaal niets op de drie tijdstippen, dan lijkt de bezorgdheid tussen 1970 en 1980 enigszins te zijn toegenomen. Of dit ook geldt voor de pendant van de bezorgdheid, te weten het relativeren van milieuproblemen moet uit de hierna volgende gegevens blijken. In de vier milieu-onderzoeken, waarin aan het relativeringsprobleem aandacht is besteed (G.M.O., 1976; Ester, 1979; Van der Meer, 1981; ILBO/Veldkamp, 1981), wordt met attitudeschalen gewerkt. Omdat deze schalen divergeren qua aantal en formulering van de items, zijn de resultaten niet direct vergelijkbaar.

Desalniettemin is er een mogelijkheid tot zinvolle vergelijking van enkele afzonderlijke items.

In alle vier de studies komt bijvoorbeeld als onderdeel van de relativeringsschaal het item 'al dat praten over milieuproblemen maakt de mensen meer ongerust dan nodig is' voor. Een vergelijking van de scores op dit item in de vier onderzoeken levert het in tabel 5 gegeven beeld op.

Concludeerde Van der Meer nog enkele jaren geleden, dat de mensen in absolute zin meer doordrongen waren geworden van de ernst van de milieuproblematiek (Van der Meer, 1981: 144), de laatst beschikbare gegevens laten deze conclusie – ondanks de beperkingen van het materiaal – niet geheel toe.

Om meer duidelijkheid te krijgen hebben wij naar overig vergelijkingsmate-

Tabel 5. Vergelijking van resultaten m.b.t. ontkenning van milieuproblemen in vier onderzoeken.

Onderzoek		Mee eens	?	Mee oneens
G.M.O.	(veldwerk '73)	45%	13%	42%
Ester	(veldwerk '77)	33%	17%	51%
V.d. Meer	(veldwerk '78)	36%	9%	55%
ILBO	(veldwerk '80)	44%	16%	40%
ILBO	(veldwerk '81)	44%	–	–

Tabel 6. Vergelijking van resultaten inzake milieurelativering vanaf 1977 (percentage (helemaal) eens).

Item	Ester '77	V.d. Meer '78	ILBO '80	ILBO '81
– Als we goed willen verdienen zullen we nu eenmaal wat milieuproblemen op de koop toe moeten nemen.	35%	44%	25%	29%
– het probleem van de milieuvervuiling wordt wel wat overdreven.	28%	34%	35%	35%
– de meeste mensen die voortdurend klagen over milieuvervuiling, zijn de mensen die altijd en overal iets op aan te merken hebben.	–	51%	52%	50%
– het probleem van milieuvervuiling is eigenlijk maar een stokpaardje van een aantal mensen.	24%	–	23%	20%

riaal gezocht. Dit hebben wij gevonden in enkele relativeringsitems bij Ester, Van der Meer en het ILBO/Veldkamp-onderzoek, doch het betreft slechts de periode 1977 – 1981 (tabel 6).

Met uitzondering van de scores ten aanzien van het eerste relativeringsstatement schijnen de antwoorden op de overige items erop te duiden, dat de ontkenning van de milieuproblematiek in de laatste vijf jaren niet of nauwelijks is toe- of afgenomen.

4.4. *Opvattingen over maatregelen tegen milieuproblemen*

De derde index voor milieubesef wordt gevormd door metingen over de noodzaak en wenselijkheid om de milieuproblematiek tot een oplossing te brengen en door het als (on-)voldoende kwalificeren van maatregelen, die tegen milieuverontreiniging worden genomen.

Het Sociaal Cultureel Planbureau legt aan een representatieve steekproef van N = ca. 1500 personen iedere vijf jaar de vraag voor of men vindt dat er genoeg tegen milieuvervuiling wordt gedaan.

Tabel 7 geeft de percentuele verdeling van de antwoorden weer over het tijdvak 1970 tot 1980.

Het aantal mensen dat vindt dat er teveel tegen milieuverontreiniging wordt gedaan neemt in de loop van de jaren weliswaar toe, doch daar staat weer te-

Tabel 7. Wordt er genoeg tegen de milieuvervuiling gedaan?

	1970	1975	1980
te veel	0,5%	1,7%	2,3%
genoeg	12,3%	25,9%	22,3%
te weinig	87,2%	72,4%	75,3%

Bron: S.C.P., Sociaal Cultureel Rapport, 1982.

Tabel 8. Meningen over maatregelen van diverse actores tegen milieuverontreiniging.

	Industrie	Overheid	Mensen Zelf
te veel	1%	1%	1%
voldoende	11%	24%	7%
te weinig	76%	61%	88%
weet niet	13%	15%	5%

Bron: Ester, 1979.

genover dat ook het percentage, dat de mening is toegedaan dat er te weinig gebeurt, na een terugval in 1975 vijf jaar later weer is toegenomen.

Differentieert men de betreffende vraag naar onderscheiden actores, dan levert dit voor 1977 het beeld op dat in tabel 8 is weergegeven.

De differentiatie naar diverse actores levert een veel genuanceerder beeld op, dan met de vraagstelling van het SCP wordt bereikt. Met name de opinies over de rol van de individuele burgers zijn opvallend te noemen.

Pleegt men een herberekening van de percentages na weglating van de categorie 'weet niet', dan komen de percentages voor de overheid grotendeels overeen met de SCP-gegevens over 1975; die voor de industrie liggen dan op het SCP-niveau van 1970 (!).

Hoewel de onderzoeksresultaten van Van der Meer (1981) door een anderssoortige vraagstelling niet met de bovenstaande te vergelijken zijn, geven wij en toch enkele weer, omdat hij een differentiatie aanbrengt naar onderdelen van de totale milieuproblematiek (tabel 9).

Hoewel een vergelijking tussen deze gegevens en die van het SCP en Ester moeilijk is te maken, kan men de richting van de percentages van Van der Meer overeenstemmend noemen met die van de andere twee studies. Dit geldt te meer, indien men bedenkt dat geluidsoverlast nauwelijks als milieuprobleem wordt gezien (zie par. 3) en de betreffende percentages dus eigenlijk niet in de beschouwing moeten worden betrokken.

Tabel 9. Noodzaak tot oplossing van milieuproblemen (1978; N=927).

	Een oplossing is:				
	noodzakelijk	gewenst	niet zo belangrijk	niet nodig	kan niet
– het onnodig sterven van dieren	73	21	2	3	1
– verdwijnen van plantensoorten	67	24	4	3	2
– bodemvervuiling door industriëel afval	69	26	2	3	1
– luchtvervuiling door schadelijke stoffen	67	26	3	3	1
– geluidsoverlast	38	47	10	5	1

Bron: Van der Meer, 1981.

4.5. *Andere indices voor ontwikkelingen in milieubesef*

In de vorige paragraaf werd als vierde index voor milieubesef onderscheiden: de mate waarin een persoon het milieu als cognitieve coderingsdimensie hanteert bij het waarnemen van de omgeving.

Deze index, ook wel milieusaillantie of milieugevoeligheid genaamd, is weliswaar zeer relevant, doch in dit kader voor ons bij nader inzien niet bruikbaar, omdat hij alleen in de studie van Van der Meer (1981) is gebruikt. Vergelijkingen in de tijd zijn daarom niet mogelijk. Dit geldt wel voor andere indices als offerbereidheid en milieurelevante activiteiten.

Zowel in het GMO (1972), door Ester (1977) als in het Longitudinaal Milieubelevingsonderzoek (1980) werd een nagenoeg overeenkomstige attitudeschaal gehanteerd, die de mate meet waarin mensen bereid zijn (financiële) offers te brengen om de milieuverontreiniging een halt toe te roepen. Een vergelijking van de percentuele verdeling van de scores op deze offerbereidheidsschaal leert, dat de milieu-offerbereidheid op deze wijze gemeten in bijna een decennium nauwelijks is veranderd.

Consultatie van andere bronnen m.b.t. offerbereidheid leert dat:

- in 1970 34% van de ondervraagden niet bereid was voor het behoud van een gezond leefmilieu financiële offers te brengen, bijna 30% hiervoor minder dan *f* 10,— over had, terwijl 37% bereid was tot een financiëel offer van meer dan *f* 10,— (CRM, 1970);
- in 1974/1975 de meerderheid van de Nederlandse bevolking van mening was, dat het zich getroosten van bepaalde offers ten behoeve van het milieu niet ten koste mag gaan van de stijging van het eigen inkomen en de werkgelegenheid (S.M.O., 1975);
- in 1982 74% het acceptabel vindt om loon af te moeten staan om de milieuverontreiniging te bestrijden en 85% bereid is om meer voor een liter benzine te betalen om het loodgehalte daarin te verminderen (Intomart, 1982).

De hoge percentages uit de Intomart-enquête moeten naar onze mening worden toegeschreven aan de aard van de vraagstelling: bij de vraag naar het inleveren van loon werd in de introductie over 'iets minder loon' gesproken; bij de prijsstijging van de (schonere) benzine werd het bedrag van 3 cent per liter genoemd.

Op grond van het bovenstaande moet de voorzichtige conclusie worden getrokken dat de offerbereidheid tussen 1970 en 1982 nauwelijks is veranderd: ongeveer 40% van de Nederlandse bevolking blijft, naar men zelf meedeelt, bereid enigszins substantiële financiële offers voor de bestrijding van milieuproblemen te brengen.

Hoewel het mogelijk is deze schets van ontwikkelingen in het milieubesef van de Nederlandse bevolking in het laatste decennium aan de hand van andere indices als actiebereidheid (G.M.O.; Ester, Van der Meer; ILBO/Veldkamp) en milieugedragingen (Van der Meer, 1974 en 1978; Ester, ILBO/Veldkamp) te continueren, zien wij hiervan af. Screening van het betreffende materiaal leert ons, dat de toegevoegde waarde van een dergelijke operatie als laag moet worden bestempeld. De verschillen in vraagstelling c.q. operationalisering van de variabelen zou ons tot welhaast onverantwoord

kunst- en vliegwerk, om niet te spreken van halsbrekende acrobatische toeren dwingen.

5. CONCLUSIES EN DISCUSSIE

Tien jaar geleden publiceerde Anthony Downs een theoretisch model, met behulp waarvan hij een antwoord trachtte te geven op de vraag, waarom een maatschappelijk probleem op het ene tijdstip in het centrum van de publieke belangstelling staat en op een volgend tijdstip qua aandacht naar de periferie is verdwenen, terwijl het probleem als zodanig niet is opgelost. In het model, door de auteur de 'issue-attention' cyclus genoemd, onderscheidt hij vijf stadia in de publieke belangstelling voor maatschappelijke vraagstukken. Deze stadia of fasen werden in paragraaf 2 besproken.

Downs heeft verder aangegeven, welke kenmerken maatschappelijke vraagstukken moeten hebben om de gehele cyclus te doorlopen.

Als kenmerken noemt hij (Ester en Schreurs, 1981: 155):

1. de meerderheid van de bevolking wordt niet direct en in zeer geringe mate door het vraagstuk getroffen;
2. het vraagstuk genereert sociale regelingen die grote voordelen met zich meebrengen voor de meerderheid of een invloedrijke minderheid van de bevolking;
3. het vraagstuk heeft niet langer – of geen – intrinsieke opwindende kwaliteiten.

Ook de maatschappelijke aandacht voor de milieuproblematiek in Nederland laat zich in termen van Downs' model beschrijven. Uit een vijftal mogelijke indicatoren, aan de hand waarvan deze belangstelling kan worden beschreven, kozen wij voor de publieke opinie over en de publieke belangstelling voor het milieu, ook wel milieubesef genaamd. Het milieubesef van de Nederlandse bevolking is sinds 1970 op uiteenlopende wijzen in opiniepeilingen en sociaal-wetenschappelijke onderzoeken gemeten. Als indices voor milieubesef werden door ons gehanteerd:
- het gepercipiëerde belang van milieuproblemen;
- relativering van en bezorgdheid over het vraagstuk;
- opvattingen over (de noodzaak van) milieubeleidsmaatregelen;
- offerbereidheid m.b.t. het milieu.

Traceren wij met behulp van deze indices veranderingen in het milieubesef van de bevolking in het laatste decennium, dan blijkt in de ogen van de Nederlander het *relatieve* belang van het milieuvraagstuk te zijn afgenomen: na een toppositie tot begin 1973 neemt de problematiek thans qua belangstelling ongeveer de vierde plaats in. Het *absolute* belang van het vraagstuk blijft echter hoog: het percentage mensen dat milieuproblemen zeer belangrijk vindt ligt in 1981 boven de zeventig (ILBO/Veldkamp, 1981).

De bezorgdheid over milieuproblemen ligt, na een afname in 1975, in 1980 hoger dan in 1970. De mate waarin het vraagstuk wordt gerelativeerd is ten aanzien van tien jaar geleden redelijk constant, hoewel ook hier een daling in

het midden van de jaren zeventig kan worden geconstateerd. Het aantal personen, dat de noodzaak tot het nemen van maatregelen tegen milieuproblemen onderschrijft, is in 1980 geringer dan in 1970 doch hoger dan in 1975. Bovendien blijkt, dat een differentiëring naar actores (overheid, industrie, de mensen zelf) in deze geen overbodige luxe is (Ester, 1979). De milieuofferbereidheid tenslotte ligt tien jaar na het veldwerk (1972) van het Geïntegreerd Milieu-onderzoek (O.L. Rijnmond, 1976) in 1982 op het zelfde peil; ook nu weer valt een tijdelijke teruggang rond 1975 – 1976 waar te nemen.

Op grond van de zojuist beschreven onderzoeksresultaten kan slechts in beperkte mate van een afname in het milieubesef van de Nederlandse bevolking sinds 1970 worden gesproken. Weliswaar wordt het relatieve belang van het milieuvraagstuk lager ingeschat dan tien jaar geleden, de andere indices doen veel eerder een redelijk constante mate van milieubesef vermoeden. De tot drie keer toe waargenomen terugval van het milieubesef in de periode 1975 – 1977 in vergelijking met 1970 – 1972 én 1980 – 1982 vormt zelfs een indicatie voor een – zij het lichte – opleving van de publieke belangstelling voor de milieuproblematiek in de afgelopen drie à vier jaren. Dit zou betekenen, dat de ontwikkeling van de aandacht voor het milieuvraagstuk niet conform het door Downs geschetste stramien verloopt.

Deze afwijking ten aanzien van zijn eigen model werd ook door Downs zelf reeds bevroed. Als argumenten waarom het aannemelijk is, dat het milieuvraagstuk niet naar de periferie van de publieke belangstelling zal verdwijnen, noemt de auteur:

a. het milieuvraagstuk is meer zichtbaar en bedreigend dan veel andere sociale problemen;
b. milieuverontreiniging bedreigt een ieder en niet slechts een kleine minderheid;
c. het mileuvraagstuk heeft geleid tot een grote privé-industrie die gespecialiseerd is in bestrijdingstechnologieën.
   Deze industrieën – Downs spreekt zelfs van een 'environmental-industrial complex' zullen er zeker zorg voor dragen dat de publieke belangstelling voor het milieuvraagstuk niet wegebt.
d. het milieuvraagstuk is dermate veelomvattend, dat er altijd wel aspecten van de 'issue-attention' cyclus zullen doordringen.

Deze argumenten komen ons redelijk plausibel voor.

Ondanks het streven onzerzijds om de bovenstaande conclusies niet met een al te grote stelligheid te poneren, bestaat het gevaar dat aan de resultaten een grotere hardheid wordt toegekend dan op grond van het beschikbare materiaal verantwoord is. De onderzoeken waaraan de gegevens zijn ontleend verschillen immers naar:

a. wijze van steekproeftrekking en omvang van de steekproef;
b. aard van de vraagstelling: algemene versus specifieke vragen; open versus gesloten vragen; enkelvoudige vraagstelling versus schaalconstructies;
c. het aantal antwoordcategorieën per relevante vraag;
d. de volgorde van de vragen;

e. presentatie van het onderwerp van onderzoek;
f. benaderingswijze: mondeling, schriftelijk, telefonisch.

Hoezeer deze onderzoeksaspecten de resultaten (kunnen) beïnvloeden, hebben wij wat betreft onderzoek naar geluidshinder elders beschreven (Schreurs, 1979: 89, 90).

Werkt de onderzoeksopzet reeds door bij het vergelijken van studies die in ongeveer dezelfde periode (zijn) worden uitgevoerd, dit geldt a fortiori wanneer het een vergelijking van onderzoeksresultaten in de tijd betreft.

Voor een longitudinale, diachronische analyse van (veranderingen in) de belangstelling voor een maatschappelijk vraagstuk is het hanteren van een nagenoeg identieke vraagstelling een eerste vereiste.

Maar ook de omvang van de steekproef, de wijze van steekproeftrekking en de benaderingswijze dienen zoveel mogelijk overeen te stemmen.

Van de door ons gehanteerde bronnen voldoen alleen de SCP-gegevens over culturele veranderingen in Nederland (Sociaal Cultureel Rapport, 1982) en de Veldkamp/FEM-enquêtes aan deze criteria.

Beide bronnen maken het mogelijk op redelijk betrouwbare wijze *nettoveranderingen* in milieubesef, d.w.z. veranderingen in de marges (op groepsniveau) aan te geven.

Wil men naast veranderingen op groepsniveau *ook ontwikkelingen op individueel niveau achterhalen én verklaren,* dan zal in het onderzoek noodzakelijkerwijs van een vaste steekproef (een panel) gebruik moeten worden gemaakt. Het enige ons bekende Nederlandse voorbeeld op het gebied van de milieuproblematiek is het reeds eerder geciteerde Longitudinaal Milieubelevingsonderzoek van ILBO/Veldkamp (1981).

Ter illustratie van de toegevoegde waarde van panelonderzoek laten wij enkele resultaten uit dit onderzoek, dat een methodisch-technische doelstelling had en om die reden van een ad-hoc-panel (geen panelwerving e.d.) gebruik maakte, tot slot de revue passeren.

Toegespitst op veranderingen in milieubesef kon over een periode van slechts zeven maanden (november 1980 tot juni 1981) worden vastgesteld dat:

- het absolute belang van de milieuproblematiek zowel op groeps- als op individueel niveau significant is toegenomen.
  De resultaten van de controlegroepen wezen uit dat hier niet van een systematische vertekening sprake was;
- dit absolute belang naar woonplaats sterk kan verschillen;
- het relatieve belang van de vier onderscheiden milieucomponenten (water, bodem, lucht, geluid) naar gemeente is gedifferentieerd en in een kort tijdsbestek kan veranderen;
- mensen, die aan de kwaliteit van het milieu of op de onderscheiden componenten een lager waarderingscijfer toekennen, het milieuprobleem vaker belangrijk gaan vinden. Met betrekking tot lucht was reeds aangetoond, dat dit waarderingscijfer ten nauwste met de feitelijke vervuilingsgraad van de lucht samenhing;
- de schaalscores op de relativeringsschaal op individueel niveau drastisch konden wijzigen zonder dat dit op groepsniveau tot significante verande-

ringen aanleiding hoefde te geven. (Individuele veranderingen kunnen elkaar opheffen);
- eenzelfde opmerking ten aanzien van de milieuofferbereidheid kan worden gemaakt.

Beperkt men zich louter tot het achterhalen van veranderingen in milieubesef op het niveau van de bevolking door wisselende representatieve steekproeven dezelfde vragen herhaald te laten beantwoorden, dan wordt de mogelijkheid ontnomen om aan te geven welke respondenten met welke kenmerken (opleiding, leeftijd, inkomen, woonplaats, regio e.d.) het de facto betreft. Het aanreiken van concrete handvaten voor milieu-beleidsmaatregelen wordt daardoor uiterst moeilijk.

BIBLIOGRAFIE

Bronner, A.E., L.J.M. Schreurs en K.C.P. Slootman, 1981. *The perception and evaluation of environmental quality.* Proceedings 34th ESOMAR Congress. Amsterdam, 175-203.

Bronner, A.E., L.J.M. Schreurs en D. Verzijden, 1981. *Longitudinaal Milieubelevingsonderzoek.* Amsterdam: ILBO/Veldkamp Marktonderzoek.

Centraal Bureau voor de Statistiek, 1980. *Algemene Milieustatistiek.* 1977 – 1978. Den Haag: Staatsuitgeverij.

Centrale Raad voor de Milieuhygiëne, 1982. *Milieu van jaar tot jaar, 1981.* Den Haag: Staatsuitgeverij.

Craik, K.H. en E.H. Zube (ed.), 1976. *Perceiving environmental quality.* New York: Plenum.

Downs, A., 1972. Up and down with ecology, the issue-attention cycle. *The Public Interest,* 28, 38-50.

Ester, P., 1979a. *Milieubesef en Milieugedrag.* Amsterdam: Instituut voor Milieuvraagstukken, Vrije Universiteit.

Ester, P. (red.), 1979b. *Sociale aspecten van het Milieuvraagstuk.* Assen: Van Gorcum.

Ester, P. en L.J.M. Schreurs, 1981. 'Het maatschappelijk draagvlak voor een effectief milieubeleid'. In: P. Nijkamp e.a. (red.), *Milieu-effectrapportering.* Assen: Van Gorcum. pp. 146-162.

Hattum, B. van, L.J.M. Schreurs, P. de Voogt en J.F. Feenstra, 1980. *Milieu, milieubeleving en gezondheid in Uithoorn.* Amsterdam: Instituut voor Milieuvraagstukken, Vrije Universiteit.

Intomart, 1982. *Milieu en economie*; een enquête ten behoeve van KRO/Brandpunt. Hilversum: april.

McKennell, A.C., 1970. 'Noise complaints and community action'. In: J. Chalupnik (ed.), *Transportation Noises,* pp. 238-242.

Meadows, D., 1972. *Grenzen aan de groei.* Utrecht: Spectrum.

Meer, F. van der, 1981. *Attitude en milieugedrag.* Leiden: Rijksuniversiteit.

Meer, F. van der, J. Bastiaansen en H. Weys, 1977. *Tussenverslag van het projekt Achtergronden van Milieugedrag.* Leiden: Rijksuniversiteit.

Ministerie van Cultuur, Rercreatie en Maatschappelijk Werk, 1970. *Milieubesef,* een enquête naar de betrokkenheid van de Nederlandse bevolking met de problematiek voortvloeiend uit de aantasting van het natuurlijk leefmilieu. Rijswijk.

Openbaar Lichaam Rijnmond, 1976. *Geïntegreerd Milieuonderzoek.* Rotterdam.

Samuels, S.W., 1971. 'Assessment of perception of air pollution'. In: W.R.D. Sewell en J. Burton (eds.), *Perceptions and Attitudes in Resources Management.* Ottawa: Information Canada.

Schreurs, L.J.M., 1979. 'Geluidhinder in sociaal-wetenschappelijk perspectief'. In: P. Ester (red.), *Sociale aspecten van het milieuvraagstuk.* Assen: Van Gorcum, pp. 79-93.

Schreurs, L.J.M. en N.J.M. Nelissen, 1975. *Het meten van milieubesef.* Maastricht: Raad der Europese Gemeenten.

Schreurs, L.J.M. en N.J.M. Nelissen, 1977. Het meten van milieubesef. *Recreatievoorzieningen,* nr. 12, pp. 520-523.

Sociaal Cultureel Planbureau, 1982. *Sociaal Cultureel Rapport.* Den Haag: Staatsuitgeverij.

Stichting Maatschappij en Onderneming, 1975. *Zorgen voor morgen.* Den Haag.

Tellegen, E. en J. Willems (red.), 1978. *Milieu-aktie in Nederland.* Zutphen: De Trommel.
Veldkamp/Marktonderzoek, 1978. FEM-enquêtes, 1973-1978. Amsterdam.
Veldkamp Marktonderzoek /ILBO, 1980. *Interimverslag van de eerste meting van het Longitudinaal Milieubelevingsonderzoek.* Amsterdam.
Weinstein, N.D., 1976. *'Human evaluation of environmental noise'.* In: K.H. Craik en E.H. Zube, o.c., pp. 229-252.

7. Non-response en de rol van de interviewer daarbij

A. ZIJDENBOS

SAMENVATTING

Non-response is een veelvuldig onderzocht verschijnsel in het markt- en opinie-onderzoek. De meeste studies hebben tot dusverre betrekking gehad op de relatie tussen respondent en non-response.

Non-response laat zich evenwel ook verklaren door de interviewereffecten en de interactie interviewer-respondent. In dit artikel wordt nader ingegaan op de relatie tussen weigeringsnon-response en socio-economische en psychologische kenmerken van de interviewer.

Voorts wordt aandacht besteed aan de non-response op inkomen als zijnde een representant van taboe-vragen en de rol van de interviewer hierbij.

1. INLEIDING

De non-response in mondelinge enquêtes neemt nog steeds toe. Onderzoeken met een response van 65% zijn tegenwoordig eerder regel dan uitzondering. Non-response kan men in twee hoofdcategorieën indelen, te weten 'niet mee willen werken (doorgaans weigeringen in al zijn gradaties)' en 'onbereikbaarheid (niet thuis)'.

Aan het non-response fenomeen zijn al vele studies gewijd, zowel vanuit de universitaire- als vanuit de marktonderzoekwereld. Bij de meeste van deze studies wordt de nadruk gelegd op de eigenschappen van de te ondervragen persoon. Zo stelt 't Hart (1) vast, dat de urbanisatiegraad van de woonplaats van de respondent een sterk verklarende factor is en dat het aantal weigeraars hoger is bij ouderen en vrouwen dan bij jongeren en mannen. Deze uitkomsten worden min of meer bevestigd door Inter/view (2), die daarbij ook constateert dat er een duidelijke samenhang bestaat tussen het weigeren en de sociale klasse; hoe hoger de sociale klasse, des te groter de kans, dat men weigert.

Ook het onderwerp van het onderzoek speelt een belangrijke rol. Zo is de non-response in onderzoeken, waarbij een maatschappelijk probleem aan de orde wordt gesteld aanzienlijk lager dan bij onderzoeken over bijvoorbeeld financiële zaken.

Als men al deze non-response studies bekijkt, blijkt er tot dusverre relatief weinig aandacht te zijn besteed aan de rol van interviewer, terwijl toch algemeen bekend is, dat – onafhankelijk van de al eerder geconstateerde samenhangen – de ene interviewer beter 'binnen komt' dan de ander.

De oorzaak van het feit, dat er zo weinig studies deze kant van de non-response hebben belicht, ligt mijns inziens enerzijds aan het feit dat men niet altijd bereid is informatie te verstrekken over de grootte en samenstelling van

het enquêteursapparaat* en anderzijds omdat men niet over een geschikt onderzoekkader beschikt om dergelijke studies uit te voeren.

Zoals bekend is de rol van de interviewer echter wel degelijk belangrijk bij het verkrijgen van informatie. Durbin en Stuart (3) stellen bijvoorbeeld vast, dat ervaren interviewers aanzienlijk beter binnenkomen dan onervaren, terwijl Berent (4) stelt dat de ideale interviewer een 'pedante schoolmeester moet zijn die gauw binnenkomt'. Dit laatste is uiteraard makkelijker gezegd dan in de praktijk uit te voeren.

Op een heel andere wijze hebben Van den Berg en Vermunt (5) aangetoond, dat er verband bestaat tussen de wijze, waarop men de respondenten benadert (beïnvloedt) en het 'succes' dat men hiermee behaalt.

Voorzover bekend is er weinig gepubliceerd over de invloed van de interviewer, met andere woorden, waarin verschilt de interviewer die weinig weigeringen tegenkomt van degene die juist vaak met weigeringen wordt geconfronteerd.

In dit artikel wordt nader ingegaan op de socio-economische kenmerken van de interviewer in relatie tot het aantal weigeringen.

Ook wordt de beantwoording van *een* representant van een taboevraag (de inkomens-vraag) aan de orde gesteld. Ook voor deze vraag wordt de rol van de interviewer aan de orde gesteld.

2. HET ONDERZOEKMODEL

Het Woningbehoeften-onderzoek 1981 (WBO 1981) is de basis geweest voor deze studie. Ten behoeve van dit onderzoek zijn er in het najaar 1981 door het NIPO, INTOMART en de NSS in totaal 81789 adressen bezocht, waaruit 58973 volledige gesprekken tot stand zijn gekomen. De response was derhalve 72,1%.

De populatie werd gevormd door personen van 18 jaar en ouder. De namen en adressen van deze personen zijn middels bemiddeling van het CBS door de gemeenten zèlf geleverd.

Alle interview(st)ers zijn kort voor aanvang van het veldwerk mondeling geïnstrueerd, terwijl voorts alle respondenten een introductie-brief hebben ontvangen.

Voor de volledigheid vermelden wij dat de non-respondenten van het najaar 1981 (eerste golf) in het voorjaar 1982 herbezocht zijn. De gerealiseerde response na twee veldwerkgolven was 80,2%.

Ik merk op dat dit resultaat aanzienlijk hoger is dan gebruikelijk bij mondelinge enquêtes; dit vindt zijn oorzaak in het feit dat:

- het onderwerp voor de respondenten interessant is (huisvestingsprobleem)
- de interview(st)ers mondeling geïnstrueerd zijn; dit werkt responseverhogend
- de respondenten vooraf een aankondigingsbrief hebben ontvangen; ook dit heeft een positieve invloed op de response.

* Zo wordt het niet als een pré beschouwd als men veel studenten als enquêteur in dienst heeft.

Aan de eerste veldwerkgolf hebben 667 NSS interview(st)ers deelgenomen; gemiddeld hebben zij 41 (met een minimum van 21) adressen bezocht. Per interviewer zijn de volgende gegevens berekend:

- weigeringspercentage $\left(= \frac{\text{aantal weigeringen}}{\text{aantal bezochte adressen}} \times 100\%\right)$
- onbereikbaarheidspercentage
  $\left(= \frac{\text{aantal adressen waar men niet thuis is}}{\text{aantal bezochte adressen}} \times 100\%\right)$
- non-response op inkomensvragen
- non-response op 'bedragen'-vragen, zoals huur, koopsom, huursubsidie, huurgewenningsbijdrage
- overige partiële non-response.

In dit artikel beperk ik mij tot het weigeringspercentage en de non-response op inkomensvragen.

In juni 1982 is vervolgens aan deze interviewers een schriftelijke vragenlijst toegezonden. In de vragenlijst waren naast de socio-economische vragen ook vragen opgenomen over het werken als interviewer (wijze van werken, aantal gesprekken per maand, enquêteren bij andere bureaus, enz.).

Voorts hebben wij aan de interviewers de S.I.G.-schaal (6) voorgelegd. Dit is de Schaal voor Interpersoonlijk Gedrag. Deze meet onder andere assertief gedrag; een eigenschap die mijns inziens belangrijk is om succesvol als interviewer te kunnen opereren.

In totaal zijn 535 volledig ingevulde vragenlijsten teruggestuurd. De non-respondenten bleken op de in dit artikel aan de orde gestelde variabelen niet significant af te wijken van de respondenten.

3. WEIGEREN BIJ ENQUÊTES

Er is al geconstateerd dat het weigeren bij enquêtes in sterke mate samenhangt met de urbanisatiegraad van de woonplaats van de respondent. Desalniettemin spelen ook andere factoren een rol. Ieder hoofd van de buitendienst van een marktonderzoekbureau kan vertellen dat interviewer A 'overal' binnenkomt en interviewer B (voorzover hij/zij nog in dienst is) 'nergens'. Als men naar het 'waarom' vraagt, moet men doorgaans het antwoord schuldig blijven. Het blijkt in de praktijk lang niet altijd te liggen aan uiterlijke vriendelijkheid of verzorging. De vraag is dan ook welke factoren wèl en welke géén rol spelen bij het realiseren van gesprekken. Het zal duidelijk zijn dat inzicht in deze problematiek kan leiden tot een effectievere werving van interviewers en kwaliteitsverhoging van het veldwerk.

De totale non-response in de eerste golf van het WBO 1981 was over drie bureaus gemeten 28,9%; waarvan 15,4% '*wil* niet meedoen' (weigeringen, e.d.).

De overige non-response heeft als oorzaak het feit dat men niet heeft *kunnen* meedoen vanwege 3 × niet thuis, ziekte, werkzaamheden, etc.

De navolgende resultaten hebben betrekking op het niet mee willen doen (voortaan het weigeringspercentage genoemd).

3.1. *Geslacht en leeftijd van de interviewer*

In een eerste analyse is aandacht besteed aan de relatie tussen leeftijd/geslacht en het weigeringspercentage. De resultaten van deze analyse tonen aan dat leeftijd een dominantere factor is dan geslacht. Interview(st)ers boven de 40 jaar worden significant minder vaak geconfronteerd met weigeringen dan hun jongere collega's, hetgeen hieronder wordt geïllustreerd:

	totaal	weigeringspercentage			gemiddeld
		minder dan 10%	10-20%	20% of meer	weigerings-percentage
		HORIZONTAAL GEPERCENTEERD			
totaal	535 = 100%	28	44	28	15,4
mannen					
totaal	98 = 100%	33	48	19	13,9
jonger dan 40 jaar	27 = 100%	26	33	41	17,3
40 jaar of ouder	71 = 100%	35	54	11	12,6
vrouwen					
totaal	437 = 100%	27	43	30	15,8
jonger dan 40 jaar	223 = 100%	23	40	37	17,2
40 jaar of ouder	214 = 100%	31	45	24	14,3

Ik merk op dat mannen boven de 40 jaar succesvoller zijn dan vrouwen boven de 40. Op grond van dit onderzoekmateriaal is moeilijk vast te stellen wat hiervan de oorzaak is.

Er zijn evenwel indicaties dat juist de wat oudere interviewers (55 jaar en ouder) nauwelijks worden geweigerd.

Zij ogen blijkbaar vertrouwd en men ziet in hen mogelijk een (groot)vaderfiguur aan wie men met een gerust hart informatie kan en/of wil verstrekken.

3.2. *Schoolopleiding van de interviewer*

Men zou verwachten dat er een samenhang bestaat tussen het schoolopleidingsniveau van de interviewer en het weigeringspercentage. Op grond van dit onderzoek kan zulks nauwelijks worden aangetoond.

Interview(st)ers met een hogere schoolopleiding hebben nauwelijks een lager weigeringspercentage dan degenen met lagere schoolopleiding.

3.3. *Interview-ervaring*

Durbin en Stuart (3) stelden vast dat ervaren interviewers een hogere response

(minder weigeringen) hadden dan minder ervarenen. Dit resultaat wordt door dit onderzoek bevestigd getuige het volgende overzicht:

in dienst sedert	gemiddeld % weigeringen
1980/1981	17,4
1978/1979	16,4
1976/1977	15,2
1975 of eerder	11,4

Op grond van deze uitkomsten kan men de conclusie trekken dat er in een enquêteursapparaat sprake is van de 'survival of the fittest', met andere woorden de besten (qua response) blijven over.

Interviewers met veel weigeringen vallen vanzelf af. Enerzijds omdat het zeer ontmoedigend is om steeds voor gesloten deuren te komen, anderzijds omdat men daardoor weinig verdient; doorgaans worden interviewers namelijk betaald op basis van stukloon.

3.4. *Assertiviteit*

Naast socio-economische kenmerken spelen ook andere factoren een rol in het succesvol opereren als interviewer. Uiteindelijk wordt het leggen van een effectief contact in belangrijke mate bepaald door de wijze waarop men de respondenten weet te overtuigen. Er is dus sprake van interactie tussen interviewer en respondent. De vraag is nu of er een instrument is dat op enigerlei wijze dit interpersoonlijk contact kan meten, zodanig dat het een voorspellende waarde heeft ten aanzien van het slagen in het leggen van contacten.

Daartoe is aan alle interviewers de Schaal voor Interpersoonlijk Gedrag voorgelegd, voorzover het de items betrof die betrekking hadden op de dimensies 'uiten van onzekerheid en eigen ontoereikendheid' en 'jezelf kenbaar maken: contact leggen, mening geven'. In totaal moest men op 23 items scoren.

De items besonden uit situatieschetsen, waarvan men op een 5-puntsschaal moest aangeven hoe gespannen men zich zou voelen als zo'n situatie zich voor zou doen. Per dimensie is een Likert-(somscore)schaal geconstrueerd, die bestond uit vijf klassen, te weten:

1 (zeer ontspannen)
2
3
4
5 (zeer gespannen).

Voor de weigerings-problematiek heb ik mij beperkt tot de dimensie 'jezelf kenbaar maken: contact leggen, mening'. De items van deze schaal benaderen namelijk het meest de veldsituatie.

De relatie tussen enerzijds het weigeringspercentage en anderzijds de mate van gespannenheid is als volgt:

mate van gespannenheid bij het leggen van contacten			percentage van de interviewers, met een weigeringspercentage van meer dan 20%
zeer ontspannen	(score 1)	n = 106	25,5
neutraal	(score 2/4)	n = 369	26,3
zeer gespannen	(score 5)	n = 60	38,3
totaal		n = 535	27,5

Interviewers die zich op de schaal 'contact leggen' als 'zeer gespannen' laten kwalificeren worden gemiddeld met significant meer weigeringen geconfronteerd dan zij die op de schaal 'neutraal' of 'zeer ontspannen' scoren.

Op grond van dit resultaat kan men van interview(st)ers, die op deze schaal als 'zeer gespannen' worden gekenmerkt reeds in een vroegtijdig stadium voorspellen of hij/zij relatief vaak of relatief weinig met weigeringen zal worden geconfronteerd.

4. DE NON-RESPONSE BIJ INKOMENSVRAGEN

Het zal duidelijk zijn dat een interviewer niet alleen beoordeeld kan en mag worden op het feit dat hij of zij makkelijk binnenkomt.

Er spelen ook andere factoren een rol. Een van deze factoren is het krijgen van response op 'gevoelige' informatie. Hieronder kan men informatie verstaan waarop min of meer een taboe heerst. Een van deze taboe-onderwerpen is 'geld' en dan in het bijzonder het 'inkomen'. Niet iedereen is genegen daarover informatie te verstrekken. Maar ook hier kan men zich weer afvragen, waarom de ene interviewer wèl die informatie krijgt die een ander niet krijgt.

Ik wil wel opmerken, dat inkomens-vragen *een* representant van 'gevoelige' vragen is, maar niet *de* representant. De beantwoording op vragen over een ander taboe-onderwerp (bijv. sexualiteit) kan wel eens een heel ander beeld te zien geven.

4.1. *Algemeen*

In het WBO 1981 werd naar het exacte bedrag aan netto-inkomen gevraagd; er is dus niet gevraagd naar inkomensklassen. De non-response op de inkomensvraag was, ondanks de begeleiding en interesse in het onderwerp toch nog 32,7%. Ik merk wel op dat deze non-response bestaat uit de componenten 'wil niet zeggen' en 'weet niet'; deze laatste categorie komt nogal vaak bij zelfstandigen voor.

Per interviewer is nagegaan hoe vaak hij/zij een inkomensvraag moest stellen (voor de respondent zèlf, zijn/haar partner en/of zijn/haar thuiswonende kinderen). De (non)response op inkomensvragen is uiteindelijk op interviewer-niveau ingedeeld. De aldus berekende inkomens-response coëfficiënt (I.R.C.) was gemiddeld 0,172.

De interviewers zijn ten aanzien van de beantwoording in twee klassen inge-

deeld, te weten 'relatief hoge response op inkomen' en relatief lage response (rep. 58% en 42%).

4.2. *Urbanisatiegraad*

De belangrijkste variabele die de non-response op inkomen verklaart is de urbanisatiegraad van de woonplaats van de respondent. Hoe kleiner de gemeente des te hoger is de non-response. In onderstaande tabel wordt dit duidelijk geïllustreerd:

	totaal	urbanisatiegraad woonplaats interviewer*	
response op inkomensvragen		steden	platteland
relatief hoog	50	61	49
relatief laag	42	39	51
totaal	100%	100%	100%
n =	535	399	136
i.r.c.	0,172	0,163	0,199

* In vrijwel alle gevallen was de woonplaats van de respondent gelijk aan de woonplaats van de interviewer.

Er zijn mijns inziens drie redenen om deze verschillen te verklaren:

a. Op het platteland is men meer gesloten dan in de grote stad; om die reden zal men ook niet responderen op dergelijke taboe-vragen.
b. Omdat de interviewer in dezelfde gemeente woont als de respondent, kennen zij elkaar nogal vaak ook al is dat doorgaans alleen van naam en van gezicht. Het feit dat men door een min of meer bekende wordt ondervraagd zal bij de respondent een zekere schroom teweeg brengen.
c. Op het platteland wonen veel agrariërs en het is bekend dat zij hun inkomen vaak niet weten.

Enerzijds komt men dus op het platteland gemakkelijker binnen dan in de grote stad, maar anderzijds krijgt men aanzienlijk minder informatie op 'taboe'-vragen.

Eenzelfde soort tegenstelling ziet men ook bij de wijze waarop de respondenten benaderd worden. Interviewers die doorgaans telefonisch afspraken maken hebben een groter aantal weigeringen, maar daarentegen behalen zij een hogere response op inkomensvragen dan interviewers die doorgaans de zg. 'overval'-methode toepassen (i.r.c. resp. 0,123 en 0,180).

4.3. *Geslacht/leeftijd*

Het blijkt dat vooral oudere mannen een hoge response op inkomen behalen. Een verklaring voor deze uitkomst is moeilijk te geven.

Het zou kunnen zijn dat men in hen een soort vader-figuur ziet in wie men vertrouwen heeft (zie ook 3.1). Kortom, in deze zal de leeftijdsverhouding

respondent/interviewer vermoedelijk een heel belangrijke rol spelen.

De i.r.c.-resultaten waren:

mannen –	totaal	n= 98	0,149
	jonger dan 40 jaar	n= 27	0,163
	40 jaar of ouder	n= 71	0,144
vrouwen –	totaal	n=437	0,178
	jonger dan 40 jaar	n=223	0,172
	40 jaar of ouder	n=214	0,184

Oudere vrouwen scoren relatief laag op inkomensvragen. Ik heb de indruk – en dat is ook door sommigen van hen bevestigd – dat zij het erg vervelend vinden om naar het inkomen te vragen. Met andere woorden er is hier niet alleen sprake van een taboe bij de respondent, maar ook bij de interviewer. Nader onderzoek zou deze hypothese evenwel moeten bevestigen.

Uit het onderzoek blijkt de non-response op inkomensvragen nauwelijks samen te hangen met opleiding, ervaring en assertiviteit.

5. CONCLUSIES

1. Bij de contactlegging tussen interviewer en respondent spelen drie factoren een rol:
- de medewerking van de respondent
- de interviewer
- de interactie tussen de respondent en de interviewer.

Het non-response verschijnsel is tot dusverre voornamelijk geanalyseerd vanuit de respondent. Uit de tot dusverre uitgevoerde studies komt naar voren dat de mate van non-response, veroorzaakt door weigeringen, samenhangt met de urbanisatiegraad van de woonplaats, de sociale klasse van de respondent en de interesse in het te onderzoeken onderwerp.
De mate van weigeren hangt evenwel ook samen met bepaalde socio-economische en psychologische eigenschappen van de interviewer.

Uit de onderhavige studie blijkt dat leeftijd, interview-ervaring en assertief gedrag sterk samenhangen met de mate waarin men met weigeringen wordt geconfronteerd.

Op grond van deze resultaten zal het in de toekomst mogelijk worden reeds vooraf in zekere mate te voorspellen of een interviewer op dit aspect (confrontatie met relatief veel weigeringen) wel of niet succesvol zal opereren.

2. De non-response op inkomensvragen is hoog op het platteland. Dit wordt in hoge mate veroorzaakt door de geslotenheid van de 'plattelander' en het feit dat de interviewer de respondent vaak kent ook al is het alleen maar van gezicht.

In veel mindere mate dan bij de non-response door weigeringen spelen

socio-economische en psychologische kenmerken van de interviewer hierbij een rol.

NOTEN

1. Hart, H. 't: Selektie en zelf-selektie van informanten in enquêtes, Amsterdam 1974.
2. Inter/viewgroep: Rondje non-response van de inter/viewgroep, Amsterdam 1978.
3. Durbin, J. en A. Stuart; Differences in response rates of experienced an inexperienced interviewers; Journal of the Royal Statistical Society. A 114, 1951, pp 163-205.
4. Paul H. Berent op het ESOMAR-Congres, Evian 1962 (Interviewer recruitment and Training).
5. Gerda van den Berg en Riet Vermunt, weigeringen bij enquêtes, Jaarboek van de Nederlandse Vereniging voor Marktonderzoeker, 1980.
6. S.I.G.-schaal; Schaal voor Interpersoonlijk Gedrag, copyright 1980, Swets en Zeitlinger B.V. te Lisse.

8. Marktonderzoek: een onmisbaar instrument bij de introduktie van nieuwe produkten

C.J.M. POOTERS

SAMENVATTING

Dit artikel geeft een systematisch overzicht van de verschillende fasen die de produktontwikkeling doorloopt en de taak en functie, die marktonderzoek binnen elk van deze fasen vervult. Als voorbeeld is gekozen voor 'DRIEHOEK WASBUILTJES', dat eind 1981 op de Nederlandse markt geïntroduceerd. Het produkt wordt gevolgd vanaf het eerste conceptonderzoek tot en met de situatie ongeveer een half jaar na introductie. Een totale periode van ongeveer 7,5 jaar produktontwikkeling.

1. INLEIDING

In dit artikel wordt een beschrijving gegeven van de taak en functie die marktonderzoek vervult als onmisbaar instrument in het proces van produktontwikkeling en de uiteindelijke introductie van nieuwe produkten. Een constatering, die zowel geldt voor introducties in de consumentenmarkt als in de industriële markt. Succesvolle produktintroducties zijn in deze moeilijke economische tijd voor elke ondernemer van levensbelang.

Immers falen kost veel geld en de (financiële) speelruimte van fabrikanten is beperkt!

Dat marktonderzoek een belangrijk hulpmiddel kan zijn om de kans op falen te verkleinen, is niet alleen een mooi verkooppraatje van bureaumensen. Systematisch onderzoek in alle stadia van het produktontwikkelingsproces is bittere noodzaak.

Als voorbeeld om een en ander toe te lichten, is gekozen voor 'DRIEHOEK WASBUILTJES', een wasmiddel dat eind 1981 op de Nederlandse markt geïntroduceerd is. Uit een enquête van het levensmiddelenvakblad 'Distrifood', kwam naar voren dat de introductie van dit produkt naar de mening van de handel de meest succesvolle van 1981 was. (1)

Vanaf het allereerste begin in het ontwikkelingsproces, is dit produkt middels marktonderzoek begeleid en ook nu, na de introductie worden nog steeds op een systematische manier marktgegevens verzameld. De methoden en technieken van onderzoek, die daarbij zijn toegepast, zijn uiteraard niet alleen voorbehouden aan onderzoek voor wasmiddelen. Ook bij andere produkten in de consumentensfeer en de industriële markt zijn vergelijkbare procedures mogelijk.

Ik dank de fabrikant van 'DRIEHOEK WASBUILTJES', Kortman &

Schulte b.v. te Veenendaal* enerzijds voor het ter beschikking stellen van alle onderzoeksmateriaal en anderzijds voor het kritische commentaar naar aanleiding van het eerste concept van dit artikel.

In het produkt-ontwikkelingsproces en het daarmee veelal gepaard gaande, begeleidende onderzoek kan een viertal fasen worden onderscheiden (3), namelijk:
- produkt – idee – onderzoek
- produkt – concept – onderzoek
- produkt – ontwikkelings – onderzoek
- produkt – acceptatie – onderzoek

In dit artikel zal met name worden ingegaan op de laatste drie fasen. Niet dat de eerste fase minder interessant zou zijn, maar extern marktonderzoek is in deze fase niet zo systematisch aan bod geweest. Intern marktonderzoek, middels een zorgvuldige analyse van de markt én de noodzaak technische oplossingen te vinden voor fosfaatvervangers (i.v.m. de wetgeving voor fosfaathoudende wasmiddelen), vormden de basis voor het idee een nieuw wasmiddel te ontwikkelen met soda als belangrijkste bestanddeel.

Daarnaast bleken de produkt-idee-fase en de concept-fase vrij ongemerkt in elkaar over te vloeien. Steeds weer werd de behoefte gevoeld interne inzichten te toetsen aan de wensen en behoeften van de doelgroep van straks: de Nederlandse huisvrouw.

Vervolgens zullen we stilstaan bij het moment van introductie, oktober 1981 en wat er daarna met het produkt gebeurd is. Dat onderzoek daarbij eveneens een belangrijke functie heeft vervuld, zal de lezer niet vreemd voorkomen na het doorlezen van dit verhaal.

Alvorens echter op de gehanteerde onderzoekmethoden en -technieken binnen elk van de onderscheiden fasen in te gaan, zal allereerst een korte beschrijving worden gegeven van de wasmiddelensituatie in Nederland alsmede van de specifieke produkteigenschappen van 'DRIEHOEK WASBUILTJES'.

2. WASMIDDELENSITUATIE IN NEDERLAND

In Nederland beschikken ca. 3.9 miljoen huishoudens over een trommelwasmachine (wasautomaat). Per huishouding worden per jaar in deze trommelwasmachines gemiddeld bijna 100 wassen op temperaturen van 60° C en hoger uitgevoerd. Het volume aan wasmiddelen dat daarbij jaarlijks wordt gebruikt, kan worden geraamd op ca. 60.000 ton. Deze temperatuurbeperking is in zoverre van belang, dat 'DRIEHOEK WASBUILTJES' tot op dit moment (ten tijde van het schrijven van dit artikel) alleen gebruikt kon worden voor machine-wassen van 60°C of hoger en derhalve niet voor wassen op 30-40°C in de machine. Een aspect, dat enige marktbeperkingen heeft, maar waaraan gezien de technische problemen, die men daar in 1981 nog mee had, niet het

* Door een interne reorganisatie is de naam in de loop van dit jaar gewijzigd in Kortman Redipro b.v.

hoofd kon worden geboden.

Binnen deze afgebakende totale markt opereert een groot aantal aanbieders met als belangrijkste merken: All, Ariel, Dash, Dixan, Dobbelman, Omo en Witte Reus. Fosfaatloze en fosfaatarme wasmiddelen zoals Klok C5, Klok Hagelwit en Tricel hadden marktaandelen, die nauwelijks iets voorstelden in de totale wasmiddelenmarkt. Bovendien hadden fosfaatloze wasmiddelen het negatieve image, dat ze nauwelijks schoon wasten.

3. DRIEHOEK WASBUILTJES – EEN INNOVATIE

In het ondernemingsbeleid van Kortman & Schulte b.v. is een wezenlijke plaats ingeruimd voor produktinnovatie, d.w.z. het ontwikkelen van werkelijk nieuwe produkten.

Vanuit deze beleidsfilosofie begon men in de eerste helft van de jaren '70 te denken over de mogelijkheden om een nieuw fosfaatvrij wasmiddel te ontwikkelen. De afdeling Research en Development meende dat een formule zou kunnen worden ontwikkeld voor een volstrekt nieuw, op soda gebaseerd wasmiddel. Een nadeel van het gebruik van soda bestond echter uit het achterblijven van kalk- en magnesiumaanslag op het wasgoed en in de wasmachine. Dit aspect vereiste een behandeling met een mildzure fractie. De indruk bestond dat beide elementen gecombineerd moesten worden in één wasmiddel, wilde het concept overlevingskansen bezitten. Op basis van een uitwerking van verschillende mogelijkheden (tabletten, toverballen, etc.) werd uiteindelijk gekozen voor het wasbuiltje. Dit is een verpakte dosering voor één wasbehandeling, waarin het mild-zure en het alkalische bestanddeel van elkaar gescheiden zouden zijn. Een bijzondere lijmlaag moest ervoor zorgen dat het alkalische gedeelte pas in het water zou kunnen komen nadat het milde zuur zijn werk had gedaan. De wijze waarop het wasmiddel werkt, maakt het onmogelijk om voor te wassen. Een activiteit, die in de 70-er jaren op grote schaal door huisvrouwen werd verricht. Deze werkwijze van het wasmiddel kan zowel in positieve zin worden uitgelegd (minder waterverspilling) als in negatieve zin (mindere vlekverwijdering).

Dat de nieuwheid van het concept (zowel in technische zin als in communicatieve zin naar de consument toe) zonder twijfel van invloed is geweest op de aard en omvang van het onderzoeksprogramma is een constatering die zeker voor de hand ligt.

Immers alle andere wasmiddelen bestonden tot de introductie van 'DRIEHOEK WASBUILTJES' uit poeder, dat in het doseerbakje van de wasmachine door de huisvrouw gebruikt kon worden. Pas de introductie van de vloeibare wasmiddelen, eind 1982 (Vizir en Sil) hebben een nieuwe dimensie aan het wasmiddelengebruik toegevoegd.

Ten behoeve van de produktontwikkeling was een aantal basisvoorwaarden geformuleerd, waaraan het nieuwe wasmiddel zou moeten voldoen.

Het wasmiddel zou:
- fosfaatvrij moeten zijn
- even goed moeten wassen als een fosfaathoudend wasmiddel
- in de wasautomaat gebruikt moeten kunnen worden

– een voor de consument aanvaardbare prijs moeten hebben (gerekend per wasbeurt en in vergelijking tot fosfaathoudende wasmiddelen).

Aan de introductie van 'DRIEHOEK WASBUILTJES' is in totaal bijna een periode van 7 jaar ontwikkelingswerk en voorbereidingen vooraf gegaan, met een totaal geïnvesteerd bedrag van 17 miljoen gulden. Op de plaats en functie van onderzoek daarbinnen zal nu worden ingegaan.

4. ONDERZOEK TER VOORBEREIDING VAN DE INTRODUCTIE

In het onderzoek dat voorafging aan de introductie van 'DRIEHOEK WASBUILTJES' op de Nederlandse markt, kan een drietal fasen worden onderscheiden, namelijk:

1. het *concept-onderzoek,* waarin een veelheid van afzonderlijke elementen onderzocht moest worden teneinde reacties van consumenten op een volledig nieuwe manier van wassen na te gaan;
2. het *produktontwikkelingsonderzoek,* dat inzicht moest verschaffen in de werking van het wasbuiltje, de 'performance' van het wasmiddel alsmede aspecten als verpakking, merknaam en communicatie-uitingen;
3. het *produkt-acceptatie-onderzoek,* waarin het uiteindelijke antwoord op de vraag moest worden verkregen of het produkt levensvatbaar was in de Nederlandse marktsituatie.

De verschillende onderzoeksvormen, die binnen elk van deze fasen aan bod kwamen, zullen in deze paragraaf in het kort de revue passeren. Daarbij is niet gekozen voor een behandeling, waarin elk onderzoekstype diepgaand wordt beschreven naar vorm en inhoud, maar voor een opsomming van de aspecten die onderzocht zijn en de stapsgewijze voortgang naar het moment van introductie. Gezien de omvang van het onderzoeksmateriaal zal een en ander niet anders dan in vogelvlucht aan de orde kunnen komen.

4.1. *Het concept-onderzoek*

Middels een aantal kwalitatieve onderzoeken (diepte-interviews met huisvrouwen) werd gestart met het verzamelen van gegevens omtrent de behoefte aan informatie over een fosfaatvrij wasmiddel in verschillende verschijningsvormen (bolletjes, tabletten, etc.), de preferentie van consumenten en de elementen die bijdragen tot die preferentie. Uit deze eerste fase werd al snel duidelijk, dat op een uitgebreide manier informatie moest worden gegeven over de werking van het wasmiddel, waaronder: oplossen, verdelen, doseren, voorwas- en hoofdwaswerking. Als randvoorwaarden voor acceptatie kunnen worden vermeld (aspecten die in latere fasen op uitgebreidere schaal onderzocht en gekwantificeerd zijn):

– het middel moet schoon wassen
– de offerbereidheid voor het milieu lijkt niet groot (zowel v.w.b. de eerste randvoorwaarde als t.a.v. de prijs, die men bereid is te betalen voor een fosfaatvrij wasmiddel)

Een ander aspect, dat nadere onderzoeksgegevens vereiste, was een inventarisatie van het *doseergedrag* van huisvrouwen: gebeurt dat middels het doseerbakje of rechtstreeks in of op de trommel. Uit een kwantitatief gebruiksonderzoek kwam vast te staan, dat het rechtstreeks doseren in of op de trommel reeds op uitgebreide schaal voorkwam, met name bij wasmachines die als bovenlader getypeerd kunnen worden. Deze bevinding was van groot belang voor de verdere ontwikkeling van het produkt (i.c. de wasbuiltjes).

In een volgende fase werden de reacties van consumenten op de *verpakkingsvorm* (een zich openend sachet) en het gebruik (minitest in een proefwasserij) vastgelegd, waarbij de gegevens zijn verkregen middels observatie, interviews en groepsdiscussies.

In deze fase werd een bevestiging gevonden van eerdere constateringen, dat men het gebruiksgemak waardeert en het fosfaatvrije karakter ondergeschikt is aan het wasresultaat. De prijs van het wasmiddel en de temperatuurbeperking (60°C of hoger) worden als belangrijke nadelen gezien. Immers dit laatste aspect vereist de aanwezigheid van een tweede wasmiddel binnen elke huishouding (voor de bonte en fijne was op lagere temperaturen). De aanwezigheid van meerdere wasmiddelen binnen een gezin werd middels een pantrycheck (telefonisch én mondeling) nog separaat onderzocht. Hieruit kwam naar voren, dat driekwart van de huishoudingen wasmiddelen uit 2 of meer segmenten (wit, bont of fijn) in huis heeft, met een gemiddelde van 2,3 wasmiddelen/merken.

Tenslotte bestond het concept-onderzoek uit een aantal afzonderlijke, meer kwalitatief gerichte deel-onderzoeken, waarin werd nagegaan in hoeverre het mogelijk is de consument met een beperkte hoeveelheid *informatie* de werking en het gebruik van het wasmiddel duidelijk te maken. Tevens werd daarbij getracht een relatie te leggen tussen de variabelen: begrip van de werking en gebruik en de acceptatie van het wasmiddel. De belangrijkste conclusie was, dat het geven van extra informatie over het produkt zowel het begrip als de acceptatiegraad verhogen. Dat een korte tekst daarover niet voldoende is en ook het gebruik van tekeningen niet bijdraagt tot een beter begrip, bemoeilijken het communicatie-beleid. De elementen, die in de communicatie naar voren moeten worden gebracht, zijn een combinatie van convenience en fostaatvrije en niet elk van deze aspecten afzonderlijk.

Ondanks de communicatieve problemen, bleek uit een 'in-home-test' (waarin huisvrouwen gedurende 2 weken met het testprodukt wasten), dat er duidelijk sprake is van een leerproces dat zich vrij snel voltrekt. Huisvrouwen waren reeds na een eerste ervaring met het wasmiddel in staat om het produkt bij een tweede keer bijna routinematig te gebruiken. Deze constatering is van groot belang voor het promotionele plan tijdens en na de introductie (sampling), waarop in het vervolg nog zal worden teruggekomen.

Op basis van de resultaten van de onderzochte elementen in de conceptfase (fosfaatvrij, verpakkingsvorm, temperatuurbeperking, informatiebehoefte) kon deze fase worden afgerond en werd overgegaan naar de fase van het produktontwikkelingsonderzoek.

4.2. *Het produktontwikkelingsonderzoek*

In deze fase van de produktontwikkeling kan een onderscheid worden gemaakt tussen het verpakkingsonderzoek van het sachet (wasbuiltje) en de verpakkingsvorm van meerdere wasbuiltjes in een doos, het reclame- en advertentieonderzoek en het testen van de merknaam. Bij elk van deze elementen zal in het vervolg van dit artikel summier worden stilgestaan.

4.2.1. *Verpakkingsonderzoek m.b.t. sachet (wasbuiltje)*

Een drietal deel-onderzoeken was gewijd aan de voorkeur van de huisvrouw voor verschillende soorten materiaal, waarvan het wasbuiltje gefabriceerd kon worden. Wordt het plastic sachet vergeleken met papier, dan lijkt de voorkeur duidelijk in de richting van het eerste soort materiaal te gaan. Er bestaat namelijk grote angst voor verpulveren (pluizen) van papieren sachets (vrijwel elke huisvrouw lijkt wel negatieve ervaringen te hebben opgedaan met papieren zakdoekjes die in de was terecht zijn gekomen).

Vergelijkt men echter het plastic sachet met non-woven materiaal, dan gaat de voorkeur duidelijk in de richting van het tweede. Deze voorkeur wordt enerzijds ingegeven door de specifieke eigenschappen van het non-woven sachet (milieuvriendelijker, laat beter water door) en anderzijds door een afkeer van het plastic sachet (algemene afkeer van plastic, niet milieu-vriendelijk, kan smelten). Op basis van de uitkomsten van deze onderzoeken is uiteindelijk gekozen voor het vervaardigen van de sachets van non-woven materiaal.

4.2.2. *Verpakkingsonderzoek doos met wasbuiltjes*

We hebben al geconstateerd, dat een combinatie van convenience en fosfaatvrij op de verpakking het beste het produkt, de eigenschappen en het gebruik communiceren. In het verpakkingsonderzoek heeft de nadruk duidelijk gelegen op de communicatieve aspecten van de verpakking en minder op de fysieke verpakkingsvorm van de wasbuiltjes. Een aspect waarvan later bleek, dat het in relatie tot de prijs van het wasmiddel van groot belang is. Immers een kleinere doos waspoeder dan de vergelijkbare merken (2 kg) met een afgepaste hoeveelheid (12 wassen), veronderstelt bij de huisvrouw een prijsnadeel. Men is namelijk van mening, dat men met een 2 kg verpakking van een ander merk meer dan 12 wassen doet (gemiddeld 16 wassen), terwijl het aantal wassen dat men met 'DRIEHOEK WASBUILTJES' kan doen per definitie 12 is.

Meerdere ontwerpverpakkingen werden vervolgens getest op produkt- en merkidentificatie, waaruit bleek dat de onderlinge verschillen niet groot waren. Een bevinding, die werd bevestigd in een onderzoek dat voorafging aan de testmarktoperatie. Op basis van eigen judgment werd tenslotte een keuze gemaakt voor een bepaald verpakkingsontwerp.

Het verpakkingsonderzoek werd op een tweetal momenten in de ontwikkeling van het produkt uitgevoerd, namelijk:

a. bij de keuze van de merknaam Ecotex, die heel lang als te hanteren merknaam in de markt, is gebruikt;
b. bij de wisseling naar de merknaam 'DRIEHOEK WASBUILTJES'.

4.2.3. *Communicatie-onderzoek*

Het advertentie-onderzoek bestond uit een kwalitatief onderzoek, waarin middels diepte-interviews een drietal alternatieve concepten werd getest, op: begrijpelijkheid, onduidelijke elementen en de mate waarin de drie claims 'gemakkelijk', 'fosfaatvrij' en 'goed schoon' overgebracht kunnen worden.

De ontworpen advertenties bleken snel en zonder problemen te communiceren:

- dat er een nieuw wasmiddel bestaat in de vorm van een zakje;
- dat in de trommel gelegd moet worden;
- dat geschikt is voor wassen van 60-95°C;
- dat het om een fosfaatvrij produkt gaat.

Ook de ontwikkelde t.v.-commercials werden aan een oriënterend onderzoek onderworpen voor wat betreft communicatieve waarde, op basis waarvan één commercial werd gekozen voor de testmarkt.

De uitkomsten van deze fasen overziend, bestond de indruk dat men een afgerond produkt bezat, dat geschikt was voor een proefmarktintroductie. Het produkt-acceptatie-onderzoek kon dan ook van start gaan.

4.2.4. *De merknaam*

Alvorens een beschrijving te geven van de volgende fase, dient eerst nog nader te worden ingegaan op de merknaam die voor dit nieuwe produkt gehanteerd wordt. In de testmarkt-operatie werd gekozen voor de naam 'Ecotex', een merknaam die een duidelijk verband legt met de milieu-claim ('ecologie') en het nieuwe karakter (revolutionaire) van het produkt. Tijdens de zich elkaar snel opvolgende produkt-tests, werd om concern-politieke overwegingen gekozen voor de merknaam 'DRIEHOEK WASBUILTJES'.

Een onderzoek naar de voorkeur van de consument voor één van beide merknamen, gaf bovendien een voorkeur te zien voor 'DRIEHOEK WASBUILTJES' boven Ecotex. Aangezien de merknaam Driehoek daarnaast geen negatieve associaties leek te hebben met de nieuwe produktpropositie, werd uiteindelijk voor deze merknaam gekozen.

4.3. *Het produkt-acceptatie-onderzoek*

In het produkt-acceptatie-onderzoek is middels een aantal produkt-tests inzicht verkregen in de waswerking van het produkt en het oordeel van de huisvrouw over 'DRIEHOEK WASBUILTJES' en in een 'simulated' testmarkt omtrent het te verwachten marktaandeel.

In een *eerste* produkt-test (in de vorm van een kwantitatieve pilot-studie) werd nagegaan hoe het nieuwe produkt in de praktijk beoordeeld wordt, met name t.a.v. het functioneren van de sachets. Tevens werd in deze test getracht een eerste inzicht te verkrijgen in de prijsstelling van het produkt. Met dat doel hebben huisvrouwen gedurende een periode van 8 weken met het produkt gewassen. De belangrijkste conclusies die op basis van deze test konden worden getrokken, zijn:

- onvolkomenheden aan de sachets;

- men niet bereid was meer te betalen, dan de gemiddelde prijs van andere wasmiddelen;
- het gemiddeld aantal wassen, dat men met een ander wasmiddel van 2 kg doet wordt ingeschat op 16 (4 meer dan met het testprodukt, waardoor bijkomende prijsnadelen ontstaan);
- een in het algemeen positief oordeel over het testprodukt en in vergelijking met het eigen merk de volgende positieve aspecten (geur, schoon/wit worden van de was en het zacht aanvoelen van het gewassen goed).

Op basis van de resultaten van deze pilot-studie werd besloten om in een *tweede* produkt-test op grotere schaal inzicht te verkrijgen in de waswerking van het vernieuwde produkt (op specifieke onderdelen) en een totaal-beoordeling. Uit deze test resulteerde een aantal produktonvolkomenheden, dat het noodzakelijk maakte deze test voortijdig te beëindigen. In de fabricage werd een aantal produktaanpassingen doorgevoerd, waardoor het produkt geschikt leek om in een proefmarkt te worden opgenomen.

Na een uitgebreide studie van de onderzoeksmogelijkheden van een proefmarkt, werd besloten het produkt niet in een testmarkt uit te zetten. De volgende overwegingen hebben aan deze beslissing ten grondslag gelegen (5):
- de duur en de kosten van een testmarktoperatie;
- het vrijgeven van het totale produktconcept aan de concurrentie en de mogelijkheden die de concurrentie heeft om de test te verstoren (positieve verstoring, actieve verstoring en manipulatie) (2);
- het vinden en nader afbakenen van geschikte gebieden in Nederland (met name in termen van media-afbakening);
- de problemen, die een testmarkt onderzoekstechnisch met zich meebrengt.

Als alternatief voor introductie van het produkt in een testmarkt werd gekozen voor een in Amerika ontwikkeld model, dat als 'gesimuleerde testmarkt' onder de naam Assessor bekend is (4). Dit model geeft de fabrikant primair inzicht in het marktaandeel van het testprodukt na afsluiting van de introductiecampagne, dus in een gestabiliseerde marktsituatie.

Daarnaast wordt inzicht verkregen in de marktbewegingen die ontstaan als gevolg van de introductie van een nieuw merk, de reacties van de consument op het produkt zelf, de advertentie-campagne en de prijsstelling. Het model gaat ervan uit, dat de fabrikant het produkt, de verpakking en de advertentie in zijn meest finale stadium klaar heeft en de totale introductie in een marketing-plan heeft neergelegd.

De voorspelling van het marktaandeel van het nieuwe produkt vindt in Assessor zijn oorsprong in twee basismodellen, namelijk enerzijds het preferentiemodel en anderzijds het 'trial-repeat'-model. In het eerste wordt een relatie gelegd tussen merkenpreferentie en koopwaarschijnlijkheid terwijl het tweede te werk gaat volgens het bekende principe van aankoop- en herhalingsaankoopcijfers.
Additioneel werd in dit onderzoek de variabele 'prijs' ingebracht, teneinde enig inzicht te verkrijgen in de prijselasticiteit van het produkt.

Op basis van de resultaten van deze gesimuleerde testmarkt, waaruit bleek dat er zeker een markt voor dit produkt bestond, werd tevens geconcludeerd dat nadere verfijningen ten aanzien van de werking van het wasmiddel (i.c. het wasbuiltje) noodzakelijk waren.

In volgende produkt-tests werden de aangebrachte produktveranderingen nauwkeurig onderzocht en werd steeds de relatie gelegd met de totaalbeoordeling van het produkt. Achtereenvolgens zijn de volgende tests uitgevoerd, met bij elk steeds aangegeven de produktveranderingen die op basis van de uitkomsten van het onderzoek zijn aangebracht:

1. een *5-wassen-test*, waaruit een aantal produkt-onvolkomenheden naar voren kwam, dat te maken had met de 'performance', de geur en open zakjes in de verpakking. Met name het onderscheid in beoordeling van het wasresultaat tussen huisvrouwen die normaliter met enzymatische en niet-enzymatische wasmiddelen wasten, had grote gevolgen. Besloten werd immers om enzymen aan de wasbuiltjes toe te voegen, die de performance van het produkt (m.n. vlekverwijdering) verhoogden. Daarnaast werden veranderingen in de 'trigger' (de lijmnaad) en in het 'parfum' aangebracht;
2. een *tussen-test,* waarin een viertal produktie-alternatieven werd onderzocht. Er deden zich bij de produktie namelijk problemen voor, die betrekking hebben op het persen van het produkt tot korrels. Nagegaan werd of deze alternatieven geen oplosproblemen gaven bij het wassen. Op basis van de uitkomsten van dit onderzoek werd voor één van de alternatieven gekozen;
3. een *langdurige produkt-test,* waarin huisvrouwen gedurende een periode van 10 weken met het testprodukt moesten wassen. Doel was het verkrijgen van diepgaand inzicht in de waswerking van het vernieuwde produkt en de beoordeling, nadat men een langere periode ermee heeft gewassen. Uit de resultaten van deze test kan worden geconcludeerd, dat de 'performance' (waswerking) van het produkt positief beoordeeld wordt zowel voor de witte als voor de bonte was en de geur eveneens positief beoordeeld wordt. Niets stond introductie van 'DRIEHOEK WASBUILTJES' op de Nederlandse markt nog in de weg.

5. DE INTRODUCTIE VAN 'DRIEHOEK WASBUILTJES'

In oktober 1981 werd 'DRIEHOEK WASBUILTJES' op de Nederlandse markt geïntroduceerd (bij de consument), omgeven met een uitgebreide p.r. en reclame-inspanning. Het totale budget aan media-reclame bedroeg in dat jaar ruim 1 miljoen gulden, waarvan globaal genomen 1/3 deel aan t.v.-reclame, 1/3 in dagbladen en 1/3 in vrouwenbladen besteed werd.

De opening van de fabriek in Nijmegen gebeurde door de toenmalige staatssecretaris voor Volksgezondheid en Milieuhygiëne en haalde het NS-journaal van acht uur. In diverse dag- en weekbladen werd op uitgebreide schaal redactionele aandacht aan het produkt besteed.

Zeer snel werd een hoog distributieniveau bereikt, waarbij in de winkels speciaal display-materiaal voor extra aandacht zorg droeg. Daarnaast werden ruim 1 miljoen proefpakjes met 3 wasbuiltjes verkocht, teneinde zoveel mogelijk

huisvrouwen in staat te stellen kennis te nemen van het produkt. Als belangrijkste produktclaims werden in reclame-uitingen en op de verpakking genoemd: handig, wast goed schoon en is fosfaatvrij.

6. ONDERZOEK NA DE INTRODUCTIE

Het onderzoek, dat het produkt begeleidt na de introductie kan worden onderverdeeld, in:
- een *bekendheidsmeting* van 'DRIEHOEK WASBUILTJES' op meerdere tijdstippen in het eerste jaar na de introductie;
- een voortdurende cijfermatige begeleiding middels Attwood- en Nielsen *panel-data*;
- het volgen van verschillende groepen van koopsters in de loop van de tijd, middels *telefonisch ad-hoc onderzoek*. Met name inzicht in de trouw aan 'DRIEHOEK WASBUILTJES' en de motivatie die daaraan ten grondslag ligt, is van groot belang voor verdere marketing-inspanningen;
- het meten van het effect van een *'cash-refund' aktie,* waarbij op elk pak 'DRIEHOEK WASBUILTJES' een bedrag van f 5,– werd terugbetaald.

Aangezien deze onderzoeksaanpak een grote overeenkomst vertoont met andere produkten in de consumentensfeer, zal in dit artikel niet verder hierop worden ingegaan.

7. BESLUIT

In dit artikel is een beschrijving gegeven van het ontwikkelingsproces van 'DRIEHOEK WASBUILTJES' en de rol, die marktonderzoek daarbinnen gespeeld heeft. Het produkt is begeleid vanaf het prille begin (ongeveer 7 jaar geleden) tot de huidige situatie. In bepaalde fasen van dit ontwikkelingsproces heeft marktonderzoek de beslissing gerechtvaardigd om op de ingeslagen weg door te gaan, in andere fasen zijn uitkomsten van onderzoeken aanleiding geweest tot een nadere bezinning op het produkt-concept. Het samenspel tussen marktonderzoekers aan de ene kant en technici aan de andere kant heeft uiteindelijk geleid tot een binnen de bestaande technische mogelijkheden, optimaal produkt. De systematische integratie van marktonderzoek in het ontwikkelingsproces heeft tenslotte op zijn minst ervoor gezorgd, dat de risico's die de fabrikant (in dit geval: Kortman & Schulte) liep bij de introductie van 'DRIEHOEK WASBUILTJES' tot een minimum gereduceerd waren. Het nauwgezet blijven volgen van het produkt in de eerste jaren na de introductie zal echter noodzakelijk zijn.

Een wasmiddel vertoont ondanks zijn specifieke eigenschappen natuurlijk grote overeenkomsten met andere produkten in de consumentensfeer, terwijl zelfs parallellen te trekken zijn met industriële goederen. Binnen elk van deze segmenten kan marktonderzoek, mits op het juiste moment en verantwoord toegepast, een belangrijke bijdrage leveren aan produkt-ontwikkeling, introductie en het volgen van nieuwe produkten gedurende de gehele 'product-lifecycle'.

BIBLIOGRAFIE

1. *Distrifood,* 'Toppers scoren hoog in een grossiersenquête; wasmiddel nummer één', 23 augustus 1982, pag. 1.
2. Eunen, E.A. van, 'Het verstoren van de testmarkten', *Tijdschrift voor Marketing,* juni 1982, nr. 6, 47-51.
3. Kotler, Ph., 'Marketing management' *Prentice/Hall International Editions,* London, 1976.
4. Silk, A.J. en Urban, G.L., 'Pre-Test-Market Evaluation of New Packaged Goods: A Model and Measurement Methodology', *Journal of Marketing Research,* vol. XV (May 1978) 171-191.
5. Vélu, G., 'De voor- en nadelen van testmarketing', *Tijdschrift voor Marketing,* juni 1977, nr. 6, 137-141

9. Non response bij schriftelijk onderzoek

R.E. VAN DEN BERG

SAMENVATTING

In dit artikel wordt nagegaan in hoeverre de datum van binnenkomst een valide inschatter is om non response te meten bij schriftelijk onderzoek. Ook wordt bekeken of er verband is tussen de datum van binnenkomst en de antwoorden op een aantal in het marktonderzoek gebruikelijke variabelen.

1. INLEIDING

Niet-onderzoekers zien als functie van onderzoek vaak niet meer dan een op maat gesneden mini-census met voor hen relevante gegevens. Uitkomsten blazen zij vaak op naar bepaalde populaties. De eisen die zij aan onderzoek stellen zijn:

- Zeer grote steekproeven waarbij ieder lid van de onderzochte populatie een van te voren bekende kans heeft gehad om in de steekproef te worden opgenomen.
- Eenduidigheid in en begrijpelijkheid van vragen.
- De onderwerpen die zij relevant vinden zijn over het algemeen registrerend van aard. In de praktijk komt dit vaak neer op het meten van de bekende socio-economische- en geografische kenmerken, activiteiten en meningen. Een bekend voorbeeld waar men al dit soort elementen in kan terug vinden is het Nationaal Onderzoek Persmedia.
- Uitkomsten hebben vaak geen directe invloed op het te voeren beleid op korte termijn. Zij dienen om bepaalde conclusies te voorzien van achtergrondinformatie.

Ik heb het idee dat men vaak ten onrechte veel belang hecht aan omvang en representativiteit van onderzoek. In de gevallen dat men door middel van kwalitatief onderzoek een inventarisatie wil maken van meningen, problemen, voorkeuren en weerstanden over bepaalde zaken, spelen kwantiteiten geen enkele rol. In dit verband wil ik ook vragenlijsten noemen, die men in tijdschriften en dergelijke afdrukt. De response, uitgedrukt in percentages van het aantal lezers, is vaak niet hoger dan 5%. Toch kan dit soort materiaal zeer goed dienen als inventarisatie van meningen, problemen, gedragingen en hun onderlinge samenhang zonder kwantitatieve uitspraken te doen.

Wil men kwantitatieve uitkomsten inventariseren, waarbij enige procenten meer of minder voor een beleid er niet op aan komen, met andere woorden een

globale indruk krijgen, dan kan men de eisen ten aanzien van representativiteit en steekproefomvang wat minder scherp stellen.
Indien vrij exact kwantificeerbare onderzoekresultaten gewenst zijn, dan dient men zeer hoge eisen te stellen aan representativiteit, steekproefomvang en formulering van de vragen.

Over één aspect van representativiteit namelijk de invloed van de non response op de onderzoekresultaten bij schriftelijk onderzoek wil ik het in deze beschouwing hebben.

2. DOEL VAN HET ONDERZOEK

Vaak hoort men bij schriftelijk onderzoek het verwijt: 'de non response is zo groot, daarom is er een redelijke kans dat de uitkomsten niet representatief zijn'. Ook zou de kwaliteit van de ingevulde vragenlijsten te wensen over laten. Als reden van non response wordt dan vermeld dat de betrokkenheid bij het onderzoek laag zou zijn.
Bij de ANWB doet zich de laatste jaren het verschijnsel voor, dat de response bij schriftelijk onderzoek afneemt onder gelijkblijvende omstandigheden. Weliswaar niet dramatisch, maar toch ieder jaar één of twee procent. Het response percentage komt nu nauwelijks boven de 70% na 2 reminders. Terwijl vijf jaar geleden het response percentage bijna 80% was.
Een en ander is aanleiding geweest om onderstaande hypothesen te formuleren.

a. Er is verband tussen tijdstip van inzenden en non response bij schriftelijk onderzoek met betrekking tot:
   - socio-economische-, regionale- en demografische kenmerken van de ondervraagden;
   - betrokkenheid bij onderzoek;
   - betrokkenheid bij de organisatie waarvan de enquête uitgaat;
   - mate waarin men bepaalde in het onderzoek genoemde activiteiten verricht.
b. Naarmate men later inzendt, heeft men steeds meer overeenkomst met degenen die in het geheel niet terugzenden.
c. Naarmate men later inzendt, wordt de kwaliteit van de ingevulde vragenlijsten relatief minder.

In dit verslag zal achtereenvolgens aan de orde komen
- het response verloop
- de kwaliteit van het ingezonden materiaal in relatie tot het responseverloop
- resultaten van een aantal gegevens die ook van de non-response bekend zijn
- beantwoording van een aantal vragen in het onderzoek in relatie tot het response verloop
- conclusies met betrekking tot de geformuleerde hypothesen.

3. RESPONSEVERLOOP

Jaarlijks voert de ANWB een schriftelijk onderzoek uit onder eigen leden. Het

doel van dit onderzoek is om cijfermatige inzichten te krijgen in:
- de samensteling van het ledenbestand wat socio-economisch en geografische aspekten betreft;
- vakantiepatroon;
- aankoop- en rijgedrag van automobilisten;
- mening over de ANWB;
- gebruik- en aankoopgedrag van een aantal produkten en diensten van de ANWB.

De onderwerpen zijn vrij algemeen van aard en de populatie ANWB-leden dekt meer dan de helft van de Nederlandse huishoudingen.
Via een systematische steekproef werden 1.500 ANWB-leden getrokken uit het ANWB-ledenbestand. Van het responseverloop werden de volgende groepen samengesteld.

	Aantal ondervraagden			
	abs.		*%*	
1. bruikbare response vóór verzending v.d. 1e reminder in bezit	550		37	
2. bruikbare response ná verzending 1e reminder en vóór verzending 2e reminder in bezit	298		20	
3. bruikbare response ná verzending 2e reminder en vóór verzending 3e reminder in bezit	234	1.138	15	76
4. bruikbare response ná verzending 3e reminder in bezit	56		4	
5. onbruikbare response (adres onjuist, niet ingevuld retour etc.)		41		3
6. non response (geen enkele reactie bekend)		321		21
Totaal		1.500		100%

Wat opvalt is het hoge responsepercentage van 76% in vergelijking met ander schriftelijk onderzoek.
Dit percentage werdt veroorzaakt door:
- begeleidende brieven die motiveren tot inzenden
- het stellen van vragen waarvan bekend is dat men ze graag invult
- een extra 3e reminder die 4% response opleverde.

Hoewel de inzenders op de 3e reminder wel in de verdere analyse worden betrokken moeten we gezien het kleine aantal van 56 rekening houden met grote afwijkingen.

De onbruikbare response is wel opgenomen in de totale steekproef, maar niet in de detailleringen omdat ze over het hele responseverloop verspreid voorkwam.

De eerste reminder werd niet verzonden met een vragenlijst, de 2e en 3e re-

minder wel. De reminders werden verzonden 2 weken, 5 weken en 8 weken na de eerste verzending. Het veldwerk werd gehouden van donderdag 2 december 1982 tot vrijdag 18 februari 1983. Eventueel later binnengekomen formulieren werden tot de categorie onbruikbaar gerekend.

4. KWALITEIT VAN HET INGEZONDEN MATERIAAL

Schriftelijk onderzoek heeft als nadeel dat men de potentiële invuller niet kan leren hoe hij vragenlijsten moet beantwoorden dit in tegenstelling tot een bekwame enquêteur. Dit houdt in dat er
- vragenlijsten geheel onbruikbaar zijn
- vragenlijsten voor een deel onbruikbaar zijn
- vragenlijsten op arbitraire wijze worden verbeterd bij controle

Dit laatste houdt in dat men bij incorrect gegeven antwoorden nagaat of men aan de hand van het antwoordenpatroon van de ondervraagde de incorrecte invulling kan herstellen.

We moeten hierbij denken aan bepaalde doorverwijzingen die niet opgevolgd worden (bijvoorbeeld: iemand is in een bepaald jaar niet op vakantie geweest, ziet de doorverwijzing over het hoofd, en gaat toch een vakantie uit een andere periode invullen). Dit soort vragen zijn wel logischer wijze te herstellen maar 100% zeker is men niet. Te veel van dit soort aanpassingen zijn dus niet bevorderlijk voor de kwaliteit van het onderzoek.

Uit staat 1 blijkt duidelijk dat naarmate men later verzendt het aantal ten onrechte niet ingevulde vragen en het aantal incorrect gegeven antwoorden op vragen toeneemt.
In feite zijn de verschillen nog extremer, omdat de latere inzenders over het algemeen minder vragen hadden te beantwoorden.

Staat 1. Gemiddeld aantal ten onrechte niet ingevulde vragen en gemiddeld aantal incorrect gegeven antwoorden

			response verloop			
		Totaal bruikbare respons	1 vroege inzenders	2	3	4 late inzenders
Gemiddeld aantal ten onrechte niet ingevulde vragen per inzender		2,0	1,7	1,8	2,4	4,5
Gemiddeld aantal inkorrekt gegeven antwoorden		2,4	1,8	2,5	3,2	3,4
aantal ondervraagden		1138	550	298	234	56
aantal vragen dat men moest beantwoorden	minimaal		28			
	maximaal		70			
	gemiddeld		60			

5. RESULTATEN VAN EEN AANTAL GEGEVENS DIE OOK VAN DE NON RESPONSE BEKEND ZIJN

In deze paragraaf zullen wij nagaan in hoeverre er verbanden bestaan tussen response en non response ten aanzien van kenmerken die uit de administratie bekend zijn. Al te bedrijfsgevoelige kenmerken worden niet met name genoemd.

Uit staat 2 blijkt dat er verbanden bestaan tussen response en non response ten aanzien van
- urbanisatiegraad
- lidmaatschapsduur
- het bezitten van bepaalde overige kenmerken

Staat 2. Kenmerken van de leden die uit de administratie bekend zijn

	Totale steek proef	Bruikbare response					non response exclusief onbruikbaar retour
		Totaal bruikbaar	response verloop				
			1 vroege inzenders	2	3	4 late inzenders	
	in % van het aantal ondervraagden per categorie						
A *regionaal gebied*							
Noord (Gr. Fr. Dr.)	9	8	10	8	8	2	9
Oost (Ov. Gel.)	18	19	20	17	19	7	19
West (Utr. NH. ZH.)	51	50	49	49	49	77	51
Zuid (Zeel. NB. Lim.)	22	23	21	26	24	14	21
Totaal	100%	100%	100%	100%	100%	100%	100%
B *urbanisatie graad*							
Platte land (A1 t/m A4+B1 gemeenten)	15	16	16	17	16	12	15
Kleine gemeenten (B2 + C1 t/m C3 gemeenten)	30	31	32	30	31	27	25
Grote steden + forensen gemeenten (C4, C5 + B3 gemeenten)	55	53	52	53	53	61	60
Totaal	100%	100%	100%	100%	100%	100%	100%
C *lidmaatschapsduur*							
2 jaar of korter lid	11	10	9	9	18	9	16
3 t/m 5 jaar lid	21	20	18	17	25	25	25
6 t/m 10 jaar lid	25	25	24	29	22	23	23
11 t/m 15 jaar lid	19	19	19	23	15	21	18
langer dan 15 jaar lid	24	26	30	22	20	22	18
Totaal	100%	100%	100%	100%	100%	100%	100%

Staat 2 (Vervolg).

	Totale steek proef	Bruikbare response: Totaal bruikbaar	response verloop: 1 vroege inzenders	2	3	4 late inzenders	non response exclusief onbruikbaar retour
	in % van het aantal ondervraagden per categorie						
D *overige kenmerken*							
Kenmerk I							
ja	83	84	84	82	83	95	81
nee	17	16	16	18	17	5	19
Totaal	100%	100%	100%	100%	100%	100%	100%
Kenmerk II							
ja	33	35	37	39	25	34	25
nee	67	65	63	61	75	66	75
Totaal	100%	100%	100%	100%	100%	100%	100%
Kenmerk III							
ja	51	54	57	55	49	48	40
nee	49	46	43	45	51	52	60
Totaal	100%	100%	100%	100%	100%	100%	100%
Kenmerk IV							
ja	9	10	12	10	6	11	5
nee	91	90	88	90	94	89	95
Totaal	100%	100%	100%	100%	100%	100%	100%
aantal ondervraagden	1500	1138	550	298	234	56	321

We zien dat de non response vergeleken met de totale bruikbare response relatief vaak wordt aangetroffen bij
- de grote gemeenten
- de ANWB leden die korter dan 5 jaar lid zijn
  en relatief minder vaak bij de overige kenmerken II, III en IV.

Dit soort kenmerken beschouwen wij als een vorm van ledenbinding. Leden met ledenbindende kenmerken, waartoe ook het langer lid zijn van de ANWB gerekend kan worden, zenden sneller in.

6. BEANTWOORDING VAN EEN AANTAL VRAGEN IN HET ONDERZOEK IN RELATIE TOT HET RESPONSEVERLOOP

Uit staat 3 blijkt dat inzenders
– boven de 50 jaar – met hoger beroeps of universitair onderwijs – die vaker op vakantie gaan – met een nieuw gekochte auto sneller reageren

Staat 3. Een aantal vragen uit het onderzoek naar response verloop

	Totaal bruikbare response	response verloop			
		1 vroege inzenders	2	3	4 late inzenders
	in % van het aantal ondervraagden per categorie				
A *geslacht*					
man	83	85	83	80	84
vrouw	17	15	17	20	16
Totaal	100%	100%	100%	100%	100%
B *leeftijd*					
jonger dan 25 jaar	6	5	6	9	9
25 t/m 34 jaar	25	23	25	31	25
35 t/m 49 jaar	35	37	32	34	44
50 t/m 64 jaar	25	25	28	21	20
65 jaar en ouder	9	10	9	5	2
Totaal	100%	100%	100%	100%	100%
C *hoogst genoten schoolopleiding*					
lager onderwijs of lager beroepsonderwijs	34	29	39	36	38
uitgebreid lager of middelbaar onderwijs	40	41	37	43	40
hoger beroeps of universitair onderwijs	26	30	24	21	22
Totaal	100%	100%	100%	100%	100%
D *aantal malen op vakantie okt. 1981 t/m sept. 1982*					
0×	19	16	22	24	15
1×	52	51	52	53	56
2× of vaker	29	33	26	23	29
Totaal	100%	100%	100%	100%	100%
aantal inzenders	1138	550	298	234	56
E *wijze waarop de enige of meest gebruikte auto in het gezin werd aangeschaft*					
nieuw	54	58	52	45	53
gebruikt	46	42	48	55	47
Totaal	100%	100%	100%	100%	100%
aantal inzenders dat over één of meer auto's beschikt	1098	536	287	222	53

Jongeren onder de 35 jaar en zij met lager onderwijs of lager beroepsonderwijs hebben iets meer aansporing nodig om te reageren.

Uit het onderzoekmateriaal zijn een aantal vakantiesegmenten geconstrueerd op basis van een door het IVA, in opdracht van de ANWB, ontwikkelde methode.

Een gedetailleerde beschrijving van de segmentatiemethode in de vakantiesegmenten staat in het IVA rapport: "Vakantiegangers en autobezitters, een marktsegmentatieonderzoek: Een onderzoek naar gedragingen en houdingen van vakantiegangers, autobezitters, recreanten, leden van milieu-organisaties en ANWB-leden". (Oppedijk van Veen, maart 1982).

Geen van de vakantiesegmenten bleek samenhang te vertonen met het responseverloop behalve het strandvakantiesegment. Van degenen die tot dit segment behoren is reeds bekend dat:

- ze het landschap en de natuur niet zo belangrijk vinden
- ze wel op het strand liggen, zonnebaden en zwemmen
- ze niet wandelen in de bossen of bergen.

We kunnen nu nog toevoegen dat ze ook niet van vragenlijsten invullen houden.

De betrokkenheid bij dit onderzoek is aan het eind van de vragenlijst gemeten door middel van de volgende 6 uitspraken.

A Ik had het eigenlijk niet willen doen, maar ik heb het toch maar gedaan.
B Ik was wat aarzelend en heb even overwogen om het niet te doen.
C Ik vond het niet bepaald leuk, maar overwegen om het niet te doen was er niet bij.
D Ik dacht waarom niet, als ik er iemand een plezier mee doe.
E Ik vond het wel eens aardig om te kijken wat er zoal gevraagd werd.
F Ik vond het leuk dat er ook eens naar mijn mening werd gevraagd.

Deze uitspraken zijn gebaseerd op het artikel van J.J.M. van Tulder 'Op de grens van non response' gepubliceerd in het jaarboek van de NVM 1977.

In de nu volgende staat 4 zijn de uitspraken samengevoegd tot 3 groepen t.w. negatief (A + B), neutraal (C + D) en positief (E + F) en vergeleken met de uitkomsten uit het jaarboek.

Uit staat 4 blijkt in de eerste plaats het grote verschil in betrokkenheid tussen de uitkomsten in het NVM artikel en dit onderzoek. Dit zal voor een deel te wijten zijn aan het feit dat bij schriftelijk onderzoek van enige omvang nogal een beroep wordt gedaan op de zelfwerkzaamheid van de ondervraagden. Bij mondeling onderzoek kunnen de ondervraagden zich passiever opstellen.

Verder zien we dat naarmate men later inzendt, men negatiever over het onderzoek denkt. Een en ander neemt niet weg, dat er niet onaanzienlijke aantallen vroege inzenders zijn die negatief oordeelden en niet onaanzienlijke aantallen late inzenders die het invullen van de vragenlijst toch als een positieve ervaring beleefden.

Van de groep inzenders die op de meest negatieve uitspraak scoorde is voor wat betreft de kenmerken zoals vermeld in staat 2 nagegaan in welke mate de uitkomsten leken op die van de overige response. Dit was voor een aantal ken-

Staat 4. Reakties die het dichtst komen bij hetgeen men dacht toen men de vragenlijst invulde naar response verloop

Betrokkenheid bij onderzoek	NVM 1977 (v Tulder)	Totaal bruikbare respons	response verloop 1 vroege inzenders	2	3	4 late inzenders
	in % van het aantal ondervraagden per categorie					
negatief	15	26	9	32	47	60
neutraal en geen antwoord	49	46	54	46	35	24
positief	36	28	37	22	18	16
Totaal	100%	100%	100%	100%	100%	100%
aantal ondervraagden	8523	1138	550	298	234	56

merken het geval, met name bij die kenmerken welke een binding met het ANWB lidmaatschap aangaven.

7. CONCLUSIES EN AANBEVELINGEN

Wanneer we terugkeren naar de geformuleerde hypothesen aan het begin van dit verslag dan kunnen we concluderen dat:

1e Er verband is tussen tijdstip van inzenden en non response bij schriftelijk onderzoek voor wat betreft urbanisatiegraad en binding met de ANWB.

2e Er verband is tussen het tijdstip waarop men inzendt en leeftijd, opleiding, betrokkenheid bij onderzoek, vakantieparticipatie, soort vakantie (met name de strandvakantie) en aankoopgedrag auto.

3e De kwaliteit van de ingevulde vragenlijsten minder wordt naarmate men later inzendt.

Bij ad hoc schriftelijk onderzoek lijkt het aan te bevelen rekening te houden, o.a. met behulp van herwegingsprocedures, met het volgende

- non response analyse voor wat betreft regionale kenmerken
- analyse van de datum van binnenkomst in kombinatie met de bijna non response analyse gebaseerd op de door Van Tulder ontwikkelde 'Op de grens van non response'-variabelen
- het versturen van 2 reminders. Een derde reminder levert nauwelijks iets op in aantallen, de kwaliteit is onder de maat en men ondervindt onevenredig veel weerstanden.

Deze zaken zijn bij ieder schriftelijk onderzoek probleemloos in het onderzoekmodel op te nemen.

Tot slot nog een kleine waarschuwing.

De hier vermelde responsepercentages zijn uitzonderlijk hoog voor schrifte-

lijk ad-hoc onderzoek. Ze zijn toe te schrijven aan de binding met de organisatie van wie de enquête uitgaat en de grote belangstelling voor de onderwerpen in de vragenlijst. In de vele gevallen zal men aan deze voorwaarden niet kunnen voldoen.

10. Een vergelijkende analyse van besluitvorming bij aankopen in gezinsverband*

BRONISLAW VERHAGE, ROBERT T. GREEN
EN JEAN-LOUIS CHANDON

SAMENVATTING

Marktonderzoekers hebben slechts in beperkte mate getracht de culturele dimensie in de kennis van het consumentengedrag te integreren. Dit artikel richt zich op de cross-culturele dimensie van de rol die beide echtgenoten spelen bij de aankoop van produkten en diensten. Eerst wordt een theorie van Rodman (1972) besproken die probeert om de uitkomsten van onderzoek naar de machtsverhoudingen tussen man en vrouw in een culturele context te plaatsen. Vervolgens wordt een analytische techniek – correspondence analyse – toegelicht die zich goed leent voor het analyseren van besluitvormingspatronen waarvan de gegevens in verschillende landen verzameld zijn. Tenslotte presenteren wij de resultaten van een onderzoek naar de aankoopbeslissingen van huishoudens in vijf landen om te illustreren op welke wijze de theorie van Rodman en de correspondence analyse techniek kunnen worden toegepast in internationaal consumentenonderzoek.

1. INLEIDING

Een belangrijk onderzoeksveld voor markt- en consumentenonderzoek is de besluitvorming binnen het gezin (Davis 1976). Een deel van de belangstelling richt zich op de rol die beide echtgenoten spelen bij de aanschaf van produkten en diensten (zie bijvoorbeeld Burns and Granbois 1979, Green, Verhage and Cunningham 1981, Jenkins 1979). Immers, door allerlei maatschappelijke ontwikkelingen is het rolpatroon tussen man en vrouw aan verandering onderhevig (Verhage and Green 1981, Arndt and Crane 1974). Zo zal het recente voorstel om de arbeidstijd in Nederland op den duur met een tiende te verminderen werkende gezinsleden meer vrije tijd bezorgen. Die extra vrije tijd – dit blijkt uit onderzoek van het Sociaal Cultureel Planbureau (Een week tijd, 1977) – wordt gedeeltelijk besteed aan het huishouden en gezin**. Vooral als de toename van vrije tijd gepaard gaat met een daling van het gezinsinkomen kan dit leiden tot een andere taakverdeling tussen man en vrouw bij regelmatig terugkerende aankoopbeslissingen. Voor ondernemingen die consumentengoederen op de markt brengen is een juist inzicht hierin essentieel. Als de aankoopgewoonten van de consument zich wijzigen moeten talloze bedrijven immers hun marketingstrategie aanpassen. Dat geldt zowel voor de Nederlandse markt als voor bedrijven, die buiten onze landsgrenzen actief zijn.

* Dit onderzoek werd uitgevoerd met medewerking van I.C.M. Cunningham, J.P. Leonardi en A. Strazzieri.

** Zie ook het commentaar op de verkorting van de arbeidstijd van A. de Swaan, 'De vrijetijdsverlenging', in *NRC-Handelsblad* van 16 en 17 juni 1982.

In de literatuur over aankoopbeslissingen in gezinsverband is één van de voornaamste tekortkomingen het ontbreken van internationaal georiënteerd onderzoek. Het gros van het bestaande onderzoek komt uit de Verenigde Staten. Bovendien is het onderzoek dat niet van Amerikaanse oorsprong is, meestal zo beperkt van omvang dat het zich er niet voor leent om in een internationale context te worden geplaatst (Christopher 1975, Davis and Rigaux 1974, Hempel 1974, Michel 1971).

Dit artikel vormt een poging om één van de meer fundamentele aspecten van het aankoopgedrag in gezinshuishoudingen – de rolverdeling tussen man en vrouw bij aankoopbeslissingen – in een cross-culturele context te plaatsen. Daarbij valt te verwachten dat het rolpatroon tussen de echtelieden mede afhankelijk is van de ontwikkelingsfase, waarin een maatschappij zich bevindt; zo is de emancipatie van de vrouw in sommige landen een onbekend verschijnsel. Voor dit onderzoek hebben wij dan ook aansluiting gezocht bij de bestaande sociologische theorie over gezinsbeslissingen in verschillende typen maatschappijen. Bij onze analyse ligt de nadruk op structurele verschillen in de rolverdeling bij gezinsaankoopbeslissingen tussen landen in verschillende stadia van economische ontwikkeling.

2. ACHTERGROND VAN HET ONDERZOEK

De hypothesen in dit verkennend onderzoek werden ontwikkeld op basis van een indeling van verschillende maatschappijen door Hyman Rodman (1972). Rodman heeft getracht om maatschappijen te rangschikken op grond van de machtsverhouding tussen de echtelieden. Deze sociologische graadmeter blijkt nogal wat samenhang te vertonen met het stadium van economische ontwikkeling, waarin een land zich bevindt. De classificatie is weliswaar onvolledig – Rodman beschikte niet over vergelijkbare gegevens uit alle landen – maar brengt toch enige lijn in de bestaande onderzoeksgegevens over het rolpatroon tussen de echtgenoten in diverse culturen.

De indeling van Rodman – die gebaseerd is op onderzoeksresultaten uit tien landen – bestaat uit vier ideaal-typen maatschappijen die elk een verschillend stadium van maatschappelijke ontwikkeling vertegenwoordigen. Elk van deze typen zullen wij nu kort toelichten.

Fase 1: Patriarchy

Maatschappijen in de patriarchale fase onderscheiden zich door '...strong patriarchal family norms, with a high level of paternal authority, and with no variation in paternal authority from one stratified group to another'. (Rodman, 1972, blz. 63). De maatschappijen in Fase 1 hechten een hoge waarde aan tradities en staan wat economische ontwikkeling betreft grotendeels nog in de kinderschoenen. Als voorbeeld van een land, dat zich in Fase 1 bevindt – hoewel zich daar in versneld tempo belangrijke ontwikkelingen afspelen – noemt Rodman India.

Fase 2: Modified Patriarchy

Maatschappijen in Fase 2 zijn die '...with patriarchal family norms that have been modified by equalitarian norms at the upper strata...' (Rodman,

1972, blz. 64). Het gaat hier voornamelijk om landen die korte tijd geleden nog sterk patriarchaal waren. In maatschappijen van dit type wordt de macht van de man in het gezin als vanzelfsprekend ervaren; die heeft hij dan ook louter te danken aan het feit dat hij man is. Moderne invloeden van buitenaf winnen echter snel terrein in deze landen, meestal het eerst bij de hogere sociale klassen. Dat geldt ook voor het principe van gelijkwaardigheid van man en vrouw bij echtelijke beslissingen. Die gelijkwaardigheid wordt dan het meest geaccepteerd in de hogere sociale klassen, althans in vergelijking met de andere lagen van de bevolking (en niet zozeer in vergelijking met andere landen). Als voorbeelden van landen in Fase 2 worden Griekenland en Joegoslavië genoemd.

Fase 3: Transitional Equalitarianism

De derde fase in de indeling van Rodman heeft betrekking op maatschappijen '...in which equalitarian family norms are replacing patriarchal norms and in which there is normative flexibility about marital power...' (Rodman, 1972, blz. 64). Deze landen bevinden zich in een overgangsfase. Ze onderscheiden zich vooral van de landen in Fase 1 en 2 door het feit dat de man zijn macht en status niet als vanzelfsprekend in de schoot geworpen krijgt; hij moet die zelf verwerven voordat zijn rol geaccepteerd wordt. Dat lukt hem meestal alleen als hij hoog scoort op zaken, waaraan de maatschappij waarde hecht, zoals opleiding, een beroep dat in aanzien staat en inkomen. De Verenigde Staten en West-Duitsland noemt Rodman als voorbeelden van landen in Fase 3.

Fase 4: Equalitarianism

De laatste fase omvat landen met '...strong equalitarian family norms and a high level of husband-wife sharing of power...' (Rodman, 1972, blz. 65). De gelijkwaardigheid van man en vrouw geldt hier niet alleen voor bepaalde sociale klassen, maar wordt erkend in alle lagen van de bevolking (ongeacht de status van de man). Zweden en Denemarken zijn volgens Rodman de landen waarvoor dit principe het meest opgaat.

Het concept van de machtsverhouding tussen de echtgenoten dat Rodman in zijn classificatie hanteert, stoelt op de onderlinge verdeling van beslissingsbevoegdheden. Hoewel daarbij alle mogelijke beslissingen werden betrokken, die in gezinsverband genomen worden, leent de indeling van maatschappijen zich uitstekend voor de beslissingen waarop dit onderzoek zich richt – aankoopbeslissingen. Het valt, in grote lijnen, te verwachten dat de rolverdeling bij aankoopbeslissingen overeenstemt met de algemene besluitvormingspatronen die kenmerkend zijn voor de vier fasen van de indeling van Rodman.

3. DOELSTELLINGEN

Het voornaamste doel van dit onderzoek is een poging om kennis over aankoopbeslissingen in gezinsverband in een cross-culturele context te plaatsen. Daartoe vergelijken wij de onderlinge verdeling van beslissingsbevoegdheden voor vier soorten aankoopbeslissingen bij zeven produkten en diensten in vijf

landen. Bij twee van die landen, Gabon en Venezuela, gaan wij ervan uit dat die onder één van de twee eerste fasen van de indeling van Rodman vallen. De overige landen, Nederland, Frankrijk en de Verenigde Staten, plaatsen we in de laatste twee fasen. De landen zijn hier in twee in plaats van in vier groepen ingedeeld omdat een splitsing in vier groepen een zeer gedetailleerde analyse van diverse maatschappelijke kenmerken in elk land zou vereisen. Een vergelijking van de besluitvormingspatronen binnen en tussen de landen in deze twee groepen levert een bijdrage aan de ontwikkeling van een cross-culturele invalshoek in het onderzoek naar aankoopbeslissingen in het gezin.

Het tweede doel van het onderzoek is om de indeling van Rodman met empirische gegevens te toetsen. De onderzoeken waarop zijn indeling is gebaseerd zijn slechts in grote lijnen vergelijkbaar. Zij verschillen in de onderzoeksmethode, zowel wat de vragenlijst als de steekproeven betreft. Bovendien lag er nogal wat tijd tussen het eerste en het laatste onderzoek. De vijf projecten die deel uitmaken van dit onderzoek stemmen voor wat betreft de steekproeftrekking en de vragenlijst grotendeels overeen (hierop komen wij zo terug) en werden – op één uitzondering na – binnen een periode van 14 maanden uitgevoerd. Als de resultaten daarvan overeenkomen met wat uit de indeling van maatschappijen te verwachten valt, dan ondersteunen die de validiteit van Rodman's theorie.

4. HYPOTHESEN

Er worden in dit onderzoek vier hypothesen getoetst. De eerste en tweede hebben betrekking op (1) de algemene verschillen die te vinden zijn tussen de landen in Fase 1 en 2 en de landen in Fase 3 en 4; en (2) de overeenkomsten die te verwachten zijn binnen de landen in Fase 1 en 2 en de landen in Fase 3 en 4. De nulhypothesen luiden als volgt:
H1: Er zijn geen verschillen in de besluitvormingspatronen tussen landen in Fase 1 en 2 en landen in Fase 3 en 4.
H2: Er zijn geen overeenkomsten in de besluitvormingspatronen binnen landen in Fase 1 en 2 en binnen landen in Fase 3 en 4.

De derde en vierde hypothesen zijn wat meer specifiek en gaan over de aard van de verschillen die we tussen de twee groepen landen denken te vinden. De derde hypothese is gebaseerd op de veronderstelling dat er in de landen die gekenmerkt worden door *equalitarianism* meer syncratische besluitvorming is dan in de meer traditionele, patriarchale maatschappijen. Tenslotte toetsen we met de vierde hypothese de veronderstelling dat in landen in Fase 1 en 2 de echtgenoot zelf meer beslissingen neemt dan in de laatste twee typen maatschappijen.
H3: Er komt in landen in Fase 3 en 4 niet meer syncratische besluitvorming voor dan in landen in Fase 1 en 2.
H4: In landen in Fase 1 en 2 worden door de echtgenoot niet meer beslissingen genomen dan in landen in Fase 3 en 4.

5. ONDERZOEKSMETHODE

5.1. *Landen en steekproeven*

De gegevens over de rollen in de besluitvorming werden in de vijf eerder genoemde landen verzameld: Nederland, Frankrijk, Gabon, Venezuela en de Verenigde Staten. Dat gebeurde in elk land in één van de grotere steden, respectievelijk Rotterdam, Marseille, Libreville, Valencia en Houston. Iedere steekproef bestond uit getrouwde vrouwen uit de midden- of hogere sociale klasse tussen 18 en 70 jaar oud. Het Amerikaanse onderzoek werd begin 1980 uitgevoerd; de onderzoeken in Nederland, Frankrijk en Gabon in 1979, en dat in Venezuela in 1977.

De methode om getrouwde vrouwen als respondenten te laten fungeren in onderzoek naar de rollen die gezinsleden in het beslissingsproces spelen is in de literatuur uitvoerig aan de kaak gesteld. Sommigen beweren dat antwoorden van zowel man als vrouw vereist zijn om een betrouwbare uitspraak te kunnen doen over het besluitvormingspatroon in het gezin (Safilios-Rothschild 1969). Daarentegen betoogt een andere groep onderzoekers dat de informatie die de echtgenote verstrekt wel degelijk voldoende is (Haberman and Elison 1967). In elk geval werd de betrouwbaarheid van door de vrouw gegeven antwoorden onderzocht en vervolgens bevestigd in een tweetal studies waarin de gestelde vragen specifiek betrekking hadden op produktaankopen (Becherer 1973, Davis 1971).

Dat de steekproeven getrokken werden uit getrouwde vrouwen van de sociale middenklasse had in het kader van dit onderzoek twee redenen. Ten eerste is het in cross-cultureel onderzoek belangrijk om steekproeven te gebruiken die zoveel mogelijk bij elkaar passen (Berrien 1967, Berry 1969, Brislin and Baumgardner 1971, Frida and Jahoda 1966, Green and White 1976). Als de steekproeven niet uit gelijksoortige segmenten van de bevolking afkomstig zijn, is het zelfs bij een exploratief cross-cultureel onderzoek onmogelijk vergelijkingen te trekken. In deze studie werden de steekproeven dan ook op elkaar afgestemd op het vlak van drie voorname demografische kenmerken: geslacht, burgerlijke staat en sociale klasse. Ten tweede vormen de sociale middenklassen een belangrijke groep in de theorie van Rodman, vooral in maatschappijen die in ontwikkeling zijn. De middenklasse vertegenwoordigt meestal een modern krachtenveld in ontwikkelingslanden, daar deze grotendeels aan het ontwikkelingsproces haar bestaan dankt. Ergo: als er 'moderne' waarden in een maatschappij aanwezig zijn, zijn ze waarschijnlijk in de middenklasse te vinden.

5.2. *Gegevensverzameling*

De Amerikaanse gegevens werden verzameld door middel van een schriftelijke enquête, waarbij de respondenten de vragenlijsten zelf invulden en terugstuurden. In Nederland, Frankrijk en Venezuela werd aan iedere respondent persoonlijk een vragenlijst overhandigd, die enige tijd later weer werd opgehaald. De respondenten woonden in elk land overwegend in stadswijken met voorna-

melijk gezinnen in de modale inkomensgroep; binnen die wijken werd de steekproef op a-selecte wijze getrokken.

In het Afrikaanse land Gabon kon deze procedure niet worden gevolgd. Het is daar namelijk ongebruikelijk om mensen in hun woning voor een onderzoek te benaderen. Daarom werd in Gabon gebruik gemaakt van een 'convenience' steekproef, waarbij de onderzoekgegevens verzameld werden in de buurt van supermarkten en andere drukbezochte plaatsen in de stad. De grootte van de steekproeven bedraagt 110 in Gabon; 127 in Frankrijk; 158 in de V.S.; 170 in Venezuela en 240 in Nederland.

5.3. *Vragenformulier*

Het vragenformulier voor dit onderzoek werd eerst in het Engels geschreven en vervolgens vertaald voor gebruik in de verschillende landen. De respondenten werd gevraagd om de voornaamste beslisser in het gezin aan te wijzen bij de aankoop van een zevental produkten en diensten: levensmiddelen, elektrische huishoudelijke apparaten, meubilair, auto's, sparen, vakanties en levensverzekeringen. Bij elke categorie produkten en diensten werden vier specifieke vragen gesteld die met de aankoop verband houden: wie beslist (1) wanneer het produkt wordt gekocht? (2) waar het wordt gekocht? (3) wat er wordt gekocht? en (4) hoeveel eraan wordt uitgegeven? Bij de beantwoording van elke vraag gaven de respondenten aan of de beslissing gewoonlijk genomen wordt door de vrouw, door de man, of door hen gezamenlijk. Deze procedure voor het meten van de invloed van gezinsleden is de meest gebruikte methode in de vakliteratuur en vergemakkelijkt daardoor een vergelijking met voorgaand onderzoek (Szybillo, Sosanie and Tenenbein 1979).

6. ANALYSE

In het geval van de onderzoeken naar besluitvorming in gezinsverband wordt met nominale data gewerkt. Hoewel dat soort data geen analytische problemen oplevert als het onderzoek tot één land beperkt blijft, kan dat in vergelijkend onderzoek wel het geval zijn, met name als de vergelijking meer dan twee groepen omvat. De *correspondence* analysetechniek, die in Frankrijk ontwikkeld werd (Benzecri 1973 en 1979, Caillez en Pages 1976, Hill 1974) maakt het mogelijk nominale data die bij verschillende groeperingen verzameld zijn met elkaar te vergelijken. De techniek wordt gebruikt om de verhouding tussen nominale data visueel weer te geven.

De correspondence analyse is in feite een bijzondere vorm van de meer traditionele principale componentenanalyse, met als verschil de afstandseenheid die gebruikt wordt bij het analyseren van de gegevens. Gegeven de nominale schaal van de data wordt er bij de correspondence analyse een ruimte gecreëerd die wordt gestructureerd door de chi-kwadraatsafstand, terwijl de gegevens vervolgens in een gereduceerde ruimte worden weergegeven gebaseerd op de gevonden principale componenten (assen). Het voordeel van een dergelijke analyse is dat, onafhankelijk van enig a priori probabilistisch model, de chi-kwadraatsafstand de enige metriek is die in overeenstemming is met het 'distributional equivalence' concept, een algemeen principe dat is gebaseerd

op de symmetrische verhouding tussen nominale gegevens (Benzecri 1979).

Gegeven twee variabelen, I en J, is de eerste stap bij de correspondence analyse om de onafhankelijkheid van de variabelen te testen met behulp van een kruistabel. Als de nulhypothese van onafhankelijkheid wordt verworpen met behulp van de gebruikelijke chi-kwadraatstoets voor kruistabellen, wordt er een nieuwe tabel samengesteld die proporties (of conditionele waarschijnlijkheden) bevat. Vervolgens kunnen we een grafische voorstelling van deze gegevens maken, en wel op zo'n manier dat de hogere proporties die met de verschillende niveaus van I en J verbonden zijn dicht bij elkaar staan, en de lagere proporties verder uit elkaar. De eerste principale component kunnen we dan aan de hand van deze cluster van punten vinden, terwijl we tevens een weergave in een verkleinde dimensie zien. Tegelijk met deze weergave krijgen we een aanduiding van de hoeveelheid variantie die door deze dimensie wordt verklaard. Ditzelfde kunnen we doen voor een tweede (derde, enz.) principale component, net zo lang tot er geen significante stijging meer verkregen wordt in het percentage te verklaren variantie.

Hoe moeten we de dimensies die uit de correspondence analyse naar voren komen nu interpreteren? Dit kan het beste door te kijken naar die component (as) die de grootste bijdrage levert aan de verklaring van de totale variantie. Vaak is dit mogelijk door te kijken naar de punten die aan beide kanten van de betreffende as clusters vormen. Net als bij principale componentenanalyse wordt er bij correspondence analyse het grootste deel van de variantie verklaard als er slechts een paar dimensies (assen) zijn met sterke dichotomen. Bovendien, als er wellicht meer dan twee sets van gegevens gerelateerd zijn ($I \times J \times T$), kunnen we volstaan met een tweetal van die sets te bekijken, bijvoorbeeld met een kruising van I en T $(I \times T) \times J$ en toch een zinvolle weergave krijgen. Op zo'n manier kunnen we drie nominale variabelen reduceren tot een binaire tabel om de verschillende relaties nader te bekijken.

De inputdata in de correspondence analyse van dit onderzoek bestaan uit de gegeven antwoorden over wie de *beslissingen* neemt (man, vrouw of gezamenlijk voor de vier soorten beslissingen) per *produktcategorie* (levensmiddelen, auto's enz.) en per *land* (Nederland, Frankrijk, Verenigde Staten, Venezuela en Gabon). De kaart die van de uitkomsten is afgeleid geeft de relaties weer die tussen deze variabelen bestaan.

Bij de bespreking van de onderzoeksresultaten besteden we eerst aandacht aan de kaart die de correspondence-analyse oplevert, en vervolgens aan de uitkomsten van de enkelvoudige chi-kwadraatstoetsen op de antwoorden. Deze chi-kwadraten zijn berekend over de antwoorden op elk van de vragen tussen de afzonderlijke landen in de twee groepen en binnen de landen in de twee groepen. Als de nulhypothesen worden verworpen dan valt te verwachten dat er weinig verschillen in de rollen bij de besluitvorming zijn binnen de twee groepen landen, en aanzienlijke verschillen tussen de groepen.

Meer specifiek werden de volgende chi-kwadraattoetsen uitgevoerd. Voor iedere beslissing (d.w.z. wanneer, waar, wat, hoeveel) die met elk produkt verbonden is (huishoudelijke apparaten, auto's enz.) werden chi-kwadraten berekend bij alle combinaties van landen voor het aantal beslissingen genomen door de man, door de vrouw en door hen beiden. Bij de presentatie van de resultaten kijken we eerst naar de grootte van de verschillen die werden aange-

troffen tussen de landen in Fase 1 en 2 en de landen in Fase 3 en 4. Daarna onderzoeken we de chi-kwadraatresultaten tussen de landen binnen elk van de groepen.

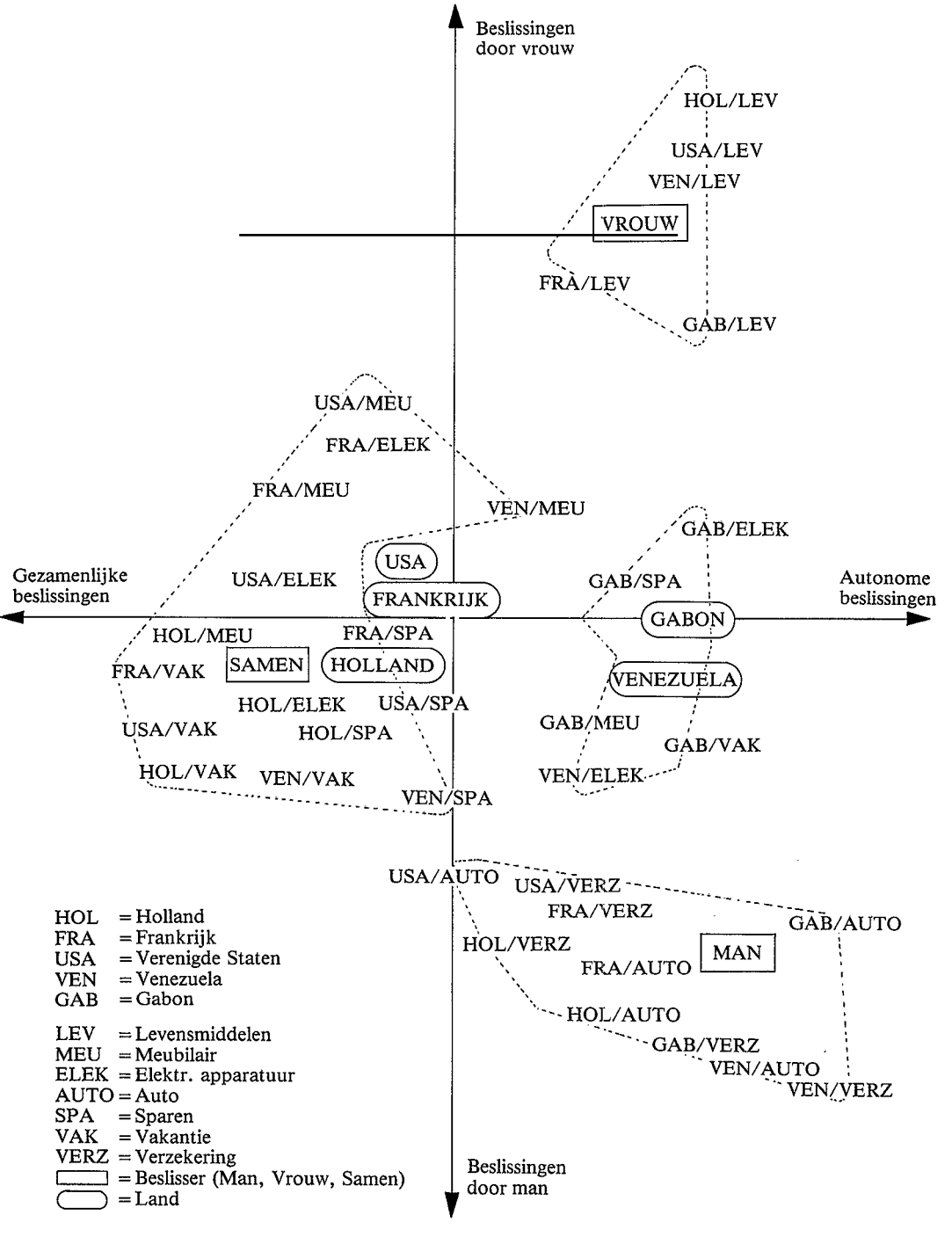
6.1. *Methodologische beperkingen*

De resultaten van deze studie moeten worden geïnterpreteerd binnen de grenzen die worden opgelegd door de methodologische beperkingen van het exploratief onderzoek. Dat zijn in eerste instantie de reeds genoemde beperkingen die gelden voor het gros van het empirisch onderzoek naar de machtsverhoudingen tussen echtgenoten (Davis 1971, Safilios-Rothschild 1969, Burns and Granbois 1979). Ten tweede zijn er beperkingen die voortkomen uit het feit dat het onderzoek cross-cultureel van opzet is. Daarbij treden soms problemen op door vertalingen die niet voor honderd procent gelijkwaardig zijn of door het ontbreken van functionele gelijkwaardigheid van bepaalde begrippen (Berrien 1967, Berry 1969, Brislin 1970, Brislin and Baumgardner 1971, Frijda and Jahoda 1966, Green and White 1976). Ten derde zijn de steekproeven getrokken in een bepaald deel van elk land; het is dan ook niet mogelijk om de resultaten te generaliseren naar de landen in totaliteit. En tenslotte vormen de verschillen in de gevolgde procedures voor het verzamelen van de gegevens in de vijf landen een beperking, met name het gebruik van de 'convenience' steekproef in Gabon. Vanzelfsprekend dicteren de uiteenlopende omgevingsfactoren in landen vaak het gebruik van verschillende methoden van gegevensvergaring, en er is enig onderzoek voorhanden dat aangeeft dat dit niet noodzakelijkerwijs de vergelijkbaarheid van de resultaten belemmert (Webster 1966). Dat neemt niet weg dat de resultaten van dit onderzoek geïnterpreteerd dienen te worden binnen de begrenzingen van de vermelde methodologische beperkingen.

7. RESULTATEN

7.1. *Correspondence-analyse kaart*

De kaart die werd afgeleid uit de correspondence-analyse staat in Figuur 1. Op de verticale as zien we de dimensie die aangeeft of de beslissing gedomineerd wordt door de man of door de vrouw; op de horizontale as staat de dimensie waaruit blijkt in hoeverre de beslissingen gezamenlijk of autonoom worden genomen. De punten die op de linkerzijde van de kaart staan vertegenwoordigen dus de beslissingen die in het algemeen gezamenlijk worden genomen, terwijl de punten op de rechterkant van de kaart voor beslissingen staan die in het algemeen door de man of door de vrouw alleen worden genomen. Op dezelfde wijze geven de punten in het bovenste deel van de kaart beslissingen weer, die overwegend door de vrouw worden genomen, terwijl die onderaan op de kaart duiden op beslissingen die de man in het algemeen neemt. Het punt dat bijvoorbeeld helemaal rechts-bovenaan op de kaart staat is HOL/LEV (Holland, Levensmiddelen). Die locatie duidt aan dat het zeer waarschijnlijk is dat beslissingen over levensmiddelen in ons land door de vrouw alleen worden genomen.



Figuur 1. Correspondence analyse voor aankoopbeslissingen in Nederland, Frankrijk, de Verenigde Staten, Venezuela en Gabon.

De bevindingen voor elk van de afzonderlijke beslissingen die met ieder produkt verbonden zijn (wanneer, waar, wat, hoeveel) staan niet op de kaart, omdat die niet wezenlijk bleken af te wijken van het gehele beeld dat elk van de produkten oplevert. De kaart verklaart 100% van de variantie van de gegevens; de verticale as (man/vrouw) 68% en de horizontale as (samen/autonoom) 32%.

Wat blijkt nu uit deze kaart? Drie van de produkten bevinden zich voornamelijk langs de verticale as, drie langs de horizontale as, en één (sparen) min of meer in het midden van de kaart. Van de eerste groep worden bij levensmiddelen de beslissingen sterk gedomineerd door de vrouw, terwijl bij de aanschaf van een auto en het afsluiten van een levensverzekeringspolis de echtgenoot de voornaamste beslisser is. Als we nu de horizontale as bekijken, dan blijkt dat met betrekking tot vakanties, meubilair en huishoudelijke apparaten de beslissingen veel meer door de echtgenoten samen worden genomen. Deze as laat ook zien dat de neiging tot autonome besluitvorming in Gabon en Venezuela het grootst is. Daarentegen zien we dat er meer gezamenlijke beslissingen genomen worden in Nederland, Frankrijk en de Verenigde Staten.

Bekijken we de produkt configuraties die op de kaart verschijnen, dan zien we dat er in de meeste gevallen nogal wat overeenkomsten tussen de vijf landen bestaan. Voor alle vijf landen geldt dat de vrouw de meeste beslissingen neemt voor levensmiddelen en de man voor de auto en verzekeringen. Beslissingen over vakanties, huishoudelijke apparatuur, meubilair en, tot op zekere hoogte, het sparen, worden in bijna alle landen door man en vrouw gezamenlijk genomen. Deze drie groepen of 'clusters' komen duidelijk uit de kaart naar voren.

Maar we zien ook dat er een vierde cluster op de kaart te vinden is, die we als de 'ontwikkelingslandencluster' zouden kunnen beschouwen omdat er alleen beslissingen uit Gabon en Venezuela in voorkomen. In Gabon worden beslissingen die verband houden met huishoudelijke apparaten, sparen, meubilair en vakanties meer autonoom genomen dan in de andere landen; over vakanties en meubelen wordt meer beslist door de man, en over sparen en apparatuur meer door de vrouw, dan het geval is in de economisch ontwikkelde landen. In Venezuela worden beslissingen over huishoudelijke apparaten meer op autonome basis door de man genomen, in vergelijking met de meer ontwikkelde landen. Het bestaan van deze vierde cluster lijkt de a priori indeling van de landen, die eerder besproken werd, te ondersteunen.

7.2. *Chi-kwadraat-resultaten*

De kaart van de correspondence-analyse vergemakkelijkt, door een grafische weergave van de resultaten, een vergelijking van de vijf landen. De configuraties van punten op de kaart sluiten aan bij de hypothesen over de verschillen tussen de twee groepen landen. De resultaten van de chi-kwadraatsanalyse, die nu aan bod komen, geven een beknopt beeld van een aantal van de statistische berekeningen die aan het ontwikkelen van de kaart ten grondslag liggen. In verband met de omvang van de uitgevoerde berekeningen volstaan wij hier met het geven van een samenvattend overzicht.

Om een nadere indicatie te geven van de plaats die de landen in de indeling

van Rodman innemen, werden er twee reeksen chi-kwadraatsanalyses uitgevoerd, waarin we de antwoorden tussen de landen telkens paarsgewijs vergeleken. Deze analyses hadden betrekking op het aantal beslissingen die ieder van de echtelieden had genomen (of die gezamenlijk werden genomen) in de verschillende landencombinaties over *wanneer* de produkten werden gekocht, *waar* ze werden gekocht, *wat* er precies werd gekocht en *hoeveel* er aan werd besteed. Zo werd bijvoorbeeld het aantal door de man genomen 'wanneer'-beslissingen opgeteld voor de steekproef uit de V.S. en voor de steekproef uit Venezuela; daarop werden dan chi-kwadraatsanalyses toegepast om eventuele significante verschillen op te sporen. Dit proces werd herhaald voor het aantal beslissingen genomen door de vrouw en door hen beiden voor alle typen beslissingen, zodat er voor elke combinatie van twee landen telkens in totaal 12 analyses werden uitgevoerd*.

In Tabel 1 staan de chi-kwadraatsresultaten met betrekking tot de paarsgewijze vergelijkingen van de landen in Fase 1 en 2 versus de landen in Fase 3 en 4. Deze resultaten zijn consistent met hypothesen 1, 3 en 4: er zijn significante verschillen tussen de landen in de twee groepen voor wat betreft het aantal beslissingen door de man en het aantal gezamenlijke beslissingen.

Een analyse van de 72 oorspronkelijke tabellen, die deel uitmaken van de chi-kwadraatsanalyse, leverde een reeks verschillen op die, zonder uitzondering, consistent zijn met de hypothesen. In alle gevallen blijkt in de ontwikkelingslanden de man de voornaamste beslisser te zijn en worden er minder beslissingen gezamenlijk genomen dan in de ontwikkelde landen.

Van de 72 chi-kwadraatsresultaten die in Tabel 1 staan vermeld, zijn er 61 statistisch significant. Alle niet-significante toetsen vallen onder de beslissingen van de vrouw, en zijn dus niet met de hypothesen verbonden. Toch verdient hiervan één aspect vermelding. De tabel laat zien dat er tussen Venezuela en de ontwikkelde landen erg weinig verschillen zijn in de mate waarin beslissingen door de vrouw worden genomen. Dat is anders voor Gabon, waar er in 9 van de 12 gevallen dit soort verschillen te vinden zijn. De oorspronkelijke tabellen (die hier dus niet zijn afgedrukt) geven aan dat in – op één na – alle gevallen waarin er verschillen zijn in de besluitvorming door de echtgenote tussen Gabon en de Westerse landen, de vrouwen in Gabon meer geneigd zijn om autonoom over een aantal zaken te beslissen dat hun collega's in de Westerse wereld. Dat zagen we trouwens ook al op de kaart van de correspondence analyse, waarop uit de configuraties bleek dat de besluitvorming in Gabon meer autonoom was dan in de ontwikkelde landen. De resultaten in Tabel

* Als illustratie volgt hier de chi-kwadraatsanalyse op het aantal 'wanneer'-beslissingen, genomen door de man in de Verenigde Staten en Venezuela. Eerst wordt het aantal door de man genomen 'wanneer'-beslissingen in elk land voor de zeven produktcategorieën opgeteld. Vervolgens wordt het aantal 'wanneer'-beslissingen dat *niet* door de man genomen is in elk land (d.w.z. genomen door de vrouw of door man en vrouw samen) voor de verschillende produkten opgeteld. Deze berekeningen leveren dan de getallen op die worden ingevuld in de cellen van de 2×2 chi-kwadraatsanalyse:

	Man	*Niet door man*
V.S.	480	698
Venezuela	221	877

$\chi^2 = 113.4$, d.f. $= 1$, $p \leqslant .01$

Tabel 1. Verschillen in besluitvorming tussen landen in Fase 1 en 2 en landen in Fase 3 en 4.

		V.S./ Gabon	V.S./ Venezuela	Frankrijk/ Gabon	Frankrijk/ Venezuela	Holland/ Gabon	Holland/ Venezuela
'Wat' beslissingen	Man	××	××	××	××	××	××
	Samen	××	××	××	××	××	××
	Vrouw	××		×		××	××
'Waar' beslissingen	Man	××	××	××	××	××	××
	Samen	××	××	××	××	××	××
	Vrouw	×		××		××	××
'Wanneer' beslissingen	Man	××	××	××	××	××	××
	Samen	××	××	××	××	××	××
	Vrouw			×		××	
'Hoeveel' beslissingen	Man	××	××	××	××	××	××
	Samen	××	××	××	××	××	××
	Vrouw	××	××			×	

×: $p \leqslant ,05$
××: $p \leqslant ,01$

1 wijzen er nu op dat deze autonomie in aankoopbeslissingen zich zowel uitstrekt naar de vrouw als naar de man.

In Tabel 2 staan de uitkomsten van de chi-kwadraats analyse die werd uitgevoerd tussen de landen *binnen* elk van de twee groepen. Hoewel deze enige ondersteuning bieden voor hypothese 2, geldt dit niet in dezelfde mate als voor de andere hypothesen. Er zijn zeer weinig verschillen tussen de V.S. en Frankrijk en tussen Nederland en Frankrijk, hetgeen de hypothese ondersteunt. Aan de andere kant werden er tussen de Verenigde Staten en Nederland in acht van de twaalf categorieën significante verschillen gevonden. Bovendien is de richting waarin die verschillen zich bewegen, niet consistent. De 'wat' en 'waar' beslissingen werden in de Nederlandse steekproef veel meer genomen door de man, en veel minder door de vrouw dan in de Amerikaanse steekproef het geval was. Doch voor de 'waar' en 'hoeveel' beslissingen bleek er in de Amerikaanse steekproef meer besluitvorming door de vrouw alleen en minder gezamenlijke besluitvorming voor te komen dan in ons land.

In de laatste kolom van Tabel 2 staan de verschillen tussen Venezuela en Gabon; die blijken in zeven van de twaalf categorieën voor te komen. Ze duiden ook op een consistent patroon: in de meeste gevallen vonden we in Gabon meer beslissingsmacht bij de vrouw, terwijl er in Venezuela een sterke neiging tot gezamenlijke besluitvorming optrad.

De verschillen tussen Venezuela en Gabon lijken erop te wijzen dat Venezuela tot de landen in Fase 2 van de indeling van Rodman moet worden gerekend en Gabon tot de landen in Fase 1. Venezuela heeft een hoger peil van economische ontwikkeling bereikt dan Gabon, en dit feit uit zich wellicht ook in de mate van sociale ontwikkeling waardoor de twee landen worden gekenmerkt. Een gezamenlijke besluitvorming is in het algemeen kenmerkend voor *'equali-*

Tabel 2. Verschillen in besluitvorming binnen de landen in Fase 1 en 2 en de landen in Fase 3 en 4.

		V.S./ Frankrijk	V.S./ Holland	Frankrijk/ Holland	Gabon/ Venezuela
'Wat' beslissingen	Man		× ×		
	Samen				× ×
	Vrouw		× ×	× ×	×
'Waar' beslissingen	Man		× ×		
	Samen				× ×
	Vrouw		× ×		×
'Wanneer' beslissingen	Man				
	Samen		×		× ×
	Vrouw		× ×	×	× ×
'Hoeveel' beslissingen	Man				× ×
	Samen	× ×	× ×		
	Vrouw	× ×	× ×		

× : p ⩽ ,05
× × : p ⩽ ,01

tarian' landen, en Rodman stelt dat de meer moderne sociale waarden in het algemeen het eerst worden overgenomen door de hogere sociale klassen in de maatschappij (Rodman 1972, blz. 64). Daar de steekproeven in dit onderzoek voornamelijk uit de midden- en hogere sociale klassen werden getrokken geven onze bevindingen dat verschijnsel wellicht weer. Uit Tabel 1 blijkt echter, dat de verschillen tussen Venezuela en Gabon meer verschillen in gradatie zijn dan in aard. In vergelijking met de meer ontwikkelde landen is de besluitvorming in Venezuela en Gabon opmerkelijk patriarchaal: het principe van gelijkwaardigheid en wederzijdse inspraak van de echtgenoten bij beslissingen is zeker niet karakteristiek voor die landen.

8. DISCUSSIE EN CONCLUSIES

Duidelijk blijkt uit de onderzoeksresultaten dat de hypothesen 1, 3 en 4 dienen te worden verworpen. Er zijn verschillen in de besluitvormingspatronen bij de aankoop tussen landen in Fase 1 en 2 en landen in Fase 3 en 4. De aard van de verschillen in de besluitvormingsstructuur gaat in de richting die in de hypothesen is aangegeven: meer patriarchale besluitvorming in Fasen 1 en 2, en meer gezamenlijke besluitvorming in Fasen 3 en 4.

De onderzoeksresultaten geven geen aanleiding tot het verwerpen van hypothese 2, die betrekking had op de overeenkomsten in de aankoopbeslissingspatronen bij de landen binnen elk van de groepen. Er zijn diverse verschillen naar voren gekomen tussen Gabon en Venezuela, hoewel die lijken te passen in het onderscheid tussen maatschappijen in Fase 1 en 2. Moeilijker is het om de verschillen tussen Nederland en de Verenigde Staten te verklaren. Het meest opvallende verschil is dat de Amerikaanse vrouw veel meer geneigd is

om *autonoom* over de aankoop van produkten en diensten te beslissen dan de Nederlandse echtgenote. Toch blijken sommige Nederlandse aankoopbeslissingen weer sterk te worden gedomineerd door de man – meer dan in de V.S. – terwijl er in andere gevallen meer sprake is van gezamenlijke besluitvorming in het Nederlandse gezin. We mogen daaruit nou niet gelijk afleiden, dat Nederland en de V.S. ondergebracht moeten worden in verschillende fasen van de classificatie van Rodman. De onderzoeksresultaten wijzen eerder op structurele verschillen in de besluitvormingspatronen onder de landen in de verschillende fasen.

Concluderend kunnen we stellen dat de twee doelstellingen van dit onderzoek, die in het begin van het artikel werden vermeld, zijn verwezenlijkt. Ten eerste zijn de eerste schreden gezet op de weg naar het plaatsen van onze kennis over aankoopbeslissingen in het gezin in een cross-culturele context. De onderzochte landen liggen op vier verschillende continenten en variëren van sterk geïndustrialiseerd tot bijzonder weinig geïndustrialiseerd, en van relatief modern tot relatief traditioneel. Wel zijn alle eerder genoemde beperkingen op het onderzoek van toepassing. Bovendien is er maar één van de vele aspecten van aankoopbeslissingen in het gezin in het onderzoek aan bod gekomen, en is er slechts aansluiting gezocht bij één van de verscheidene theorieën die worden gebruikt bij het verklaren van aankooppatronen. Toch is deze studie, binnen die beperkingen, erin geslaagd om een aantal relatief stabiele patronen van verschillen te onderkennen in huishoudelijke aankoopbeslissingen in maatschappijen op verschillende niveaus van ontwikkeling. Daarmee heeft dit onderzoek enig inzicht verschaft in de gevarieerdheid in besluitvormingsstructuren die we, voor zover het koopbeslissingen betreft, in verschillende culturen aantreffen. Het onderzoek heeft ook aannemelijk gemaakt dat deze variaties voor een deel kunnen worden verklaard door de fase van sociale ontwikkeling waarin een maatschappij zich bevindt.

De tweede doelstelling van de studie was om de indeling van landen, die Rodman ooit heeft geopperd, te valideren, of door empirisch onderzoek te doen bevestigen. Ook hieraan blijkt nu te zijn voldaan. Hoewel de vier categorieën van Rodman hier tot een tweetal zijn samengevoegd, waren de verschillen die tussen de twee groepen werden gevonden, consistent met de verschillen die via de indeling te voorspellen waren. Daarnaast zijn de gevonden verschillen tussen de twee minder ontwikkelde samenlevingen in deze studie te verenigen met de verschillen die in de indeling van Rodman zouden voorkomen tussen de groepen van Fase 1 en Fase 2. Wel moet worden opgemerkt, dat dit onderzoek slechts kan worden beschouwd als een gedeeltelijke validatie van Rodman's indeling, omdat één van de voornaamste variabelen hierin, nl. sociale klasse, constant is gehouden.

Tenslotte spreken wij de hoop uit, dat dit onderzoek een indicatie geeft van de bijdrage die cross-culturele studies kunnen leveren aan onze kennis over aankoopbeslissingen in gezinsverband, of meer in het algemeen, op het terrein van het koopgedrag. Veel van de kennis op dit gebied is sterk cultuurgebonden, en is voornamelijk gebaseerd op Amerikaans onderzoek. Daarmee willen wij de waarde ervan geenszins tenietdoen, doch dit feit legt wel ernstige beper-

kingen op bij het ontwikkelen van een theorie die het mogelijk maakt om te generaliseren. Cross-cultureel onderzoek biedt de mogelijkheid om onze huidige inzichten te verdiepen en ze in een breder gezichtsveld te plaatsen.

BIBLIOGRAFIE

Arndt, Johan and E. Crane, 1974. 'Marital Roles in Intrafamilial Decision-making on Spending Matters', *Combined Proceedings,* American Marketing Association, 63-6.

Becherer, R.C. et. al., 1973. 'Spousal perceptions of household purchasing influence: a multiperson-multiscale validation', *Proceedings* of the American Marketing Association Educator's Conference, 289-92.

Benzecri, J.P., 1973. *L'Analyse des Donnees,* Tome 2, Dunod (17, rue Remy-Dumoncel 75680 Paris Cedex 4).

Benzecri, J.P., 1979. 'Practical Aspects of Correspondence Analysis', Bulletin de l'ADDAD no. 6, Association pour le Developpement et la Diffusion de l'Analyse des Donnees Laboratoire de Statistique Tour, 45-55.

Berrien, F., 1967. 'Methodological and Related Problems in Cross-Cultural Research', *International Journal of Psychology,* 2, 33-43.

Berry, J., 1969. 'On Cross-Cultural Comparability', *International Journal of Psychology,* 4, 119-28.

Brislin, R., 1970. 'Back-Translation for Cross-Cultural Research', *Journal of Cross-Cultural Psychology,* 1, 185-216.

Brislin, R., and S. Baumgardner, 1971. 'Non-Random Sampling of Individuals in Cross-Cultural Research', *Journal of Cross-Cultural Psychology,* 2, 397-400.

Burns, Alvin C., and Donald H. Granbois, 1979. 'Advancing the Study of Family Purchase Decision Making', *Advances in Consumer Research,* 7, 221-6.

Caillez, F., and J.P. Pages, 1976. *Introduction a l'Analyse des Donnees,* Paris: Smach-Buro.

Christopher, M., 1975. 'Household Decision-Making', in: *Customers in Action,* Midgley, D. and M. Christopher, Bradford: MCB European Marketing and Customer Studies, Ltd, 33-75.

Davis, Harry L., 1976. 'Decision Making Within the Household', *Journal of Consumer Research,* 2, 241-260.

Davis, H.L., 1971. 'Measurement of Husband-Wife Influence in Consumer Purchase Decisions, *Journal of Marketing Research,* 8, 305-13.

Davis, H.L., and Benny P. Rigaux, 1974. 'Perception of Marital Roles in Decision Processes', *Journal of Consumer Research,* 1, 51-62.

Frijda, N., and G. Jahoda, 1966. 'On the Scope and Methods of Cross-Cultural Research', *International Journal of Psychology, 116-28.*

Green, Robert T., Bronislaw J. Verhage and Isabella C.M. Cunningham, 1981. 'Household Purchasing Decisions: How do American and Dutch Consumers Differ?', in: *Comparative Marketing in Action,* Lionel Mitchell et al., ed., (European Journal of Marketing, 15, 1), Bradford: MCB Publications, 68-77.

Green, Robert T., and Phillip D. White, 1976. 'Methodological Considerations in Cross-National Consumer Research', *Journal of International Business Studies,* 7, 81-7.

Haberman, P.W. and J. Elinson, 1967. 'Family income reported in surveys: husbands versus wives', *Journal of Marketing Research,* 2, 191-94.

Hempel, J., 1974. 'Family Buying Decisions: A Cross-Cultural Perspective', *Journal of Marketing Research,* 11, 259-302.

Hill, M.O., 1974. 'Correspondence Analysis: a neglected Multivariate Method', *Applied Statistics,* 23, London, 340-354.

Jenkins, Roger L., 1979. 'Contributions of Theory to the Study of Consumer Decision Making', *Advances in Consumer Research,* 7, 207-11.

Michel, Andree, 1971. 'Interaction and goal attainment in Parisian working wives' families', in Michel, A. (ed.), *Family Issues of Employed Women in Europe and America,* Leiden: E.J. Brill, 43-65.

Rodman, Hyman, 1972. 'Marital Power and the Theory of Resources in Cross-Cultural Context, *Journal of Comparative Family Studies,* 1, 50-67.

Safilios-Rothschild, Constantina, 1969. 'Family Sociology or Wives' Family Sociology: A Cross-

Cultural Examination of Decision Making', *Journal of Marriage and the Family,* 31, 290-301.
Szybillo, G.J., A.K. Sosanie and A. Tenenbein, 1979. 'Family Member Influence in Household Decision Making', *Journal of Consumer Research,* 6, 312-16.
Verhage, Bronislaw J., and Robert T. Green, 1981. 'Purchasing Roles in the Family: an analysis of instrumental and expressive decision making', *European Research,* Journal of the European Society for Opinion and Marketing Research, 1, 3-10.
Webster, Lucy, 1966. 'Comparability in Multi-Country Surveys', *Journal of Advertising Research,* 6, 14-8.

11. Het effect van herhaalde waarneming op consumentenaffect

T. POIESZ

SAMENVATTING

In deze bijdrage wordt een verslag gegeven van een onderzoek naar de effecten van de herhaalde waarneming van een stimulus op de affectieve beoordeling van die stimulus. Hierbij gaat het niet om de herhaalde waarneming van overredende boodschappen zoals reclames of advertenties, maar om stimuli als een produkt, een merknaam of een merksymbool. In tegenstelling tot de marketingliteratuur schenkt de sociaalpsychologische literatuur veel aandacht aan de relatie tussen herhaalde waarneming en affect, zij het niet specifiek op consumentengebied. Daarom wordt in deze bijdrage eerst gerefereerd naar de sociaalpsychologische literatuur om vervolgens te trachten duidelijkheid te verkrijgen omtrent herhaling-affectrelaties bij consumenten. Een verklaring van deze relaties die afwijkt van de verklaringen die in de literatuur worden gesuggereerd wordt getoetst in verschillende laboratoriumexperimenten en in een veldexperiment. De resultaten van de laboratoriumexperimenten, waarin onbekende merknamen als stimuli werden gebruikt, bevestigen de hypothese dat een positief effect van herhaling op affect alleen voorkomt wanneer een hoog noch laag niveau van (produkt-)risico wordt waargenomen. Bij een laag èn bij een hoog waargenomen risico is er geen positieve herhaling-affectrelatie. De resultaten van het veldexperiment bieden echter geen bevestiging voor de hypothese.

1. HERHAALDE WAARNEMING EN CONSUMENTENGEDRAG

Hoewel de macro- en micro-economische omgeving van de consument aan voortdurende verandering onderhevig is, bevinden zich in deze omgeving vele invariante elementen. Deze elementen komen herhaaldelijk voor in verschillende situaties en op verschillende tijdstippen. Deels is herhaling het gevolg van (massa)produktie en marketingactiviteiten, deels vloeit herhaling voort uit het in een aantal opzichten betrekkelijk invariante media-, informatie-, koop- en consumptiegedrag van de consument. Zo kan een bepaalde stimulus herhaaldelijk worden waargenomen omdat de stimulus een hoge penetratiegraad heeft, bijvoorbeeld een auto van een veel verkocht merk, en/of doordat de consument herhaaldelijk gedrag vertoont in relatie tot dezelfde stimulus, bijvoorbeeld het schrijven met steeds dezelfde pen.

Sommige stimuli in de omgeving van de consument komen vaker voor dan andere. Hierdoor, en ook als gevolg van de motieven, preferenties en gewoonten van de consument, variëren stimuli in de mate waarin ze herhaaldelijk worden waargenomen. Dit doet de vraag rijzen of, en zo ja: hoe en waarom, herhaling van stimuli een autonome invloed heeft op het consumentengedrag. De nadruk valt hier op het autonome effect, dus het effect van herhaling als zelfstandige factor. Interacties met andere factoren (b.v. het effect van het herhalen van overredende argumenten in reclameboodschappen) vormen een

onderwerp apart, waarover evidentie te vinden is in de consumentengedrags- en marketingliteratuur (zie bijvoorbeeld Belch, 1982). Er wordt hier niet nader op ingegaan.

Er zijn twee fasen of aspecten van het evaluatie- of besluitvormingsproces van de consument die met name relevant lijken in relatie tot de invloed van herhaling: informatieverwerking en affect- of attitudevorming.

De functie van herhaling voor de informatieverwerking ligt voor de hand. Het reeds eerder waargenomen hebben van een bepaalde stimulus mag verondersteld worden een positieve invloed te hebben op de juiste en snelle identificering, categorisering, geheugenopslag en herinnering van die stimulus (zie ook Sawyer, 1974).

Het effect van herhaling op affect is minder eenduidig. Hoewel er verschillende empirisch ondersteunde psychologische theorieën zijn die een antwoord zouden kunnen geven op de vraag of, en zo ja: hoe en waarom, herhaling affect beïnvloedt, is hun antwoord zelden meer dan een partieel antwoord: ze zijn alleen in staat om in bepaalde combinaties van stimulus-, omgevings- en consument-variabelen het effect van herhaling op affect te verklaren en te voorspellen. Als gevolg hiervan voorspelt dezelfde theorie positieve en negatieve herhaling-affectrelaties, afhankelijk van de combinatie van de drie soorten variabelen waarvan op het betreffende moment sprake is.

Twee voorbeelden. De leertheorie geeft aan dat herhaalde waarneming voorwaarde kan zijn voor het ontdekken van nieuwe stimulusaspecten. Al naar gelang deze aspecten positief of negatief worden ervaren zal zich een positieve of een negatieve herhaling-affectrelatie ontwikkelen (zie bijvoorbeeld Hansen, 1972). Op basis van de attributietheorie (Kelley & Michela, 1980) kan men de hypothese stellen dat het herhaald waarnemen van een merk bij de consument de indruk wekt dat het merk kwalitatief goed is (groot marktaandeel – goede kwaliteit). Een alternatieve hypothese binnen deze theorie zou zijn dat de consument uit de herhaalde waarneming afleidt dat het merk goedkoop zal zijn en dientengevolge van matige kwaliteit (vele kopers – lage prijs – matige kwaliteit). In het eerste geval gaat de hypothese uit van een positieve herhaling-affectrelatie, in het tweede geval van een negatieve. Met andere woorden, psychologische theorieën die toegepast kunnen worden op herhalingseffecten, maar die niet specifiek op deze effecten zijn gericht, zijn niet in staat om een meer algemene verklaring of een eenduidige voorspelling te bieden met betrekking tot de invloed van herhaling op affect. Zoals reeds gesteld is het ook op basis van de in de marketingliteratuur gepresenteerde theorie en empirische bevindingen niet mogelijk om te komen tot een meer algemene verklaring. In de eerste plaats richt deze literatuur zich meer op stimuli die te beschrijven zijn als overredende boodschappen.

Conclusies met betrekking tot deze stimuli zijn niet zonder meer generaliseerbaar naar andere typen stimuli. In de tweede plaats valt daarbij de nadruk op effectiviteit in termen van attitude, herinnering en koopintentie en niet op de onderliggende psychologische processen. Inzicht in deze processen is noodzakelijk om herhaling-affectrelaties te begrijpen en daardoor zo goed mogelijk te voorspellen (zie Belch, 1982).

2. DE 'MERE EXPOSURE'-HYPOTHESE

Een als meer algemeen bedoelde verklaring wordt geboden met de 'mere exposure'-hypothese van Zajonc (1968). Deze, in de sociaalpsychologische literatuur geïntroduceerde, hypothese stelt dat de herhaalde blootstelling (exposure) aan een stimulus een voldoende voorwaarde is voor de toename van affectieve gevoelens ten opzichte van die stimulus. Deze hypothese richt zich dus expliciet op de autonome invloed van herhaling op affect en gaat uit van alleen positieve herhaling-affectrelaties. Zajonc staaft de hypothese met verwijzingen naar correlationeel en experimenteel onderzoek. (Bij het eerste type onderzoek moet volledigheidshalve worden opgemerkt dat juist wanneer affect één van de geassocieerde variabelen is, het aangeven van de richting van de causaliteit problematisch kan zijn. Met andere woorden, correlationele evidentie kan in dit verband strikt genomen niet als hypothese-ondersteunende evidentie worden beschouwd.)

De experimentele evidentie die door Zajonc (1968) wordt gerapporteerd duidt op een algemeen positieve invloed van herhaalde waarneming op affect.

In het navolgende zal nader worden ingegaan op de validiteit van de 'mere exposure'-hypothese en op de hypotheses ter verklaring van het verschijnsel dat door de hypothese wordt beschreven. Dit geschiedt aan de hand van een overzicht van de in de sociaalpsychologische literatuur gepresenteerde verklaringen en empirische bevindingen. Deze hebben niet rechtstreeks betrekking op het consumentenaffect. Het is echter noodzakelijk om een start te maken bij de literatuur waarin aan het onderwerp van deze bijdrage de meeste aandacht is besteed om vervolgens te trachten de koppeling met het consumentenaffect tot stand te brengen.

De introductie van de 'mere exposure'-hypothese in 1968 werd gevolgd door een grote wetenschappelijke belangstelling voor frequentie-affectrelaties (FA-relaties). Hoewel de hypothese een verklaring inhoudt van het verschijnsel maakt ze het verschijnsel niet begrijpelijk. In de sociaalpsychologische literatuur na 1968 wordt vooral aandacht besteed aan de verklaring van positieve verbanden tussen frequentie en affect. Van de belangrijkste hypotheses volgt hier een korte beschrijving:

– De 'response competitie'-hypothese (Harrison, 1968). Deze hypothese veronderstelt dat (nieuwe) stimuli op het moment van waarneming verschillende, vaak tegengestelde, responses oproepen. Dit resulteert in een als negatief ervaren psychologische toestand (response competitie). Door de herhaling van een stimulus versterken bepaalde response-tendenzen en verzwakken andere. De responsecompetitie, en daarmee de onaangename psychologische ervaring, vermindert. Dit uit zich in een meer positief affect voor de meer frequente stimuli.

– 'Arousal'-hypotheses (Berlyne, 1960, e.a.). Verschillende hypotheses komen voort uit de arousaltheorie van Berlyne (1960). Deze theorie veronderstelt een optimaal niveau van fysiologische activatie of arousal. Afwijkingen van dit niveau (in twee richtingen – te veel activatie en te weinig activatie) worden als onaangenaam ervaren. Door de waarneming en verwerking van een stimulus kan het activatieniveau veranderen. Stimuli die het niveau verhogen tot boven het optimale niveau of verlagen tot beneden dat niveau worden geasso-

cieerd met negatief affect. Een nieuwe stimulus heeft een hoge 'excitatory potential', verhoogt het activatieniveau en wekt, indien het niveau hierdoor boven het optimale punt geraakt, minder positieve affectieve gevoelens op. Door herhaling verliest een stimulus zijn 'nieuwheid' en daarmee zijn excitatory potential. Een reeds herhaaldelijk waargenomen stimulus werkt niet verstorend op het optimale activatieniveau en genereert daarom meer positief affect dan een nieuwe stimulus (afhankelijk van het activatieniveau op dat moment).
– Er zijn theorieën die uitgaan van de werking van twee factoren – de zogenaamde 'twee-factoren'-theorieën (Berlyne, 1970; Stang, 1973). Bij de gelijktijdige werking van de twee factoren, gewenning en verveling (resp. positieve habituatie en tedium) zal de ene of de andere domineren. De relatieve sterkte van de habituatiefactor zal groter zijn bij een gering aantal herhalingen; bij een groter aantal herhalingen domineert de verveling. Welke factor dominant is hangt af van variabelen die een invloed hebben op het activatie- of arousalniveau. Zo zal bij een arousal-verhogende stimulus, zoals bijvoorbeeld een visueel complexe stimulus, de habituatie-factor bij een toenemend aantal herhalingen langer domineren dan bij een niet-complexe stimulus.
– Hypotheses betreffende de toe-/afname van semantische betekenis (Grush, 1976; Jacobovits, 1968). De 'toename'-hypothese veronderstelt dat door herhaling van een stimulus de betekenis van die stimulus zal polariseren: initieel positief wordt bij toenemende herhaling meer positief; negatief wordt meer negatief. De 'afname'-hypothese gaat uit van het tegenovergestelde.
– De 'expectancy-arousal'-hypothese (Crandall, 1967) stelt dat stimuli die gematigd sterke verwachtingen oproepen geprefereerd worden boven stimuli die geassocieerd zijn met zeer sterke of zeer zwakke verwachtingen. Naarmate een stimulus vaker wordt waargenomen roept hij sterkere verwachtingen op.

Bovengenoemde hypotheses, die elkaar gedeeltelijk overlappen, richten zich met name op de verklaring van positieve frequentie-affectrelaties. Daarbij houden sommige verklaringen rekening met de mogelijkheid van een omkeringseffect bij hoge frequenties – een boemerang-effect waarmee de eerste stijging van affect geheel of gedeeltelijk teniet gedaan wordt. Voor al de vermelde hypotheses is ondersteunende empirische evidentie gerapporteerd. Daarbij is in de literatuur de meeste aandacht besteed aan de 'response competitie'-hypothese.

Probleem is dat het niet mogelijk lijkt om één van de verklaringen als de meest valide aan te wijzen, mede omdat in de literatuur naast positieve FA-relaties ook niet-positieve en negatieve relaties worden gerapporteerd en naast lineaire ook niet-lineaire. Grush (1976): 'In summary, the mere exposure literature is replete with inconsistencies'. Vergelijkbare conclusies treffen we ook aan bij andere auteurs (Sawyer, 1974; Harrison, 1977; Vanbeselaere, 1981).

3. DE 'FUNCTIONAL EXPOSURE'-HYPOTHESE

Nu het niet mogelijk is om uit de literatuur enige consistentie te extraheren, zal op de volgende manier getracht worden te komen tot een meer algemene verklaring van FA-relaties: de in de literatuur aangeboden conclusies die betrekking hebben op een verzameling experimenten worden in verband gebracht

met een interpretatie van de psychologische betekenis van de experimentele situatie en stimuli tijdens een laboratoriumexperiment over FA-relaties. Dit laatste kan worden gezien als een stap terug naar de basis: nu de empirische evidentie op zich niet een algemene verklaring kan bieden is het wellicht mogelijk deze evidentie te herinterpreteren in het licht van de omstandigheden waaronder ze is verkregen. Voor de duidelijkheid volgt nu eerst een korte beschrijving van wat men zou kunnen beschouwen als een standaard laboratoriumexperiment rond FA-relaties: Na binnenkomst in de experimentele ruimte krijgen proefpersonen, die niet geïnformeerd zijn over aard en doel van het onderzoek, het verzoek plaats te nemen voor een projectiescherm en aandachtig te kijken naar hetgeen zal worden getoond. Nadere informatie wordt niet gegeven (hoewel de proefleider in sommige experimenten een geheugentaak aankondigt). De proefpersonen zien verschillende hen onbekende stimuli, zoals bijvoorbeeld nonsense woorden, symbolen, Chinese lettertekens of patronen in verschillende frequenties. Elke stimulus wordt (per keer) gedurende enkele seconden getoond.

Na afloop van de (dia)presentatie krijgen de proefpersonen het verzoek elk van de getoonde stimuli te beoordelen op een affectschaal. Hiertoe wordt elke stimulus nog één keer aangeboden.

In 1974 is door Stang de evidentie van meer dan honderd laboratoriumexperimenten geanalyseerd (Stang, 1974). In zijn analyse waren procedurele kenmerken van deze experimenten de onafhankelijke variabelen en het al of niet vóórkomen van positieve FA-relaties de afhankelijke variabele. Hij vond drie werkzame factoren: type stimulus, de grootte van het tijdsinterval tussen de aanbiedings- en beoordelingsfase, en de wijze van aanbieding van stimuli in hun verschillende frequenties. (De laatste twee factoren zijn mogelijk 'confounded', d.w.z. ze hebben mogelijk geen onafhankelijk effect.)

Positieve FA-relaties worden vaker geobserveerd indien:

- de stimuli onbekende nonsense woorden, symbolen en portretten zijn, dan wanneer het abstracte visuele of auditieve patronen betreft.
- het tijdsinterval tussen aanbieding en beoordeling (betrekkelijk) lang is.
- de stimuli met hun verschillende frequentieniveaus in gemengde volgorde worden aangeboden, bijvoorbeeld A-B-A-C-A-D-B-A-B-C (in plaats van een wijze van aanbieding waarbij het frequentieniveau wordt bereikt alvorens over te gaan naar een andere stimulus met een andere frequentie, bijvoorbeeld A-A-A-A-B-B-B-C-C-D).

Als we deze factoren in beschouwing nemen naast de constatering dat proefpersonen in FA-experimenten gewoonlijk geen uitleg krijgen over de te verwachten procedure, dan lijkt het mogelijk om ze tot een consistente interpretatie te integreren. Een proefpersoon die geen informatie krijgt over wat hem te wachten staat zal proberen zich een indruk te vormen op basis van wat hem wordt geboden. De veronderstelling is dat hij door het zien van hem onbekende nonsense woorden in verschillende frequenties verwacht een geheugen- of intelligentietaak te zullen krijgen, of een prestatietaak waarbij zijn gedrag te beoordelen zal zijn op een goed-slecht dimensie. Daar hij onzekerheid ervaart met betrekking tot het kunnen tonen van het gewenste gedrag/het kunnen behalen van een acceptabele score zal hij zich gesteund voelen door het vaker

vóórkomen van bepaalde stimuli. De vaker aangeboden stimuli reduceren zijn onzekerheid en worden dientengevolge positiever gewaardeerd. De twee andere factoren van Stang (1974) laten zich op analoge wijze interpreteren: de onzekerheid van een proefpersoon zal groter zijn indien de wachttijd na de stimuluspresentatie lang(er) is en indien de stimuli in gemengde volgorde worden gepresenteerd.

Als daarentegen in een experimentele situatie abstracte visuele stimuli als moderne schilderijen worden gepresenteerd, lijkt het gerechtvaardigd te veronderstellen dat proefpersonen niet een prestatie- maar een esthetische evaluatie-taak verwachten. Bij een dergelijke taak is het getoonde gedrag niet met een goede of slechte score te evalueren, eenvoudig omdat men geacht wordt niet over smaken te twisten. De experimentele procedure genereert geen onzekerheidsgevoelens en derhalve hebben vaker getoonde stimuli (per definitie) geen onzekerheidsreducerende werking. Ze zullen dan ook niet positiever worden beoordeeld dan de minder vaak getoonde stimuli.

Indien deze herinterpretatie correct is kunnen we stellen dat hogere aanbiedingsfrequenties zullen leiden tot meer positieve affectieve gevoelens indien en in de mate dat ze functioneel of instrumenteel zijn: dat ze in staat zijn onzekerheid te reduceren. In plaats van de 'mere exposure'-hypothese die uitgaat van een onvoorwaardelijke positieve invloed van frequentie op affect wordt nu de 'functional exposure'-hypothese voorgesteld die aan de ontwikkeling van positieve FA-relaties wel een voorwaarde verbindt. Het is niet de bedoeling om met deze hypothese weer een hypothese toe te voegen aan de toch al lange lijst van mogelijke verklaringen van FA-relaties. De 'functional exposure'-hypothese kan gezien worden als een samenvattende hypothese die niet in tegenspraak is met de empirisch ondersteunde eerder voorgestelde (en hierboven beschreven) deelhypotheses maar die tracht een meer algemene verklaring te bieden.

Nu een voorstel gedaan is tot een meer algemene verklaring is het wenselijk deze verklaring aan een empirische toets te onderwerpen. Hierbij is het in principe mogelijk direct de stap te maken naar het affect van consumenten als afhankelijke variabele. Omdat de verklaring echter nauw verbonden is aan een herinterpretatie van de betekenis van hogere aanbiedingsfrequenties in het in de sociaalpsychologische literatuur beschreven FA-experiment, is een eerste toetsing in de meer conventionele experimentele situatie meer op zijn plaats.

4. EMPIRISCH ONDERZOEK BETREFFENDE DE 'FUNCTIONAL EXPOSURE'-HYPOTHESE

4.1. Sociaalpsychologische laboratoriumexperimenten

In de beschrijving van de experimenten zal geen aandacht worden besteed aan de technische details, maar zal de nadruk liggen op hypotheses, algemene bevindingen en conclusies. Het eerste experiment was gebaseerd op het uitgangspunt dat hogere aanbiedingsfrequenties niet functioneel zijn/geen onzekerheid reduceren en dus niet leiden tot meer positief affect indien

- de proefpersoon een prestatietaak verwacht waarvan hij de indruk heeft dat die zeer moeilijk of zeer gemakkelijk zal zijn *ongeacht* het aantal keren

dat hij, vóór de taak, de aan de taak gerelateerde stimuli kan waarnemen;
- de proefpersoon een niet-prestatietaak verwacht, dus een taak waarbij een goed/slechtbeoordeling van zijn gedrag irrelevant is;
- de belangstelling van de proefpersoon voor zijn score op een prestatietaak kan worden geëlimineerd.

Positieve FA-relaties zullen wel worden geobserveerd als:
- de proefpersoon een prestatietaak verwacht;
- de proefpersoon in een conventionele FA-experimentele situatie wordt geplaatst. In het laatste geval zou dan volgens de eerder beschreven herinterpretatie van deze situatie moeten blijken dat proefpersonen meer een prestatietaak verwachten dan een niet-prestatie- (of een esthetische evaluatie-) taak.

De respectievelijke hypotheses werden getoetst in vijf experimentele condities:

Conditie 1: de helft van de proefpersonen (ppn) kreeg een 'zeer moeilijke taak'-instructie; de andere helft een 'zeer gemakkelijke taak'-instructie.

Conditie 2: Ppn kregen het verzoek de kwaliteit van de dia's te beoordelen.

Conditie 3: Ppn kregen te horen dat ze alleen 'voor spek en bonen' mee konden doen, omdat het experiment als mislukt moest worden beschouwd; vervolgens ontvingen ze de instructie van Conditie 5.

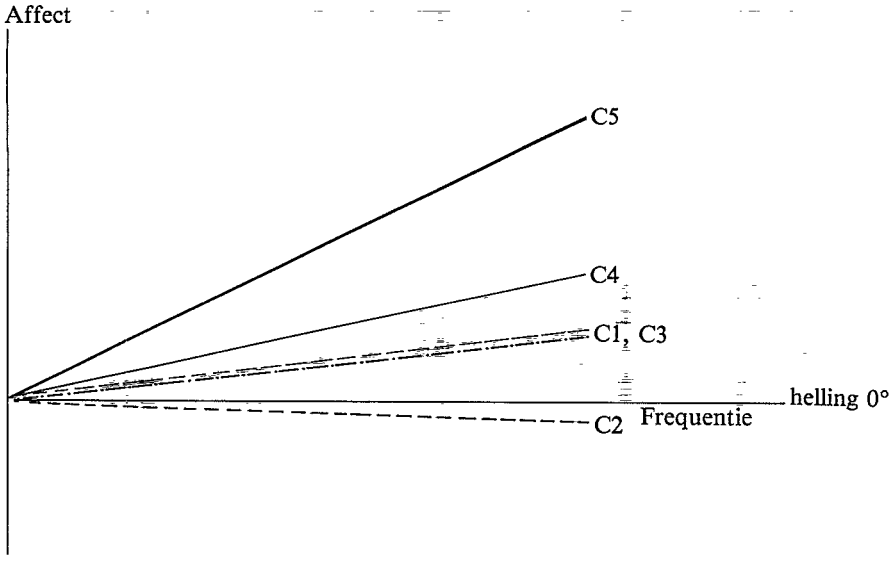
Conditie 4: Instructie: probeer een zo hoog mogelijke score te halen op de taak na de diaprestatie.

Conditie 5: Instructie: wil je de dia's die we dadelijk laten zien aandachtig bekijken.

Nonsense woorden zoals 'HEWONAT' en 'ZODEGIN', die in een vooronderzoek affectief neutraal waren gebleken, werden vervolgens 0, 1, 3, 6, 10 of 15 keer gepresenteerd. Elke dia verscheen twee seconden met een interval tussen dia's van vier seconden. Stimuli werden geroteerd over frequenties.

De affect-beoordeling op een 9-punts schaal vond plaats vrijwel onmiddellijk na de diapresentatie. Vóór de affectbeoordeling kregen ppn in Conditie 5 het verzoek aan te geven welk van twee typen taken zij verwachtten (een algemene omschrijving werd gegeven van een prestatietaak en van een niet-prestatietaak). Het tijdsinterval tussen de manipulatie en de affect-meting werd voor alle condities gelijk gemaakt.

Resultaten: In Conditie 5 verwachtte de meerderheid van de ppn (14 van de 18) een prestatietaak. Er is een significant positievere FA-relatie in de set van Conditie 4 en 5 gezamenlijk dan in de set van de eerste drie Condities gezamenlijk ($F = 6.00$, $df = 1,88$, $p < .05$). (Getoetst is met een lineaire trendanalyse, een variant op de variantie-analyse die aan kan geven of er een significant verschil bestaat tussen de hellingshoeken van de best-fitting rechte lijnen door de affect-scores behorende bij de oplopende frequenties.)



Figuur 1. Hellingen van de best-fitting rechte lijnen van FA-relaties in de condities (C's) van Experiment 1. (60° = gemiddelde van de 10% hoogste FA-scores in Experiment 1. Zowel 60° als 10% zijn willekeurig gekozen en dienen als relatief te worden geïnterpreteerd).

De helling van de FA-relaties verschilt niet binnen deze twee sets. Daarbij is de helling van de gecombineerde Condities 1, 2 en 3 niet significant verschillend van 0° ($F = 1.0$, $df = 1,51$, n.s.) en de helling van de gecombineerde Condities 4 en 5 wel ($F = 16.11$, $df = 1,34$, $p < .001$). (De helling of hellingshoek van een rechtlijnige relatie in een assenstelsel met de frequentieniveaus op de horizontale en het affect op de verticale as geeft aan in welke mate het affect toeneemt met een toenemende aanbiedingsfrequentie).

Per conditie zijn de FA-relaties conform de hypotheses, behalve in Conditie 4, waarin het resultaat weliswaar tendeert in de juiste richting, maar niet significant is. Zie ook Figuur 1.

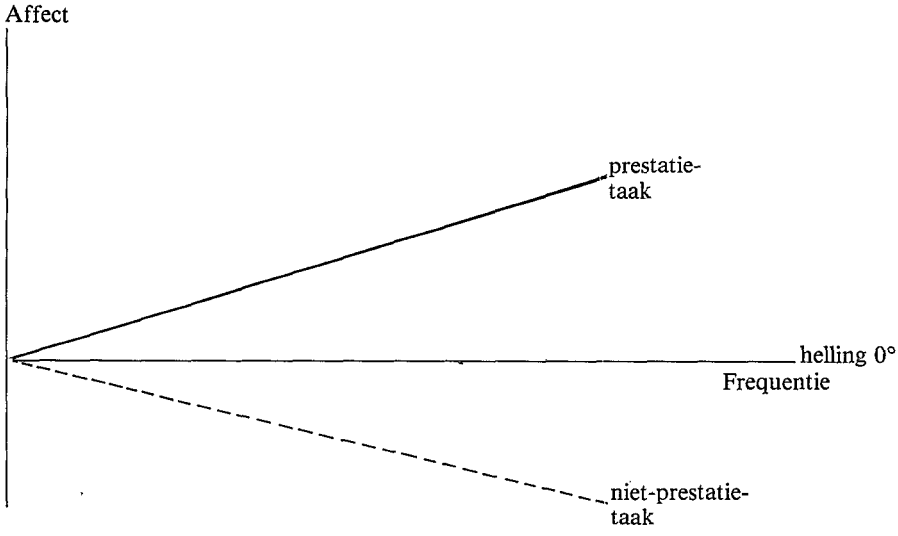
In het algemeen kunnen de hypotheses echter als bevestigd worden beschouwd. Enkele onduidelijkheden zijn:

- de functionaliteit van hogere aanbiedingsfrequenties is slechts *impliciet* gemanipuleerd door situaties te scheppen waarin we verschillen in functionaliteit vermoeden.
- op soortgelijke wijze is de onzekerheid van proefpersonen impliciet gemanipuleerd.
- Waarom was het resultaat in Conditie 4 slechts indicatief?

Het eerste en derde punt worden bekeken in Experiment 2; het tweede punt in Experiment 3.

Experiment 2

In dit experiment werden de Condities 2 (niet-prestatietaak) en 4 (prestatietaak) van Experiment 1 herhaald, met dit verschil dat in de laatste conditie minder nadruk werd gelegd op de 'zo hoog mogelijke score' (het vermoeden bestond dat door deze nadruk de instructies van deze conditie meer leken op



Figuur 2. Hellingen van de best-fitting rechte lijnen van FA-relaties in de condities van Experiment 2. (Voor nadere informatie zie omschrijving Figuur 1).

die van Conditie 1 in Experiment 1, de 'zeer moeilijke taak'-instructies, als gevolg waarvan juist geen positieve FA-relaties werden verwacht). Ook de procedure was identiek aan die van het eerste experiment, behalve dat proefpersonen tijdens de stimuluspresentatie waren verbonden aan een huidweerstandsmeter. De huidweerstand is gebleken af te nemen met conflict, onzekerheid en verwarring die door een proefpersoon wordt ervaren (de transpiratie neemt toe).

Als nu de vaker getoonde stimuli een onzekerheidsreducerende functie hebben dan zou dit uit de toegenomen huidweerstand kunnen blijken. De hypothese is dat bij proefpersonen die een positieve FA-relatie te zien geven de huidweerstand meer toe zal nemen bij de vaker getoonde stimuli dan bij proefpersonen die geen positieve FA-relatie hebben. De analyse van huidweerstandverschillen tussen de eerste en de tweede helft van de serie aanbiedingen werd alleen verricht voor de stimuli die 10 en 15 keer werden aangeboden. De hypothese werd bevestigd: proefpersonen met een positieve individuele FA-relatie gaven een grotere toename van de huidweerstand te zien bij de presentatie van de frequente stimuli dan proefpersonen met een niet-positieve FA-relatie ($t(25) = 1.19$, $p < .05$, waarbij de helling in de prestatietaakconditie significant verschilde van 0° ($F = 5.10$, $df = 1,14$, $p < .05$) en die van de niet-prestatietaak niet ($F < 1.0$, $df = 1.11$, n.s.). Zie ook Figuur 2.

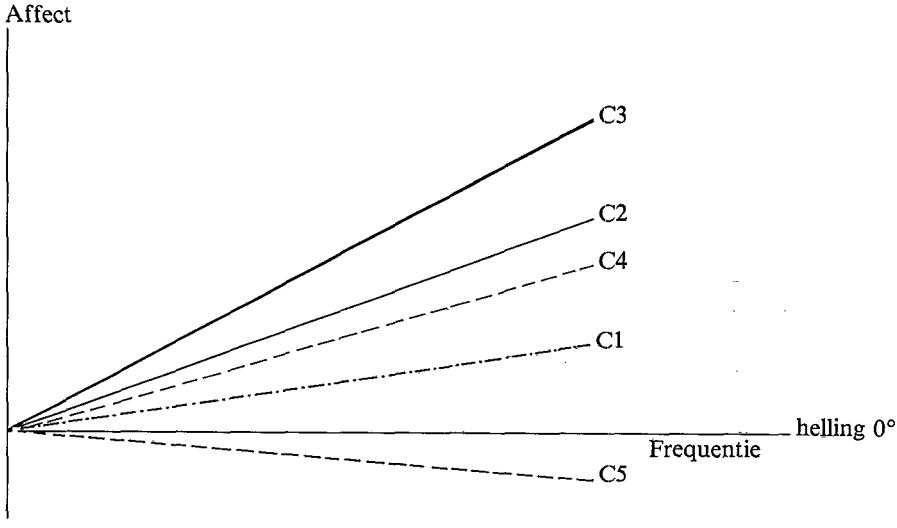
Met behulp van Experiment 3 werd getracht de onzekerheidsdimensie expliciet te maken. Het begrip onzekerheid houdt twee componenten in: onvoorspelbaarheid en bezorgdheid omtrent de gevolgen daarvan (zie bijv. Kagan, 1972). Bij constanthouding van de laatste component is onzekerheid te variëren met behulp van de (on)voorspelbaarheid, die uit te drukken is in de kans op het zullen voorkomen van een bepaalde situatie/het verkrijgen van een bepaald resultaat. De verwachting is dat vaak getoonde stimuli niet-functioneel kun-

nen zijn op beide extremen van de zekerheidsas: wanneer zekerheid bestaat omtrent het wel of niet kunnen verkrijgen van een bepaald resultaat (kans 100% of 0%).

Ter bepaling van het kansniveau waarop frequente stimuli het meest functioneel zijn (waarop de meest positieve FA-relaties gevonden zullen worden) wordt gerefereerd naar theorieën die zich bezighouden met de motivatie tot het verkrijgen van een bepaald positief resultaat. Bij toenemende motivatie neemt het belang van het functionele toe en dus ook de functionaliteit (van vaak getoonde stimuli) – een persoon die volstrekt niet gemotiveerd is tot het verkrijgen van een bepaald resultaat zal elk potentieel hulpmiddel per definitie niet als functioneel ervaren. Met andere woorden: als we de hellingshoek van FA-relaties willen voorspellen, dienen we de functionaliteit van de vaak getoonde stimuli te weten, welke laatste varieert met de motivatie (tot het verkrijgen van een positief resultaat) zoals die voorkomt bij verschillende kansniveaus. De theorie van Atkinson (1964) bepaalt de motivatie aan de hand van de formule: kans $\times$(1.0 – de kans) waarbij de laatste component de aantrekkelijkheid van het resultaat weergeeft (hoe kleiner de kans, hoe aantrekkelijker). Resultante is een omgekeerd U-vormig verband tussen kans en motivatie met een maximum bij 50% kans en een minimum bij 0% en 100%. Ten aanzien van dit laatste voorspelt Vroom (1964) precies het tegenovergestelde: de aantrekkelijkheid van een resultaat neemt toe met de kans het te verkrijgen. Dus: een positief lineair verband tussen kans en motivatie. Aangezien beide theorieën empirische ondersteuning hebben gekregen en omdat ze effecten beschrijven die elkaar aanvullen of corrigeren, worden ze hier geïntegreerd tot één hypothese: de motivatie tot het verkrijgen van een positief resultaat is het geringst bij 0% en 100% kans en hoger tussen 50% en 100% kans dan tussen 0% en 50% kans (met andere woorden: bij 0% kans en bij 100% kans worden geen positieve FA-relaties verwacht. Dergelijke relaties zullen worden geobserveerd tussen beide extremen met de meest positieve relatie niet op 50%, maar tussen 50% en 100% kans). In deze redenering kan de 'motivatie'-term worden vervangen door 'functionaliteit van vaak getoonde stimuli', en dit op zijn beurt weer door 'het optreden van positieve FA-relaties'.

Experiment 3 bestond uit 5 condities, waarin de kans op het verkrijgen van een positief resultaat werd gevarieerd. Proefpersonen kregen de instructie dat in het verleden 25% (of 50%, 75% of 100%) van de onderzoekdeelnemers erin geslaagd was de taak tot een goed einde te brengen. Conditie 4 (100% kans) en Conditie 5 (geen frequentiemanipulatie) waren controle-condities. De instructies kondigden een fijne-motoriektaak aan; stimuli waren afbeeldingen van pennen waarmee de taak moest worden uitgevoerd. Deze afbeeldingen werden in verschillende frequenties getoond waarna een affect-beoordeling plaatsvond. De resultaten tendeerden in de richting van de hypotheses. Er was een indicatief overall-effect van de kansmanipulatie binnen de set van Condities 1, 2 en 3 ($F = 2.67$, $df = 2,52$, $p < .10$): Over deze 3 condities werd een toenemende stijging van FA-relaties verwacht. Deze stijging bleek uit de resultaten: de (positieve) hellingshoek van Conditie 1 (25%) verschilde niet significant van die van controle-conditie 5 ($F < 1.0$, $df = 1,30$, n.s.), waarin de hellingshoek bijna gelijk was aan 0°.

Er was een bijna significant verschil tussen de hellingshoeken van Conditie



Figuur 3. Hellingen van de best-fitting rechte lijnen van FA-relaties in de condities (C's) van Experiment 3. (Voor nadere informatie zie omschrijving Figuur 1).

2 (50%) en Conditie 5 ($F = 3.58$, $df = 1,32$, $p < .10$) en een significant dito verschil bij de Condities 3 (75%) en 5 ($F = 7,34$, $df = 1,17$, $p < .05$). Alleen het effect in Conditie 4 was niet als verwacht: de FA-relatie was vergelijkbaar met dat van Conditie 2. Zie ook Figuur 3.

Een mogelijke verklaring hiervoor is dat proefpersonen een instructie waarin een taak met 100% slaagkans werd aangekondigd 'verdacht' of ongeloofwaardig hebben geacht, hetgeen de onzekerheid zou hebben vergroot, en daarmee ook de functionaliteit van de vaak getoonde pennen en de hellingshoek van de FA-relatie. Toch lijkt het gerechtigd te stellen dat, wanneer men de resultaten van de drie experimenten gezamenlijk in beschouwing neemt, de 'functional exposure'-hypothese empirische ondersteuning heeft gekregen.

4.2. *Laboratoriumexperimenten 'consumenten'-gedrag*

De overgang naar meer op het consumentengedrag gerichte experimenten vereist geen wezenlijk andere theoretische benadering: het onzekerheidsbegrip in de sociaalpsychologische literatuur is direct vergelijkbaar met het begrip waargenomen risico in de consumentengedragsliteratuur (Bauer, 1960; voor een uitgebreide bespreking van dit concept zie Todd, 1982).

Waargenomen risico omvat, net als onzekerheid, twee componenten: onvoorspelbaarheid en consequenties. Nu het begrip onzekerheid behulpzaam is geweest bij de verklaring van FA-relaties in het sociaalpsychologisch laboratorium, is waargenomen risico wellicht een relevante variabele bij de verklaring van de relaties tussen frequentie en consumentenaffect (in het laboratorium). In Experiment 3 is de (on)voorspelbaarheid gemanipuleerd en werd de component consequenties constant gehouden. Het omgekeerde of een gelijktijdige manipulatie is uiteraard ook mogelijk. Aangenomen wordt dat het toevoegen van de mogelijkheid van negatieve consequenties naast positieve consequen-

ties de functionaliteit van vaker getoonde stimuli vermindert. In geval van mogelijke negatieve consequenties zal een persoon minder geneigd zijn terug te vallen op frequente stimuli ter vermindering van onzekerheid/waargenomen risico. Hoog (waargenomen) risico (HR) wordt hier geoperationaliseerd als het in een keuzesituatie tegelijkertijd beschikbaar zijn van goede en slechte keuze-alternatieven waarbij niet bekend is *welke* alternatieven goed of slecht zijn. Middelmatig (waargenomen) risico (MR) is een vertaling van Conditie 2/3 van Experiment 3 (50%/75% kans op een positief resultaat). De hypothese is dat geen positieve FA-relaties zullen voorkomen in de LR- (laagrisico) en HR-conditie en wel in de MR-conditie. In de HR-conditie van Experiment 4 werd de instructie gegeven dat het onderzoek de beoordeling van 'roller pennen' betreft. Na een korte beschrijving van het produkt volgde de risico-instructie waarin gesuggereerd werd dat een consumentenorganisatie verschillende merken had getest en dat sommige merken daarbij van slechte kwaliteit bleken te zijn. 'Al met al een zeer riskant produkt'.

Bovendien kregen proefpersonen via de instructie de waarschuwing zeer voorzichtig te zijn bij het kiezen van een roller pen. In de MR-conditie was het (zogenaamd) resultaat van de test van de consumentenorganisatie dat 'er geen kans is op een miskoop omdat er geen slechte merken roller pennen zijn. Toch kan het geen kwaad enige voorzichtigheid te betrachten bij het maken van een keuze, dit vanwege de kwaliteitsverschillen. Alle merken zijn ongeveer even duur'.

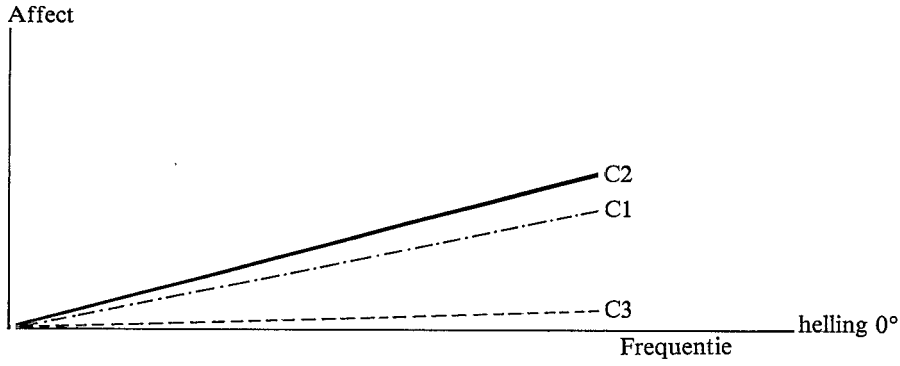
In de LR-conditie tenslotte was de conclusie: 'alle merken hebben onze test uitstekend doorstaan. Er zijn enige kwaliteitsverschillen, maar een miskoop is volstrekt onmogelijk'.

De instructies gaven aan dat na kennismaking met de verschillende merknamen – de frequentiemanipulatie: stimuli werden 0, 1, 3, 6, 9 of 12 keer getoond – een keuze gemaakt zou moeten worden uit de verschillende merken en dat men de gekozen pen zou mogen houden. (Om het waargenomen risico in de LR-conditie zo laag mogelijk te houden werd aan de instructie van proefpersonen in deze conditie nog toegevoegd dat men na de keuze van een pen deze pen eventueel nog kon ruilen tegen een andere). Stimuli, de 'merknamen', waren de namen van gehuchten in Frankrijk, die in een vóórtest affectief neutraal bleken.

Na het lezen van de instructies kregen proefpersonen vragen ter vaststelling van hun waargenomen risico. De stimuluspresentatie werd gevolgd door een affect-beoordeling van de merknamen en door het verzoek één van de merken te kiezen (N.B. proefpersonen hadden alleen de merk*namen* – in verschillende frequenties – gezien).

De resultaten op de risico-schalen laten zien dat proefpersonen in de LR-Conditie niet minder risico waarnemen dan proefpersonen in de MR-Conditie. De risico-manipulatie heeft geen significant effect op de keuze: na een MR-instructie kozen proefpersonen significant meer de frequent getoonde merknamen dan na een LR- en een HR-instructie. Daarbij was, zoals verwacht, het verschil tussen de LR- en de MR-Conditie niet significant. Zie ook Figuur 4.

Omdat de LR-Conditie niet het verwachte effect had op het waargenomen risico en mogelijk daardoor niet op de hellingshoek van de FA-relatie is in een



Figuur 4. Hellingen van de best-fitting rechte lijnen van FA-relaties in de condities (C's) van experiment 4. (Voor nadere informatie zie omschrijving Figuur 1).

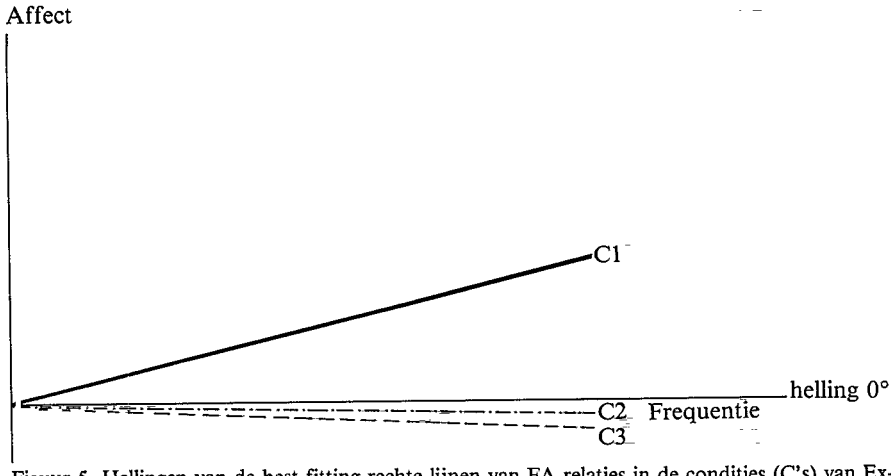
aanvullend experiment getracht een lager niveau van waargenomen risico te genereren. Dit werd geprobeerd door uit de instructie de aankondiging van de eigen keuze van de consument te elimineren. In plaats van de instructie dat de proefpersoon zelf een keuze zou dienen te maken gaf de instructie aan dat de proefleider één van de merken zou *geven*.

In Experiment 5 werden drie condities gecreëerd: Conditie 1: Zeer laag risico (ZLR) instructie met keuze; Conditie 2: Middelmatig risico instructie, geen keuze; Conditie 3: ZLR-instructie, geen keuze.

De ZLR-instructie luidde: 'Alle merken roller pennen hebben onze test uitstekend doorstaan. Een miskoop is volstrekt onmogelijk' (geen vermelding van kwaliteitsverschillen). De MR-instructie was identiek aan die in het vorige experiment. Hoewel proefpersonen in de Condities 2 en 3 tijdens de stimuluspresentatie geen keuze verwachtten kregen ze na de affect-beoordeling het verzoek het merk aan te wijzen dat ze zouden nemen als ze tóch zouden moeten kiezen ('Stel, dat u toch...' etc.). De hypothese was dat in de condities zonder keuze geen positieve FA-relaties zullen worden geobserveerd en dat de FA-relatie in Conditie 1 minder positief zal zijn dan in de MR-conditie van het vorige experiment. (Een vergelijking tussen experimenten is mogelijk omdat het onderzoek plaatsvindt onder identieke omstandigheden). Met betrekking tot het gekozen merk werden overeenkomstige hypotheses geformuleerd: in de keuzecondities zal het gekozen merk vaker waargenomen zijn dan in de condities zonder keuze. Bij de berekening van de resultaten werd een vergelijking gemaakt met de MR-conditie van Experiment 4.

De resultaten voor de FA-relaties wijzen op een significant effect voor de keuzefactor ($F = 4.70$, $df = 1.60$, $p < .05$, met meer positieve relaties in de keuze-conditie) maar niet voor de factor risico-instructie, noch voor de interactie. De frequentie van het gekozen merk werd echter significant beïnvloed door de factor keuze ($F = 13.08$, $df = 1,58$, $p < .001$) waarbij de meer frequente merken werden gekozen in de conditie met de keuze én met de MR-instructie. Er was geen significant autonoom effect van de risico-instructies. Zie ook Figuur 5.

Hoewel de resultaten de hypotheses niet in alle opzichten bevestigden tendeerden ze, voor zover ze niet significant waren, in de voorspelde richting. Als



Figuur 5. Hellingen van de best-fitting rechte lijnen van FA-relaties in de condities (C's) van Experiment 5. (Voor nadere informatie zie omschrijving Figuur 1).

we de bevindingen van de Experimenten 4 en 5 gezamenlijk in beschouwing nemen dan mogen we wellicht constateren dat het mogelijk blijkt om door manipulatie van een bij de bestudering van het consumentengedrag gehanteerde concept – waargenomen risico – een invloed uit te oefenen op de (positieve) hellingshoek van frequentie-affectrelaties.

In een laatste laboratoriumexperiment van deze serie, waarop hier niet nader wordt ingegaan, bleek dat de positie van de risico-instructie ten opzichte van de frequentiemanipulatie (=vóór of na) geen significant effect had op de resultaten. Met andere woorden, indien de risico-instructie wordt gegeven (korte tijd) ná de frequentiemanipulatie, vindt er geen significante wijziging in de resultaten plaats.

Deze laatste bevinding vergemakkelijkt het doen van veldexperimenten. Door eerst de frequentiemanipulatie uit te voeren en daarna de risico-instructie te geven is het namelijk eenvoudiger dan bij de omgekeerde volgorde te voorkomen dat mensen zich tijdens de frequentiemanipulaties al 'proefpersoon' voelen.

4.3 *Een veldexperiment*

Het bezwaar van laboratoriumexperimenten is dat de conclusies op basis van de verkregen resultaten niet zonder meer vertaalbaar zijn in praktische conclusies en implicaties. Naarmate een experimentele situatie meer gelijkt op de reële situatie neemt de mogelijkheid toe om gerechtvaardigde uitspraken te doen over het consumentengedrag in werkelijkheid. In het onderhavige veldexperiment werd getracht de reële situatie te benaderen door te voorkomen dat proefpersonen op de hoogte waren van hun rol als proefpersoon, door een frequentiemanipulatie te hanteren zoals die ook in werkelijkheid kan voorkomen en door ook bij de affect/preferentiemeting te vermijden dat proefpersonen de verbinding zouden leggen met de frequentiemanipulatie als onderdeel van

een experimentele procedure.

In het experiment werd een 2×2×2 design gehanteerd. De eerste factor betrof de frequentiemanipulatie, met twee niveaus: 1 en 7 aanbiedingen. De tweede factor was de hoeveelheid waargenomen risico: middelmatig en geen risico; factor 3 was de hoeveelheid aandacht besteed aan de stimuli: betrekkelijk veel tegenover betrekkelijk weinig. Een controleconditie buiten dit design diende ter vaststelling van de nulmeting (frequentieniveau = 0).

De stimulus was een kleine vierkante advertentie met daarin, diagonaal geplaatst, de reeds in het laboratorium aangeboden merknaam 'TRESSON' met daaronder, in kleine letters, de produktaanduiding: 'pennen'. In een advertentieblad werd deze advertentie 7 maal geplaatst, verspreid over de pagina's. Proefpersonen kregen het verzoek deel te nemen aan een telefonisch onderzoek ter evaluatie van advertentiebladen.

Geruime tijd vóór de onderzoeksdatum werden ze geïnformeerd over de ter beschikking te hebben advertentiebladen. Door proefpersonen te vragen naar 1 of meerdere bladzijden te kijken kon de frequentie worden gemanipuleerd. De helft van hen diende gedurende 15 seconden de betreffende bladzijde(n) te bekijken (ruime aandacht); de andere helft moest alleen het aantal advertenties op de bladzijde(n) *tellen* (geringe aandacht). De risico-factor werd op dezelfde wijze gemanipuleerd als in het laboratorium, met dit verschil dat de risico-instructie mondeling was en niet schriftelijk.

Na het bekijken van de advertentiepagina's konden de proefpersonen een keuze maken uit twee merken pennen (waarde ±5 gulden). De gekozen pen zou na afloop van het onderzoek als dank voor de deelname worden opgestuurd. Eén van de merknamen was de in de advertentie verschenen naam 'TRESSON'. De andere naam ('VERGONS') mocht onbekend worden verondersteld bij de proefpersonen.

De verwachting dat het frequent (7×) getoonde merk vaker gekozen zou worden in de middelmatig-risico- dan in de zeer-laag-risico-condities werd niet bevestigd door de resultaten. Er was zelfs een effect dat in de tegengestelde richting wees: van de minste voorkeur voor de meest vaak getoonde merknaam was sprake in de conditie met 7 aanbiedingen, MR-instructie en ruime aandacht.

De (achteraf!-)verklaring die hiervoor plausibel lijkt is dat een reactantieeffect is opgetreden. Het meermalen (7×) herhalen van dezelfde advertentie in dezelfde krant kan door de ontvanger als te indringend worden ervaren als gevolg waarvan hij er zich tegen teweer stelt en juist tegengesteld gedrag vertoont. Het was niet mogelijk om deze interpretatie te toetsen, althans niet binnen de context van het laatste experiment.

5. CONCLUSIE

Hoewel in laboratoriumexperimenten bevestiging werd gevonden voor de hypothese dat de frequentie-affectrelaties (mede) afhankelijk zijn van het waargenomen risico, bevestigden de resultaten van een veldexperiment deze hypothese niet. Het zou daarom voorbarig zijn nu reeds te speculeren omtrent mogelijke praktische implicaties. Daar staat echter tegenover dat we nu op het spoor lijken te zijn van een tamelijk consistent verschijnsel en dat het niet

moeilijk is om de voorwaarden voor het optreden van dit verschijnsel in verband te brengen met 'real life' consumentengedrag. Een logisch vervolg op het hier gerapporteerde onderzoek is onderzoek naar het effect van de aard van de stimulus (bijvoorbeeld advertentie - niet advertentie) en naar de omstandigheden waaronder verschillende niveaus van risico worden waargenomen.

BIBLIOGRAFIE

Atkinson, J.W. An introduction to motivation. Princeton, N.J.: Van Nostrand, 1964.

Bauer, R.A. Consumer behavior as risk taking. Proceedings, American Marketing Association. December 1960. In: D.F. Cox (Ed.). Risk taking and information handling in consumer behavior. Boston: Harvard University, 1967, 23-33.

Belch, G.E. The effects of television commercial repetition on cognitive response and message acceptance. *Journal of Consumer Research,* 1982, *9,* 56-65.

Berlyne, D.E. Conflict, arousal and curiosity. New York: McGraw-Hill, 1960.

Berlyne, D.E. Aesthetics and psychobiology. New York: Appleton-Century-Crofts, 1971.

Crandall, R. Familiarity, Preference and Expectancy Arousal. *Journal of Experimental Psychology,* 1967, 73, *3,* 374-381.

Grush, J.E. Attitude formation and mere exposure phenomena: A non-artifactual explanation of empirical findings. *Journal of Personality and Social Psychology,* 1976, *33,* 282-290.

Hansen, F. Consumer choice behavior. A cognitive theory. New York: The Free Press, 1972.

Harrison, A.A. Mere Exposure. In: Z. Berkowitz (Ed.). Advances in Experimental Social Psychology. Vol. 10, 1977, Academic Press.

Harrison, A.A. Response competition, frequency, exploratory behavior, and liking. *Journal of Personality and Social Psychology,* 1968, Vol. 9, *4,* 363-368.

Jacobovits, L.A. Effects of Mere Exposure. *Journal of Personality and Social Psychology,* 1968, *9,* 2, 2.

Kagan, J. Motives and development. *Journal of Personality and Social Psychology,* 1972, *22,* 51-66.

Kelley, H.H. en J.L. Michela, Attribution theory and research. *Annual Review of Psychology,* 1980, *31,* 457-501.

Sawyer, A.G. The effects of repetition: conclusions and suggestions about laboratory research. In: G.D. Hughes and M.L. Ray (Eds.). Buyer/Consumer Information Processing. Chapel Hill: University of North Carolina Press, 1974, 190-219.

Stang, D.J. Six theories of repeated exposure and affect. Catalog of Selected movements in Psychology, 1973, *3,* 126.

Todd. Ph.P. Waargenomen risico. Jaarboek van de Nederlandse Vereniging van Marktonderzoekers, 1982, 159-181.

Vanbeselaere, N. Mere repeated exposure and experimental stress as determiners of stimulus evaluations. *Motivation and Emotion,* 1980, *4,* 1, 61-69.

Vroom, V.H. Work and motivation. New York: Wiley, 1964.

Zajonc, R.B. Attitudinal effects of mere exposure. *Journal of Personality and Social Psychology Monograph Supplement,* 1968, Vol. 2, *2.*

12. Meting van consumenten-ontevredenheid en -klachten

D.A. FRANCKEN

1. INLEIDING

In de zeventiger jaren was er een bloei in het onderzoek en de meting van sociaal welzijn. Diverse vertakkingen vonden plaats naar deelterreinen, zoals woonomgeving, werksituatie en gezinsleven.

Consumenten(on-)tevredenheid kan eveneens als zo'n deelterrein worden gezien. Handy (1972) heeft in dit verband opgemerkt:
'While we can measure the growth of our economy, efficiency of resource use, and product flows, we have virtually no information indicating the extent consumers feel the products and services available to them actually conform to their preferences and needs (Handy, 1972, p. 739).

Handy doelt hier op metingen van consumenten(on-)tevredenheid. De mogelijke gebruikers van deze metingen zijn overheid, bedrijfsleven en consumentenorganisaties. De overheid kan deze metingen bijvoorbeeld gebruiken als leidraad voor de urgentie van invoering van wettelijke bepalingen (Bernhardt, 1981), het bedrijfsleven voor de toetsing van haar marketingactiviteiten (Kasper, 1982b) en de consumentenorganisaties voor het stellen van prioriteiten bij hun activiteiten (Hunt, 1982).

In het afgelopen decennium zag het er inderdaad naar uit dat onderzoek en theorie op het terrein van consumenten(on-)tevredenheid en -klachten een hoge vlucht nemen, vooral in de V.S. Onder invloed echter van de 'Reaganomics' lijkt daar de vlucht (voorlopig?) te zijn afgelast (Hunt, 1982).

In deze literatuurstudie wordt geprobeerd een beeld te geven van het onderzoek en de meting van consumenten(on-)tevredenheid en klachten, dat in de afgelopen 10 jaar heeft plaatsgevonden.

Onderwerpen die aan bod komen zijn o.a. (1) de toepassingen en beperkingen van metingen van consumenten(on-)tevredenheid en klachten, (2) de verschillende soorten van deze metingen, (3) de ontwikkeling van een index van consumententevredenheid, (4) de uitkomsten van studies naar consumentenproblemen (welke problemen zijn er?, welke consumenten hebben deze?, hoe loopt het met deze problemen/klachten voor een consument af?)

2. EEN KORTE HISTORISCHE SCHETS

2.1. *Angelsaksische literatuur*

Ongeveer 18 jaar geleden verschenen de eerste publikaties over wat werd genoemd 'customer satisfaction' (Cardozo, 1965a en b).

Hierna bleef het nog verscheidene jaren stil tot 1972. In dat jaar werden diverse artikelen gepubliceerd, o.a. van Handy (1972), Pfaff (1972), Lingoes en Pfaff (1972), die betrekking hadden op onderzoek naar een index van consumententevredenheid. Dit onderzoek werd uitgevoerd in opdracht van het Amerikaanse Ministerie van Landbouw met als doel de tevredenheid van consumenten met voedselprodukten vast te stellen.

De populariteit van het onderwerp is in de V.S. vooral gegroeid in de jaren 1973 – 1975. De aanzet hiervoor kwam van de hierboven reeds vermelde studie in opdracht van het Amerikaanse Ministerie van Landbouw, maar vooral ook door een opdracht in 1975 van de Federal Trade Commission tot het houden van een nationaal onderzoek naar consumenten(on-)tevredenheid en -klachten wat heeft geresulteerd in de TARP-studie (1979). Deze studie geeft een inventarisatie van (1) de in de V.S. meest voorkomende consumentenklachten over produkten en diensten en (2) de al dan niet adequate maatregelen die door het bedrijfsleven en overheid hiervoor zijn getroffen.

Andere factoren speelden in op het succes en droegen verder bij tot de bloei van dit thema van onderzoek. Zo werd in 1976, in Bloomington, U.S.A., de eerste conferentie gehouden over consumenten(on-)tevredenheid en -klachten, georganiseerd door Keith Hunt. Hij heeft in samenwerking met Ralph Day deze conferenties tot een jaarlijkse gebeurtenis weten te maken, met elk jaar meer deelnemers.

De 'proceedings' van deze conferenties dekken nu naar ruwe schatting 80% van alle relevante wetenschappelijke onderzoeken en theorieën uit de jaren 1976 – 1982 op het terrein van CS/D & CB, zoals het onderwerp in de angelsaksische literatuur afgekort wordt aangeduid: Consumer Satisfaction/Dissatisfaction and Complaining Behavior.

In alle belangrijke marketing-, huishoudkunde- en consumentengedragtijdschriften in de Verenigde Staten en Europa zijn inmiddels artikelen verschenen over CS/D & CB.

Het totaal aantal wetenschappelijke publikaties bedroeg begin 1978 reeds ongeveer 150 artikelen (Robinson, 1979); dit aantal zal inmiddels (1983) naar schatting tot boven de driehonderd zijn opgelopen.

Maar zoals reeds gezegd: er lijkt een duidelijke hapering te zijn opgetreden in de produktiviteit in de V.S. als gevolg van de geheel andere prioriteiten die door de Reagan-regering in haar beleidsprogramma's worden gesteld.

2.2. *Nederland*

In Nederland is in 1974 een nationaal onderzoek naar consumentenklachten gedaan door IPM-Inter/View (1974), Dit onderzoek beperkte zich niet alleen tot een registratie van consumentenklachten, ook vragen met betrekking tot probleembewustheid voor marketingactiviteiten, klachtenactiviteiten (wat

heeft u met uw klacht gedaan?), wensen van consumenten (wat moet verbeterd worden?, en wie moet dat doen?) en meer van deze consumentistische onderwerpen waren in het onderzoek opgenomen.

Van Westendorp en Knecht (1974) maken melding van een onderzoek waarin aan een representatieve groep Nederlanders onder andere is gevraagd aan welke marketingactiviteiten men zich het meest heeft geërgerd in de voorgaande jaren. Prijs, kwaliteit en colportage bleken het meest aanleiding te geven tot ergernis. Bediening wordt eveneens vaak spontaan genoemd in dit verband. Inmiddels is het IPM-Inter/View onderzoek in 1981 herhaald, waarbij het data-materiaal van 1974 is geactualiseerd.

In 1980 is door SWOKA een lokaal onderzoek verricht naar consumentenklachten (Francken, 1983a). Dit onderzoek verschilt van de IPM-Inter/View onderzoeken doordat het minder op het consumentisme als zodanig ingaat en meer specifiek is gericht op consumentenproblemen zoals die door consumenten zelf worden ervaren. Het betreft vragen als: welke problemen zijn er?, welke consumenten hebben deze problemen?, wat zijn de oorzaken en gevolgen? en hoe loopt een klacht voor een consument af?

In Nederland zijn ook reeds diverse studies verricht waarbij een bepaalde produktgroep voorwerp van studie is, bijvoorbeeld levensmiddelen (Beukenkamp en Gründemann, 1979; Sanders, 1981).

Kasper (1982b) heeft in zijn proefschrift het effect van consumentenklachten op het marketing-beleid van ondernemingen belicht.

Tot zover een kort chronologisch overzicht van de activiteiten in het afgelopen decennium op het terrein van consumenten(on-)tevredenheid en -klachten.

In het volgende wordt ingegaan op deze onderzoeksinspanningen. Allereerst worden de mogelijke toepassingen en beperkingen van metingen van consumenten(on-)tevredenheid belicht.

3. TOEPASSINGEN EN BEPERKINGEN VAN METINGEN VAN CONSUMENTEN(ON-)TEVREDENHEID

Bij het gebruik van metingen van consumenten(on-)tevredenheid moeten een drietal punten nader worden uitgewerkt. Deze punten betreffen meettechnische en conceptuele aspecten.

3.1. *Specificiteit van de meting*

Voor het beleid zal een algemene index van consumenten(on-)tevredenheid weliswaar nuttig kunnen zijn (vergelijk dit met bijv. het prijsindexcijfer of de index van het consumentenvertrouwen), maar voor diverse doeleinden is zo'n index toch ook te globaal en generaliserend. Voor een adequate afstemming van beleidsmaatregelen op 'zere' consumptieplekken is het van belang precies te weten welke categorieën/merken/attributen van produkten en diensten meer respectievelijk minder bijdragen tot consumenten(on-)tevredenheid. Dit vereist meer specifieke metingen.

Renoux (1973) maakt in dit verband onderscheid tussen ontevredenheid met het 'macro-marketing systeem' en met het 'micro-marketing systeem'. De

eerste soort van ontevredenheid betreft een meer algemene sociale indicator en is niet gericht op specifieke produkten, distributiekanalen, producenten.

Het tweede soort ontevredenheid deelt Renoux op in ontevredenheid met:
1. de verkrijgbaarheid ('shopping system'), dit betreft problemen met de verkrijgbaarheid en beschikbaarheid van produkten.
2. de koopkeuze ('buying system'), dit betreft problemen met de keuzemogelijkheden bij de koop en met de feitelijke aanschaf.
3. de consumptie ('consuming system'), dit heeft betrekking op de problemen als gevolg van het gebruik van goederen en diensten.

Czepiel, Rosenberg en Akerele (1975) hanteren een ander onderscheid, nl.: tevredenheid met 1. het systeem; 2. het bedrijf ('enterprise') en 3. produkten/diensten. Ook in dit onderscheid zijn duidelijk niveaus van toenemende specificiteit te herkennen.

Naast het probleem van de specificiteit van de meting kleven er conceptueel en meettechnisch nog meer problemen aan de meting van consumenten(on-)tevredenheid.

3.2. *Dimensionaliteit van de meting*

Een ander probleem met het concept consumenten(on-)tevredenheid betreft de dimensionaliteit van het begrip. Is iemand die tevreden is, dezelfde als iemand die niet ontevreden is? En omgekeerd.

Herzberg e.a. (1959) merken hierover op:
'The opposite of job satisfaction is not job dissatisfaction but, rather *no* job satisfaction; and, similarly, the opposite of job dissatisfaction is not job satisfaction but *no* job dissatisfaction (Herzberg, Mausner en Snyderman, 1959, p. 108).

Bij de tevredenheid met consumptie(-goederen) doet zich dit probleem evenzo voor (Leavitt, 1977).

Een hypothese is dat de (on-)tevredenheid met alledaagse consumptie-artikelen meer berust op 'dissatisfiers' (schaal: niet ontevreden – ontevreden), terwijl dit met de luxe-goederen meer berust op 'satisfiers' (schaal: tevreden – niet tevreden).

3.3. *Conceptuele kanttekeningen*

Ölander (1976) stelt de kritische vraag of consumententevredenheid wel een na te streven, geschikte doelstelling is voor consumentenbeleid. Hij wijst op diverse psychologische processen, die het meten van 'werkelijke' tevredenheid van individuen met een object of verschijnsel bemoeilijken, zo niet onmogelijk maken:
1. *Vergelijkingsprocessen:* omdat tevredenheid een concept is dat altijd wordt beoordeeld in relatie tot een bepaald criterium kan de hoogte (het niveau) van dit criterium voor individuen verschillen. Dit als gevolg van individuele verschillen in sociale, ruimtelijke en volgtijdelijke vergelijkingsprocessen. Bij een hoger niveau zal men minder snel tevreden zijn.

2. *Kennis van alternatieven:* verschillen in kennis over wat-anders-mogelijk-is leidt tot individuele verschillen in tevredenheid: meer kennis over gunstige alternatieven leidt minder snel tot tevredenheid en omgekeerd.
3. *Cognitieve dissonantie:* onaangename zaken worden vaak gerationaliseerd, en hierdoor minder negatief, wat de tevredenheid ermee kan vergroten.
4. *Hier-en-nu percepties:* beoordelingen van zaken hebben veelal alleen betrekking op korte-termijn aspecten, lange-termijn aspecten zijn vaak ondoorzichtig en worden over het hoofd gezien.
5. *Verwachtingen:* verschillen in verwachtingen leiden tot individuele verschillen in tevredenheid. Bij hogere verwachtingen zal men minder snel tevreden zijn.

De strekking van de bovenstaande punten is duidelijk: een consument is meer tevreden naarmate hij minder weet, minder eist, minder verwacht. En dit kan natuurlijk nooit de hoofddoelstelling van een consumentenbeleid zijn.

Ölander komt tot de volgende, uitdagende stelling:
Somewhat provocatively, one might maintain that the goal of consumer policy ought to be to maximize *dissatisfaction* since, presumably, only dissatisfied consumers wil voluntarily pass on whatever faults they find in products and service (Ölander, 1976, p. 22).

Men kan Ölander echter gemakkelijk tegenwerpen dat de bezwaren die hij heeft tegen het hanteren van een concept als tevredenheid evenzo gelden voor ontevredenheid. Als zodanig is zijn stelling snel ondergraven (Pfaff, 1977b).

De bovenstaande punten van Ölander maken echter wel duidelijk dat men goed oog moet hebben voor de diverse valkuilen. Op het gevaar af dat een onderzoek anders precies datgene aan conclusies oplevert, wat men zelf impliciet als vooronderstellingen in de vragen heeft gestopt.

4. EEN TAXONOMIE VAN METINGEN

Andreasen (1977) geeft een taxonomie van metingen van consumenten(on-)tevredenheid. Hij onderscheidt acht soorten van metingen op basis van de volgende indeling:
a. het onderwerp: tevredenheid ('maximizing') of ontevredenheid ('minimizing');
b. de methode: subjectieve of objectieve meting;
c. het tijdstip: voor of na het moment dat de leverancier/verkoper in staat is gesteld iets aan het probleem te doen; hierop is gebaseerd het onderscheid tussen aanvankelijke en uiteindelijke tevredenheid.

De meeste metingen van consumenten(on-)tevredenheid die worden verricht vallen in de volgende drie categorieën.

1. *Objectieve indicatoren van uiteindelijke tevredenheid,* bijvoorbeeld: verkochte hoeveelheden, marktaandelen, herhalingsaankopen. Deze metingen worden vooral in het bedrijfsleven gehanteerd.
   De beperkingen van deze metingen zijn dat achteraf wordt gemeten, zodat gegevens vaak laat worden verkregen (vooral bij duurzame consumptiegoe-

deren) en dat veel verstorende factoren kunnen optreden (bijv. concurrentie).

2. *Subjectieve indicatoren van aanvankelijke tevredenheid,* bijvoorbeeld: tevredenheidsschalen zoals die veelvuldig door marktonderzoekbureaus worden gebruikt.
   De bezwaren hiertegen zijn al in de vorige paragraaf beschreven.
3. *Subjectieve indicatoren van ontevredenheid,* subjectief wil hier zeggen: door de consument zelf als zodanig ervaren. Deze metingen zijn het meest relevant voor het consumentenbeleid.

Twee soorten van subjectieve indicatoren van consumentenontevredenheid zijn te onderscheiden (Cavusgil en Kaynak, 1980; Pfaff, 1977a):
- indexering van consumenten(on-)tevredenheid.
- inventarisatie van consumentenproblemen.

Indexering van consumenten(on-)tevredenheid en inventarisaties van consumentenproblemen kunnen complementair worden gebruikt bij de meting van consumenten(on-)tevredenheid.

De eerste meetmethode onderzoekt 'waar' (produktcategorie) iets fout zit, zij is registrerend en meer kwantitatief van aard. De tweede meetmethode onderzoekt 'wat' er fout zit, deze meetmethode is meer verklarend en kwalitatief van aard. Het gaat om enerzijds het constateren dat de tevredenheid met een bepaald(e) produkt(-categorie) is toe- of afgenomen ten opzichte van een voorgaand meettijdstip en anderzijds het vinden van een oorzaak (verklaring) hiervoor.

5. INDEXERING VAN CONSUMENTEN(ON-)TEVREDENHEID

5.1. *Indicaties*

Consumentenproblemen vormen slechts een deel van de maatschappelijke problemen in onze tijd. Hoe groot en belangrijk dit deel is, is niet precies aan te geven.

In een nationaal onderzoek in 1971 in de V.S. (Barksdale en Darden, 1972) is aan consumenten gevraagd in hoeverre zij het eens zijn met de stelling dat consumentenproblemen relatief *on*belangrijk zijn in vergelijking met de andere zaken waarmee een gemiddeld gezin te maken heeft. Tweederde deel van de ondervraagden is het hiermee niet eens, een kwart is het hiermee eens en 10 procent houdt het antwoord in het midden (van de schaal). Dezelfde stelling is zes jaar later (1977) nogmaals voorgelegd aan consumenten in een internationaal onderzoek in de V.S. en in Zweden. Het percentage personen dat het niet eens is met deze stelling is dan gedaald met ongeveer de helft: zowel in de V.S. als in Zweden is ongeveer éénderde van de consumenten in 1977 het niet eens met de stelling dat consumentenproblemen betrekkelijk *on*belangrijk zijn in vergelijking met de andere problemen in een gemiddeld gezin (Wikström, 1983).

Met de stelling dat het in het algemeen voor consumenten in de Verenigde Staten gemakkelijk is om problemen hersteld te krijgen, is 70 procent het *on-*

eens (Barksdale en Darden, 1972).

In een meer recent onderzoek in Nederland (IPM-Inter/View, 1981) is gevraagd over welke dingen men zich persoonlijk zorgen maakt. Bijna de helft maakt zich voortdurend zorgen over zaken als schade aan de natuur en gevaren van kernenergie. Consumentenzaken blijken in verhouding hiermee een iets minder grote bron van zorg te zijn. Over de invloed van de consument is een kleine 10 procent voortdurend bezorgd en 35 procent is af en toe bezorgd.

Over de macht van de grote internationale bedrijven en concerns is 20 procent voortdurend bezorgd, ongeveer 35 procent is af en toe bezorgd en 45 procent is dit zelden.

5.2. *Ontwikkelingen*

Betrouwbare en valide meetmethoden van consumenten(on-)tevredenheid, die gemakkelijk te hanteren zijn en toch ook beantwoorden aan te stellen eisen van fijngevoeligheid bij meting, kunnen een belangrijke graadmeter zijn bij het maken van consumentenbeleid.

Pfaff (1977a) heeft een diepgaande verhandeling gegeven over welke punten belangrijk zijn bij de ontwikkeling van een index van consumenten(on-)tevredenheid. Zijn werk en dat van anderen heeft geresulteerd in een indexering van consumenten(on-)tevredenheid met voeding (Handy en Pfaff, 1975; Handy, 1977).

Indexering van consumenten(on-)tevredenheid staat verder nog in haar kinderschoenen zowel voor marktgoederen als voor door de overheid verstrekte goederen. Denk hier ook aan bijvoorbeeld gemeentediensten (Bernhardt, 1981).

Pfaff (1977a) geeft duidelijk de functie van zo'n index aan.

Particularly as a complement to the consumer price index, which focuses more on the quantity of goods consumed and the prices at which these goods are available, the Index of Consumer Satisfaction provides useful insight into qualitative aspects of which price is only one.

At a time when both companies and the government are anxious to avoid crises of confidence between themselves and the consumer, an Index of Consumer Satisfaction could provide an early warning signal at the micro level of problems that have arisen. These then could serve as a starting point for more detailed and exhaustive analyses (Pfaff, 1977a, p. 62).

Pfaff onderscheidt hier dus ook 2 soorten metingen: een indexering van consumenten(on-)tevredenheid en, waar nodig en in aanvulling hierop, diepergaande inventarisaties van consumentenproblemen.

6. INVENTARISATIES VAN CONSUMENTENPROBLEMEN

6.1. *Het nut*

De inventariserende studies van consumentenproblemen zijn nagenoeg alle survey-onderzoeken waarin is geprobeerd consumentenproblemen subjectief

te meten, wat inhoudt dat het gaat om problemen die door de consument zelf als zodanig worden ervaren.

Dit is ook de reden om hier het woord problemen te gebruiken en niet klachten. Klachten zijn immers te omschrijven als mondelinge of schriftelijke uitingen van consumentenproblemen. Ze worden aan een ander geuit. Deze andere partij kan zijn een huisgenoot ('eerste partij'), de leverancier ('tweede partij') of een consumentenorganisatie ('derde partij').

Iemand die een consumentenprobleem heeft, maar hieraan geen uiting geeft, kent wat Beukenkamp en Gründemann (1979) noemen: een *actie-drempel.* Dit uit zich in een gebrek aan 'actievermogen' als gevolg van bijvoorbeeld passiviteit of onzekerheid over de vraag of een klacht wel gegrond en op zijn plaats is. Naast actie-drempels onderscheiden Beukenkamp en Gründemann ook entree-drempels en communicatiedrempels.

Entree-drempels hebben betrekking op bijvoorbeeld kennis van klachtenkanalen en op de (fysieke) moeite die een consument ondervindt als hij zijn klacht bij een persoon of instantie wil deponeren, bijvoorbeeld: onwillige winkelier, overbelaste telefoonlijnen bij consumentenorganisaties, hoge klachtengelden bij geschillencommissies en ingewikkelde procedures bij kantongerechten.

Communicatie-drempels betreffen met name de verbale en schriftelijke vaardigheden van de consument om zijn problemen in gerechtvaardigde klachten te vertalen en zijn recht te halen.

Deze indeling in drie soorten drempels is ontleend aan een klachtenmodel van Day en Landon (1977, p. 432).

Dit model kent de volgende (beslissings-)stadia tussen het ervaren van ontevredenheid als gevolg van een consumentenprobleem en de uiteindelijke specifieke (klachten-)acties.

Stadia van het klachtenproces

1. ontevredenheid als gevolg van een consumentenprobleem
2. overweging al dan niet actie te ondernemen — actie-drempels
3. afweging welke soort actie: in de privé-sfeer of meer publiekelijk — entree-drempels
4. keuze van specifieke klachtenactie bijvoorbeeld: — communicatie-drempels
   a. *privé-sfeer* (informeel)
      - besluiten produkt niet meer te kopen
      - anderen waarschuwen voor betreffend produkt
   b. *publiekelijk* (formeel)
      - klagen bij leverancier ('tweede' partij)
      - consumentenorganisatie inschakelen ('derde' partij)
      - klacht voorleggen aan geschillencommissie ('derde' partij)

De drempels die een consument kan ondervinden als hij een probleem wil vertalen in een concrete klacht, die aan een 'tweede' of een 'derde' partij wordt geuit, zijn velerlei.

Het rechtvaardigt de conclusie dat inventariserende studies van consumentenklachten bij bedrijven, consumentenorganisaties, klachteninstanties e.d. niet meer laten zien dan het bevroren topje van de ijsberg (Kasper, 1982a; Schutz, 1979).

Het geeft tevens de noodzaak aan van subjectieve metingen van consumentenontevredenheid.

6.2. *De kenmerken*

De inventariserende studies van consumentenproblemen centreren zich in hoofdzaak om de volgende vragen:

1. wie zijn de gebruikers van een bepaald(e) produkt of dienst?
2. wie heeft problemen hiermee en welke problemen zijn dit?
3. hoe ontevreden is men over het produkt?
4. wat wordt met het probleem gedaan, welke acties worden ondernomen?

De onderzoeken van o.a. Andreasen en Best (1977) en de TARP-studie (1979), waarvan het veldwerk in 1976 is uitgevoerd, behoren tot dit soort metingen.

De TARP-studie heeft een uitstralingseffect gehad naar diverse (vervolg-)-studies in Californië (Schutz, 1980), in Bloomington (Day en Bodur, 1978; Leigh en Day, 1979) en in Canada (Ash, 1982).

In 1980 heeft Andreasen geprobeerd een dergelijk onderzoeksproject op te zetten voor diverse Noordwesteuropese landen.

Omdat hij de financiering van dit project bij een centrale EEG-instantie niet rond kreeg, is dit project uiteengevallen in diverse min of meer zelfstandige onderzoeken in Duitsland (Bruhn), Zweden (Julander), Denemarken (Hansen, Kristensen), Nederland (Francken en Van Raaij).

Deze onderzoeken zijn, in zoverre ze ook daadwerkelijk zijn uitgevoerd, te kenmerken als:

a. *grootschalig:* veel personen worden benaderd om voldoende consumenten te hebben met een bepaald probleem, voor verdere analyses.
b. *uitgebreid:* het hele consumptieproces wordt doorgelicht van winkelen tot het afdanken van een artikel.
c. *subjectief:* de nadruk ligt op wat individuele consumenten zelf als een probleem ervaren.
d. *gestructureerd:* duidelijk en expliciet wordt aangegeven welke consumentenproblemen zich op diverse terreinen zoal kunnen voordoen.

De studies van Bruhn (1982) en Francken (1983a) zijn voorbeelden van dit soort onderzoeken.

Het onderzoeksproject zoals Andreasen dat voor ogen had is in zijn oorspronkelijke, cross-culturele opzet mislukt. De moeilijkheden die een Amerikaan ondervindt als hij in Europa cross-cultureel onderzoek wil doen, gefinancierd door een centrale EEG-instelling, zijn door Andreasen op schrift gesteld (Andreasen en Manning, 1980).

Andreasen had reeds zijn sporen verdiend op het terrein van consumentenklachtenonderzoek (Andreasen, 1977; Andreasen en Best, 1977).

6.3. *Inventarisaties*

Op diverse (deel-)terreinen zijn inmiddels inventarisaties van consumentenproblemen verricht:
- *auto*(-reparaties): Bearden, Crockett en Graham (1980); Bruhn (1982); Francken (1983a);
- *chemisch reinigen:* Francken (1983a);
- *commerciële diensten:* Day en Bodur (1977);
- *consumptief krediet:* Sauer, Chaiy en Schweitzer (1982);
- *garanties:* Darden en Rao (reparaties, 1977); Ash, Kennedy en Thirkell (auto's, 1980); Bernacchi, Kono en Willette (1980);
- *gezondheidszorg:* Swan en Carroll (ziekenzorg, 1980); Chaiy en Saxton (tandartsen, 1982);
- *huizen/wonen:* Kaynak en Stevenson (1982);
- *keukenmachines:* Westbrook (1977);
- *kleding:* Kelly (1979); Francis en Dickey (1982); Francken (1983a);
- *klachtenbureaus:* Geistfeld en Choy (1978); Ittig (1978); Bernacchi, Kono en Smith (1979); Grainer, Goodman en Hanson (1979); Dickinson en Lang (1982);
- kleurentelevisies: Bruhn (1982), Francken (1983a);
- *nutsbedrijven/publieke diensten:* Bearden e.a. (1977); Ash en Quelch (1980); Bodur, Borak en Kurtulus (publieke diensten, 1980); Marlowe (vuilnisophaaldienst, 1982); Spalding en Marcus (PTT-diensten, 1982);
- *onderwijs:* Krishnan en Grønhaug (1980); Bodur en Osdiken (1982);
- *reclame:* Olson (1979); Scammon en Miller (1979); Van Westendorp en Knecht (1974);
- *vakantiereizen:* Francken (1983a);
- *voeding:* Handy en Pfaff (1975); Handy (1977); Beukenkamp en Gründemann (1979); Sanders (1981);
- *winkelbedrijven:* Maddox (1977); Krishnan en Mills (1979); Rodgers en Sweeney (1980).

6.4. *Bevindingen van onderzoek*

6.4.1. *Consumentenproblemen*

Uit onderzoek over 1979 en 1980 in Nederland (Francken, 1983a en IPM-Inter/View, 1981) blijkt dat de te hoge prijzen het meest voorkomende consumentenprobleem vormen. In 1974 kwamen Van Westendorp en Knecht tot een zelfde conclusie.

Vier van de vijf consumenten ondervonden één of meermalen in een jaar problemen met te hoge prijzen (IPM-Inter/View, 1981). Andere marketingactiviteiten waarmee meer dan de helft van alle consumenten één of meermalen in een jaar problemen ondervonden zijn: service en voorlichting door personeel; kwaliteit van goederen en diensten; gebrekkige gebruiksaanwijzing resp. prijsaanduiding; misleidende reclame/verkooppraktijken (IPM-Inter/View, 1981; Francken, 1983a).

De produkten en diensten waarmee zich in 1980 in Nederland de meeste

problemen hebben voorgedaan zijn:

1. Voedingsartikelen (>50%); 2. Kleding en schoenen (>50%); 3. Autoreparaties (>25%) en 4. Geluidsapparatuur en kleurentelevisies (>25%).

Tussen haakjes staan de percentages consumenten, die één of meermalen in een jaar problemen met de betreffende produkten ondervonden (IPM-Inter/View, 1981).

Het percentage consumenten dat jaarlijks een probleem met een goed of dienst ervaart, is vanzelfsprekend afhankelijk van wat men onder een probleem verstaat. Opgaven van het bedrijfsleven hierover hebben veelal alleen betrekking op de (technische) kwaliteit van een goed.

Uit het bovenstaande blijkt, dat wanneer andere elementen van de marketing-mix (prijs, promotie, distributie, personeel) erbij worden betrokken, het totaal aantal consumentenproblemen immense proporties aanneemt.

Het percentage consumenten dat jaarlijks een of meer problemen ondervindt is niet precies aan te geven.

Uit onderzoek komen nogal sterk verschillende percentages naar voren: 32% (TARP, 1979), 54% (Francken, 1983a) en 75% (Shuptrine en Wenglorz, 1980).

Het eerste en laatste percentage hebben betrekking op studies in de V.S.

6.4.2. *Klachtenacties*

Lang niet iedere consument, die een probleem heeft met een goed of dienst, zal hiermee naar de leverancier (tweede partij) of naar een consumentenorganisatie e.d. (derde partij) stappen.

Uit onderzoek van Francken (1983a) bij een zestal produkten en diensten blijkt het percentage consumenten dat helemaal niets onderneemt naar aanleiding van een probleem, te variëren van 14 procent bij autoreparaties tot 47 procent bij kleine kledingstukken (sokken, ondergoed, panties, T-shirts, e.d.) Zie tabel 1.

Slechts 18 procent van de consumenten die een probleem hebben met een

Tabel 1. Soorten van klachtenacties die consumenten ondernemen bij een probleem met een produkt of dienst (Francken, 1983a)

Produkt/dienst	(1) geen actie %	(2) privé actie %	(3) leverancier %	(4) derde partij %	(5) gecombineerde actie %	aantal personen
a. autoreparaties	14	11	71	18	31	93
b. bovenkleding	29	22	48	5	24	96
c. 'kleine' kledingstukken	47	35	18	1	9	89
d. chemisch reinigen	25	13	61	9	26	85
e. pakketreizen	18	22	58	26	44	50
f. kleurentelevisie	37	6	56	2	18	98
aantal consumentenacties	150	90	263	45	122	T=511 T=565*

* Consumenten kunnen naar aanleiding van één probleem verscheidene acties hebben ondernomen.

klein kledingstuk, klaagt bij de leverancier of verkoper. Wel onderneemt 35 procent een privé-actie, zoals niet meer hetzelfde merk of in dezelfde winkel kopen en/of anderen waarschuwen voor het produkt of de winkel/zaak.

Bij problemen met autoreparaties (71%) gaan de consumenten het meest naar de leverancier om te klagen.

Acties waarbij een derde partij (consumentenorganisatie, geschillencommissie, deurwaarder, kantongerecht) is betrokken komen het meest voor bij problemen met pakketreizen (geboekte vakanties) – 26%. Bij dit soort problemen komt ook het meest voor dat een consument meerdere acties onderneemt (44%). Zie tabel 1.

Uit een studie van Day en Bodur (1978) blijkt dat mensen die zeer ontevreden zijn over een niet-duurzaam produkt in 50% van de gevallen geen enkele consumentenactie onderneemt, voor duurzame produkten is dit 30% en voor diensten 23%.

Uit een onderzoek van IPM-Inter/View (1981) blijkt dat de volgende percentages consumenten een bepaalde actie hebben ondernomen in het voorafgaand jaar (1980) als gevolg van ergernissen en klachten met produkten of diensten. Duidelijk komt naar voren hoeveel 'actie' er zit in consumentenproblemen.

Consumentenacties

bij leverancier	
– produkten waaraan iets mankeerde teruggebracht naar de leverancier	63%
– telefonisch of schriftelijk geklaagd bij de leverancier of fabrikant	53%
– bij de leverancier geklaagd over een verkeerde prijs of een verkeerde afrekening	38%
bij consumentenorganisatie	
– geklaagd bij consumentenorganisatie of overheidsinstelling over bepaalde produkten of diensten	6%
– lid geworden van een consumentenorganisatie	11%
in de privé-sfeer (informeel)	
– produkt of merk tijdlang niet gekocht	37%
– bij een winkel waarover men ontevreden was niet meer gekocht	49%
– met andere mensen gezamenlijk protest aangetekend	5%
– meegedaan aan een boycot voor een goed doel	13%
– geklaagd bij vrienden, kennissen, buren e.d. over de kwaliteit of prijzen	58%
– niet betalen	7%

N.B.: Consumenten kunnen naar aanleiding van één probleem verscheidene acties hebben ondernomen.

6.4.3. *Kenmerken van klagers*

De onderzoeken naar kenmerken van klagers hebben zich in hoofdzaak geconcentreerd op (1) socio-demografische kenmerken, (2) houdingen (attitudes) en (3) persoonlijkheidskenmerken.

Socio-demografische kenmerken

Liefeld e.a. (1975) concluderen uit een onderzoek onder consumenten, die een klachtenbrief hadden geschreven aan een consumentenorganisatie of overheidsbureau in Canada, dat de brieven meer dan gemiddeld behoren tot de beter opgeleide en rijkere bovenlaag van de bevolking, werkzaam in leidinggevende functies en/of met een vrij beroep.

Mason en Himes (1973) melden dat in hun studie consumenten, die een klachtenactie ondernemen, deel uitmaken van grotere huishoudingen, een gemiddeld hoger inkomen hebben en hoofdzakelijk van middelbare leeftijd zijn.

Miller (1970) vindt in een onderzoek dat consumenten die een klacht hebben jonger zijn en meer mobiel. Zaltman et al. (1978) vinden eveneens bij jongere mensen een grotere geneigdheid tot klachtenacties. Studies van Pfaff en Blivice (1977) en Zaichkowsky en Liefeld (1977) bevestigen bovenstaande uitkomsten.

Maar niet alle studies geven éénduidige uitkomsten met betrekking tot de socio-demografische kenmerken van de klagers. Grønhaug (1977) en Granbois et al. (1977) concluderen uit hun studies dat socio-demografische variabelen geen beschrijvende of voorspellende waarde hebben voor klaaggedrag (klachtenacties). Ook Andreasen (1977) betwijfelt de verklarende waarde van socio-demografische variabelen met betrekking tot klachtenacties. Volgens hem zijn vooral het soort produkt/dienst, waarmee zich een probleem voordoet, en de aard van het probleem bepalend voor het al dan niet ondernemen van een klachtenactie door een consument. Socio-demografische kenmerken zijn volgens Andreasen hiervan slechts een afgeleide, dat wil zeggen zij hangen samen met het soort produkt dat mensen kopen en met de produkteigenschappen die belangrijk worden gevonden; dit is dan bepalend voor het al dan niet ervaren van consumentenproblemen en eventueel hieruit voortkomende klachtenacties.

Houdingen

Warland, Herrmann en Willits (1975) zijn min of meer de eersten geweest die hebben onderzocht in hoeverre het ondernemen van een klachtenactie niet alleen vanuit socio-demografische kenmerken van consumenten is te verklaren, maar ook vanuit houdingen die een consument heeft.

Uit hun studie blijkt, dat mensen die actie ondernemen, wanneer zich een consumentenprobleem voordoet, weliswaar jonger zijn en meer dan gemiddeld behoren tot de groep met een hoger inkomen, hogere opleiding en hogere sociale status, maar deze mensen zijn ook vanuit hun houding meer geïnteresseerd in consumentisme en meer ontevreden over diverse praktijken in het bedrijfsleven. Deze personen blijken bovendien meer actief te zijn in het dagelijkse sociale en politieke leven om hen heen (Warland en Herrmann, 1981).

Grønhaug en Zaltman (1981) voeren 'marketplace participation' aan als verklaring voor het al dan niet ondernemen van klachtenacties. Zij wijzen er-

op dat één van de rollen die een persoon in onze maatschappij moet kennen betrekking heeft op de consumptie van goederen en diensten. Sommige personen beheersen deze rol als consument (koper) beter dan anderen, dankzij meer koopervaring. Meer koopervaring verwijst naar het vaker kopen van een produkt in het verleden, wat de kans op een probleem groter maakt. Dit zou dan een verklaring kunnen zijn voor het vaker ondernemen van klachtenacties bij een hogere 'marktparticipatie'.

Als wordt bedacht dat personen in de hogere klassen van beroep, inkomen en opleiding veelal over meer geld kunnen beschikken, en dus meer produktaankopen kunnen doen, en dus meer consumentenproblemen kunnen hebben, dan is ook te verklaren dat deze socio-demografische kenmerken volgens diverse onderzoeken (zie hierboven) samenhangen met het al dan niet vaker ondernemen van klachtenacties. Zo loopt ongeveer de redenering van Grønhaug en Zaltman.

Francken (1983a) komt tot een min of meer gelijke redenering, maar dan gebaseerd op een kritische consumenteninstelling. Uit zijn onderzoek blijkt dat personen die een betere rechtspositie van de consument voorstaan, meer consumentenproblemen hebben. Hierentegen komen minder consumentenproblemen voor bij personen die weinig of geen regulering van het marktmechanisme door de overheid voorstaan (in hoeverre hier sprake is van oorzaakgevolg, of omgekeerd, laat zich raden). Wellicht heeft de eerste groep een lagere drempel voor probleemervaring, of anders gezegd: heeft zij een meer kritische consumenteninstelling dan de tweede groep. In dit onderzoek blijkt verder dat meer consumentenproblemen worden ervaren door personen die jonger zijn, een hogere opleiding hebben en deel uitmaken van grotere huishoudens. Minder consumentenproblemen komen voor bij oudere personen (m.n. boven 65 jaar) en personen met een hoger inkomen en beroep.

In tegenstelling tot Grønhaug en Zaltman is een verband tussen socio-demografische kenmerken en het al dan niet ondernemen van klachtenacties hier dus niet zozeer te verklaren vanuit een grotere kans op problemen bij meer produktaankopen, maar vanuit een meer kritische houding als consument. Hierbij wordt dan min of meer verondersteld dat het ervaren van meer consumentenproblemen tot meer klachtenacties leidt.

Het IPM-Inter/View onderzoek (1981) toont aan dat consumentisten aanzienlijk meer consumentenacties ondernemen in een jaar dan mensen die niet (minder) consumentistisch zijn ingesteld.

Richins (1982) noemt uit onderzoek een drietal basisfactoren die ten grondslag zouden liggen aan houdingen ten aanzien van klachtenacties.

1. afweging van baten en kosten (materieel en psychologisch) met betrekking tot de vraag of klagen loont (persoonlijk nut);
2. de individuele normen ten aanzien van klagen;
3. het maatschappelijk nut dat men ziet in klagen.

Persoonlijkheidskenmerken

Houdingen van consumenten, al dan niet causaal samenhangend met diverse socio-demografische kenmerken, zijn niet de enige onderzochte persoonskenmerken van klagende consumenten.

Ook persoonlijkheidskenmerken komen in de onderzoekliteratuur aan bod: maatschappelijke vervreemding, interne-externe beheersing (locus of control), zelfvertrouwen, sociaal isolement, politieke handigheid ('political efficacy'), assertiviteit, introvertie-extravertie, attributie.

Robinson (1979) geeft een beknopt overzicht van de uitkomsten van diverse studies op dit terrein. Consumenten zouden sneller tot een klachtenactie overgaan als zij meer zelfvertrouwen hebben en meer 'intern gecontroleerd' zijn, d.w.z. als zij de oorzaak van wat hen in het dagelijkse leven overkomt eerder toeschrijven aan eigen presteren of falen dan aan toeval of invloedrijke derden, de zogenaamde 'powerful others'.

Politieke handigheid zou eveneens klachtenacties bevorderen. Minder klachtenacties komen voor bij mensen die sociaal geïsoleerd en vervreemd zijn (met name minderheidsgroeperingen).

De uitkomsten van deze studies zijn echter niet éénduidig. Diverse onderzoekers hebben de verklarende waarde van persoonlijkheidskenmerken voor het al dan niet ondernemen van klachtenacties niet of nauwelijks kunnen aantonen (Zaichkowsky en Liefeld, 1977; Fornell en Westbrook, 1978).

Succes wordt wel gemeld door Grabicke, Schaetzle en Staubach (1982). Met een uitgebreide testbatterij voor het meten van de geneigdheid tot klagen ('propensity to complain'), komen zij op een totaal verklaard percentage van 36 procent van klachtenacties voor consumenten.

In hun studie komen zij tot de volgende constatering:

'...it is evident that most studies show that only a relatively small portion of the total variance of consumer behavior can be explained through personality characteristics. The most important reason for these unsatisfactory results is found in the set of inadequate or insufficiently validated and reliable test instruments. The same thing applies, except for a few special cases, to the description of personality characteristics by complaining behavior researchers' (Grabicke, Schaetzle, Staubach, 1982, p. 50).

Eén van de meest onderzochte en relevant bevonden psychologische variabelen met betrekking tot klachtenacties is *attributie*. Deze variabele slaat op de toeschrijving van de oorzaak van een probleem door een consument aan bijvoorbeeld de leverancier, het toeval of eigen fouten (Hunt, 1980).

Francken (1983b) gebruikt de veronderstelde werking van deze variabele om meer inzicht te geven in de soort klachtenacties die een consument zal ondernemen. Hiermee wordt tegemoet gekomen aan een beperking van nagenoeg alle studies. De studies beperken zich tot de vraag welke consumenten *wel* actie ondernemen en welke consumenten dit *niet* doen. Geen enkele studie geeft een specificatie naar de soort actie die zal worden ondernomen.

De hoeveelheid stress (narigheid) die een consument van een probleem ondervindt en de toeschrijving ('attributie') van de oorzaak van het probleem vormen de twee ingangen van figuur 1.

Aan de hand van deze tweedeling wordt inzichtelijk gemaakt in welke situaties een bepaald soort consumentenactie het meest mag worden verwacht. Zie figuur 1.

In figuur 1 zijn twee soorten van uitkomsten onderscheiden die een consument met een actie kan beogen, psychologische en materiële compensatie.

AANZET KLACHTENACTIE (actie: ja of nee)	RICHTING KLACHTENACTIE (welke actie?)	
	GEEN TOESCHRIJVING VAN OORZAAK AAN LEVERANCIER	WEL TOESCHRIJVING VAN OORZAAK AAN LEVERANCIER
MINDER STRESS – minder hinder, overlast – minder financieel verlies – minder aantasting humeur	*geen actie* – verdringen – vergeten – dissonantie-reductie	*direct verhaal* (bij tweede partij) – mondelinge regeling – brief aan leverancier/verkoper/fabrikant
	PSYCHOLOGISCHE COMPENSATIE	MATERIËLE COMPENSATIE
MEER STRESS – meer hinder, overlast – meer financieel verlies – meer aantasting humeur	*privé-actie* – niet meer kopen – kwaad gerucht	*wettelijk verhaal* (bij derde partij) – consumentenorganisatie – geschillencommissie – kantongerecht

Figuur 1. Theoretische indeling van consumentenklachtenacties (Francken, 1983b).

Verondersteld wordt dat een consument minder gericht zal zijn op psychologische compensatie en meer op materiële compensatie, naarmate hij meer overtuigd is dat de oorzaak van het probleem aan de leverancier (tegenpartij) kan worden toegeschreven. De hoeveelheid stress (narigheid) die een consument als gevolg van het probleem ervaart, wordt als een belangrijke motor voor het al dan niet ondernemen van een klachtenactie gezien, met name voor de 'zwaarte' ervan.

7. SLOT

Veel van de aangehaalde onderzoeksliteratuur in dit artikel is van Amerikaanse hand. De afgelopen 10 jaar heeft een 'boom' laten zien in het onderzoek en de meting van consumenten(on-)tevredenheid en klachten in de Verenigde Staten. In tegenstelling tot in de Verenigde Staten is in Nederland nog nauwelijks geprobeerd metingen van consumenten(on-)tevredenheid als graadmeter bij het consumentenbeleid te hanteren. De kansen hierop lijken echter op dit moment in Nederland niet ongunstig te zijn, wellicht zelfs gunstiger dan in de Verenigde Staten, waar de 'Reaganomics' hard toeslaat op sommige terreinen van wetenschappelijk onderzoek. Denk aan de hoofddoelstelling van het consumptiebeleid in Nederland: 'het verbeteren van de mogelijkheden om door middel van consumptie behoeften te bevredigen' (Nota Consument en Con-

sumptie, 1979, p. 9).

Bij toetsing van een dergelijke doelstelling zal men moeilijk om de meting van een concept zoals consumenten(on-)tevredenheid heen kunnen, hoe dit concept dan ook geoperationaliseerd moge worden. Consumenten(on-)tevredenheid kan immers worden gezien als de resultante van behoeftenbevrediging door middel van consumptie (-goederen).

In dit artikel zijn de moeilijkheden die samenhangen met de operationalisatie van het concept consumenten(on-)tevredenheid aangegeven. De (subjectieve) tevredenheid met een produkt kan per consument verschillen, gegeven de werkelijke (objectieve) kwaliteiten van dat produkt.

Hiernaast kunnen zich bij meting van consumenten(on-)tevredenheid problemen voordoen met betrekking tot de dimensionaliteit van het concept en de specificiteit van de meting.

Deze problemen kunnen grotendeels worden ondervangen – zo is in dit artikel aangegeven – door het complementair gebruik van twee soorten metingen, (1) indexering van consumenten(on-)tevredenheid en (2) kwalitatieve inventarisatie van consumentenproblemen. Met de methode van indexering wordt tegemoet gekomen aan het probleem van de specificiteit van de meting. Van een brede range produkten en diensten kan met een betrekkelijk simpele vraag de tevredenheid worden vastgesteld. De vereiste dimensionaliteit van de meting kan hierbij op de afzonderlijke produkten worden afgestemd.

Ook het probleem van de momentane invloeden op een meting wordt met indexering ondervangen, terwijl meettechnische problemen die samenhangen met het gebruik van verscheidene metingen in de tijd kunnen worden opgelost, analoog aan de meetmethode bij de Index van Consumentenvertrouwen.

Het subjectieve (relatieve) karakter van tevredenheidsindices is eveneens te ondervangen. Bij de vaststelling van consumenten(on-)tevredenheid met behulp van indexering dient niet te worden uitgegaan van de (absolute) waarden van tevredenheidsindices op een bepaald meettijdstip, maar van de (relatieve) verschillen in deze waarden tussen twee verschillende meettijdstippen. Deze relatieve verschillen kunnen dan als maatstaf dienen voor het al dan niet uitvoeren van de meer kwalitatieve, inventariserende studies van consumentenproblemen om te onderzoeken waardoor de tevredenheid met een bepaald(e) produkt-(categorie) is afgenomen (of toegenomen) ten opzichte van een vorig meettijdstip.

BIBLIOGRAFIE

Andreasen, A.R., Consumer dissatisfaction as a measure of market performance, *Journal of Consumer Policy,* 1, 1977.

Andreasen, A.R. en A. Best, Consumers complain – does business respond? *Harvard Business Review,* 55, 1977, 93-101.

Andreasen, A.R. en J.M. Manning, Conducting cross-national consumer policy research. In: J.C. Olsen, *Advances in consumer research,* Vol. VIII. Ann Arbor: Association for Consumer Research, 1980.

Ash, S.B., *Consumer satisfaction, dissatisfaction and complaining behaviour: major findings and directions for use.* University of Western Ontario, London, Ontario, 1982.

Ash, S.B., J.R. Kennedy en P. Thirkell, Consumer satisfaction with products warranties: a study of Canadian automobile and appliance owners. In: H.K. Hunt en R.L. Day, *Refining concepts and measures of consumer satisfaction and complaining behavior,* Indiana University Blooming-

ton, Indiana, 1980.
Ash, S.B. en J.A. Quelch, Consumer satisfaction, dissatisfaction and complaining behavior: a comprehensive study of rentals, public transportation and utilities. In: H.K. Hunt en R.L. Day, *Refining concepts en measures of consumer satisfaction and complaining behavior,* Indiana University, Bloomington, Indiana, 1980.
Barksdale, H.C. en W.R. Darden, Consumer attitudes towards marketing and consumerism. *Journal of Marketing,* 36 (4), 28-35. 1972.
Bearden, W.O., M. Crockett, en S. Graham, Consumers' propensity-to-complain and dissatisfaction with automobile repairs. In: H.K. Hunt en R.L. Day, *Refining concepts and measures of consumer satisfaction and complaining behavior,* Indiana University, Bloomington, Indiana, 1980.
Bearden, W.O., R.M. Durand, J.B. Mason en J.E. Teel, *Consumer satisfaction with utilities.* Proceedings 23rd annual conference of the American Council on Consumer Interests, ACCI, University of Missouri, Columbia, 1977.
Bernacchi, M.D., K. Kono en J.E. Smith, The satisfaction of consumer complainers with consumer protection agencies. In: R.L. Day en H.K. Hunt, *New dimensions of consumer satisfaction and complaining behavior,* Indiana University, Bloomington, Indiana, 1979.
Bernacchi, M.D., K. Kono en G.L. Willette, An analysis of automobile warranty service dissatisfaction. In: H.K. Hunt en R.L. Day, *Refining concepts and measures of consumer satisfaction and complaining behavior,* Indiana University, Bloomington, Indiana, 1980.
Bernhardt, K.L., Consumer research in the federal government. In: M.P. Mokwa en S.E. Permut, *Government marketing: theory and practice,* New York: Praeger Publishers, 1981.
Beukenkamp, P.A. en P.Th. Gründemann, *Behandeling van klachten van levensmiddelen – consumenten,* Centrum voor Bedrijfstechnisch Onderzoek, rapport 7920 M, Erasmus Universiteit Rotterdam, 1979.
Bodur, M., E. Borak en K. Kurtulus, A comparative study of satsfaction/dissatisfaction and complaining behavior with consumer services: Bloomington and Instanbul. In: R.L. Day en H.K. Hunt, *New findings on consumer satisfaction and complaining.* Indiana University, Bloomington, Indiana, 1982.
Bodur, M. en B. Osdiken, Dimensions of satisfaction with educational services. In: H.K. Hunt en R.L. Day, *Conceptual and empirical contributions to consumer satisfaction and complaining behavior,* Indiana University, Bloomington, Indiana, 1982.
Bruhn, M., Konsumentenzufriedenheid und Beschwerden. In: H. Meffert. *Schriften zum Marketing,* Band 4. Verlag Peter Lang, Frankfurt am Main, 1982.
Cardozo, R.N. (a), An experimental study of customer effort, expectations and satisfaction. *Journal of Marketing Research,* August 1965, 244-249.
Cardozo, R.N. (b), Customer Satisfaction: Laboratory study and marketing action. In: L.G. Smith, *Reflections on Progress in marketing.* Proceedings of the December, 1964, American Marketing Association Educators Conference, 1965, 284-289.
Cavusgil, S.T. en E.A. Kaynak, A framework for cross-cultural measurement of consumer dissatisfaction. In: R.L. Day en H.K. Hunt, *New findings on consumer satisfaction and complaining.* Indiana University, Bloomington, Indiana 1982.
Chaiy, S. en M.J. Saxton, Consumer satisfaction with orthodontic treatment. In: H.K. Hunt en R.L. Day, *Conceptual and empirical contributions to consumer satisfaction and complaining behavior,* Indiana University, Bloomington, Indiana, 1982.
Czepiel, J.A., L.J. Rosenberg en A. Akerele, Perspectives on consumer satisfaction. In: R.C. Curhan, *Combined proceedings Chicago:* American Marketing Association, 1975.
Day, R.L. en M. Bodur, A comprehensive study of satisfaction with consumer services. In: R. Day, *Consumer satisfaction, dissatisfaction and complaining behavior,* Indiana University, Bloomington, 1977.
Day, R.L. en M. Bodur, Consumer response to dissatisfaction with durable products. In: W.L. Wilkie, *Advances in Consumer Research,* Vol. 6. Miami, Association for Consumer Research, 1978, 263-272.
Day, R.L. en E.L. Landon, Toward a theory of consumer complaining behavior. In: A.G. Woodside e.a., *Consumer and industrial buying behavior,* New York, Elsevier North-Holland, 1977.
Darden, W.R. en C.P. Rao, Satisfaction with repairs under warranty and perceived importance of warranties for appliances. in: R. Day, *Consumer satisfaction, dissatisfaction and complaining*

behavior, Indiana University, Bloomington, 1977.
Dickinson, V.H. en B. Lang, Consumer satisfaction with complaint handling: an Oregon analysis, *Proceedings 28th annual conference of the American Council of Consumer Interests,* ACCI, University of Missouri, Columbia, 1982.
Francken, D.A., Oorzaken en gevolgen van consumentenproblemen, *SWOKA-onderzoeksrapporten,* no. 22, 1983a.
Francken, D.A., Postpurchase consumer evaluations, complaint actions and repurchase behavior, *Journal of Economic Psychology,* 4, 1983b.
Francis, S.K. en L.E. Dickey, Correlates of women's satisfactions with their purchases of selected outerwear: implications for satisfaction theory. In: H.K. Hunt en R.L. Day, *Conceptual and empirical contributions to consumer satisfaction and complaining behavior,* Indiana University, Bloomington, Indiana, 1982.
Geistfeld, L.V. en R.K.M. Choy, An analysis of third party consumer complaint handling mechanisms: the Hawaii Office of Consumer Protection, *Journal of Consumer Affairs,* 12, 1978, 116-123.
Grainer, M., J. Goodman en E. Hanson, An update evaluation of federal agency consumer complaint-handling practices. *Proceedings 25th annual conference of the American Council on Consumer Interests,* ACCI, University of Missouri, Columbia, 1979.
Granbois, P., J. Summers en G.L. Frazier, Correlates of consumer expectations and complaining behavior. In: R.L. Day, *Consumer satisfaction, dissatisfaction and complaining behavior,* Bloomington, Indiana University, 1977, 18-25.
Grabicke, K., T. Schaetzle en F. Staubach, The influence of personality factors on complaining behavior. In: H.K. Hunt en R.L. Day, *Conceptual and empirical contributions to consumer satisfaction and complaining behavior,* Indiana University, Bloomington, Indiana, 1982.
Grønhaug, K., Exploring consumer complaining behavior: a model and some empirical results. In: W.D. Perreault, *Advances in Consumer Research,* 1977, Vol. IV.
Grønhaug, K. en G. Zaltman, Complainers and noncomplainers revisited: another look at the data. *Journal of Economic Psychology,* 1, 1981, 121-134.
Handy, C.R., Applications of the index of consumer satisfaction for public policy pertaining to market performance. In: M. Venkatesan, *Proceedings of the 3rd Annual Conference of the Association for Consumer Research,* 1972, 738-741.
Handy, C.R., Indexes of consumer satisfaction with food products: 1974 and 1976 survey results. In: R. Day, *Consumer satisfaction, dissatisfaction and complaining behavior,* Indiana University, Bloomington, 1977.
Handy, C.R. en M. Pfaff, Consumer satisfaction with food products and marketing services. *Agricultural Economic Report no. 281.* Economic Research Service, USDA, 1975.
Herzberg, F., B. Mausner en B.B. Snyderman, *The motivation to work.* 2nd Edition, New York: John Wiley & Sons, 1959.
Hunt, H.K., Consumer satisfaction/dissatisfaction: of continuing interest? To whom? Why? In: P.N. Bloom, *Consumerism and beyond: research perspectives on the future social environment.* Cambridge, Massachusetts: Marketing Science Institute, 1982.
Hunt, J.M., Consumer responses to inequitable economic exchanges: an attribution theoretical analysis. In: H.K. Hunt en R.L. Day, *Refining concepts and measures of consumer satisfaction and complaining behavior,* Bloomington, Indiana University, 1980.
I.P.M.-Inter/View, *Consumentisme, I.P.M.* – Schiedam, Inter/View – Amsterdam, 1974.
I.P.M.-Inter/View, *Consumentisme, I.P.M.* – Schiedam, Inter/View – Amsterdam, 1981.
Ittig, K., Consumer evaluation of local redress alternatives. *Proceedings 24th annual conference of the American Council on Consumer Interests,* ACCI. University of Missouri, Columbia, 1978.
Kasper, J.P.D., Klachten van de consument: een internationale inventarisatie. In: P.A. Beukenkamp e.a. *Consument, onderneming en overheid.* Kluwer/NIMA, Deventer/Den Haag, 1982a.
Kasper, J.P.D., *Marketing en consumentensoevereiniteit,* Amsterdam, VU Boekhandel/Uitgeverij, 1982b.
Kaynak, E. en L. Stevenson, Consumer satisfaction/dissatisfaction with single-detached houses in Atlantic Canada. In: H.K. Hunt en R.L. Day, *Conceptual and empirical contributions to consumer satisfaction and complaining behavior,* Indiana University, Bloomington, Indiana, 1982.
Kelly, J.P., Consumer expectations of complaint handling by manufacturers and retailers of clothing products. In: R.L. Day en H.K. Hunt, *New dimensions of consumer satisfaction and complaining behavior,* Indiana University, Bloomington, Indiana, 1979.

Krishnan, S. en K. Grønhaug, A multi-attribute approach to consumer satisfaction with a professional program. In: H.K. Hunt en R.L. Day, *Refining concept and measures of consumer satisfaction and complaining behavior,* Indiana University, Bloomington, Indiana, 1980.
Leavitt, C., Consumer satisfaction and dissatisfaction: bipolar or independent. In: H.K. Hunt, *Conceptualization and measurement of consumer satisfaction and dissatisfaction,* Cambridge, Mass.: Marketing Science Institute, 1977.
Krishnan, S. en M.K. Mills, Dissatisfaction with retail stores and repatronage behavior. In: R.L. Day en H.K. Hunt, *New dimensions of consumer satisfaction and complaining behavior,* Indiana University, Bloomington, Indiana, 1979.
Leigh, T.W. en R.L. Day, Satisfaction/dissatisfaction and complaint behavior with nondurable products. In: R.L. Day ane H.K. Hunt, *New dimensions of consumer satisfaction and complaining behavior,* Indiana University, Bloomington, 1979.
Lingoes, J.C. en M. Pfaff, The index of consumer satisfaction: Methodology. In: M. Venkatesan. *Proceedings of the 3rd Annual Conferences of the Association for Consumer Research,* 1972, 689-712.
Liefeld, J.P., F.H.C. Edgecomb en L. Wolfe, Demographic characteristics of Canadian consumer complainers. *Journal of Consumer Affairs,* 9, 1975, 73-80.
Maddox, R.N., Consumer satisfaction with supermarket: a factor-analytic study. In: R. Day, *Consumer satisfaction, dissatisfaction and complaining behavior,* Indiana University, Bloomington, 1977.
Marlowe, J., Citizen satisfaction with garbage removal service: public vs. private provision. *Proceedings 28th annual conference of the American Council on Consumer Interests,* ACCI, University of Missouri, Columbia, 1982.
Mason, J.B. en S.H. Himes, An explanatory behavioral and socio-economic profile of consumer action about dissatisfaction with selected household appliances. *Journal of Consumer Affairs,* 7, 1973, Vol. 7, 121-127.
Miller, J.A., Who is the discontented consumer. In: T.V. Greer, *Combined proceedings,* Chicago, American Marketing, Association, 1970, 486-490.
Ölander, F., *Consumer satisfaction – a skeptic's view.* Aarhus School of Business Administration, Institut for Markedøkonomi, Aarhus, Denmark, 1976.
Olson, D., The influence of advertising on consumer satisfaction with new products. In: R.L. Day en H.K. Hunt, *New dimensions of consumer satisfaction and complaining behavior,* Indiana University, Bloomington, Indiana, 1979.
Pfaff, A., An index of consumer satisfaction, M. Venkatesan, e.a., *Proceedings of the 3rd Annual Conference of the Association for Consumer Research,* 1972, 713-737.
Pfaff, M., The index of consumer satisfaction. Measurement problems and opportunities. In: H.K. Hunt, *Conceptualization and measurement of consumer satisfaction and dissatisfaction,* Cambridge, Mass.: Marketing Science Institute, 1977a.
Pfaff, M. Who decides what is good for whom? *Journal of Consumer Policy,* 2, 1977b, 138-142.
Pfaff, M. en S. Blivice, Socio-economic correlates of consumer and citizin dissatisfaction and activism. In: R. Day, *Consumer satisfaction, dissatisfaction and complaining behavior,* Indiana University, Bloomington, 1977.
Renoux, Y., Consumer dissatisfaction and public policy. In: F.C. Alloine. *Public policy and marketing practices.* Chicago: American Marketing Association, 1973.
Richins, M.L., An investigation of consumers' attitudes toward complaining. In: A. Mitchell, *Advances in Consumer Research.* Ann Arbor: Association for Consumer Research, 1982, Vol. IX, 502-506.
Robinson, L.M., Consumer complaint behavior: a bibliography of research findings. In: R.L. Day en H.K. Hunt, *New dimensions of consumer satisfaction and complaining behavior,* Indiana University, Bloomington, 1979.
Rodgers, A.L. en D.J. Sweeney, Satisfaction with retail stores as reflected in consumers' opinions. In: H.K. Hunt en R.L. Day, *Refining concepts and measures of consumer satisfaction and complaining behavior,* Indiana University, Bloomington, Indiana, 1980.
Sanders, L., *Klachtenbehandeling bij levensmiddelen en de rechtspositie van de consument.* Doctoraal scriptie, landbouwhogeschool, Wageningen, 1981.
Sauer, W.J., S. Chaiy en R. Schweitzer, Consumer complaint behavior among credit users. In: H.K. Hunt en R.L. Day, *Conceptual and empirical contributions to consumer satisfaction and complaining behavior,* Indiana University, Bloomington, Indiana, 1982.

Scammon, D.L. en K.E. Miller, Assessing consumer satisfaction with informations in advertising: a pilot study. In: R.L. Day en H.K. Hunt, *New dimensions of consumer satisfaction and complaining behavior,* Indiana University, Bloomington, Indiana, 1979.
Schutz, H.G., A California study of consumer satisfaction, problem, perception and complaint action. In: *Proceedings 25th annual conference of the American Council on Consumer Interests,* ACCI, University of Missouri, Columbia, 1979.
Schutz, H.G., *California consumers' satisfaction with goods and services: problems, actions and attitudes.* California Department of Consumer Affairs, Sacramento, California, 1980.
Shuptrine, F.K. en G. Wenglorz, Product satisfaction/dissatisfaction: how content are consumers with products they buy and use. In: H.K. Hunt en R.L. Day, *Refining concepts and measures of consumer satisfaction and complaining behavior,* Indiana University, Bloomington, Indiana, 1980.
Spalding, J.B. en N. Marcus, Postal and telephone complaint handling procedures: a comparative study of the U.S. and U.K. In: H.K. Hunt en R.L. Day, *Conceptual and empirical contributions to consumer satisfaction and complaining behavior,* Indiana University, Bloomington, Indiana, 1982.
Swan, J.E. and M.G. Carroll, Patient satisfaction: an overview of research – 1965 – 1978. In: H.K. Hunt en R.L. Day, *Refining concepts and measures of consumer satisfaction and complaining behavior,* Indiana University, Bloomington, Indiana, 1980.
TARP, Technical Assistance Research Programs Inc. Consumer Complaint handling in America: final report United States Office of Consumer Affairs, 1979.
Warland, R.H. en R.O. Herrmann, Consumer complaining and community involvement: theoretical and empirical linkages. In: *Proceedings 27th annual conference of the American Council on Consumer Interests,* ACCI, University of Missouri, Columbia, 1981.
Warland, R.H., R.O. Herrmann en J. Willits, Dissatisfied consumers: who gets upset and who takes action. *Journal of Consumer Affairs,* 9, 1975, 148-163.
Westbrook, R.A., Correlates of post-purchase satisfaction with major household appliances. In: R. Day. *Consumer satisfaction, dissatisfaction and complaining behavior,* Indiana University Bloomington, 1977.
Westendorp, van J. en J. Knecht, Marketing and the consumer: an empirical investigation of some attitudes in the Netherlands. *Proceedings ESOMAR congress,* Amsterdam: ESOMAR, 1974.
Wikström, S., Another look at consumer dissatisfaction as a measure of market performance, *Journal of Consumer Policy,* 6, 1983, 19-35.
Zaichkowsky, J. en J. Liefeld, Personality profiles of consumer complaint letter writers. In: R. Day, *Consumer satisfaction, dissatisfaction and complaining behavior,* Indiana University, Bloomington, 1977.
Zaltman, G., R.K. Srivastava en R. Deshpande, Perceptions of unfair marketing practices: consumerism implications. In: H.K. Hunt, *Advances in consumer research.* Ann Arbor: Association for Consumer Research, 1978, 247-253.

13. Voorspelling van het seizoenbeloop

Nieuwe ontwikkelingen in methoden van seizoencorrectie

M.M.G. FASE

SAMENVATTING

Seizoenanalyse en voorspelling hangen nauw samen. De laatste jaren valt in de literatuur een toenemende belangstelling voor seizoencorrectiemethoden waar te nemen. Dit artikel biedt een overzicht van een aantal van deze methoden, die vrijwel alle hun uitgangspunt vinden bij de bekende Census-X-11 methode en het univariate ARIMA-model van Box en Jenkins. Kenmerkend voor de besproken methoden is dat ze operationeel zijn en dat de desbetreffende programmatuur in beginsel beschikbaar is.

1. PROGNOSE EN SEIZOENCORRECTIE

Vanouds pogen macro-economen en marktonderzoekers aan economische data signalen te ontlenen op grond waarvan voorspellingen worden afgegeven. Deze bezigheid wordt met wisselend succes beoefend en met een gezonde scepsis begeleid. De scheiding van de verantwoordelijkheden van de voorspeller en de beslisser lijkt hierbij van principieel belang voor de geloofwaardigheid van beiden.

Wetenschappelijk verantwoorde prognoses dienen scherp te worden onderscheiden van de befaamde koffiedikkijkerij of, de niet zonder reden veel gesmade, orakeltaal. Evenzo dienen prognoses niet te worden vereenzelvigd met de al dan niet bijbelse profetieën. Deze behelzen, zoals bekend mag worden verondersteld, eerder een waarschuwende vermaning tot inkeer dan een ondubbelzinnige prognose ten behoeve van de besluitvorming. De ervaringen van de profeet Jonas zijn in dit verband in verschillende opzichten leerzaam.

Besluitvorming vraagt van de beslissende instantie dat deze anticipeert op toekomstige ontwikkeling en inzicht heeft in de gevolgen van zijn beleidsbeslissing. Een wetenschappelijk verantwoorde prognose is daarbij hoogst welkom. De vroegere Amsterdamse hoogleraar Van Dantzig heeft eens gepoogd criteria te formuleren voor een wetenschappelijke voorspelling [Van Dantzig (1952)]. Een eerst criterium is *ondubbelzinnigheid.* Een prognose dient duidelijk te zijn en niet te lijden aan het euvel dat je er alle kanten mee uit kan. Zoals bekend, was dit vaak het geval met de voorspellingen van het Orakel van Delphi. Diens boodschap aan Croesus, de laatste koning van Lydië, dat hij een groot rijk te gronde zou richten bij oversteek van de rivier de Halys, is zonder twijfel het meest dramatische voorbeeld uit de literatuur. Daarnaast verlangt Van Dantzig dat de voorspeller zich rekenschap geeft van

de *nauwkeurigheid* en *betrouwbaarheid* van de predictie. Deze eis vertoont verwantschap met de in de statistiek ingeburgerde gewoonte toleranties, omkleed met een waarschijnlijkheid, te kwantificeren. In dit verband wordt – geheel analoog aan de vereisten voor schatters – vaak gestreefd naar een optimale voorspeller [vgl. Granger & Newbold (1977, p. 111)]. In de hedendaagse economische literatuur treffen we dit streven aan bij de zgn. rationele verwachtingenhypothese [vgl. Fase (1981a)]. Een laatste, door Van Dantzig geformuleerde, eis betreft een Popperiaans verlangen naar *falsifieerbaarheid.* Dit houdt in dat de aan een prognose ten grondslag liggende procedure verteld moet kunnen worden, zodat de gebezigde argumenten ook achteraf op hun feitelijke merites kunnen worden beoordeeld.

Er zijn vele wetenschappelijke voorspelprocedures. Het valt buiten het bestek van dit overzicht hierop nader in te gaan. Van belang is de vaststelling dat de seizoencorrectiemethode onder de voorspelprocedures een klasse apart vormt, die het laatste decennium een interessante en voor de praktijk belangrijke ontwikkeling heeft laten zien. Het volgende beoogt van deze ontwikkeling een schets te geven en deze zo nu en dan met een praktisch voorbeeld te illustreren. Uitgangspunt daarbij is, dat het uiteindelijke oogmerk van elke seizoencorrectiemethode de raming of voorspelling van de seizoeninvloed is. Kennis ervan betekent dat deze als ongewenst signaal in beginsel uit een tijdreeks kan worden verwijderd. Met Pierce (1980, p. 125) menen we dat:

> 'The principal reason for this interest is that our ability to recognize, interpret, or react to the important nonseasonal movements in a series... is hindered by the presence of seasonal movements'.

Voorspelling van dit laatste is aldus bittere noodzaak.

2. DECOMPOSITIE VAN TIJDREEKSEN

Economische tijdreeksen kenmerken zich door een veelheid van bewegingen, die zich gemakkelijker laten onderscheiden dan scheiden. Niettemin is het laatste het onderwerp van het onderhavige opstel. De literatuur [vgl. Tinbergen (1943), pp. 1 – 10; Delfgaauw (1973), pp. 116 – 172] biedt een typologie van verschillende economische bewegingen. In de eerste plaats kunnen worden onderscheiden zich herhalende en niet-herhalende bewegingen. Economische groei is een voorbeeld uit de laatste categorie; incidentele fluctuaties tengevolge van b.v. een natuurramp of een door een kartel – b.v. OPEC – opgelegde eenmalige prijsverhoging of -verlaging eveneens. Ten tweede is er binnen de groep van de zich min of meer herhalende bewegingen een nader onderscheid te maken. Delfgaauw onderscheidt in dit verband de ritmisch vrije en de ritmisch gebonden bewegingen. De eerste betreffen de conjuncturele schommelingen. De ritmisch gebonden bewegingen hebben betrekking op de seizoenschommelingen, welke hun oorsprong vinden bij klimatologische of maatschappelijke factoren. Voorbeelden zijn de tijdens de winter oplopende werkloosheid of de groeiende bankbiljettenomloop in december.

In de seizoenanalyse tracht men dit laatste type fluctuaties – de seizoenschommeling – te isoleren. Hierbij kunnen causale methoden of mechanische methoden worden gehanteerd. De eerste, b.v. bepleit door Mendershausen

(1937, 1939), zijn intellectueel aantrekkelijk maar op pragmatische gronden vaak niet uitvoerbaar in de praktijk. Om die reden zijn de mechanische methoden meer ingeburgerd. Hierbij is het gebruikelijk een reeks te splitsen in een aantal componenten, t.w.:

a. de trendcomponent. Deze bestaat uit de algemene tendentie van de reeks die zich uitstrekt over vele jaren. Deze beweging is zich niet herhalend. De lange of Kondratieffgolven zijn hieronder begrepen.
b. de conjunctuurcomponent. Deze betreft schommelingen van 3 à 11 jaar. Dit zijn ritmische vrije bewegingen, die in de economie dikwijls ultrakorte golven (Kitchin-golven of minor cycles) of korte golven (Juglar golven) worden genoemd.
c. de seizoencomponent. Deze omvat de schommelingen binnen het jaar. Dit betreft de ritmisch gebonden fluctuaties die hun oorsprong vinden bij klimatologische, traditionele dan wel institutionele invloeden.
d. de onregelmatige component of toevallige schommeling. Dit betreft zich niet herhalende bewegingen, waarvoor geen bepaalde oorzaak valt aan te geven. Men klasseert ze daarom als toevallige schommelingen. Dit ter onderscheiding van de hierna volgende component;
e. de trendbreuken. Deze zijn, anders dan de onregelmatige component, uiting van structuurbreuken of andere eenmalig optredende wijzigingen in de beschouwde cijfers, bijv. ten gevolge van een definitieverandering of een administratieve maatregel.

Het scheiden van tijdreeksen in bovengenoemde componenten is het voornaamste oogmerk van de seizoenanalyse of seizoenuitschakeling van economische of andere reeksen*. Gelet op de gebruikelijke lengte van de reeksen die worden geanalyseerd, worden de trend en conjunctuur meestal samengenomen, terwijl de trendbreuk op *ad hoc* basis wordt bewerkt (bijv. door ophoging).

Aan alle methoden van seizoenuitschakeling ligt steeds het volgende model ten grondslag.

$$X_{ij} = T_{ij} + S_{ij} + I_{ij}$$

met X_{ij} de oorspronkelijke reeks in maand of kwartaal i en jaar j.
T_{ij} de trend-conjunctuurcomponent in periode (of moment) i, j.
S_{ij} de seizoencomponent in periode (of moment) i, j.
I_{ij} de incidentele component in periode (of moment) i, j.

Deze splitsing gaat uit van een additieve opbouw der oorspronkelijke reeks. Daarnaast is er de multiplicatieve modelformulering die luidt:

$$X_{ij} = T_{ij} \times S_{ij} \times I_{ij}$$

Opgemerkt zij dat de multiplicatieve methode door een logaritmische transformatie wordt herleid tot het additieve model. De voorgaande formulering

* Hier zij uitdrukkelijk vermeld dat ondanks de gegeven classificatie, het ontbinden van reeksen niet noodzakelijk beperkt is tot economische gegevens. Zo hield de bekende meteoroloog Buys Ballot zich in de 19e eeuw reeds met seizoenanalyse bezig. Vergelijk hiervoor Nerlove, Grether en Carvalho (1979).

laat zien dat de waargenomen reeks X_{ij} wordt ontbonden in drie niet-waarneembare componenten. Dit kan op velerlei manieren gebeuren en hierin vinden de verschillende methoden van seizoenuitschakeling hun ontstaansgrond. Daarnaast is er de methodologie die zich richt op het modelleren van een tijdreeks met behulp van een stochastisch model. Deze benadering sluit nauw aan op de methodologie van Box & Jenkins (1970), maar vertoont tal van raakvlakken met de decompositiemethoden, zoals uit de recente literatuur naar voren komt. De volgende paragraaf beoogt een aantal van deze ontwikkelingen voor het voetlicht te brengen, mede omdat dit terrein vanouds de aandacht heeft gehad in het marktonderzoek [vgl. Stridiron (1941), Verdoorn (1950), Van der Zwan & Verhulp (1980)], maar de laatste 10 jaren een opmerkelijke activiteit laat zien.

3. NIEUWE ONTWIKKELINGEN

De in het licht van het voorgaande te bespreken methoden vormen een keuze uit de sinds 1970 ontwikkelde seizoencorrectiemethoden. Deze ontwikkeling hangt zonder twijfel samen met de sindsdien toegenomen aandacht voor het tijdreeksenmodel als prognosemiddel en de daarmede gepaard gaande ruime beschikbaarheid van rekenautomaten en programmatuur. Dit laatste is onze leidraad geweest bij de geboden keuze. Voor een vollediger overzicht zij verwezen naar het werk van W.P. Cleveland (1972), Bourdeau de Fontenay (1979), Pierce (1980) en W.S. Cleveland (1982).

3.1. *Het seizoenmodel van Box-Jenkins*

Het stochastische seizoenmodel van Box-Jenkins is een uitbreiding van hun tijdreeksenmodellen*. Evenals het tijdreeksenmodel vindt het seizoenmodel zijn uitgangspunt in de correlatiestructuur van de tijdreeks die als het ware het onderliggende stochastische model weerspiegelt. Een seizoenbeweging kenmerkt zich door herhaling en vaste ritmiek. Dit betekent dat waarnemingen in overeenkomstige weken, maanden of kwartalen – m.a.w. over het jaar – samenhang bezitten. Op grond daarvan kan bijv. gelden:

$$x_t = \mu + \Phi x_{t-12} + \varepsilon_t \quad \text{of} \quad (1 - \Phi B^{12})(x_t - \mu') = \varepsilon_t \tag{1}$$

met B de vertragingsoperator zodanig dat $B^{12}x_t = x_{t-12}$.
Voorts zou tussen de maanden kunnen gelden:

$$\varepsilon_t = \phi\varepsilon_{t-1} + a_t \quad \text{of} \quad (1 - \phi B)\varepsilon_t = a_t \tag{2}$$

waarbij a_t in de tijd ongecorreleerd is met verwachting 0 en constante variantie. Substitutie van (2) in (1) resulteert in:

$$(1 - \phi B)(1 - \Phi B^{12})(x_t - \mu') = a_t \tag{3}$$

Dit eenvoudige voorbeeld kan gemakkelijk worden uitgebreid en veralgemeend tot de ruime formulering die Box & Jenkins (1970, hoofdstuk 9) bezigen. Deze luidt:

* Voor een beknopt overzicht met praktische illustraties zij verwezen naar Fase (1981b, 1983).

$$\phi_p(B)\,\Phi_P(B^s)\,(1-B)^d(1-B^s)^D(x-\theta_0)=\theta_q(B)\,\Theta_Q(B^s)a_t \tag{4}$$

met $\phi_p(B)=1-\phi_1B-\phi_2B^2-\ldots-\phi_pB^p$, en $\theta_q(B)=1-\theta_1B-\theta_2B^2-\ldots-\theta_qB^q$ polynomen in B van resp. de orde p en q; $\Phi_P(B^s)=1-\Phi_1B^s-\Phi_2B^{2s}-\ldots-\Phi_PB^{Ps}$ en $\Theta(B^s)=1-\Theta_1B^s-\Theta_2B^{2s}-\ldots-\Theta_QB^{Qs}$ polynomen in B^s van resp. de orde P en Q. B is de vertragingsoperator zodanig dat $Bx_t=x_{t-1}$ en $B^2x_t=x_{t-2}$, enz. De D en d zijn gehele getallen van tenminste nul, θ_0 is een constante, x_t de beschouwde tijdreeks en a_t is een onafhankelijk, identiek normaal verdeelde stochastische variabele met verwachting 0 en constante variantie. De grootheid s indiceert de tijdspanne van de reeks. Zo is voor weekcijfers $s=52$. Voor maandcijfers geldt $s=12$ en voor kwartaalcijfers $s=4$. Bezien we tijdvakken van 4 weken, dan geldt $s=13$. Model (4) is een *a*utoregressive *i*ntegrated *m*oving *a*verage of ARIMA-model van de orde $(p, d, q)\times(P, D, Q)$ en wordt symbolisch dikwijls genoteerd als ARIMA-$(p, d, q)\times(P, D, Q)$.

Aan model (4) wordt concrete inhoud gegeven door de iteratieve werkwijze van Box & Jenkins op de beschouwde tijdreeks toe te passen. Deze verloopt in drie stadia, te beginnen met de identificatiefase. Dit betreft de keuze van numerieke waarden voor p, d, q, P, D en Q en geschiedt aan de hand van autocorrelatie en partiële autocorrelatie-patronen. Vervolgens vindt in de tweede fase schatting van de coëfficiënten θ, Θ, ϕ en Φ plaats. Het resultaat hiervan wordt tenslotte in de derde diagnostische fase bezien en onderworpen aan grafische inspecties en statistische criteria. Dit leidt tot aanvaarding van het aanvankelijke model of herziening, waarna de cyclus opnieuw wordt doorlopen tot een aanvaardbaar model resulteert. Dit uiteindelijke model vormt de basis voor de prognose van zowel de trend als de seizoeninvloed. Ter illustratie beschouwen we de vierweekscijfers van het volume aan personeelsadvertenties in een landelijk dagblad (uitgedrukt in kolom-millimeters) over de periode 1977 – 1982 (t/m 8e vierweeksperiode).

Figuur 1 biedt van deze reeks een grafische uitbeelding. Toepassing van de door Box & Jenkins voorgestane iteratieve werkwijze leidt tenslotte tot het volgende ARIMA-model:

$$\underset{(3,2)}{(1-0{,}4B-0{,}6B^2)}\,\underset{(5,3)}{(1-B^{13})}x_t=\underset{(8,7)}{(1-0{,}8B^{13})}\,a_t \tag{5}$$

waarbij de waarden tussen haakjes de zgn. t-waarden zijn (een t-waarde groter dan 2 betekent praktisch dat de desbetreffende coëfficiënt als statistisch significant mag worden aangemerkt) en x_t het advertentievolume per periode voorstelt. Deze schatting blijkt het reeds bij de grafische inspectie vastgestelde seizoenpatroon – het betreft vierweekscijfers, hetgeen exponent 13 verklaart! – te bevestigen. De neergaande trend in de data wordt met dit model eveneens gevat.

Model (5) is uitgeschreven een differentievergelijking van de orde 15, waarvan de oplossing een voorspelformule oplevert voor het advertentievolume in toekomstige vierweeksperioden T. Het rangnummer T ontstaat uit de datering als $T=13i+j$ met $i=0, 1, 2, \ldots, 13$ de jaarindex en $j=1, 2, \ldots, 13$ de periode-index, beide gemeten vanaf het voorspelmoment.

Figuur 1. Volume personeelsadvertenties 1977 – 1982 in landelijk dagblad (× 1000 kolom-millimeters).

De oplossing, welke volgt uit (5), luidt:

$$x(13i+j) = A_j \times (0{,}997)^{13i} + B_j \times (-0{,}609)^{13i} + C_j \qquad (6)$$

waarbij A_j, B_j en C_j constanten zijn die verschillen per vierweeksperiode en worden berekend uit 15 beginwaarden, welke zijn verkregen met (5) [zie Goldberg (1961) voor de techniek van het oplossen]. De verkregen constanten die tevens de seizoeninvloed representeren zijn opgenomen in tabel 1.

Tabel 1: Berekende constanten per 1983 (T = 0)

j	A_j	B_j	C_j
1	2497	−4,291	−2223
2	2489	2,613	−2221
3	2482	−1,591	−2250
4	2474	0,969	−2191
5	2467	−0,590	−2310
6	2459	0,359	−2411
7	2452	−0,219	−2164
8	2445	0,133	−2227
9	2437	−0,081	−2231
10	2430	0,049	−2240
11	2423	−0,030	−2335
12	2415	0,018	−2195
13	2408	−0,011	−2162

Invulling van deze waarden in (6) levert voor de weeknummers $T = 13i + j$ voorspellingen inclusief de seizoenfluctuaties.

Verschillende auteurs hebben de voorgaande benadering beproefd voor de schatting van seizoeninvloeden. Box & Jenkins zelf analyseerden in hun boek de airline data, een analyse die door velen is bekritiseerd en uitgebreid [vgl. W.P. Cleveland (1972), Mélard & Kiehm (1980)]. Belangwekkend, mede vanwege de duidelijkheid en de uitwerking van een voorbeeld, is de case study van

Thompson & Tiao (1971) over het aantal maandelijkse telefoonaansluitingen in de staat Wisconsin. Ook De Vos (1975), in zijn analyse van de werkloosheidscijfers in Nederland, zocht aansluiting bij de benadering van Box & Jenkins, zonder overigens tot bevredigender resultaten te komen dan met de gangbare methoden. Interessant is de kritische analyse van de procedure door Chatfield & Prothero (1973) en de daardoor uitgelokte repliek van Box & Jenkins (1973), die het belang van een passende transformatie der oorspronkelijke cijfers aantoonden. Hiervoor bevelen zij de Box-Cox (1964) procedure sterk aan, omdat aldus langs empirische weg de meest geschikte transformatie kan worden vastgesteld.

3.2. *De Census X-11 methode*

Een van de in de praktijk meest gebezigde methoden van seizoencorrectie is de Census X-11 methode [zie Shiskin, Young & Musgrave (1967)]. In ons land past de Nederlandsche Bank sinds eind 1972 de methode toe op de door haar verzamelde financiële maand- en kwartaalreeksen*. Ook het CBS is omstreeks dezelfde tijd overgegaan op deze methode voor reëel-economische reeksen als industriële produktie, consumptieve bestedingen, werkgelegenheid, bedrijfsinvesteringen of omzet in detailhandel. Banken en andere individuele bedrijven passen de methode – zij het op veel geringere schaal – eveneens toe.

De methode kan worden toegepast op zowel maand- als kwartaalreeksen. Vooraf dient een keuze te worden gemaakt uit de additieve of multiplicatieve variant. Dit betekent dat nagegaan moet worden of de seizoenuitslag al dan niet samenhang lijkt te vertonen met de trend. Om de gedachten te bepalen wordt hier uitgegaan van maandcijfers waarvoor een multiplicatief seizoenpatroon wordt verondersteld. De Census X-11 methode verloopt als volgt.

1. Als eerste benadering van de trend-conjunctuurlijn – de T_{ij} – wordt een gecentreerd twaalfpunts voortschrijdend gemiddelde berekend. De afwijkingen van de oorspronkelijke reeks ten opzichte van deze trend – het quotiënt O_{ij}/T_{ij} – is een eerste benadering van de seizoen- en toevallige invloeden, $S_{ij} \times T_{ij}$.
2. Voor elk der overeenkomstige maanden wordt van het quotiënt O_{ij}/T_{ij} een gewogen vijfpunts voortschrijdend gemiddelde berekend. Deze zgn. vijfpunts Henderson is een eerste benadering van de seizoenindices S_{ij} voor de beschouwde maand j.
3. Deling van de oorspronkelijke reeks door de aldus berekende seizoenindex geeft een eerste voorlopige benadering van de seizoenvrije reeks.
4. Vervolgens wordt van deze seizoenvrije reeks andermaal een gecentreerd gewogen voortschrijdend gemiddelde berekend. De lengte hiervan hangt af van de variabiliteit van de reeks en is een 9, 13 of 23 punts Henderson. Het aldus berekende voortschrijdende gemiddelde is een tweede benadering van de trend-conjunctuurlijn.
5. Van het quotiënt van de waargenomen reeks en de onder 4 berekende trend-conjunctuur wordt vervolgens voor de overeenkomstige maanden

* Vgl. hiervoor Fase, Koning & Volgenant (1973) en Koning (1974).

een gewogen zevenpunts voortschrijdend gemiddelde bepaald. Dit resulteert in een definitieve schatting van de seizoenindices.

6. Deling van de oorspronkelijke reeks door de onder 5 verkregen seizoenindex resulteert in de seizoenvrije reeks.

Hoewel de X-11 veelvuldig wordt gebruikt in de praktijk van de seizoenanalyse is ze niet vrij van kritiek gebleven, zonder nochtans haar dominerende rol te hebben prijsgegeven. De aard van deze kritiek betreft verschillende gezichtspunten. Een eerste groep critici heeft vooral statistisch-methodologische bezwaren. Ze vindt het ontbreken van een statistisch model voor de X-11 een ernstige tekortkoming die het inzicht in de resultaten belemmert. Deze kritiek heeft geleid tot interessante pogingen de X-11 procedure modelmatig te funderen. In dit verband neemt het werk van Cleveland (1972) en Cleveland & Tiao (1976) een prominente plaats in. De tweede stroom critici voert bezwaren van meer praktische aard aan die zowel een algemeen als een specifiek aangrijpingspunt hebben.

Algemeen geldt dat de X-11 een bewegend seizoen schat dat uiteindelijk resulteert uit een 145-punts voortschrijdend gemiddelde. Dit impliceert dat tijdens de berekeningen waarnemingen verloren gaan, waardoor de X-11 er toe neigt de veranderingen in het seizoenpatroon aan het begin en het eind van de reeks systematisch te onderschatten. Deze gevoeligheid voor reeksverlenging ondervindt grote bezwaren als de betreffende reeks zich mag verheugen in een ruime politieke belangstelling zoals bijv. het geval is met cijfers omtrent de werkloosheid, produktie, of – en dit in het bijzonder in de V.S. – de geldhoeveelheid. Dit heeft aanleiding gegeven tot een uitbreiding van de X-11 methode die bekend staat als de X-11-ARIMA, die door het werk van Dagum van het Canadese Bureau voor de Statistiek ruime bekendheid heeft verworven. Hierop wordt in de volgende paragraaf nader ingegaan.

De meer specifieke kritiek richt zich op de constatering dat werkloosheidscijfers zich moeilijk met de X-11 voor seizoen laten corrigeren. Deze kritiek geldt voor tal van landen en heeft geleid tot de ontwikkeling van specifieke methoden van seizoencorrectie voor werkloosheidscijfers. Voor Nederland kan in dit verband het werk van Vos (1975) worden genoemd, dat nauw aansluit op de in paragraaf 3.1. weergegeven benadering van Box & Jenkins. Hij bereikte overigens, zoals hiervoor reeds is opgemerkt, met zijn methode geen betere resultaten dan met de X-11. Thans is op het CBS een mengeling van de Planbureau- en X-11 methode in gebruik. Hierop wordt in paragraaf 3.4 nader ingegaan. Op het Engelse CBS is in dit verband de zgn. methode van Kenny & Durbin (1982) en eerder die van Durbin & Murphy (1975) ontwikkeld.

Wat het ontbreken van een modelmatige fundering van de X-11 methode betreft, leiden Cleveland en Tiao in de reeds vermelde geschriften af dat de in de X-11 methode gebezigde voortschrijdende gemiddelden bij benadering kunnen worden beschreven door een ARIMA-model voor de oorspronkelijke reeks y_t van de vorm:

$$
\begin{aligned}
(1-B)(1-B^{12})y_t = (1 - 0{,}337B &+ 0{,}144B^2 + 0{,}141B^3 \\
&+ 0{,}139B^4 + 0{,}136B^5 + 0{,}131B^6 \\
&+ 0{,}125B^7 + 0{,}117B^8 + 0{,}106B^9 \\
&+ 0{,}093B^{10} + 0{,}077B^{11} + 0{,}417B^{12} \\
&+ 0{,}232B^{13} - 0{,}001B^{20} - 0{,}003B^{21} \\
&- 0{,}004B^{22} - 0{,}006B^{23} + 0{,}035B^{24} \\
&- 0{,}021B^{25})a_t
\end{aligned} \tag{7}
$$

met B de vertragingsoperator zoals eerder gebezigd en a_t een normaal verdeelde stochastische variabele met verwachtingswaarde nul en constante variantie σ_a^2. Het bovenstaande model is ontstaan uit een model met de wiskundige vorm:

$$
\begin{aligned}
(1-B)^2 p_t &= (1 - \psi_{11}B - \psi_{12}B^2)\, b_{1t} \\
(1-B^{12}) s_t &= (1 - \psi_{21}B^{12} - \psi_{22}B^{24})\, b_{2t}
\end{aligned}
$$

met $\psi_{11} = -0{,}49$, $\psi_{12} = 0{,}49$, $\psi_{21} = -0{,}64$, $\psi_{22} = -0{,}83$, $\sigma_{b_2}^2/\sigma_{b_1}^2 = 1{,}3$ en $\sigma_e^2/\sigma_{b_1}^2 = 14{,}4$ op zodanige wijze geselecteerd dat de overeenkomstige gewichten in het Census X-11 programma zo goed mogelijk worden benaderd. Voorts is verondersteld dat het stochastische model additief is en luidt:

$$y_t = p_t + s_t + e_t$$

met e_t de toevallige component met variantie σ_e^2, s_t de seizoen- en p_t de permanente trendcomponent, welke, in tegenstelling tot de reeks y_t, niet waarneembaar zijn.

Het bovenstaande resultaat betekent dat toepassing van de Census X-11 procedure inhoudt dat aan de beschouwde reeksen telkens hetzelfde ARIMA-model wordt opgelegd. Daarnaast kan de vraag gesteld worden of vrije schatting van een ARIMA-model een resultaat oplevert dat correspondeert met het klaarblijkelijk aan de X-11 ten grondslag liggende model. Tenslotte zij vermeld dat pogingen zijn ondernomen model (7) te vereenvoudigen, bijv. door Winkel (1982), waarbij overigens kan worden opgemerkt dat ook de oorspronkelijke formulering door Cleveland (1972, p. 59) aanmerkelijk eenvoudiger was dan bij Cleveland & Tiao (1976).

3.3. *De X-11-ARIMA methode*

Van groot praktisch belang lijkt de uitbreiding van de Census X-11 methode, die vooral door mevrouw Estella Dagum van het Canadese Bureau voor de Statistiek ten uitvoer is gebracht en die thans in ruime kring navolging vindt [zie Dagum (1975, 1980a, b, 1982)]. De theoretische grondslag hiervan gaat terug op De Forest (1877) en Greville (1979) met hun aandacht voor symmetrische gewichten [zie ook Bordeau de Fontenay (1979) en W.S. Cleveland (1982)].

De gangbare seizoencorrectiemethoden gaan alle uit van voortschrijdende gemiddelden, d.w.z. van een bepaalde verzameling symmetrische gewichten. Dientengevolge kunnen de begin- en eindwaarden voor een reeks niet op dezelfde wijze worden 'gladgestreken' als de meer centraal gelokaliseerde waarnemingen. Toevoeging van nieuwe waarnemingen leidt daardoor tot herzie-

ning van de seizoenvrije reeksen. Vooral wanneer deze herzieningen van aanzienlijke omvang zijn is dit vervelend voor praktisch gebruik en schadelijk voor de betrouwbaarheid van de resultaten. Naast dit praktische bezwaar wordt ook het ontbreken van een expliciet tijdreeksenmodel als kritiek aangevoerd. Het door Dagum ontwikkelde X-11-ARIMA model ondervangt beide bezwaren in belangrijke mate. Het nieuwe van de methode is de benutting van het univariate ARIMA model binnen de X-11 opzet. Het ARIMA model in zijn algemene formulering luidt:

$$\phi_p(B)\,\Phi_P(B^s)\nabla^d\nabla_s^D x_t = \theta_0 + \theta_q(B)\,\Theta_Q(B^s)\,a_t \tag{8}$$

met $\nabla = 1 - B$, terwijl de overige symbolen de reeds eerder aangeduide betekenis hebben. De operator ∇_s^D bewerkstelligt – analoog aan ∇^d voor de reeks binnen het jaar – dat de niet-stationaire seizoencomponent wordt getransformeerd in een stationaire reeks.

Een eenmaal geschat ARIMA-model kan worden herschreven in de gedaante van een differentievergelijking op grond waarvan 'toekomstige waarden' kunnen worden geëxtrapoleerd, zodat ten behoeve van de seizoenanalyse de tijdreeks kan worden verlengd. Ter illustratie bezien we het volgende voorbeeld. Zij het verkregen ARIMA-model voor de variabele x_t – genoteerd in de schrijfwijze uit paragraaf 3.1 – van de orde $(1, 0, 1)(1, 0, 0)_4$. Er geldt derhalve:

$$(1 - \phi B)(1 - \Phi B^4)x_t = (1 - \theta B)\,a_t \tag{9}$$

of, herschreven als differentievergelijking:

$$x_t = \phi x_{t-1} + \Phi x_{t-4} - \phi\Phi x_{t-5} + a_t - \theta a_{t-5} \tag{9'}$$

Uitdrukking (9′) beschrijft de wijze waarop de kwartaalreeks wordt geëxtrapoleerd. Voor details, mede m.b.t. de behandeling van a_t, zij verwezen naar de literatuur [bijv. Box & Jenkins (1970), Hoofdstuk 5].

Er zijn tenminste twee wijzen om in de context van de X-11-ARIMA methode tot de keuze van het geschikte ARIMA-model te komen. Ten eerste is er de werkwijze die voor elke individuele reeks een model schat. Dit gebeurt op de gebruikelijke iteratieve wijze, waarbij achtereenvolgens de identificatie-, schattings- en diagnostische verificatierondes zolang worden doorlopen tot een bevredigend model resulteert. Deze werkwijze is typisch en aanbevelenswaard voor seizoencorrectie van een enkele reeks. Het gekozen ARIMA-model wordt dan in het seizoencorrectieprogramma geïncorporeerd op *ad hoc* basis.

De andere werkwijze, die vooral bij massale seizoencorrectie wordt gevolgd, maakt aan de hand van een aantal in het programma opgenomen criteria een keuze uit een assortiment van 12 in de praktijk van de seizoenanalyse beproefde ARIMA-modellen. De desbetreffende modellen – symbolisch genoteerd – luiden:

$(1,1,1)(1,1,1)_s$
$(2,1,2)(0,1,1)_s$
$(2,0,1)(0,1,2)_s$
$(1,1,2)(0,1,2)_s$

$(2,0,0)(0,1,1)_s$
$(1,1,2)(1,0,2)_s$

$(2,1,1)(0,1,2)_s$ log
$(0,1,2)(1,1,2)_s$ log
$(0,1,1)(0,1,1)_s$ log
$(0,1,1)(0,2,2)_s$ log
$(0,2,2)(0,1,1)_s$ log
$(2,1,1)(0,1,1)_s$ log

Model $(1,1,1)(1,1,1)_s$ betekent bijvoorbeeld in concreto:

$$(1-\phi B)(1-\Phi B^s)\nabla x_t=(1-\theta B)(1-\Theta B^s)a_t$$

terwijl het zevende model, met de variabele logaritmisch getransformeerd, luidt:

$$(1-\phi_1 B-\phi_2 B^2)\nabla\nabla_s \log x_t=(1-\theta B)(1-\Theta_1 B^s-\Theta_2 B^{2s})a_t$$

met de index $s=4$ voor kwartaalcijfers, enz. Evenzo kunnen de overige 10 modellen worden uitgeschreven.

Opgemerkt zij dat bovenvermelde 12 modellen zijn opgenomen als mogelijke opties in het programma en dat hieraan een uitgebreide empirische analyse van 174 Canadese tijdreeksen ten grondslag ligt. Niettemin kan het heel wel voorkomen dat het voor een bepaalde reeks meest passende ARIMA-model zich niet onder de gegeven 12 modellen bevindt. De analyse door Den Butter, Coenen en Van de Gevel (1982) biedt hiervan een illustratie voor een zestal economische of financiële reeksen en een synthetische reeks. Ook in de door ons beproefde analyse van de gasafzet in Nederland kon niet worden volstaan met een van de bovenvermelde modellen, maar diende een afzonderlijk model te worden geschat [vgl. Fase (1982)].

3.4. *De CPBX11-methode*

Tot december 1975 gebruikte het CPB de Census X-11 methode voor de seizoencorrectie van arbeidsmarktgegevens. Vooral voor de bouwsector echter bleek deze methode tot een systematische overschatting van de seizoeninvloed te leiden bij een hoog peil van werkloosheid. Om die reden werd in december 1975 overgegaan op de CPB-methode die een gemengd additieve-multiplicatieve seizoencorrectiemethode is [vgl. Den Haan (1979)]. De resultaten bleven echter onbevredigend en nader onderzoek resulteerde in een synthese tussen de CPB-methode en de X-11, genaamd de CPBX11-methode [vgl. Van der Hoeven & Hundepool (1980)].

Uitgangspunt van de CPBX11-methode is – overigens in navolging van Wald (1936) – dat de seizoencomponent minder dan evenredig toeneemt met de trend. Hierbij wordt de aard – lineair of kromlijnig – en intensiteit van deze samenhang voor elke reeks afzonderlijk langs empirische weg vastgesteld. Het model voor de CPBX11 is als volgt:

$$Y_{ij} = T_{ij} + S_{ij} + I_{ij}$$

$$S_{ij} = m(T_{ij}) \cdot S_{ij}$$

met Y_{ij} de oorspronkelijke reeks in jaar i en maand j
T_{ij} de trend-conjunctuurcomponent
S_{ij} de seizoencomponent
$m(T_{ij})$ de vermenigvuldigingsfactor in afhankelijkheid van de trend-conjunctuur
S_{ij} de exogene seizoenfactor
I_{ij} de onregelmatige component.

Voor zover $m(T_{ij})$ een lineaire functie is, wordt deze gespecificeerd als:

$$m(T_{ij}) = \lambda + (1-\lambda)\frac{T_{ij}}{\bar{T}}$$

voor $0 < \lambda < 1$ en $\bar{T}$ het gemiddelde trendniveau in de basisperiode. Indien en voor zover $m(T_{ij})$ kromlijnig is, luidt de specificatie:

$$m(T_{ij}) = \left[(1-\lambda)\left(\frac{T_{ij}}{\bar{T}}\right)^{-\beta} + \lambda\left(\frac{tgh\,\frac{T_{ij}}{\bar{T}}}{tgh\ 1}\right)^{-\beta}\right]^{-1/\beta}$$

met $\beta = 1{,}5$ en $\lambda = 0{,}1,\ 0{,}2,\ \ldots,\ 0{,}9$. De parameter β is gekozen op grond van uitvoerige experimenten. De λ is bepaald met behulp van de kleinste kwadraten methode. Hiertoe moet op a priori gronden worden gekozen voor een kromlijnig of lineair verband, terwijl de exogene seizoenfactor door middel van 3- of 5-punts voortschrijdende gemiddelden wordt berekend. De volgende stappen worden doorlopen:

1. Berekening van de intitiële trend-conjunctuur, T_{ij}, door 12-maands voortschrijdende gemiddelde van de oorspronkelijke reeks, waarbij rekening wordt gehouden met de trendafwijking.
2. Berekening van $\bar{T}$ en een eerste ruwe schatting van $S_{ij} = Y_{ij} - T_{ij}$.
3. Schatting van λ en S_{ij} langs iteratieve weg waarbij uitbijters worden geëlimineerd en vervangen.
4. Bepaling van de seizoenfactoren, waarbij de Census gewichten worden gebruikt om deze te modificeren.
5. Correctie voor seizoeninvloeden en extremencorrectie.
6. Bepaling van de trend-conjunctuur door toepassing van een 13-punts voortschrijdend Henderson gemiddelde.
7. Vervang de beginwaarde van T_{ij} door die onder 6 verkregen en herhaal de stappen 4 en 5 één maal.

De eerste resultaten van de CPBX11-methode voor werkloosheidscijfers zijn bevredigend en tonen een aanmerkelijke verbetering ten opzichte van zijn voorgangers de Census X-11- en de CPB-methode [vgl. CBS (1976)]. Het bezwaar dat een duidelijke stochastische modelstructuur ontbreekt, blijft overigens onverkort gelden en opent wegen voor verder onderzoek. Een aanknopingspunt daarbij biedt mogelijk de door Durbin & Murphy (1975) ontwikkelde regressiebenadering waaraan eveneens een gemengd additief-multiplicatief model ten grondslag ligt.

3.5. *De SABL-methode*

De SABL (Seasonal Adjustment Bell Laboratories)-methode is een correctieprocedure voor maandcijfers, waarbij in het bijzonder aandacht wordt geschonken aan de transformatie der data en de grafische presentatie der resultaten. Het is een operationele methode die nog steeds in ontwikkeling is, zoals blijkt uit de reeks publikaties van een der geestelijke vaders van de methode: W.S. Cleveland [vgl. bijv. W.S. Cleveland, Dunn & Terpenning (1978), W.S. Cleveland & Devlin (1980), W.S. Cleveland & Terpenning (1982)].

Zoals de meeste seizoencorrectiemethoden is SABL een algoritme voor de berekening van de trend-conjunctuur en de seizoencomponent. Deze decompositie is in grote mate geïnspireerd op de X-11 methode, wier filosofie van het gladstrijken (smoothing) in SABL wordt gevolgd. Er zijn echter drie belangrijke verschillen.

In de eerste plaats zijn de filters robuuster dan in de X-11, waardoor overcorrecties aan het eind (en begin) van de reeks worden vermeden. Een tweede belangrijk verschil tussen de X-11 en SABL betreft de aard van de decompositie. De X-11 kent twee mogelijkheden: de additieve of de multiplicatieve variant, waarbij de laatste door een logaritmische transformatie is herleid tot een speciaal geval van de additieve variant. De SABL-methode biedt een veel ruimere klasse van mogelijkheden waarin de additieve en multiplicatieve modellen als speciale gevallen voorkomen. Dit wordt bereikt door toepassing van de machtstransformatie van Box & Cox (1964). De oorspronkelijke reeksen x worden daarom onderworpen aan deze klasse van transformaties, $x^{(p)}$, zodanig dat:

$$x^{(p)} = \begin{cases} x^p \text{ indien } p > 0 \\ \ln x \text{ indien } p = 0 \\ -x^p \text{ indien } p < 0 \end{cases}$$

waarbij p empirisch wordt bepaald. (Het middel van de machtstransformatie wordt overigens ook wel gebruikt in de context van de Box-Jenkinsmodellen). Een derde onderscheidend kenmerk van de SABL-methode is de uitgebreide benutting van grafische methoden voor diagnose en interpretatie, welke door de programmatuur op afroep worden verschaft. Vooral bij de seizoenanalyse van grote aantallen reeksen is dit grafisch hulpmiddel erg handig. Wegens de gelijkenis met de X-11 wordt er hier kortheidshalve van afgezien de afzonderlijke stappen in het algoritme te vermelden. Hiervoor zij verwezen naar de reeds geciteerde auteurs. Opgemerkt zij dat de grafische presentatiemogelijkheden aan SABL een bijzondere aantrekkelijkheid verlenen voor gebruik op grote schaal.

4. BEOORDELING DER METHODEN

Nieuwe methoden worden ontwikkeld omdat gebruikers onvoldaan zijn over de resultaten van de bestaande. Illustratief hiervoor is de CPBX11-methode die door het CBS werd ontwikkeld om het probleem van de overcorrectie in de werkloosheidscijfers te vermijden. Hiermede is niet gezegd dat de laatst ontwikkelde methode toegepast op een willekeurige reeks tevens de meest optima-

le werkwijze garandeert. Uiteindelijk zal de onderzoeker zich in zijn keuze laten leiden door eigen inzicht, ervaring en beschikbare programmatuur.

In het verleden is door Fase *et al.* (1973) een vergelijkend onderzoek ingesteld tussen een negental alternatieve seizoencorrectiemethoden. Deze methoden omvatten mede de eenvoudige additieve en multiplicatieve correctiemethoden uit de elementaire statistische leerboeken. De X-11 scoorde, beoordeeld naar zeven gezichtspunten als stabiliteit bij reeksverlenging en residueel seizoen, het hoogst. Kuipers (1978) herhaalde dit vergelijkend onderzoek en betrok tevens de X-11-ARIMA in zijn analyse. Hij kwam tot de slotsom dat in termen van stabiliteit deze correctieprocedure het meest te verkiezen was voor de beschouwde tijdreeksen. De beoordeling van de kwaliteit van de X-11-ARIMA methode ten opzichte van de Census X-11 en de signaal extractiemethode van Burman (1980) was doel van het onderzoek door Den Butter *et al.* (1982). Hiertoe beschouwen zij een drietal varianten van de X-11-ARIMA methode. Zij concludeerden dat, afgemeten aan de door hen gebezigde criteria als stabiliteit of afwezigheid van residueel seizoen, de X-11 en de X-11-ARIMA varianten weinig in kwaliteit verschillen. Deze uitkomst is met name verrassend omdat de X-11-ARIMA juist is ontworpen om een grotere stabiliteit in de seizoenfactoren te verkrijgen. Interessant hierbij is dat hun bevinding in het bijzonder geldt voor de X-11-ARIMA variant van Statistics Canada met een beperkt aantal mogelijke ARIMA-modellen. Het verdient daarom volgens deze auteurs aanbeveling bij toepassing van de X-11-ARIMA zelf de desbetreffende modellen te schatten.

Elke comparatieve beoordeling van seizoencorrectiemethoden wordt gehinderd door het feit dat het werkelijke seizoen onbekend is. Een oplossing voor dit probleem biedt de analyse van een reeks met een ingebrachte seizoenbeweging. Om die reden is tot slot door ons voor een gesimuleerde reeks met een bekend seizoenpatroon nagegaan hoe goed hieruit het seizoenpatroon zich laat schatten [vgl. Fase (1982)]. Ten behoeve van de X-11-ARIMA is hieruit het volgende univariate ARIMA-model geschat:

$$\underset{(5{,}8)}{(1+0{,}45\,B}+\underset{(2{,}6)}{0{,}20\,B^2)}\nabla\nabla_{12}x_t=\underset{(32{,}0)}{(1-0{,}87\,B^{12})}a_t$$

Dit model is ingevoerd in het X-11-ARIMA programma. Nadat het programma de seizoenfactoren heeft berekend, kunnen deze vergeleken worden met de werkelijke – en in dit voorbeeld bekende! – seizoenfactoren. Als beoordelingsmaatstaf gebruiken we daarbij de gemiddelde absolute afwijking, gedefinieerd als:

$$\frac{1}{12}\sum_{j=1}^{12}|S_{ij}-\hat{S}_{ij}|$$

met S_{ij} de werkelijke en $\hat{S}_{ij}$ de geschatte seizoenfactoren in jaar i en maand j.

Onderstaande tabel bevat deze maatstaf voor het X-11-ARIMA model alsmede voor een drietal andere correctiemethoden [Census X-11, Multiplicative Moving-methode (MM), Multiplicative Fixed-methode (MF)]. Gebezigd zijn de synthetische data tot zeg 1983, zodat de berekende seizoenpatronen met de bekende uit 1983 vergeleken kunnen worden (zie tabel 2).

Tabel 2: Kwaliteit seizoencorrectie volgens vier methoden

methode	gemiddelde absolute afwijking
X-11-ARIMA	0,624
Census X-11	0,839
MM	0,687
MF	1,283

De uitkomsten in tabel 2 leren dat de gemiddelde absolute afwijking voor de X-11-ARIMA methode het laagst uitvalt, maar dat deze methode op de voet wordt gevolgd door de multiplicatieve variant met bewegende seizoenfactoren (MM-variant). Dit is een verrassend resultaat waaraan overigens geen algemene geldigheid kan worden toegekend [zie ook Fase *et al.* (1973)]. Hoe dit ook zij, geconcludeerd mag voorshands worden dat de X-11-ARIMA procedure in dit, overigens zeer gekunstelde voorbeeld, als beste uit de bus komt. Dit is ook wat men zou verwachten van een methode die is geïntroduceerd als verbetering van de beproefde Census X-11 methode. Niettemin is, zoals de resultaten van Den Butter *et al.* laten zien, in de praktijk voorzichtigheid geboden met dit soort uikomsten en mag men niet te snel concluderen dat de X-11-ARIMA methode het meest geschikt is. Het resultaat is vermoedelijk mede afhankelijk van de beschouwde reeks. Een seizoenanalist zal zich daarom telkens opnieuw moeten beraden over de meest in aanmerking komende methode.

5. BESLUIT

Het voorgaande biedt een overzicht van een aantal operationele methoden voor seizoencorrectie die bovendien sterk in de belangstelling staan. Voor al deze methoden is programmatuur beschikbaar of gemakkelijk te verkrijgen [zie ook Den Butter *et al.* (1982), Fase (1980, 1982)]. De seizoenanalyse is het laatste decennium tevoorschijn gekomen uit het wat onooglijke en vaak onderschatte domein van de beschrijvende statistiek. De resultaten van de seizoenanalyse raken een ieder. Dit maakt deze activiteit tot 'one of our most important statistical tools' [Cleveland *et al.* (1982), p. 52], welker ontwikkeling is begunstigd door de opkomst van de electronica. Ook de belangstelling van de wetenschap voor de seizoenanalyse heeft hierdoor een opleving ondergaan. Het Amerikaanse Bureau voor de Statistiek organiseerde in 1976 en 1981, in samenwerking met het National Bureau of Economic Research, een tweetal bijeenkomsten, geheel gewijd aan de problematiek der seizoenanalyse [zie Zellner (1978, 1983)]. Deze bijeenkomsten hebben een belangrijke stimulans gegeven aan verder onderzoek en hebben geleid tot verdieping van inzicht.

Het in deze bijdrage gebodene beperkt zich tot de grote lijnen. Geen aandacht is geschonken aan oude methoden als die van Wald (1936) of Zaycoff (1936) of aan de Bayesiaanse benaderingswijze van bijv. Akaike (1980) van wiens hand een reeks publikaties ter zake verscheen. Evenmin is stilgestaan bij de werkdagencorrecties of de correctiemethoden die zich richten op week- of dagcijfers [zie bijv. Cleveland & Grupe (1981), Hillmer (1982), Pierce (1979)]. Daarnaast is er een lange traditie om, binnen het kader van het traditionele regressiemodel, het seizoeneffect te meten via dummy-variabelen [zie Fase

(1968) en Rosen (1979) voor toepassingen]. Belangrijk is vooral de vaststelling dat het terrein der seizoenanalyse bruist van onderzoekactiviteit, waarvan ook het marktonderzoek bij de voorspelling van toekomstige marktontwikkelingen de vruchten kan plukken. Stridiron (1941, p. 214) schreef ruim 40 jaar geleden 'Prognose van het komende seizoen biedt minder moeilijkheden dan prognose van de conjunctuur, aangezien het seizoen een grotere mate van constantheid bezit'. Niettegenstaande de kern van waarheid in deze constatering, is er sindsdien vooruitgang en verfijning in het instrumentarium der seizoenanalyse mogelijk gebleken. Deze ontwikkeling is ook voor de marktonderzoeker boeiend en nuttig.

BIBLIOGRAFIE

Akaike, H., 1980. Seasonal adjustment by Bayesian modelling, *Journal of Time Series Analysis,* 1, pp. 1-13.

Bordeau de Fontenay, A.J.M.D., 1979. *A contribution to the foundations of the seasonal analysis,* Ph.D. thesis Vanderbilt University, Nashville, Tennessee (ongepubliceerd).

Box, G.E.P. en D.R. Cox, 1964. An analysis of transformations, *Journal of the Royal Statistical Society,* series B, 26, pp. 211-243.

Box, G.E.P. en G.M. Jenkins, 1970. *Time series analysis: forecasting and control,* Holden Day, San Francisco.

Box, G.E.P. en G.M. Jenkins, 1973. Some comments on a paper by Chatfield and Prothero and on a review by Kendall, *Journal of the Royal Statistical Society,* series A, 136, pp. 337-345.

Burman, J.P., Seasonal adjustment by signal extraction, *Journal of the Royal Statistical Society,* series A, 143, pp. 321-377.

Butter, F.A.G. den, R.L. Coenen en F.J.J.S. van de Gevel, 1982. *Seizoencorrectie met behulp van ARIMA-modellen,* rapport 8207 De Nederlandsche Bank N.V.

CBS, 1976. De uitschakeling van seizoeninvloeden uit reeksen betreffende de arbeidsmarkt, *Sociale Maandstatistiek,* 24, No. 1 (januari 1976), pp. 4-12.

Chatfield, C. en D.L. Prothero, 1973. Box-Jenkins seasonal forecasting: problems in a case study, *Journal of the Royal Statistical Society,* series A, 136, pp. 295-315.

Cleveland, W.P., 1972. *Analysis and forecasting of seasonal time series,* Ph.D. thesis University of Wisconsin, Madison, Wisconsin (ongepubliceerd).

Cleveland, W.P. en M.R. Grupe, 1981. Modelling time series when calendar effects are present, *Special Studies Paper 162,* Board of Governors of the Federal Reserve System, Washington D.C.

Cleveland, W.P. en G.C. Tiao, 1976. Decomposition of seasonal time series: a model for the Census-X-11 program, *Journal of the American Statistical Association,* 71, pp. 581-587.

Cleveland, W.S., 1982. Seasonal and calendar adjustment, in: D.R. Brillinger en P.R. Krishnaiah (ed.), *Handbook of Statistics,* vol. 3, North Holland, Amsterdam.

Cleveland, W.S. en S.J. Devlin, 1980. Calendar effects in monthly time series; detection by spectrum analysis and graphical methods, *Journal of the American Statistical Association,* 75, pp. 487-496.

Cleveland, W.S., D.M. Dunn en I.J. Terpenning, 1979. SABL: a resistant seasonal adjustment procedure with graphical methods for interpretation and diagnosis, in: A. Zellner (ed.) *Seasonal analysis of economic time series,* US Department of Commerce, Washington D.C., pp. 201-231.

Cleveland, W.S. en I.J. Terpenning, 1982. Graphical methods for seasonal adjustment, *Journal of the American Statistical Association,* 77, pp. 52-62.

Dagum, E.B., 1975. Seasonal factor forecast from ARIMA models, *Proceedings of the International Institute of Statistics,* 1975, vol. 3, Warsaw, pp. 206-219.

Dagum, E.B., 1980a. *The X-11 ARIMA seasonal adjustment method,* Statistics Canada, Ottawa.

Dagum, E.B., 1980b. A note on the X-11-ARIMA method of seasonal adjustment, *Canadian Statistical Review,* pp. 6-11.

Dagum, E.B., 1982. The effects of asymmetric filters on seasonal factor revisions, *Journal of the American Statistical Association,* 77, pp. 732-742.

Dantzig, D. van, 1952. Voorspelling en profetie, *Statistica Neerlandica 6,* pp. 195-203.

De Forest, E.L., 1877. On adjustment formulas, *The Analyst,* 4, pp. 79-86, en 107-113.

Delfgaauw, G.Th.J., 1973. Inleiding tot de economische wetenschap, *II: macro economie,* Delwel, Den Haag.
Durbin, J. en M.J. Murphy, 1975. Seasonal adjustment based on a mixed additive-multiplicative model, *Journal of the Royal Statistical Society,* series A, 138, pp. 385-410.
Fase, M.M.G., 1968. De vraag naar openbaar vervoer in de drie grote steden, *De Economist,* 114, pp. 619-639.
Fase, M.M.G. (m.m.v. A.J. Pool), 1980. *Handleiding voor gebruik van een zestal rekenprogramma's voor seizoencorrectie, rapport 8019/M, vakgroep marktanalyse en bedrijfsstatistiek,* Erasmus Universiteit Rotterdam.
Fase, M.M.G., 1981a. *Op het breukvlak van macro- en micro-economie,* Stenfert Kroese, Leiden/Antwerpen.
Fase, M.M.G., 1981b. Tijdreeksenanalyse voor marktprognose en marktbeleid, *Jaarboek van de Nederlandsche Vereniging van Marktonderzoekers 1981,* pp. 175-195.
Fase, M.M.G. (m.m.v. J.J. Bouma-Schneijder), 1982. De X-11-ARIMA methode voor seizoencorrectie: een handleiding voor de programmatuur, *Rapport 8219/M, vakgroep marktanalyse en bedrijfsstatistiek,* Erasmus Universiteit Rotterdam.
Fase, M.M.G., 1983. Tijdreeksenonderzoek als informatiebron voor marktbeleid, *Tijdschrift voor Marketing.*
Fase, M.M.G., J. Koning, en A.F. Volgenant, 1973. An experimental look at seasonal adjustment, *De Economist,* 121, pp. 441-480.
Goldberg, S., 1961. *Introduction to difference equations,* Wiley, New York.
Granger, C.W.J. en P. Newbold, 1977. *Forecasting economic time series,* Academic Press, New York.
Greville, T.N.E., 1979. Moving-weighted-average smoothing extended to extremities of the data, *Technical Summary Report No. 2025,* Mathematics Research Center, University of Wisconsin, Madison, Wisconsin.
Haan, R. den, 1979. *Seizoencorrectie, wijzigingen in de methode,* notitie CPB, nr. 14.
Hillmer, S.C., 1982. Forecasting time series with trading day variation, *Journal of Forecasting,* 1, pp. 385-395.
Hoeven, H. van der en A.J. Hundepool, 1980. *A new method of seasonal adjustment: a combination of two existing methods, rapport Centraal Bureau voor de Statistiek,* Voorburg (augustus 1980).
Kenny, P.B. en J. Durbin, 1982. Local trend estimation and seasonal adjustment of economic and social time series, *Journal of the Royal Statistical Society,* series A, 145, pp. 1-41.
Koning, J., 1974. Seizoencorrectie als hulpmiddel bij de economische en monetaire analyse, *Kwartaalbericht De Nederlandsche Bank N.V.,* 1973, No. 4 (maart 1974), pp. 35-40.
Kuiper, J., A survey and comparitive analysis of various methods of seasonal adjustment, in: A. Zellner (ed.), *Sasonal analysis of economic time series,* US Department of Commerce, Washington D.C., pp. 59-76.
Mélard, G. en J.L. Kiehm, 1980. Specification d'un modèle ARIMA evolutif: approche par decomposition, *Cahier du Centre d'Etudes de Recherches Operationelles,* 22, pp. 265-277.
Mendershausen, H., 1937. Methods of computing and eliminating changing seasonal fluctuations, *Econometrica,* 5, pp. 234-262.
Mendershausen, H., 1939. Eliminating changing seasonals by multiple regression analysis, *Review of Economic Statistics,* 21, pp. 171-177.
Nerlove, M., D.M. Grether en J.L. Carvalho, 1979. *Analysis of economic time series: a synthesis,* Academic Press, New York.
Pierce, D.A., 1979. *Seasonal adjustment of daily data with special reference to the US money supply,* Economie Appliquée, 32, pp. 49-72.
Pierce, D.A., 1980. A survey of recent developments in seasonal adjustment, *The American Statistician,* 34, pp. 125-134.
Rosen, K.T., 1979. *Seasonal cycles in the housing market: patterns, costs, and policies,* MIT Press, Cambridge.
Shiskin, J., A.H. Young en J.C. Musgrave, 1967. The X-11 variant of the Census method II seasonal adjustment program, *Technical Paper No. 15,* US Bureau of the Census, Washington D.C.
Stridiron, J.G., 1941. *Handboek der bedrijfseconomische statistiek,* De Haan, Utrecht.
Thompson, H.E. en G.C. Tiao, 1971. Analysis of telephone data: a case study of forecasting seasonal time series, *Bell Journal of Economics and Management Science,* 2, pp. 515-541.

Tinbergen, J., 1943. *Economische bewegingsleer,* Elsevier, Amsterdam.
Verdoorn, P.J., 1950. *Grondslagen en techniek van de marktanalyse,* Stenfert Kroese, Leiden.
Vos, A.F. de, 1975. *Seizoencorrectie van werkloosheidscijfers: kriteria en modellen,* Krips Repro, Meppel.
Wald, A., 1936. *Berechnung und Ausschaltung von Saisonschwankungen,* Springer, Wien.
Winkel, E.G.F., 1982. Interpretation of time series models: a simplified model for the Census X-11 program, in: O.D. Anderson (ed.), *Time series analysis,* North Holland, Amsterdam.
Zaycoff, R., 1936. Über die Zerlegung statistischer Zeitreihen in drie Komponenten, *Publication of the Statistical Institute for Economic Research,* State University of Sofia IV, pp. 141-162.
Zellner, A., 1978. *Seasonal analysis of economic time series, proceedings of the conference,* september 9-10, 1976, US Department of Commerce, Washington D.C.
Zellner, A., 1983. *Applied time series of economic data, proceedings of the conference,* october 13-15, 1981, US Department of Commerce, Washington D.C.
Zwan, A. van der en J. Verhulp, 1980. *Grondslagen en techniek van de marktanalyse,* Stenfert Kroese, Leiden/Antwerpen.

14. Sensorisch onderzoek van voedingsmiddelen en de relatie tot het consumentengedragsonderzoek

J.E.R. FRIJTERS

SAMENVATTING

In het sensorisch onderzoek worden de kleur, geur, smaak en textuur van voedingsmiddelen onderzocht met behulp van proefpersonen, die deze produkten onder (quasi)-experimentele omstandigheden in een laboratoriumsituatie beoordelen. Deze vorm van onderzoek is een tussenschakel tussen het sociaal-psychologisch consumentengedragsonderzoek en het natuurwetenschappelijk voedingsmiddelen onderzoek. In dit artikel wordt in het kort ingegaan op enkele methoden die in het sensorisch onderzoek worden gebruikt.

1. INLEIDING

Voedingsmiddelen zijn economische goederen die behoeften van consumenten dienen te bevredigen, net als auto's, kleding en meubilair om maar een paar willekeurige voorbeelden te noemen. In tegenstelling tot veel andere op industriele schaal voortgebrachte produkten zijn voedingsmiddelen voor de consumenten uiteindelijk van biologische betekenis. Zij leveren namelijk de benodigde stoffen voor de opbouw van het lichaam en voor de bescherming van het lichaam tegen ziekten en infecties. Bovendien zorgen zij voor de energie die nodig is voor het verrichten van lichamelijke arbeid. Voor het echter zover is dat de nutriënten ook daadwerkelijk deze functies kunnen vervullen dienen ze wel eerst in het maag-darmkanaal te zijn opgenomen, dat wil zeggen dat het voedsel waarin deze stoffen voorkomen eerst gegeten of gedronken moet zijn.

Wat zijn nu de factoren die de voedselkeuze en voedselopnameregulatie bepalen? Deze vraag vereist een uitermate complex antwoord. Om dit te vermijden kan worden gesteld dat dergelijke factoren grofweg in drie groepen kunnen worden opgedeeld. Als eerste zijn er de factoren die kenmerkend zijn voor voedingsmiddelen of voedselsoorten zelf. De verzameling van dit soort voedselgebonden factoren wordt meestal aangeduid met het begrip *voedselkwaliteit*. Op de tweede plaats treft men de *sociaal-economische* en *culturele* factoren, zoals de beschikbaarheid van bepaalde produkten, de prijs, normen en waarden met betrekking tot het gebruik van bepaalde produkten, en dergelijke. De derde groep van factoren zijn *persoonsgebonden*. Deze zijn afhankelijk van de persoon die de voedselkeuze maakt. Men kan hierbij denken aan uitwendige factoren, zoals leeftijd, gezondheidstoestand, erfelijkheid, lichamelijk activiteitsniveau en geslacht. Daarnaast zijn er zuiver psychologische

factoren. Deze zijn weer onder te verdelen in sociaal-psychologische factoren die voornamelijk in het consumentengedragsonderzoek worden bestudeerd, en in waarnemingspsychologische factoren, die van oudsher door experimenteel psychologen worden bestudeerd. Sociaal-psychologische factoren zijn gerelateerd aan voedselattituden, terwijl waarnemingspsychologische factoren betrekking hebben op de sensorische perceptie. Het waarnemen van de geur, smaak, textuur, vorm, kleur en temperatuur van voedingsmiddelen. Het onderzoek van deze aspecten en vooral de methoden die daarbij gebruikt worden vormen het eigenlijke onderwerp van dit hoofdstuk.

2. ATTITUDE EN SENSORISCHE PERCEPTIE

Hoewel, zoals al is opgemerkt, voedingsmiddelen uiteindelijk een belangrijke biologische functie voor de consument vervullen, kan men zich toch afvragen of consumenten die rol in hun keuzegedrag expliciet onderkennen en tot uitdrukking brengen. Is het zo dat consumenten voedingsmiddelen voornamelijk zien als verzamelingen van voedingsstoffen en andere chemische bestanddelen die een bepaalde uitwerking op hun lichaam hebben? Waarschijnlijk niet, al zullen er zeker uitzonderingen te vinden zijn. De wijze waarop consumenten voedingsmiddelen percipiëren hangt af van de functies van het eetgedrag en de gemotiveerdheid voor bepaalde soorten van eetgedrag. Naast de bevrediging van de eetlust kan voedingsgedrag andere instrumentele functies hebben. Dit komt bijvoorbeeld tot uitdrukking in het eetgedrag van mensen die ziek zijn en een bepaald voedingspatroon volgen om beter te worden, of in het gedrag van iemand die zichzelf te dik vindt en door een bepaalde wijze van eten wil afvallen. Dat dergelijke motivaties die het eetgedrag sturen tot uitdrukking komen in de attituden, dus in kennis, affect en gedragstendentie ten opzichte van voedingsmiddelen ligt voor de hand. Het betekent tevens dat houdingen ten opzichte van voedingsmiddelen geheel afwijkend kunnen zijn van die welke voort komen uit een goede voedingskundige kennis van de biologische en biochemische effecten op het lichaam van de consumptie van een bepaald voedingsmiddel. Deze zienswijze impliceert dat de perceptie van voedingsmiddelen en de daarmee verbonden attituden-structuren op de eerste plaats sociaal-psychologisch bestudeerd dient te worden. Kenmerkend voor een dergelijke benadering is dat een voedingsmiddel wordt gerepresenteerd als een verzameling van psychologische attributen (eigenlijk cognities). Voor verschillende marktsegmenten of consumenten kunnen profielen worden opgesteld van dergelijke attributen die aan een bepaald voedingsmiddel worden toegeschreven en die afhankelijk van het voedingsmiddel en van het marktsegment in kwestie, een bijdrage leveren voor het maken van de voedselkeuze of het ontstaan van de koop- of eetintentie.

Wat voor soort psychologische attributen worden zoal aan voedingsmiddelen toegekend? En wat is het gewicht dat consumenten aan ieder van deze toekennen bij het maken van een voedselkeuze? Een recent leeronderzoek (Linden en Klingeberg, 1981) uitgevoerd door studenten van de Vakgroep Humane Voeding van de Landbouwhogeschool kan bij wijze van illustratie worden gebruikt om een antwoord op deze vragen te geven. Op grond van een literatuurstudie werden 16 attributen geselecteerd die het meest frequent in consu-

Tabel 1. Psychologische kenmerken die aan voedingsmiddelen worden toegeschreven.

1 Bekendheid	bekendheid met
2 beschikbaarheid	gemakkelijk/moeilijk te verkrijgen
3 Bereidingswijze	gemak/ongemak van de bereiding
4 Combineren	veel/weinig combinatiemogelijkheden
5 Eetgemak	gemakkelijk/moeilijk te kauwen
6 Feestelijkheid	weinig/zeer geschikt voor bijzondere gelegenheden
7 Gezondheidswaarde	erg gezond/erg ongezond
8 Houdbaarheid	lang/kort bewaarbaar
9 Lijnen	gunstig/ongunstig effect op lichaamsgewicht
10 Prijs	erg duur/erg goedkoop
11 Smaak/Geur	aangenaamheid voor wat betreft smaak en geur
12 Textuur	aangenaamheid voor wat betreft textuur
13 Uiterlijk	aangenaamheid voor wat betreft vorm en kleur
14 Verteerbaarheid	gemakkelijk/moeilijk te verteren
15 Verzadigingswaarde	hoge/lage verzadigingswaarde
16 Voedingswaarde	hoge/lage voedingswaarde

mentengedragsonderzoek worden gebruikt. Deze zijn vermeld in Tabel 1. De 16 kenmerken werden in paren aan ieder van 45 (dietetiek) studenten voorgelegd zodat iedere proefpersoon dus 120 paren ter beoordeling kreeg voorgelegd. De vraag die steeds beantwoord diende te worden luidde: Welke van de twee eigenschappen vind je het belangrijkst bij het kiezen van een voedingsmiddel? Na uitvoering van het onderzoek werden uit de verkregen data schaalwaarden berekend voor alle kenmerken met behulp van Thurstone's Case V van de Wet van Paarsgewijze Vergelijkingen (Thurstone, 1927). Deze waarden zijn gerangschikt van meest belangrijk naar minst belangrijk, en staan vermeld in Tabel 2. Wat opvallend is aan deze resultaten is dat voor deze groep studenten de attributen die refereren aan biologische functies van voedingsmiddelen, zoals verteerbaarheid en voedingswaarde, laag genoteerd staan. Daar staat tegenover dat attributen zoals gemak van bereiding, prijs, visueel uiterlijk, en smaak en geur als meest belangrijke kenmerken worden

Tabel 2. Resultaten paarsgewijze vergelijking produktkenmerken.

Kenmerk	Schaalwaarde	
1 Smaak/geur	0.47	meest belangrijk
2 Prijs	0.42	
3 Uiterlijk	0.41	
4 Bereidingswijze	0.34	
5 Bekendheid	0.31	
6 Beschikbaarheid	0.15	
7 Gezondheidswaarde	0.09	
8 Feestelijkheid	0.09	
9 Lijnen	0.09	
10 Textuur	0.04	
11 Eetgemak	−0.02	
12 Combineren	−0.26	
13 Houdbaarheid	−0.30	
14 Voedingswaarde	−0.51	
15 Verzadigingswaarde	−0.66	
16 Verteerbaarheid	−0.94	minst belangrijk

gezien. Smaak en geur, en visueel uiterlijk en textuur zijn kenmerken die betrekking hebben op de zintuiglijke waarneming en die gerangschikt worden onder het begrip smakelijkheid. Blijkbaar is dit aspect zeer belangrijk bij het tot stand komen van voedselkeuzen, althans als wordt afgegaan op het door deze studenten gepercipieerde belang. Dat dit niet ten onrechte is kan worden bevestigd door ervaringen van consumentengedragsonderzoekers (bijv. Solms en Hall, 1981) en door uitkomsten van experimenteel psychologisch en biologisch onderzoek (bijv. Kroeze, 1979; Barker, 1982).

Hoewel het begrip smakelijkheid bekend is uit het leven van alledag is het eigenlijk minder eenduidig dan het lijkt. De smakelijkheid van een voedingsmiddel zoals bepaald door middel van een vragenlijstmethode in een sociaal-psychologisch surveyonderzoek is in eerste instantie gerelateerd aan de attitude van de respondent(en) ten opzichte van het voedingsmiddel. Voor een belangrijk deel is een dergelijke attitude een residu van vroegere ervaringen met het product of met producten uit een bepaalde klasse. De smaak (zoet/zuur), geur en textuur (droog/sappig, taai/mals) van een sinaasappel, om maar weer een willekeurig voorbeeld te noemen, wordt tijdens het consumeren met behulp van de zintuigen waargenomen. Deze sensaties worden cognitief gecodeerd, en deze worden daarna in het langdurend geheugen van de waarnemer opgeslagen. Deze geheugensporen zijn in feite een aantal cognities die aan deze soort sinaasappel (als voorbeeld van een klasse) of aan sinaasappels in het algemeen (als productsoort) worden gekoppeld en vormen dus een bestanddeel van de psychologische representatie hiervan.

De smakelijkheid van een voedingsmiddel als attitudecomponent is niet unidimensioneel, hetgeen volgt uit het feit dat diverse zintuigen zijn betrokken bij de sensorische perceptie wanneer een bepaald voedingsmiddel wordt geconsumeerd. Zoals reeds eerder is genoemd spelen de waarneming van de geur, de smaak, het visueel uiterlijk en de textuur (met behulp van de zintuigen die in de spieren en pezen van het kauwapparaat zijn gelegen) een belangrijke rol. Maar ook sensaties die ontstaan door middel van de temperatuurzintuigen, koude en warmte, en die welke worden waargenomen met behulp van de tast en drukreceptoren in de mond- en keelholte dragen hiertoe bij. Samenvattend zou men kunnen stellen dat de smakelijkheid van een voedingsmiddel in feite de sensorische aangenaamheid is op grond van het totaal van zintuiglijk waargenomen indrukken. Een expliciet onderscheid dient te worden gemaakt tussen smakelijkheid als attitudecomponent (residu van vroegere sensorische ervaringen) en smakelijkheid als hedonische ervaring op het moment van consumptie van het voedingsmiddel. Smakelijkheid van een voedingsmiddel als attitudecomponent wordt meestal onderzocht in een sociaal-psychologisch veldonderzoek, waarbij verschillen tussen voedselcategorieën (bijv. verschillende soorten groenten, of verschillende merken limonade) centraal staan. Smakelijkheid van voedsel als hedonische ervaring welke het gevolg is van sensorische perceptie van een voedingsmiddel wordt meestal onderzocht in een situatie die te vergelijken is met een psychologisch experiment, waarbij veel eerder verschillen tussen varianten van één type product centraal staan. Het ligt voor de hand dat sensorisch onderzoek wordt uitgevoerd in de sfeer van de kwaliteitscontrole en product ontwikkeling. Ook in het onderzoek naar verschillen tussen mensen voor wat betreft de zintuiglijke waarneming van voe-

dingsmiddelen neemt het sensorisch onderzoek een belangrijke plaats in. Men kan hierbij bijvoorbeeld denken aan verschillen tussen bejaarden en jongere mensen, mensen met en zonder suikerziekte etc.

3. SENSORISCH ONDERZOEK

Kenmerkend voor het sensorisch onderzoek is het gebruik van (meestal menselijke) waarnemers (proefpersonen) aan wie één of meer voedingsmiddelen, geuren, smaakoplossingen of andere 'objecten' wordt of worden aangeboden. Aan hen wordt gevraagd aan de monsters (ook wel stimuli genoemd) te ruiken, ervan te proeven of er alleen maar op te kauwen, al naar gelang de soort van object en het zintuig dat in het geding is. Tenslotte wordt gevraagd om één of meer vergelijkende of absolute beoordelingen te geven. Dit kunnen eenvoudige ja/nee antwoorden zijn, maar er kunnen ook beschrijvingen worden gegeven of kwantitatieve uitspraken aan de hand van beoordelingsschalen worden gedaan.

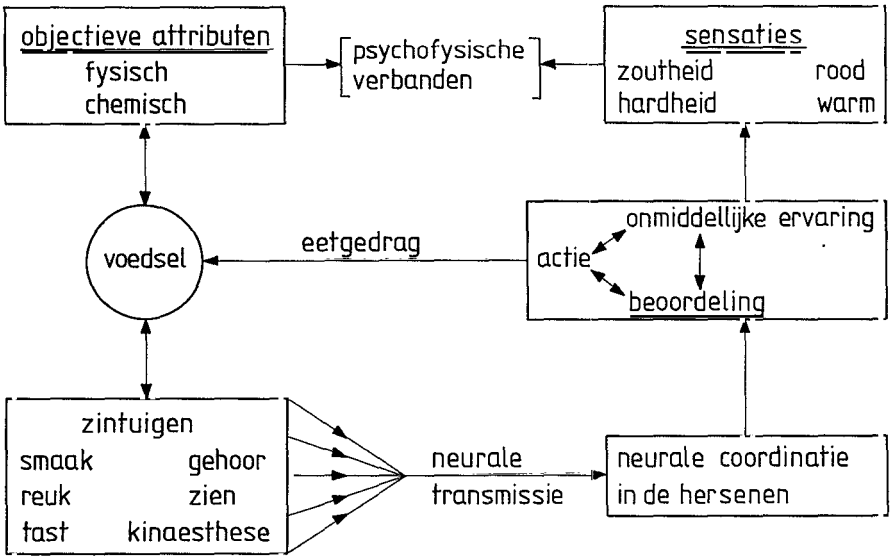
De laatste jaren is de belangstelling voor het sensorisch onderzoek toegenomen. Hiervoor kunnen een aantal redenen worden gegeven. Op de eerste plaats is het een gevolg van het toegenomen 'consumentisme'. Door de grotere invloed van de consumentenorganisaties is de zorg die de producenten aan de kwaliteit van hun produkten besteden eveneens toegenomen, waardoor de behoefte aan sensorisch onderzoek is vergroot. Daarnaast is in deze kringen ook het besef gegroeid dat een kwaliteitsbeheer en -beleid dat uitsluitend is gebaseerd op meer traditionele fysisch-chemische produktkenmerken onvoldoende is. Men is gaan inzien dat deze gegevens op zijn minst dienen te worden aangevuld met informatie over de zintuiglijk waarneembare eigenschappen. In kringen van consumentengedragsonderzoek is de belangstelling voor het sensorisch onderzoek toegenomen, voornamelijk uit de behoefte om attituden ten aanzien van sensorische aspecten van voedingsmiddelen te kunnen vertalen in werkelijk waar te nemen attributen. Tenslotte is er een vierde factor te noemen die een gunstige invloed heeft gehad op de belangstelling voor het sensorisch onderzoek. Dit is de wetenschappelijke en vooral de methodologische vooruitgang die de laatste jaren op dit terrein is geboekt. Deze is vooral toe te schrijven aan het toegenomen besef dat het sensorisch onderzoek een vorm is van toegepast experimenteel psychologisch onderzoek. Methoden die door psychofysici en psychometrici zijn ontwikkeld worden dan ook steeds vaker toegepast.

Voor buitenstaanders is het niet eenvoudig een goed inzicht te krijgen in de methoden van het sensorisch onderzoek. Dit komt met name door het feit, dat er binnen deze discipline geen algemeen geaccepteerd begrippenstelsel is ontwikkeld en dat veel van de methoden zijn geleend van diverse andere wetenschappen. Dit heeft er tevens toe bijgedragen dat de verscheidenheid aan gebruikte methoden groot is. Er zijn beschrijvende en kwantitatieve methoden, toetsings- en schattingsprocedures, sensorisch analytische en sensorische preferentiemethoden en er worden zowel ééndimensionele als meerdimensionele methoden gebruikt. De bedoeling van dit artikel is om op slechts enkele van deze (psychologische) methoden in te gaan. Noodzakelijkerwijs is de behandeling zeer globaal. De aandacht zal vooral worden gericht op de achtergronden

en de beginselen van deze methoden, zonder al teveel detaillering.

4. HET WAARNEMEN

Het onderscheid tussen attitude en zintuiglijke waarneming is al eerder besproken. Om de meetmethoden waaraan in dit betoog aandacht wordt besteed goed te kunnen begrijpen, is het nodig dat eerst wordt stilgestaan bij het waarnemingsproces zelf.
Voedingsmiddelen zijn te beschouwen als verzamelingen van fysisch-chemische eigenschappen, waarvan er een aantal door middel van instrumenten kunnen worden bepaald. Sommige eigenschappen kunnen fungeren als prikkel (stimulus) voor een bepaald zintuig. Sucrose opgelost in water bijvoorbeeld zal het smaakzintuig stimuleren wanneer door iemand van deze oplossing wordt gedronken. In het algemeen geldt dat vóór en tijdens de consumptie van een voedingsmiddel de zenuwcellen van meestal meer dan één zintuig worden gestimuleerd. In het geval van de sucroseoplossing wordt bijvoorbeeld het netvlies geprikkeld door het licht dat gereflecteerd wordt met als gevolg dat een bepaalde kleur wordt waargenomen. Ook wordt de temperatuur en viscositeit waargenomen en ook de smaak, in dit geval zoetheid. Een basisprincipe is dat ieder zintuig gevoelig is voor één specifieke soort van energie. Het gevolg van perifere stimulatie van een zintuig is dat fysiologische activiteit wordt opgewekt. Deze resulteert dan in potentiaalverschillen die langs banen van het zenuwstelsel naar de hoger gelegen gedeelten van de hersenen wordt gevoerd. Hier wordt deze activiteit (op tot nu toe onbekende wijze) getransformeerd in een bewustzijnsinhoud met een waarnemingskarakter, daarom ook wel sensorische perceptie genoemd. Deze bewustzijnsinhoud kan heel complex zijn, maar ook heel eenvoudig, afhankelijk van het aantal zintuigen en soorten prikkels die aan het ontstaan ervan hebben bijgedragen. In de (bewustzijns)psychologie wordt ervan uitgegaan dat sensorische percepties zijn opgebouwd uit kleinere elementen, die sensaties worden genoemd. Er wordt verder vanuit gegaan dat de zintuigen de informatie aanleveren in termen van elementaire sensaties die op het niveau van de hersenschors volgens waarnemingspsychologische wetmatigheden worden gesynthetiseerd tot complexe bewustzijnsinhouden, de sensorische percepties.
Uit deze schematische beschrijving kan worden opgemaakt dat het waarnemen kan worden beschouwd als een drietraps-proces met als achtereenvolgende 'het object' (fysisch-chemisch), 'de zintuigen' (fysiologisch) en de 'sensorische perceptie' (waarnemingspsychologisch). De cognitief gecodeerde sensorische waarneming kan worden geïntegreerd in al bestaande attitudesystemen (sociaal-psychologisch). Het laatste stadium van een dergelijk proces is 'het gedrag' dat de waarnemer naar aanleiding van de sensorische waarneming kan vertonen, bijvoorbeeld wanneer deze persoon zijn waarneming beschrijft. Het zal duidelijk zijn dat voor een buitenstaander slechts twee van deze stadia waarneembaar zijn. Dit zijn het door de waarnemer waargenomen object (het voedingsmiddel wanneer het daar om gaat) en het gedrag dat de proefpersoon vertoont, bijvoorbeeld het eetgedrag. De fysiologische activiteit opgewekt door het object en de resulterende bewustzijnsinhoud blijven voor de buitenstaander onzichtbaar. Echter, voor de waarnemer zelf is de bewustzijnsin-



Figuur 1: Schematische weergave van de waarneming van een voedingsmiddel.

houd wel waarneembaar, en wel voor hem alleen. Dit is nu precies de reden waarom in het sensorisch onderzoek van proefpersonen gebruik wordt gemaakt, en niet van instrumenten zoals dat wel het geval is voor fysisch-chemische eigenschappen van het voedingsmiddel. Het waarnemen voltrekt zich op unieke wijze binnen de proefpersoon en alleen hij heeft toegang tot het resultaat door middel van introspectie. De taak van de sensorisch onderzoeker is, om uit het gedrag van een proefpersoon kennis omtrent de sensorische perceptie af te leiden.

Samengevat kan uit het voorafgaande worden geconcludeerd dat het sensorisch onderzoek gericht is op de bestudering van relaties tussen fysisch-chemische kenmerken van het voedingsmiddel als externe stimulus en de daarmee corresponderende kenmerken van de sensorische gewaarwording (interne perceptie), waarvan de proefpersoon indirect blijkt geeft door middel van zijn gedrag. Dit is nog eens afgebeeld in het schema in Fig. 1

4. ENKELE HOOFDPROBLEMEN IN HET SENSORISCH ONDERZOEK

Gegeven de aard van het waarnemingsproces zoals in algemene termen is geschetst presenteert zich nu in min of meer logische volgorde een serie vraagstellingen waarmee sensorisch onderzoekers regelmatig worden geconfronteerd. Eerst zullen vier van deze problemen, die ieder een eigen aanpak en meetmethode vragen, worden aangestipt, waarna op ieder van de afzonderlijke vraagstellingen verder zal worden ingegaan.

Ieder zintuig reageert op een bepaalde soort van energie, zoals eerder al is opgemerkt. De eerste en meest voor de hand liggende vraag is dan ook: Wat is de kleinste hoeveelheid van een bepaalde soort energie die nodig is om een sen-

sorische gewaarwording op te wekken? In termen van geur- en smaakonderzoek vertaald wordt dus gevraagd wat de laagste concentratie is van een bepaalde geur- of smaakcomponent die nog juist kan worden waargenomen. Deze vraag representeert het *detectieprobleem* en de bepaling van de zogenaamde drempelwaarde.

Wanneer deze laagste concentratie, de drempelconcentratie, bekend is, kan ervan uit worden gegaan dat alle concentraties die hoger zijn in de meerderheid van de gevallen tot een sensatie zullen leiden. De vraag die zich nu voordoet is: Hoe neemt de sterkte van de sensatie toe met een toename van de concentratie? Stel dat de zoetheidssterkte die hoort bij een bepaalde sucroseoplossing 'x' wordt genoemd. Is dan de waargenomen zoetheidsintensiteit tweemaal zo groot wanneer de sucroseconcentratie wordt verdubbeld? Het meten van intensiteiten van sensaties en het beschrijven van de functionele relatie tussen een bepaald type sensatie en de bijbehorende stimulus (smaak- of geurcomponent) wordt het probleem van het *psychofysisch* meten genoemd.

Aangenomen dat het psychofysisch verband tussen waargenomen zoetheid en suikerconcentratie bekend is, dan komt onmiddellijk de vraag naar voren: Wat is het verband tussen de waargenomen intensiteit van de sensatie en de (on)aangenaamheid van de sensatie? In het voorbeeld vertaald: Wat is de relatie tussen de waargenomen zoetheidsintensiteit en de aangenaamheid van de zoetheid? Het zal duidelijk zijn dat deze vraag niet alleen refereert aan het sensorische aspect van de zintuiglijke waarneming, maar ook aan de affectieve betekenis (ook wel hedonische waarde genoemd, hedos = lust) hiervan. Om weer een voorbeeld te geven: 'Wanneer de concentratie van sucrose in water toeneemt zal ook de waargenomen sterkte van de zoete smaaksensatie toenemen. De vraag is nu hoe met een toename van de waargenomen zoetheid de aangenaamheid van zoet toe- of afneemt. Het meten van de affectieve waarde (aangenaamheid, hedonische waarde) van een sensorische gewaarwording noemt men het *hedonisch meten* of the preferentie meten.

Tot nu toe is steeds als voorbeeld gebruikt de waarneming van de zoetheid van sucrose opgelost in water. In tegenstelling tot het drinken van een oplossing van één smaakstof, is de sensorische perceptie die het gevolg is van het nuttigen van een voedingsmiddel in de meeste gevallen complex. Dit is het gevolg van de fysisch-chemische complexiteit van de meeste voedingsmiddelen. Bij grondige bestudering dient de sensorische perceptie, die opgebouwd is uit een aantal afzonderlijke geur-, smaak- en textuursensaties, in termen van deze elementen te worden geanalyseerd. Dit gebeurt vaak door middel van de *constructie van een sensorisch* geur-, smaak- of textuur*profiel.*

5. HET DETECTIEPROBLEEM

Er zijn grenzen aan de gevoeligheid van ieder zintuig. Naarmate de nog juist waarneembare concentratie van een bepaalde geur- of smaakcomponent lager is, des te gevoeliger is het zintuig. Voedingsmiddelen bevatten vaak grote hoeveelheden verschillende chemische verbindingen. Koffiegeur bijvoorbeeld bevat meer dan 500 verschillende componenten. Deze chemische substanties kunnen in principe door middel van gaschromatografische en massaspectrometrische technieken worden aangetoond. Voor de zintuiglijke waarneming

van de geur zijn ze niet alle even belangrijk. Alleen die stoffen, die in een concentratie aanwezig zijn die hoger is dan de drempelconcentratie, zijn van betekenis. Voor veel enkelvoudige componenten zijn drempelconcentraties bepaald in diverse media. Nog steeds bestaat een belangrijk onderdeel van het sensorisch onderzoek uit de bepaling van de zintuiglijke gevoeligheid voor bepaalde stoffen. Hiervoor zijn een groot aantal verschillende methoden beschikbaar. De onderlinge verschillen tussen deze methoden bestaan hoofdzakelijk uit verschillen in experimentele procedure, die wordt gevolgd bij het aanbieden van de stimuli, en uit de wijze waarop de verkregen data worden geanalyseerd. De meest gebruikte methode is de 'Methode van enkelvoudige stimuli'. Hierbij wordt een reeks van oplopende concentraties van een component in water of lucht geprepareerd, al naar gelang of het gaat om de smaak of de geur. Soms wordt aan een dergelijke reeks een 'blank' stimulus (alleen water of alleen lucht) toegevoegd. Deze concentraties worden al of niet in willekeurige volgorde aan één of meer proefpersonen aangeboden. Hen wordt verzocht van de oplossingen te proeven of aan de geuren te ruiken. Steeds dienen ze met ja of nee aan te geven of een smaak of geur wordt waargenomen. Na beëindiging van het experiment wordt per concentratie de proportie ja-antwoorden berekend. De concentratie die overeenkomt met 50% ja-antwoorden wordt de absolute drempel voor die groep proefpersonen en die component genoemd.

Voor het schatten van dit 50% punt bestaan een groot aantal verschillende methoden en modellen (bijv. Guilford, 1954). Veel gehanteerd is de Normit methode. Deze houdt in dat alle proporties ja-antwoorden worden omgezet in (lineaire) standaardnormaalafwijkingen met behulp van de cumulatieve normaalverdeling. Vervolgens wordt met behulp van simpele lineaire regressie een detecteerbaarheidscurve geconstrueerd, waarbij 'x' de concentratie (of de logarithme van de concentratie) is, en 'y' de bijbehorende normit (standaardnormaalafwijking).

Behalve de absolute drempel van een component kan ook de verschildrempel voor een bepaalde concentratie worden bepaald. Dit is de concentratie die in 75% van de aanbiedingen als sterker dan een bepaalde standaardconcentratie wordt waargenomen. Hiertoe wordt steeds één paar stimuli in plaats van één stimulus aangeboden. Eén van de twee stimuli van ieder paar is hetzelfde in alle paren die worden aangeboden en deze stimulus wordt daarom standaardstimulus genoemd (meestal de laagste concentratie van de reeks). De vraag aan de proefpersoon luidt steeds: Welke van de twee stimuli heeft de sterkste smaak? (of geur als het daar om gaat). De proefpersoon dient altijd een keuze te maken, ook al meent hij dat beide stimuli even sterk zijn. Een dergelijke procedure heet daarom 'gedwongen keuze procedure'. Bij de berekening van de verschildrempelconcentratie gaat men in grote lijnen op dezelfde wijze te werk als in het geval van de absolute drempelconcentratie. In dit geval wordt echter voor de berekening van de normit behorende bij een bepaalde proportie juiste uitspraken eerst een correctie toegepast. Dit is nodig om tegemoet te komen aan de kans dat de proefpersoon een gedeelte van zijn uitspraken juist heeft als gevolg van goed raden.

Een gedeelte van een experiment dat op de Vakgroep Humane Voeding is uitgevoerd kan als illustratie dienen voor de bepaling van de verschildrempel-

Tabel 3. Stimulusconcentraties; aantallen en proporties juiste uitspraken met bijbehorende normits (standaardnormaalafwijkingen). De discrimineerbaarheidsfunctie tussen z (normit) en ln S (logarithme van de sucroseconcentratie in water) is ook gegeven.

Sucroseconcentratie (mol/l)	Aantallen juiste uitspraken/aanbied.	Proportie juiste uitspraken	Normit (z)
0,0006	20/39	0,5128	−1,977
0,0013	20/39	0,5128	−1,977
0,0025	23/39	0,5897	−0,915
0,0050	33/39	0,8462	0,502
0,0100	33/39	0,8462	0,502
0,0200	35/39	0,8974	0,827

Discrimineerbaarheidsfunctie: z = 0,5897 ln S + 3,2556
Verschildrempel: 0,0040 Mol/l

Naar: Frijters, J.E.R. en Rasmussen-Conrad, E.L. (1982).

concentratie van sucrose opgelost in water. Aan 13 overgewichtige vrouwen werd drie maal zes paren stimuli aangeboden. Een paar bestond uit een oplossing van sucrose in water en de 'blank' stimulus: gedistilleerd water. Zij dienden steeds aan te geven welke van de twee aangeboden stimuli het zoetst was. De concentraties, de aantallen en de proporties juiste uitspraken met bijbehorende normit-waarden, die via een tabel zijn verkregen, zijn vermeld in Tabel 3. Ook de discrimineerbaarheidsfunctie is gegeven; in dit geval berekend met een gewogen Normit-procedure (Gacula & Kubala, 1978) waarop verder niet ingegaan zal worden. (Er wordt in dit verband trouwens gesproken over discrimineerbaarheidsfunctie en niet over detecteerbaarheidsfunctie omdat het gaat over het waarnemen van verschillen en niet om absolute detectie). Door deze discrimineerbaarheidsfunctie gelijk te stellen aan nul kan de verschildrempelconcentratie ten opzichte van water worden berekend. Deze bedraagt 0.0040 mol/l.

Voor de bepaling van verschildrempels is de eerste psychofysische wet van belang die door Weber (1846) is opgesteld en die daarom ook naar hem is genoemd. Deze onderzoeker ontdekte dat de grootte van de verschildrempel afhankelijk is van de concentratie van de standaardstimulus, en ook, dat er een wetmatigheid bestaat in deze afhankelijkheid (zie ook bijvoorbeeld Ekman, 1959). Weber kwam tot de volgende relatie:

$$\frac{S_2 - S_1}{S_1} = \frac{\Delta S}{S_1} = \text{Constant} . \qquad (1)$$

In deze uitdrukking is S_1 de standaardstimulusconcentratie en ΔS is de verschildrempelconcentratie die aan S_1 toegevoegd moet worden om S_2 te krijgen. De stimulus S_2 wordt dus in 75% van de aanbiedingen als sterker dan S_1 waargenomen omdat ΔS de verschildrempelconcentratie is. Uit deze wetmatigheid blijkt dat er steeds een constante fractie van de standaardstimulusconcentratie moet worden toegevoegd om een waarneembaar verschil tussen S_2 en S_1 te verkrijgen.

Voor geurstoffen bijvoorbeeld blijken de waarden voor de Weberfractie te variëren rond 0,24, voor smaakstoffen rond 0,15 (Baird & Noma, 1978). In

deze paragraaf is uitsluitend aandacht besteed aan de bepaling van absolute en verschildrempels. Deze begrippen spelen een belangrijke rol bij het sensorisch onderzoek en ook bij het chemisch onderzoek van aromastoffen. Des te opmerkelijker dat in meer geavanceerde theorieën over het waarnemen van lage fysische intensiteiten wordt afgezien van het gebruik van deze begrippen. In de 'Signaal Detectietheorie' die hier niet behandeld zal worden (Green & Swets, 1966) wordt het bestaan van dergelijke sensorische drempels zelfs ontkend. Talrijke experimenten bevestigen deze theorie in zoverre, dat er inderdaad bij toename van de concentratie geen punt wordt gevonden waarbij de proportie ja-antwoorden abrupt overgaat van 0% naar 100%. Deze overgang is altijd heel geleidelijk, wat erop duidt dat ook andere factoren een rol spelen bij het totstandkomen van een ja-antwoord (in het geval van absolute detectie). Het ziet ernaar uit dat deze geavanceerde theorieën in de toekomst ook bij het sensorisch onderzoek van voedingsmiddelen zullen worden gebruikt.

6. PSYCHOFYSISCH METEN

De absolute drempelconcentratie wordt beschouwd als het laagste energieniveau waarop het zintuig actief is. Concentraties boven deze waarde geven meestal aanleiding tot het ontstaan van een sensatie, die naarmate de concentratie toeneemt, steeds sterker wordt. Een concentratie hoger dan 0,0040 Mol/l sucrose in water geeft dus aanleiding tot een smaaksensatie, die zoet wordt genoemd. Zoetheid is de sensorische kwaliteit (hoedanigheid). De vraag is nu hoe precies de waargenomen intensiteit toeneemt met de sucrose concentratie. Een beschrijving van een dergelijk functioneel verband kan pas worden gegeven wanneer vaststaat hoe een bepaalde waargenomen zoetheidsintensiteit kan worden gemeten. Immers, zoals in een van de voorafgaande paragrafen reeds is opgemerkt, sensaties zijn persoonlijke ervaringen die niet van buitenaf kunnen worden waargenomen. Om te komen tot een schaal voor subjectieve (= sensorische) intensiteit van sensaties werd door Fechner (Adler, 1966) in 1860 gebruik gemaakt van de al eerder besproken Wet van Weber. De algemene vorm van een dergelijke schaal verkreeg hij door het doen van twee aannamen. Ten eerste definieerde hij een eenheid voor het meten van de sensatie, en wel als volgt. Twee stimuli S_2 en S_1 die één verschildrempelconcentratie van elkaar verschillen worden juist-waarneembaar-verschillend genoemd: hun sensorisch verschil is één j.n.d. (just-noticeable-difference = juist-waarneembaar-verschil). Door vervolgens de Weberfractie te integreren onder de aanname dat alle j.n.d.'s psychologisch gelijk zijn verkreeg Fechner de volgende vergelijking:

$$I = k \log \mathrm{S} + \mathrm{C}. \qquad (2)$$

Hierin is I de waargenomen intensiteit (gemeten in j.n.d.'s) van een stimulus en S de fysische sterkte van die stimulus (in het geval van geur en smaakoplossingen, de concentratie). De grootheden k en C zijn beide constanten, die empirisch bepaald dienen te worden. De vergelijking (2) staat bekend als de Wet van Fechner. In het psychofysisch onderzoek werd deze wet gebruikt tot in de jaren vijftig. Rond die tijd is er scherpe kritiek gekomen op de geldigheid ervan, met name van de kant van S.S. Stevens, hoogleraar in de psychofysi-

ca aan de Harvard Universiteit (Stevens, 1961). Deze kritiek richtte zich op een drietal aspecten. Ten eerste op de aanname dat alle j.n.d.'s subjectief gelijk zijn. Uit zijn experimentele bevindingen concludeerde Stevens dat dit niet het geval is. Ten tweede op de constantie van de Weberfractie. Het blijkt inderdaad dat bij lage en bij hoge prikkelsterkte de Weberfractie verschillend is van de fracties die bepaald kunnen worden voor stimuli van middelmatige sterkte. Het derde punt dat Stevens naar voren bracht was, dat sensaties niet persé indirect (dat is door vergelijking van twee stimuli) gemeten hoeven te worden. Integendeel, Stevens ging ervan uit dat een proefpersoon in staat is om een direct kwantitatief oordeel te geven over de intensiteit van een bepaalde sensatie. Om te komen tot directe schattingen van waargenomen intensiteiten ontwikkelde Stevens 'De Methode van Magnitude Estimation' (Stevens, 1956). Bij deze methode wordt de proefpersoon vóór de aanvang van het experiment een standaardstimulus aangeboden. Hem wordt verteld dat deze een subjectieve sterkte heeft van 10. Gedurende het experiment wordt aan de proefpersoon een serie stimuli van verschillende concentraties aangeboden. De proefpersoon dient nu de waargenomen inensiteit van de sensaties te beoordelen in verhouding met de intensiteit van de standaardstimulus. Als een bepaalde stimulus dus als 10 maal zo sterk als de standaardstimulus wordt waargenomen, dient de proefpersoon het getal 100 te geven. Is een stimulus 5 maal zo sterk, dan geeft de proefpersoon het getal 50. Is de intensiteit van de teststimulus de helft van die van de standaardstimulus, dan geeft hij een 5. Een groot voordeel van deze methode is dat de proefpersonen hun eigen getallen kunnen kiezen. Dit kunnen ook niet-gehele getallen zijn. In het geval dat de proefpersoon de bedoelde sensatie niet waarneemt (wanneer de concentratie van de teststimulus lager is dan de drempelconcentratie) dan geeft hij een nul als response.

Na beëindiging van het experiment wordt op grond van de data een zogenaamde machtsfunctie berekend. Deze wordt ook wel de Stevens functie genoemd. De algemene vorm van een dergelijke psychofysische functie is:

$$I = kS^n \,, \tag{3}$$

Waarin I de waargenomen intensiteit van een stimulus is, en S de concentratie van die stimulus. De grootheid k is een schaalconstante; de grootheid n is een constante die specifiek is voor zowel het zintuig (bijv. smaak) als voor het type stimulus (bijv. sucrose) dat is gebruikt. Wil men het feit in rekening brengen dat het zintuig pas werkzaam is boven de absolute drempelwaarde, dan wordt vergelijking (3) gewijzigd in:

$$I = k(S - S_0)^n \,, \tag{4}$$

waarin S_0 de drempelconcentratie is. De Wet van Stevens en de bijbehorende 'Methode van Magnitude Estimation' (waarvan overigens ook een aantal varianten van ontwikkeld zijn) heeft als grootste voordeel de eenvoud waarmee psychofysische functies kunnen worden bepaald. In vergelijking met de indirecte methode die door Fechner werd voorgestaan is deze methode veel minder omslachtig. De taak van de proefpersoon lijkt op het eerste gezicht erg moeilijk, maar onderzoekers die ervaring met de methode hebben zullen kunnen beamen dat de meerderheid van de proefpersonen na slechts één enkele oefen-

Tabel 4. Exponenten (n) en constanten van individuele psychofysische functies $I = kS^n$ voor sucrose opgelost in water.

Proefpersoon	n	k
1	0,74	18
2	1,68	65
3	1,03	34
4	1,59	105
5	2,00	341
6	1,56	135
7	1,39	176
8	2,45	168
9	1,92	120
10	1,25	68
11	1,84	118
12	0,99	45
13	1,24	67
Gemiddeld	1.51*	87**

* Rekenkundig gemiddelde
** Geometrisch gemiddelde

Naar: Frijters, J.E.R. en Rasmussen-Conrad, E.L. (1982).

zitting de procedure goed kunnen uitvoeren. Sinds de introductie van de Wet van Stevens en de bijbehorende methode zijn letterlijk honderden experimenten uitgevoerd volgens dit principe.

Een ander gedeelte van het hiervoor reeds genoemde sucrose-experiment kan als voorbeeld voor het psychofysisch meten met behulp van 'Magnitude Estimation' worden gebruikt.

Na een oefenzitting werd aan de 13 overgewichtige proefpersonen 3 maal 6 concentraties van sucrose in water aangeboden. De volgorde van de stimuli was willekeurig. De concentraties waren 0,0600; 0,1148; 0,2089; 0,3082; 0,6918 en 1,3000 Mol/l. Op grond van de resultaten van het experiment werd na afloop voor iedere proefpersoon een psychofysische zoetheidsfunctie berekend. Deze zijn gegeven in Tabel 4. Men vindt hierin een kolom onder *k*, de schaalconstante, en een kolom onder *n*, de exponent van de machtsfunctie. De

Tabel 5. Representatieve exponenten (n) van de psychofysische functie $I = kS^n$ voor een aantal sensaties.

Sensatie	Exponent (n)
1 Elektrische Prikkeling (Huid)	3,5
2 Visuele Verzadiging (Rood)	1,7
3 Gewicht	1,1
4 Lengte van Lijnen (Visueel)	1,0
5 Oppervlakte (Visueel)	0,8
6 Huid vibratie (250 Hz)	0,6
7 Koffiegeur	0,5
9 Luidheid (1000 Hz)	0,3
9 Helderheid (Visueel)	0,3

Tabel 6. Exponenten voor een zevental geuren.

Component	Exponent (n)
1 Aceton	0,75
2 Amylacetaat	0,21
3 Benzylacetaat	0,33
4 1-Butanol	0,52
5 Ethylacetaat	0,31
6 1-Pentanol	0,45
7 1-Propanol	0,65

Naar: Patte et al., 1975.

variatie tussen de individuele waarden van *n* is van betekenis, omdat dit het onderlinge verschil tussen de proefpersonen representeert. In dit voorbeeld is de variatie aanzienlijk. Echter, de groepsexponent die berekend is op grond van de individuele exponenten is gelijk aan 1,51, een waarde die dicht bij de waarde van 1,45 ligt wat ongeveer het gemiddelde is van de in de literatuur gerapporteerde waarden voor de exponent van sucrose in water en bepaald op ongeveer dezelfde wijze als in dit experiment.

Uit onderzoeken blijkt dat de exponent die wordt verkregen voor een psychofysische functie onder andere afhankelijk is van het type stimulus dat gebruikt is, maar ook van het zintuig waarmee deze stimulus wordt waargenomen. Tabel 5 geeft een overzicht van exponenten voor diverse soorten sensaties; het zijn gemiddelden over verschillende typen stimuli. In Tabel 6 zijn de waarden opgenomen van de exponenten van psychofysische functies van een zevental geuren (naar Patte et al, 1975). Hieruit blijkt duidelijk de variatie tussen stimuli die met hetzelfde zintuig worden waargenomen.

Hoewel vaak wordt gesproken over *de* exponent voor een bepaalde stimulus dient toch niet te worden vergeten dat allerlei experimentele omstandigheden van invloed kunnen zijn op de exacte vorm van de verkregen psychofysische functie. Zo blijken van invloed te zijn, de grootte van het bereik van de concentraties die worden gebruikt, de wijze waarop de stimuli worden aangeboden, de volgorde waarin dit gebeurt etc. Bij het toepassen van de 'Methode van Magnitude Estimation' dient de onderzoeker rekening te houden met de invloed van dergelijke factoren.

7. HEDONISCH METEN

Een sensorische waarneming is opgebouwd uit sensaties die een bepaalde kwaliteit en een bepaalde intensiteit hebben. Zij vormen de elementen van de sensorische perceptie, die als gevolg van de organisatie van de elementen volgens psychologische mechanismen tot stand komt. Sensaties dienen afzonderlijk te worden gemeten, al of niet in samenhang met de fysische stimulus.

Er is echter nog een belangrijk ander aspect aan de sensorische perceptie, zoals al eerder is uiteengezet. Dit is de affectieve beleving van datgene wat door de proefpersoon wordt waargenomen. Dit kan voor hem aangenaam, neutraal of onaangenaam zijn. De affectieve beleving van datgene wat wordt waargenomen representeert de motivationele verhouding van de waarnemer

ten opzichte van het waarnemingsobject. Daarom wordt het ook wel het hedonische aspect genoemd. Het is duidelijk dat er bij het sensorisch onderzoek altijd een onderscheid moet worden gemaakt, en dit is zeer fundamenteel, tussen dat wat wordt waargenomen en de affectieve beleving van die waarneming. De waargenomen zoetheid van een bepaalde sucroseoplossing is bijvoorbeeld niet hetzelfde als de ervaren aangenaamheid van deze zoetheid. Staat in de psychofysica de relatie tussen stimulus en sensatie centraal, in de psychohedonica probeert men wetmatigheden te vinden voor de beschrijving van het verband tussen de intensiteit van een sensatie en de hedonische waarde van deze sensatie. Dit type onderzoek is nog moeilijker dan het psychofysisch onderzoek omdat het hier gaat om de relatie tussen twee psychologische grootheden. Het meeste onderzoek dat tot nu toe is verricht is niet uitgevoerd met voedingsmiddelen, maar met relatief eenvoudige stimuli.

Reeds voor Wundt, de grondlegger van de experimentele psychologie, werd in 1874 een grafische afbeelding gegeven van een psychohedonisch verband. In zijn algemeenheid ziet een dergelijk verband er als volgt uit. Met een toename van de waargenomen intensiteit van een bepaalde sensatie zal ook de aangenaamheid toenemen, dit echter tot een bepaald punt, waarna een verdere toename van de intensiteit leidt tot een afname van de aangenaamheid. Het verloop kan dus worden voorgesteld door een functie met één piek, vandaar dat deze functie soms wel enkel-gepiekte-preferentie functie wordt genoemd. Het punt dat de maximale aangenaamheid representeert is dus het zogenaamde ideaalpunt. Met een voorbeeld kan dit worden toegelicht. Wanneer de concentratie van sucrose in water toeneemt, neemt de waarneembare zoetheid eveneens toe. Met een toename van de waargenomen zoetheid zal aanvankelijk de zoetheidsaangenaamheid toenemen en daarna vanaf een bepaald punt weer afnemen, waarbij de aangenaamheid kan overgaan in onaangenaamheid.

Het meten van hedonische waarden die aan sensaties worden toegekend gebeurt in het algemeen minder gefundeerd dan het meten van de intensiteit van een sensatie. Dit is het gevolg van het gebrek aan inzicht in psychologische processen die leiden tot het functionele verband tussen de intensiteit van een bepaalde sensatie en de (on)aangenaamheid van de sensatie. Voor het meten van de hedonische waarde wordt een aantal verschillende schaaltypen gebruikt, bestaande uit een volgorde van categorieën (ordinale schalen). Deze kunnen zowel door cijfers (bijvoorbeeld van 1 t/m 9, waarbij 1 = zeer onaangenaam en 9 zeer aangenaam) worden verankerd als door woorden (bijvoorbeeld, zeer onaangenaam, onaangenaam, een beetje onaangenaam, neutraal, een beetje aangenaam, aangenaam, zeer aangenaam). Dit type schaal noemt men vaak een Likert-schaal, naar de uitvinder ervan, en deze wordt in veel andere vormen van (sociaal)psychologisch onderzoek ook gebruikt. In het sensorisch onderzoek wordt een categorieschaal die wordt gebruikt voor het meten van de affectieve beleving vaak ook wel hedonische schaal genoemd. Naast de hiervoor genoemde typen kunnen ook wel de zogenaamde 'visual-analog' schalen of ongestructureerde schalen worden gebruikt. Een dergelijke schaal wordt gevormd door een lijnstuk van een bepaalde lengte, waarop de proefpersoon een kruisje moet zetten om zijn affectief oordeel over de sensatie kenbaar te maken.

Door Moskowitz (1976) is een algemene vorm gespecificeerd voor de functie die de (on)aangenaamheid van een sensatie relateert aan de sterkte van de sensatie. Deze functie ziet er als volgt uit:

$$A = f(I) = k_1(I+k_2)^2+k_3 \,, \tag{5}$$

waarin I de intensiteit van de sensatie is en k_1 t/m k_3 constanten zijn die ad hoc kunnen worden bepaald. (k_1 negatief, k_2 en k_3 positief).

Om via de waargenomen intensiteit het verband tussen aangenaamheid van een sensatie en de concentratie van een stimulus te kunnen leggen, moet voor ieder type sensatie een dergelijke functie worden opgesteld. Wanneer men bedenkt dat de sensorische perceptie die optreedt naar aanleiding van het waarnemen van een voedingsmiddel uit een aantal verschillende sensaties bestaat kan men zich voorstellen dat de bepaling van een aantal psychohedonische functies in een dergelijk geval zeer omslachtig en moeizaam is. Het is dan ook niet verwonderlijk dat veel onderzoekers hiertoe niet komen. In plaats van de bepaling van deze individuele psychohedonische functies, nodig voor de voorspelling van de overall-aangenaamheid van het voedingsmiddel, wordt een rechtstreeks verband gelegd tussen instrumenteel bepaalde fysische grootheden en de totale aangenaamheid van het voedingsmiddel. Dergelijke relaties worden meestal ad hoc bepaald, en het is dan ook niet verwonderlijk dat de predictieve waarde van dergelijke relaties meestal slecht is.

8. PROFILEREN

Oplossingen van één bepaalde smaak- of geurstof in water of in lucht zijn relatief eenvoudige stimuli. De relatie tussen de concentratie en de sensorische kwaliteit van de resulterende sensatie is vaak eenduidig; sucrose opgelost in water leidt tot een zoete smaakgewaarwording, keukenzout in water wordt als zout waargenomen. Er bestaat blijkbaar een duidelijke één-op-één correlatie tussen de fysische karakteristiek en de kwaliteit van de sensatie. Toch is dit niet helemaal waar. Bij hoge suikerconcentraties kunnen naast de zoetheid bijsmaken ontstaan, zoals bijvoorbeeld bitterheid. Bij lage concentraties van keukenzout is de sensatie zoet in plaats van zout.

Bij voedingsmiddelen, zoals bijvoorbeeld zwarte bessen, uien, gebraden vlees, whiskey en bier, is de flavour perceptie (=smaak plus geur) veel complexer dan bij de eenvoudige stimuli die hiervoor steeds zijn genoemd. Voordat verbanden kunnen worden gelegd tussen fysisch-chemische kenmerken van dergelijke produkten en de sensorische perceptie dient eerst te worden onderzocht tot welke sensaties de consumptie van een produkt aanleiding geeft. Dit gebeurt van oudsher door het toepassen van een zogenaamde profielmethode (alhoewel steeds vaker gebruik gemaakt wordt van multidimensionele schaaltechnieken; Schiffman et al, 1981). Hiervan bestaan een groot aantal varianten. De meest eenvoudige methode bestaat uit het aanbieden van een aantal verschillende varianten van één bepaald produkt type aan een aantal proefpersonen met het verzoek dit produkt in sensorische (en dus niet in affectieve of cognitieve) termen te beschrijven. Deze beschrijvende sensorische worden dan verzameld en de meest voorkomende worden beschouwd als representaties van afzonderlijke sensaties. Een lijst van dergelijke termen die

voor een bepaald produkt type is opgesteld noemt men een smaak-, geur-, flavour of textuurprofiel.

Een andere aanpak die kan worden gevolgd bij het profileren van een produkt is de aanbieding van een lijst van termen tegelijk met één of meer varianten van een bepaald produkt type. Een vaak gebruikte lijst voor geuren is die van Harper et al. (1968). De proefpersoon kan de beschrijvingen op de lijst aankruisen die van toepassing worden geacht voor de geurindruk. Meestal wordt op een toegevoegde schaal van 0–5 de mate aangegeven waarin de sensatie van toepassing wordt geacht. Dit laatste is afhankelijk van de waargenomen intensiteit. Bij verschillen tussen de proefpersonen onderling wordt door middel van additionele discussies getracht overeenstemming te verkrijgen over het gebruik van bepaalde termen, of men besluit de termen, waarover men het niet eens kan worden, toch te handhaven.

De profielmethoden zijn aantrekkelijk in gebruik omdat ze erg flexibel zijn. Ze kunnen gemakkelijk worden aangepast aan de eisen van het betreffende onderzoek en ze zijn gemakkelijk te begrijpen voor de proefpersonen waaruit het panel is geselecteerd; de meeste proefpersonen kunnen de methode gemakkelijk aanleren. Toch zijn er ook enige nadelen. Op de eerste plaats is het niet altijd duidelijk dat twee verschillende beschrijvende termen ook inderdaad twee verschillende sensaties representeren. Op de tweede plaats levert een dergelijke min of meer inductieve aanpak vaak grote verschillen tussen de proefpersonen onderling op. Een derde nadeel is van meer theoretische aard; vaak wordt alleen maar nominale of ordinale informatie verkregen. Profielmethoden zijn echter onmisbaar in het sensorisch onderzoek, doch zij worden steeds vaker gebruikt samen met multivariate schaaltechnieken, zoals al is opgemerkt. Technieken zoals factor- en principale-componenten analyse, kunnen worden gebruikt voor het ontwerpen van zogenaamde perceptuele ruimten, waarvan de assen de afzonderlijke sensaties representeren (zie bijv. Frijters, 1976).

9. SLOTWOORD

Detectie, discriminatie, psychofysisch en hedonisch meten en het profileren van de flavour van voedingsmiddelen is in dit hoofdstuk in de notedop aan de orde geweest.

Niet is besproken dat er vaak grote individuele verschillen tussen proefpersonen onderling bestaan. Dit is ook wel te verwachten. Mensen verschillen onderling voor wat betreft hun biologische kenmerken die belangrijk zijn voor de werking van de zintuigen. Ze verschillen ook in de sensorische ervaringen die ze met voedingsmiddelen hebben opgedaan voordat ze het laboratorium van de sensorisch onderzoeker betreden. Dit heeft met name betrekking op verschillen in sensorische preferenties en de wijze waarop met behulp van de taal hun indrukken beschrijven.

Ook is niet besproken dat sensaties van voedingsmiddelen nooit op zich worden waargenomen, maar altijd in de context van het nuttigen van produkten. Dit betekent dat beoordelingen van sensaties of sensorische percepties gemakkelijk beïnvloed kunnen worden door de attitude van de proefpersoon ten opzichte van het voedingsmiddel en de situatie waarin het wordt geconsu-

meerd. Zoals in het begin van dit hoofdstuk is betoogd kunnen deze attitudes betrekking hebben op niet-sensorische factoren zoals, prijs, bekendheid, houdbaarheid e.d. Het zal duidelijk zijn dat naast de sensorisch perceptieve factoren deze houdingsaspecten van groot belang zijn bij het verklaren van het koop of eetgedrag van de proefpersoon. Voor een goed begrip van het gedrag van de proefpersoon ten aanzien van een produkt dient naast het bestuderen van zijn sensorische perceptie aandacht besteed te worden aan de bepaling van de attitude, zoals dat in het sociaal-psychologisch consumentengedragsonderzoek wordt gedaan. Uiteindelijk is het sensorisch analytisch en preferentieonderzoek niet meer dan een schakel tussen het natuurwetenschappelijk onderzoek van voedingsmiddelen en het sociaalwetenschappelijk onderzoek van het consumentengedrag (Frijters, 1975).

BIBLIOGRAFIE

Adler, H.E., 1966. *Elements of Psychophysics.* Holt, Rinehart en Winston, New York.

Baird, J.C. en E. Noma, 1978. *Fundamentals of Scaling and Psychophysics.* Wiley, New York.

Barker, L.M., 1982. *The Psychobiology of Human Food Selection.* Avi Publishing Comp., Westport.

Ekman, G., 1959. Weber's Law and related functions. *The Journal of Psychology, 47,* 343-352.

Frijters, J.E.R., 1975. Sensory evaluation: A link between food research and food acceptance research. *Food Science & Technology, 8,* 294-297.

Frijters, J.E.R., 1976. Evaluation of a texture profile for cooked chicken breast meat by principal component analysis. *Poultry Science, 55,* 229-234.

Frijters, J.E.R. en E.L. Rasmussen-Conrad, 1982. Sensory Discrimination, intensity perception and affective judgement of sucrose-sweetness in the overweight. *The Journal of General Psychology, 107,* 223-247.

Gacula, M.C. en J.J. Kubala, 1978. Weighing coefficients for the estimation of sensory threshold. *Chemical Senses & Flavour, 3,* 105-121.

Green, D.M. en J.A. Swets (Eds.), 1966. *Signal Detection Theory and Psychophysics.* Wiley, New York.

Guilford, J.P., 1954. *Psychometric Methods.* McGraw Hill, New York.

Harper, R., 1972. Human Senses in Action. Churchill Livingstone, Edinburgh.

Harper, R., D.G. Land, N.M. Griffiths en E.C. Bate-Smith, 1968. Odour qualities: a glossary of usage. *British Journal of Psychology,* 59, 231-252.

Kroeze, J.H.A. (Ed.), 1979. *Preference Behaviour and Chemoreception.* I.R.L., London.

Linden, H. en I. Klingenberg, 1981. *Voedselkeuze: Het ontwikkelen van een meetinstrument.* Ongepubliceerde doctoraalscriptie, Vakgroep Humane Voeding, Landbouwhogeschool, Wageningen.

Moskowitz, H.R., 1976. The nature of acceptability functions. *Journal of Texture Studies, 7,* 235-242.

Patte, F., M. Etcheto en P. Laffort, 1975. Selected and standardized values of suprathreshold odor intensities for 110 substances. *Chemical Senses & Flavour, 1,* 283-305.

Schiffman, S.S., M.L. Reynolds en F.W. Young, 1981. *Introduction to Multidimensional Scaling.* Academic Press, New York.

Solms, J. en R.L. Hall (Eds.), 1981. *Criteria of Food Acceptance: How Man chooses what he eats.* Forster, Zürich.

Stevens, S.S., 1956. The direct estimation of sensory magnitudes – loudness. *American Journal of Psychology, 69,* 1-25.

Stevens, S.S., 1961. To honour Fechner and to repeal his law. *Science, 113,* 80-86.

Thurstone, L.L., 1927. A law of comparative judgment. *Psychological Review, 34,* 278-286.

Weber, E.H., 1846. Der Tastsinn und das Gemeinfühl. In: Wagner (Ed.) *Handwörterbuch der Physiologie,* deel 3. Vieweg, Braunschweig.

J.B. Amiabel volgde na een militaire carrière rond 1960 de Sociale Academie (Algemeen maatschappelijk werk). Was werkzaam als sociaal werker en personeels-chef. Heeft ruim 20 jaar ervaring als marktonderzoeker/wetenschappelijk medewerker. Was als stafmedewerker gedurende 5 jaar werkzaam bij het Instituut voor Psychologisch Markt- en Motievenonderzoek (IPM) en eveneens gedurende 5 jaar bij Interact B.V. Werkt gedurende de laatste 12 jaar als wetenschappelijk medewerker bij het Economisch Instituut voor het Midden- en Kleinbedrijf (E.I.M.). Huidige functie: plv. hoofd van de afdeling sociaal-economisch onderzoek.

J.W. van der Ben studeerde macro-economie en bedrijfseconomie aan de Nederlandse Economische Hogeschool te Rotterdam. Hij is werkzaam bij Philips Nederland waar hij verschillende functies vervuld heeft op het gebied van marktonderzoek en planning. Sinds 1980 is hij verantwoordelijk voor Strategische Beleidsvorming.

R.E. van den Berg studeerde enige jaren sociografie aan de Gemeentelijke Universiteit te Amsterdam was daarna 5 jaar werkzaam bij de N.V. v/h Nederlandse Stichting voor Statistiek. Hij is thans werkzaam als hoofd afdeling Marktonderzoek bij de Koninklijke Nederlandse Toeristenbond ANWB.

Jean-Louis Chandon is hoogleraar marketing aan het Institut d'Administration des Entreprises, dat deel uitmaakt van de universiteit van Aix-Marseilles in Frankrijk.

M.M.G. Fase studeerde aan de Universiteit van Amsterdam, waar hij in 1969 promoveerde op een proefschrift *An econometric model of age-income profiles.* Sinds 1964 is hij werkzaam geweest in Nederland en de V.S. Thans adjunct-onderdirecteur van de Nederlandsche Bank en hoofd van afdeling Wetenschappelijk Onderzoek en Econometrie, waar hij o.a. verantwoordelijk is voor de macro-economische modelbouw. Daarnaast is hij buitengewoon hoogleraar in de marktanalyse en de bedrijfsstatistiek aan de Erasmus universiteit Rotterdam.

Dick A. Francken (32 jaar) is in 1977 afgestudeerd in de economische en sociale psychologie aan de Katholieke Hogeschool in Tilburg. Hij is 2 jaar als wetenschappelijk medewerker verbonden geweest aan de vakgroep Economische psychologie van de K.H.T. en aan het N.R.I.T. in Breda. Sinds 1979 is hij werkzaam bij de Stichting Wetenschappelijk Onderzoek Konsumentenaangelegenheden (S.W.O.K.A.), te Den Haag. Binnen dit instituut houdt hij zich vooral bezig met onderzoek naar koopevaluaties door consumenten: produktwaardering, kooptevredenheid, consumentenklachten.

Jan E.R. Frijters studeerde aan de Rijksuniversiteit te Utrecht. In 1972 legde

hij het doctoraalexamen experimentele psychologie af met als subspecialisaties waarnemingspsychologie en methodologie. Gedurende de periode 1972 – 1978 was hij in dienst van de Directie Landbouwkundig Onderzoek van het Ministerie van Landbouw en Visserij alwaar hij was belast met het sensorisch en consumentengedragsonderzoek. Sind 1978 is hij als wetenschappelijk hoofdmedewerker verbonden aan de vakgroep Humane Voeding van de Landbouwhogeschool te Wageningen en verzorgt onderwijs en onderzoek naar eetgedrag en voedselkeuze. In 1980 is hij gepromoveerd op een proefschrift getiteld *Psychophysical and Psychometrical Models for the Triangular Method.*

H.H.F. Galama volgde in de eind-jaren 70 een universitaire opleiding economie in Groningen, alwaar hij in 1980 afstudeerde in de studierichting bedrijfseconomie. In hetzelfde jaar trad hij als projectleider in dienst bij het Instituut voor Toegepast Marktonderzoek (Intomart), alwaar hij thans de functie van senior-projectleider bekleedt, in welke functie hij onder andere nauw betrokken is bij de opzet en uitwerking van het in dit artikel beschreven onderzoeksinstrument.

Robert T. Green is hoogleraar marketing and international business aan de University of Texas at Austin. Hij is als gasthoogleraar verbonden geweest aan diverse Europese en Zuidamerikaanse universiteiten, en is lid van de Editorial Review Board van de Journal of Marketing.

Theo B.C. Poiesz studeerde van 1969 tot en met 1972 aan het Nederlands Wetenschappelijk Instituut voor Toerisme te Breda. Na het afstuderen in december 1972 startte hij de studie psychologie aan de Katholieke Hogeschool te Tilburg. Doctoraalexamen in mei 1977 met als hoofdvakken economische psychologie en sociale psychologie. Op basis van een Rotary Foundation Fellowship Award werkte hij gedurende het academische jaar 1977-1978 samen met Prof. dr. M.P. Friedman, Eastern Michigan University. Vanaf juli 1978 is hij als wetenschappelijk medewerker verbonden aan de subfaculteit Psychologie van de Katholieke Hogeschool Tilburg. Hij heeft een proefschrift geschreven over de relatie tussen aanbiedingsfrequentie en consumentenaffect waarop hij in oktober 1983 hoopt te promoveren.

C.J.M. Pooters studeerde sociologie aan de Katholieke Hogeschool te Tilburg. Sinds 1972 is hij werkzaam in het marktonderzoek, waarvan de laatste 5 jaar bij r + m, Research en Marketing b.v. te Heerlen. Hij bekleedt daar momenteel de functie van senior-projectleider.

L.J.M. Schreurs studeerde sociologie aan de Katholieke Universiteit te Nijmegen. Van 1974 tot medio 1975 was hij werkzaam bij de Raad der Europese Gemeenten, speciaal belast met onderzoek naar milieubesef en milieubeleid. Daarna was hij gedurende vijf jaar als coördinator van het sociaal-wetenschappelijk milieu-onderzoek verbonden aan het Instituut voor Milieuvraagstukken der Vrije Universiteit te Amsterdam. Sinds 1980 is hij directeur van het Instituut voor Longitudinaal Beleidsonderzoek (ILBO) te Amsterdam.

Bronislaw Verhage werd opgeleid aan Amerikaanse business schools. Hij werkte drie jaar voor een multinational, en doceerde marketing management aan de Graduate School of Business van de University of Texas in Austin. Sinds 1978 is hij verbonden aan de Interuniversitaire Interfaculteit Bedrijfskunde te Delft.

G.P. van der Vorst (1946) studeerde wiskunde aan de Rijksuniversiteit te Utrecht. Is thans in dienst bij de Rijksuniversiteit als Hoofd van het Centrum voor Data-Analyse, een instelling van de Faculteit der Sociale Wetenschappen, dat tot taak heeft het ondersteunen van de geautomatiseerde informatieverwerking (voornamelijk ten behoeve van onderwijs en onderzoek binnen de eigen faculteit).

Auteur van enkele kritische publikaties over de stand van zaken op computergebied, zoals 'De computer als blackbox' (Intermediair, december 1982).

Hans Wentholt (41 jaar), sociaal-psycholoog, is als onderzoeker werkzaam bij de NOS-Programmadienst Televisie. Zijn werkzaamheden als communicatiewetenschapper voor de NOS begon hij in 1966 voor de toenmalige afd. Studie en Onderzoek, later afd. Kijk- en Luisteronderzoek. Hij schreef een zeer groot aantal onderzoeksrapporten, waaronder bijvoorbeeld 'Televisie kritiek in Nederland' (1971), 'Arbeiders van Onderen' (1972), 'Meningen over (kern-)energie' (1980), 'Rayon 69' (1983). Voor de NOS-onderzoekafdeling zette hij al in de jaren 1968 – '70 een wekelijkse telefonische enquête op poten. Zijn huidige (parttime) functie bij de Programmadienst Televisie omvat naast onderzoek, ook meer beleidsmatige kanten. Zo is hij secretaris van de Redactie Advies Commissie van het Jeugdjournaal. Naast onderzoeker is hij nu ook al bijna zes jaar wethouder Welzijn in zijn woonplaats Hoevelaken.

Martin J.C. Zuurhout studeerde aan de HEAO te Utrecht in de commercieel-economische richting. Is op dit moment free-lance marktonderzoeker bij Centrum voor Marketing Informatie, Marketing Data. In 1984 hoopt hij zijn Master of Business Administration in management te behalen aan de Newport University Holland.

A. Zijdenbos, (1946), studeerde MO-A wiskunde. Trad in 1968 in dienst bij de N.V. v/h Nederlandse Stichting voor Statistiek (NSS); eerst als onderzoekassistent, later als projectleider.
Werkte van 1976-1977 als marktonderzoeker bij de Amro Bank. Keerde vervolgens weer terug naar de NSS alwaar hij thans hoofd van de afdeling Ad Hoc Onderzoek en lid van het management-team is.