1. Einde fusie-fobie in Nederland?

A.E. BRONNER

SAMENVATTING

In dit artikel wordt nader ingegaan op fusie-technieken. Omdat de gespreksduur van een enquete maar beperkt kan zijn en de onderzoeker vaak achteraf verbanden tussen vragen wil leggen die hij vooraf niet heeft bedacht, is het soms noodzakelijk om in plaats van tot single-source onderzoek de toevlucht te nemen tot fusie van data-bestanden. Fusie van data-bestanden betekent in feite van twee bestanden er één maken. Er wordt een persoon uit een steekproef gelicht en vervolgens wordt in de andere steekproef zijn of haar tweelingbroer/-zuster gezocht. Fusies worden in Nederland weinig uitgevoerd, dit in tegenstelling tot in ons omringende landen. Binnen SUMMO-verband fuseerde Bureau Veldkamp met behulp van de door dit bureau ontwikkelde VENUS-procedure Psyche '86 met de SUMMO-Scanner/Produktdeel '87. Afwegingen, controles, plus- en minpunten rond deze fusie worden aan de orde gesteld.

INLEIDING

Onderzoekers willen altijd veel zaken van mensen weten. En vervolgens wordt bekeken of er tussen al die zaken verbanden of relaties zijn, opdat zo meer naar voren komt over het hoe en waarom van het menselijk gedrag. Het meest ideaal is al die vragen te stellen aan respondenten op één tijdstip in één onderzoek: het zogeheten single-source onderzoek. Maar omdat mensen maar een beperkte cognitieve last aankunnen en de onderzoeker vaak achteraf verbanden tussen vragen wil leggen die hij vooraf niet heeft bedacht, is het soms noodzakelijk om in plaats van tot single-source onderzoek de toevlucht te nemen tot fusie van data-bestanden.

Fusie van data-bestanden betekent in feite van twee bestanden er één maken. We hebben bijvoorbeeld een steekproef van mensen gevraagd naar hun produktgebruik en een andere steekproef naar hun opinies en houdingen.
Er wordt een persoon uit het produktdeel gelicht en vervolgens wordt in de andere steekproef zijn of haar tweelingbroer/-zuster gezocht. De persoon uit het produktdeel krijgt dan de opinies en houdingen van die tweelingbroer/-zuster. Door dat zo bij iedereen uit te voeren ontstaat één steekproef met daarin zowel de antwoorden op de vragen over produktgebruik als op de opinies en houdingen. In technische termen heet iemand die gegevens van anderen krijgt toegedeeld een recipiënt (= ontvangt gegevens van tweelingbroer/-zuster) en degene die de gegevens levert een donor (= schenkt gegevens aan tweelingbroer/-zuster). De truc om tweelingen op te sporen heet het koppelingsmodel.

In dit artikel wordt een beschrijving gegeven van een recent uitgevoerde fusie. Con-

trole-procedures staan daarbij centraal. Bovendien wordt ingegaan op het feit dat fusies in Nederland zo weinig populair zijn (fusie-fobie?), dit in tegenstelling tot in veel andere landen.

1. ERVARINGEN MET FUSIES

Fusies worden in Nederland weinig uitgevoerd. In het media-onderzoek is dat voor het laatst in de 70-er jaren gebeurd en ook buiten het media-onderzoek is toepassing schaars (van Noordwijk, 1983).
In Nederland is altijd een soort fusie-fobie geweest, wellicht gevoed door het idee dat elk mens uniek is en dat daarom deze statistische 'manipulaties' ongeoorloofd zijn. Dit in tegenstelling tot een aantal andere landen zoals Frankrijk, West-Duitsland en de Verenigde Staten waar steeds meer van fusie-technieken gebruik wordt gemaakt. In West-Duitsland vond zelfs een soort fusie-wedstrijd plaats, waarbij verschillende bureaus een fusie op dezelfde data dienden uit te voeren (zie Wiegand, 1986, Scheler en Wiegand, 1987).
Wendt, één van de pioniers op dit terrein (zie bijvoorbeeld Wendt, 1976, Wendt, 1984) werd de overwinnaar. Dit leidde tot acceptatie van het 'partnership model', een model waarbinnen fusie plaatsvindt van gegevens over lezen (pers), kijken (tv) en luisteren (radio). Het werk van Antoine en Santini zorgde in Frankrijk voor een enorme opbloei van fusie-toepassingen (Antoine en Santini, 1986, Antoine, 1987, Antoine en Santini, 1987). Daarin staat centraal de koppeling van leesgegevens, radio/tv-gegevens en produktgebruik-gegevens. Voor wat betreft de Verenigde Staten kan bijvoorbeeld verwezen worden naar Frankel, 1982.
Geïnspireerd door deze nieuwe ontwikkelingen heeft SUMMO in 1987 een criteria-lijst voor fusie met de SUMMO-Scanner opgesteld. Eind 1987 voerde Bureau Veldkamp met behulp van het VENUS-model (Veldkamp's Nonresponse en Uitval Simulatieprocedure) de eerste door SUMMO geaccepteerde fusie uit. Het betrof de fusie van het SUMMO Nationaal Produkten en Interessenonderzoek/SUMMO-Scanner 1987 en Psyche 1986. De fusie werd uitgevoerd in opdracht van Admedia. Uiteraard beweert geen van alle genoemde deskundigen dat fusie beter is dan single-source meting, maar wel dat een goed uitgevoerde fusie beter is dan allerlei face-value data-verbindingen die anders toch in het hoofd van de mediaplanners en andere gebruikers van onderzoek plaats vinden.
Het goed uitvoeren van een fusie vereist uiteraard een adequaat koppelingsmodel, maar misschien nog belangrijker is een uitgebreid scala aan controleprocedures. Op beide aspecten gaan wij in de volgende paragraaf in.

2. HET VENUS-MODEL

Het VENUS-model is in 1982 ontwikkeld en aanvankelijk benut voor de simulatie van missing-data (Psyche 1982, Bronner 1983 (a), 1983 (b)), voor injectie ('the word injection should be used for situations where one wants to add information within a given file', dit in tegenstelling tot fusie, 'the word fusion should be used when one wants to build a complete file by putting together two files which do not contain the same information'). Het model is voor de eerste keer voor een fusie gebruikt bij de koppeling Psyche 1986 en Scanner/Produktdeel 1987.

Om een goede fusie te kunnen uitvoeren moet er een groot aantal gemeenschappelijke kenmerken zijn, dat wil zeggen vragen die op dezelfde manier aan beide groepen respondenten in beide bestanden zijn gesteld (in fusie-terminologie: het donor-bestand en recipiënten-bestand). Zonder een groot aantal gemeenschappelijke kenmerken (tenminste een 50 à 100) is een adequate fusie onmogelijk. Deze gemeenschappelijke kenmerken worden in twee groepen onderverdeeld:

1. koppelingskenmerken, dat zijn de kenmerken op grond waarvan bekeken wordt of iemand een geschikte tweelingbroer of -zus is. Via multivariate analyses wordt bepaald of iets een goed koppelingskenmerk is.
2. controlekenmerken, dat zijn kenmerken die niet bij de feitelijke koppeling benut worden, maar die na de uitvoering van de fusie dienst doen als criteria waarop je bekijkt of donor en recipiënt niet al te veel van elkaar afwijken.
   Is de afwijking te groot dan is de fusie niet geslaagd, is de afwijking klein dan is het blijkbaar goed gegaan.

Deze twee typen kenmerken worden in verschillende fasen van de procedure benut:

a. selectiefase
De koppelingskenmerken worden geselecteerd op basis van afhankelijkheid tussen te simuleren of fuseren gegevens en andere gegevens (in ons geval: antwoorden op vragen uit het Scanner- en Produktenonderzoek enerzijds en Psyche-variabelen anderzijds). Hiertoe wordt een groot aantal regressie-analyses uitgevoerd.

b. koppelingsfase
Als koppelingsmodel is gekozen voor het additief verschil model, city block metriek, ongewogen. Dit model wordt veel benut in meerdimensionale schaling en multi-attribuut keuze-modellen (Bronner, de Hoog, 1978). Conceptuele duidelijkheid, goede ervaringen, flexibiliteit en mathematische eigenschappen bepaalden de keuze. Aan elke recipiënt wordt die donor gekoppeld die over een aantal kenmerken gerekend de kleinste afstand tot de recipiënt heeft. Eén persoon kan meer keren als donor optreden. Bij gelijke afstand is een random-selectieprocedure ingebouwd. Bij de verschillende onderzoekers ook in het buitenland is er verschil van mening hoeveel keer iemand als donor benut mag worden. Onze analyses rechtvaardigen in onze ogen het uitgangspunt dat in principe geen maximum wordt opgelegd (in een situatie waarbij het aantal donoren ongeveer gelijk is aan het aantal recipiënten).

c. segmentatie
Toepassing van het afstandsmodel geschiedt in ons model per segment, waarbij de segmenten zijn opgebouwd op basis van een combinatie van geslacht, leeftijd en opleiding. Men zou in modeltermen kunnen zeggen dat deze drie criteria een zodanig gewicht krijgen dat geen compensatie door waarden op andere kenmerken meer mogelijk is. Een man kan nooit aan een vrouw gekoppeld worden, een jong iemand nooit aan een oud iemand, een laag opgeleide nooit aan een hoog opgeleide. Per segment kunnen verschillende koppelingskenmerken in het afstandsmodel opgenomen worden.

d. controlefase
Controle dient zowel op geaggregeerd niveau als op individueel niveau plaats te vinden. De resultaten van verschillende controleprocedures dienen gecombineerd te wor-

den. Vooral de controle op individueel niveau ontbreekt nogal eens in de in andere landen toegepaste modellen.
Als wij onze procedure vergelijken met andere internationaal geaccepteerde modellen, dan menen wij te kunnen stellen dat door ons veel meer controleprocedures worden uitgevoerd. We kunnen niet de pretentie hebben dat ons koppelingsmodel als zodanig beter is, maar wel dat een veel uitgebreider controle-scala is ingebouwd. Wat leren al die controles? In de eerste plaats dat het weinig zinvol is alleen naar de verdelingen van de scores over de variabelen te kijken (zoals zo vaak in andere landen gebeurt), maar dat het veel relevanter is op individueel niveau de scores van donor en recipiënt te vergelijken. Niet alleen: over alle Jansens (= donoren) gerekend drinkt 40% regelmatig bier, over alle Pietersens (= recipiënten) 38%. Maar ook: Jansen is gekoppeld aan Pietersen, hoe verschillend zijn ze bijvoorbeeld wat betreft roken, drinken, uitgaan?

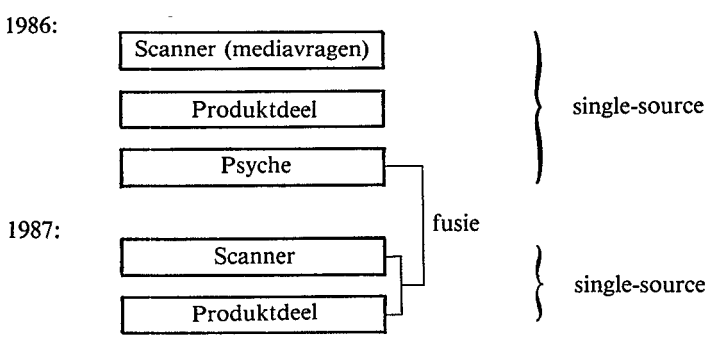
Bovenstaande abstract beschreven stappen in het model lichten wij in de volgende paragraaf toe aan de hand van een concrete fusie.

3. UITVOERING VAN EEN FUSIE

3.1. De beschikbare bestanden

Jaarlijks worden via de SUMMO-Scanner zo'n 39.000 respondenten telefonisch ondervraagd. In 1986 werd een gedeelte van deze groep respondenten schriftelijk benaderd voor het Produkten- en Interessenonderzoek en Psyche.

Aan het Produktenonderzoek 1986 deden 7.176 respondenten mee (deze respondenten beantwoordden dus ook via de SUMMO-Scanner mediavragen). Van deze groep bleek 83% bereid ook aan Psyche 1986 deel te nemen. In totaal kon in 1986 dus van 1.189 respondenten geen informatie worden verzameld wat betreft Psyche 1986.
In 1987 is wederom het Produkten- en Interessenonderzoek uitgevoerd, nu bij 7.551 respondenten. In 1987 is geen afzonderlijk Psyche-onderzoek uitgevoerd. Omdat de wens leefde om bij de beschikbare psychografische data zo recent mogelijke produktgegevens te hebben, is aan Veldkamp de opdracht verleend een fusie uit te voeren tussen Produktdeel 1987 en Psyche 1986. Daartoe is ook de VENUS-methode benut.
Een overzicht wordt gegeven in schema 1.



Schema 1: Overzicht bestanden.

3.2. Bepaling van de koppelingskenmerken

Het selecteren van de juiste koppelingskenmerken is de eerste stap. Kernpunt is dat koppelingsvariabelen worden gezocht in het Scanner- en Produktengedeelte die de sterkste samenhangen vertonen met de Psyche-variabelen en die vervolgens worden gebruikt als voorspellers van Psyche-variabelen bij personen bij wie alleen Scanner- en produktvragen ter beschikking staan. In het Produktdeel zijn ook vragen over interessen en activiteiten opgenomen en die kunnen uiteraard uitstekend als koppelingskenmerk optreden. Dat zijn in feite al vragen die in de psychografische sfeer liggen. Het zoeken naar koppelingsvariabelen kunnen we doen door in het 1986-bestand (dat wel single-source was) een aantal psychografische gegevens als afhankelijke variabelen te nemen en produkt/mediagegevens als onafhankelijke variabelen. Door regressie-analyses toe te passen verkrijgen wij inzicht in de samenhangen. Deze regressie-analyses worden per segment uitgevoerd. Er wordt daarbij gewerkt met acht segmenten, gevormd door geslacht, leeftijd (in twee klassen) en opleiding (eveneens in twee klassen). Voor het in aanmerking komen als koppelingskenmerk is het nodig dat sprake is van meer dan één significante beta-coëfficiënt. Per segment worden dan 20 à 30 koppelingsvariabelen geselecteerd, die gebruikt worden om donoren te vinden met de geringste afstand tot de recipiënten. In totaal worden over alle segmenten heen 68 verschillende koppelingsvariabelen benut.
De koppelingsvariabelen die bij minstens vijf van de acht segmenten voorkomen betreffen: het lezen van Privé, interesse levenslied, interesse bijzondere gerechten, frequentie tuinieren, frequentie bezoeken sportwedstrijden, frequentie trimmen, frequentie fotografen, frequentie treinreizen, prijsbewustheid, wel of niet roken, frequentie kerkbezoek. De koppelingskenmerken vertegenwoordigen een breed spectrum aan variabelen (interessen, activiteiten, consumptiegedrag, media-consumptie, socio-demografische gegevens).

3.3. Toepassing koppelingsmodel

Na de selectie van de koppelingskenmerken volgt de afstandsbepaling tussen donoren en recipiënten. Bij de afstandsbepaling krijgen alle geselecteerde variabelen per segment een gelijk gewicht en een gelijke lengte. Elke donor en recipiënt met de kleinste afstand tot elkaar worden aan elkaar gekoppeld.

De respondenten uit het Produktendeel 1987 zijn de recipiënten : n = 7.551.
De Psyche 1986-respondenten zijn de donoren: n = 5.987.
Theoretisch zou elke donor dus gemiddeld $\frac{7.551}{5.987} = 1.26$ keer benut dienen te worden.
Omdat het aantal donoren feitelijk kleiner is dan het aantal recipiënten kan het niet anders dan dat een aantal donoren meer dan één keer wordt gebruikt. Maar ons model laat ook toe dat een donor meer keren benut kan worden. De kleinste afstand tussen donor en recipiënt is in ons model doorslaggevend voor de koppeling en bij elke recipiënt kan er weer uit het gehele reservoir aan donoren worden geput (in feite vergelijkbaar met een steekproef met teruglegging).

Van alle donoren blijkt 59% gebruikt te worden (n = 3.526), zie voor een overzicht tabel 1.

Tabel 1. Gebruik donoren.

	(100% = 5.987)
– donoren 0 x gebruikt	41%
– donoren 1 x gebruikt	27%
– donoren 2 x gebruikt	16%
– donoren 3 x gebruikt	8%
– donoren 4 x of meer gebruikt	9%

Als we naar de uiteindelijke afstanden tussen donoren en recipiënten kijken blijkt dat het vooral bij de oudere, laag opgeleide groepen gemakkelijk is een tweelingbroer of - zuster te vinden. De meeste moeite is dat bij de oudere, hoog opgeleide vrouwen en de jongere mannen groepen.
Na uitvoering van de fusie en overbrenging van het psychografisch gedeelte van donor naar recipiënt volgt de belangrijke controlefase. Hierop gaan wij in de volgende paragraaf in.

4. CONTROLE VAN EEN FUSIE

4.1. Ccntroleprocedures

Wij hebben een aantal verschillende controles gebruikt die naar onze mening een breder beeld geven van de geslaagdheid of mislukking van de fusie dan de nu in het buitenland geldende fusie-procedures (zie bijvoorbeeld de artikelen in Readership Research: theory and practice, H. Henry (ed.), Amsterdam, 1987):

1. vergelijken van de rechte tellingen tussen Psyche 1986 en Psyche 1987;
2. de koppelingskenmerken vormen een deelverzameling van de gemeenschappelijke kenmerken van donor- en recipiëntenbestand. Uit de overige gemeenschappelijke variabelen kan een deelverzameling controlekenmerken worden gehaald.
   De controlekenmerken dienen zoveel mogelijk een representatie te zijn van de psychografische variabelen. Op de controlekenmerken dienen de scores van donoren en recipiënten vergeleken te worden;
3. relaties binnen het 1986 bestand vergelijken met relaties binnen het 1987 bestand;
4. schaalconsistentie vaststellen bij vergelijking van de schaalanalyseresultaten 1986 en 1987.

4.2. Via controlekenmerken

In dit artikel concentreren wij ons vooral op de tweede controlemethode, omdat ervaring in Nederland en andere landen leert dat de eerste controlemethode vrijwel geen verschillen oplevert.
In het Scanner- en Produktengedeelte komt een aantal vragen voor, die zowel in 1986 als in 1987 zijn gesteld, die niet door ons werden benut als koppelingsvariabelen en verwantschap tonen met Psyche-variabelen.
We kunnen deze controlekenmerken vergelijken op individueel niveau, waarbij we de data van donoren rechtstreeks koppelen aan die van de recipiënten. We kunnen veron-

derstellen dat de informatie op deze punten van persoon x lijkt op die van persoon y, met wie de koppeling heeft plaatsgevonden. Daarbij moet wel opgemerkt worden dat deze gelijkenis minder zal zijn dan op de Psyche-variabelen, omdat we immers deze variabelen bij de analyse hebben afgekeurd als koppelingsvariabelen. In totaal zijn 68 kenmerken in deze analyse betrokken (er zijn dus 68 koppelingsvariabelen en 68 controlevariabelen).

Eenvoudig gesteld: mijnheer Jansen (donor) wordt gekoppeld aan Pietersen (recipiënt). Pietersen krijgt de psychografische gegevens van Jansen toegedaan. Ze worden beschouwd in ons model als tweelingbroers. Maar hoe zit het met hun scores op andere kenmerken die dus voor beide personen bekend dienen te zijn. Roken ze allebei? Wat drinken ze? Wat zijn hun interesses? De volgende controlekenmerken zijn gebruikt:

* *Interesse-kenmerken*: interesse in opvoedkunde, ontwikkelingshulp, menselijk lichaam, natuurgeneeswijzen, politiek (Nederlands en Europees), energiebronnen, ruimtevaart, gezond eten, lijnen, samenlevingsvormen, schepen, monumentenzorg, vreemde landen en volkeren, inkomstenbelasting, schoonheidstips, antiek, planten en bloemen, man/vrouw verhouding, bejaardenzorg, moderne kunst, medische wetenschap, psychologie, astrologie, science fiction, woonideeën, effecten, goud en diamanten, beleggen in onroerend goed, munten sparen, postzegels sparen, blues en rhythm en blues, soul, cabaret, advertenties, mode, onderwerpen in krant, natuurfilms op televisie, sport op televisie, hypotheken;
* *Activiteiten-kenmerken*: voetballen, tennissen, hockey, zwemmen, zeilen, wandelen, gezelschapsspellen doen, bioscoop bezoeken, muziekinstrument bespelen, klaarmaken buitenlandse gerechten, volgen cursussen, zelf kleren maken, schouwburg bezoeken, bibliotheek bezoeken, boeken kopen, winkelen, stripboeken lezen, reizen met de bus;
* *Attituden*: prijsbewustheid;
* *Media*: Telegraaf, Volkskrant, NCRV-gids, TROS-Kompas, Veronica, VPRO, Elseviers Magazine, Viva, Het Beste, Vorsten.

De kenmerken betreffen een brede variëteit en vertonen een sterke verwantschap met de te fuseren psychografische gegevens. De toets via controlevariabelen is de belangrijkste toets (zie ook het discussiegedeelte over fusie blz. 370-376 uit het reeds geciteerde boek van H. Henry, Readership Reseach: theory and practice, Amsterdam, 1987, bijvoorbeeld V. Appel: 'thought that what Scheler had shown on the screen about the accuracy of the fusion procedure was quite trivial. It was no great achievement to duplicate the marginals, nor was it particularly surprising that the data could be duplicated by using linkages within those people characterised by the linkage variables: the real question was whether the data held up when cross-tabulated by non-linkage variables'). Bij onze controlevariabelen gaat het nu juist om dit type non-linkage variables.

Van de 68 relaties die we op deze manier bestudeerden bleken er slechts twee niet significant te zijn ($p < 0.001$). Een z-waarde (berekend op basis van tau-b, met correctie voor ties) van 3.19 is voldoende.
Bij 47 van de 68 relaties is zelfs sprake van een z-waarde die boven de 10 ligt. Maar gezien de grote n-aantallen moet daarbij gedacht worden aan correlaties boven .10.

Behalve in termen van correlaties kunnen de relaties tussen donoren en recipiënten wat

betreft de controlevariabelen ook nog strenger, namelijk per schaalpositie bekeken worden:

* bij 16 controlevariabelen zijn de scores van donor en recipiënt in 85%-95% van de gevallen gelijk of bijna gelijk (= één positie verschil) aan elkaar;
* bij 32 controlevariabelen zijn de scores van donor en recipiënt in 75%-85% van de gevallen gelijk of bijna gelijk aan elkaar;
* bij 11 controlevariabelen zijn de scores van donor en recipiënt in 65%-75% van de gevallen gelijk of bijna gelijk aan elkaar;
* bij 9 controlevariablen zijn de scores van donor en recipiënt in minder dan 65% van de gevallen gelijk of bijna gelijk aan elkaar. Dit blijken telkens min of meer algemene gedragingen te zijn die blijkbaar niet zozeer met een bepaalde psychografische levensstijl samenhangen (tot deze 9 behoren: wandelen, zwemmen, gezelschapsspelen, winkelen, bibliotheek bezoeken, Telegraaf lezen, buitenlandse gerechten klaarmaken).

4.3. Via vergelijken relatiepatronen

De derde in het begin van deze paragraaf genoemde controleprocedure betreft het vergelijken van relaties. Wanneer de relaties tussen geslacht, leeftijd en opleiding enerzijds en psychografische criteria anderzijds vergeleken worden, blijkt de overeenkomst in relatiepatronen en hoogte van de correlaties tussen 1986 (single-source) en 1987 (gefuseerd) zeer groot.
Voorts is nog nagegaan of een aantal hoge correlaties (.10 of hoger) binnen het 1986 materiaal ook na de fusie binnen het 1987 bestand hoog gebleven is.

Bijvoorbeeld: kookliefhebberij x interesse bijzondere gerechten
belang gezonde voeding x gezond eten
belang vlees x vegetarisch eten
houding roken x roken
interesse populaire sporten x kijken naar sport op televisie

Het gaat telkens om een Psyche-schaal en elementen uit het produktengedeelte. In totaal zijn op deze wijze 40 relaties nader onderzocht; de sterkte van de correlaties blijkt minder te worden, maar vrijwel alle correlaties blijven significant.

4.4. Via vergelijken schaalresultaten

Na de fusie is op het 1987 bestand een aantal schaalanalyses uitgevoerd om te bezien of de items die in 1986 deel uitmaakten van een schaal ook nu nog tot dezelfde schaal behoren. Evenals in 1986 is van de Mokken-techniek gebruik gemaakt. De H-waarde geeft aan of we met een sterke, matige of zwakke schaal te maken hebben.
De schalen zijn willekeurig gekozen met dien verstande dat verschillende sectoren er in vertegenwoordigd zijn.
De H-waarden blijven op vrijwel hetzelfde niveau liggen (zie tabel 2).
Geconcludeerd kan worden dat de fusie geen nadelig effect heeft gehad op de onderlinge schaalbaarheid van de items.

Tabel 2. Overzicht H-waarden in 1986-bestand en 1987-bestand.

schaal		1986 H-waarden	1987 H-waarden (na fusie)
1	kookliefhebberij	.55	.55
6	leuk huishoudelijk werk	.49	.55
9	merkentrouw	.44	.48
18	plezier in winkelen	.71	.70
29	interesse in mode	.54	.57
31	mode volgen	.68	.71
41	behoudendheid	.43	,46
44	rel. milieuproblemen	.50	.48
48-a	technol. ontwikkelingen	.48	.47
56	aandacht cosmetica	.59	.63
62	sociale vaardigheid	.42	.42
69	stress	.44	.46
78	politiek maatsch. interesse	.54	,53
80	culturele interesse	.41	,36
86	gastronomische interesse	.61	.67
90	interesse klassieke muziek	.50	.61

5. CONCLUSIES

In de paragraaf over ervaringen met fusies schreven wij dat fusies in Nederland weinig worden uitgevoerd, dit in tegenstelling tot in ons omringende landen. De in 1987 opgestelde criterialijst voor fusie met de SUMMO-Scanner biedt weer een opening tot uitvoering van dit soort data-koppelingen. Bureau Veldkamp fuseerde in opdracht van Admedia Psyche 1986 met de SUMMO-Scanner/Produktdeel 1987. Het feit dat weer eens een kans tot benutting van fusie-technieken werd geboden is heel belangrijk omdat fusie-verbetering vaak een kwestie is van cumulatieve opbouw van kennis. Het VENUS-model bleek een fusie op te leveren die acceptabel was en door de uitgebreide controles bovendien veel leerde over positieve en negatieve punten.

Van de 68 controlerelaties die we bestudeerden bleken er slechts twee niet significant te zijn. Bij 47 van de 68 relaties is zelfs sprake van een zeer sterke relatie. Behalve in termen van correlaties kunnen de relaties tussen donoren en recipiënten wat betreft de controle-variabelen ook per schaalpositie bekeken worden.
De koppeling blijkt dan het slechtst te lukken bij min of meer algemene gedragingen die blijkbaar niet zozeer met een psychografisch bepaalde levensstijl samenhangen: wandelen, zwemmen, gezelschapsspelen, winkelen, bibliotheek bezoeken, buitenlandse gerechten klaarmaken.

Controles op het gebied van hoogte van correlaties wijzen uit dat vrijwel altijd relaties die in het oorspronkelijke bestand significant zijn, dat na fusie ook blijven en dat niet

significante relaties nooit significant worden. Wel zakt vaak de hoogte van correlaties ('indikkingseffect'). Schaalanalyse voor en na fusie leert dat de fusie geen nadelig effect heeft gehad op de onderlinge schaalbaarheid van de items.

Op basis van de uitvoerige controles kan geconcludeerd worden dat de fusie geslaagd is, wel komen de correlaties op een wat lager niveau te liggen. Deze ervaring leert dat waar single-source meting niet mogelijk of niet betaalbaar is, fusie een redelijk tot goed substituut kan bieden. Overigens wordt beoordeling van de geslaagdheid van een fusie in de toekomst wat gemakkelijker omdat we nu met de onderhavige studie een aantal 'benchmarks' hebben neergezet.

LITERATUUR

Antoine, J. (1987), A case study illustrating the objectives and perspectives of fusion techniques, in: H. Henry (ed.), *Readership Research: theory and practice*, Amsterdam, 1987, blz. 336-351.

Antoine, J., G. Santini (1986), *An experiment to validate fusioned files obtained by the referential factorial method*, ESOMAR Seminar of New Developments in Media Research, Helsinki, April 1986.

Antoine, J., G. Santini (1987), Fusion techniques: alternative to single-source methods, *European Research*, 15,3, blz. 178-187.

Bronner, A.E. (1983a), *Improving panel research by simulating missing values*, paper presented at the 14th Europanel Forum, Deauville, June 1983.

Bronner, A.E. (1983b), A way to handle non-response: simulation of missing values, *European Research*, 11,3, blz. 86-92.

Bronner, A.E., R. de Hoog (1978), *Politieke voorkeur: oordelen en beslissen*, doct. dissertatie, Universiteit van Amsterdam.

Frankel, M.R. (1982), Ascription in magazine audience research, in: H. Henry (ed.), *Readership Research: theory and practice*, Amsterdam, 1982, blz. 373-378.

Henry, H. (1987), *Readership Research: theory and practice*, North-Holland, Amsterdam

Noordwijk, A.J. van (1983), Koppeling van databestanden, Sociaal en Cultureel Planbureau, *Stukwerk* nr. 10.

Psyche 1982: Technische Toelichting, Admedia B.V., Amsterdam

Scheler, H.E., J. Wiegand (1987), A report on experiments in fusion in the 'official' German media research, in: H. Henry (ed.), *Readership Research: theory and practice*, Amsterdam, 1987, blz. 352-360.

Veldkamp/Marktonderzoek B.V. (1988), *Technische verantwoording fusie Produktdeel '87-Psyche '86*, Amsterdam.

Wendt, F. (1976), *Beschreibung einer Fusion*, Schriftenreihe Band 21, Gruner & Jahr, Hamburg.

Wendt, F. (1984), The AG.MA model, in: H. Henry (ed.), *Readership Research*: Montreal 1983, Amsterdam, 1984, blz. 393-403.

Wiegand, J. (1986), *The combining of two separately derived data-sets into an integrated intermedia planning system: the German 'model of partnership'*, ESOMAR Seminar on New Developments in Media Research, Helsinki.

2. Computergestuurd conjunct meten

M.A. BRADLEY, E.P. KROES EN B.J. VOGELS

SAMENVATTING

In dit artikel worden de ontwikkelingen van computergestuurd conjunct meten geschetst. Achtereenvolgens wordt ingegaan op conjunct meten en computergestuurd interviewen in het algemeen waarna een vergelijking wordt gemaakt tussen twee bestaande pakketten, te weten ACA en MINT. Voor dit laatste programma worden de diverse veldwerkmethoden belicht, en wordt via een praktijkvoorbeeld nader ingegaan op een computergestuurde conjuncte meting. Tenslotte wordt nog kort ingegaan op de toekomstverwachtingen.

1. INLEIDING

Het is alweer zo'n 10 jaar geleden dat het conjunct meten in de Verenigde Staten gemeengoed werd (Green en Srinivasan 1978) en ook in Nederland werd geïntroduceerd (bijv. Van Raaij 1978). Sindsdien is conjunct meten tot de standaarduitrusting van de marktonderzoeker gaan behoren, hetgeen wordt geïllustreerd door het zeer grote aantal commerciële toepassingen dat vooral in de Verenigde Staten werd uitgevoerd (zie Cattin en Wittink 1982, Cattin en Wittink 1986).

In die tien jaar werd veel onderzoek verricht naar de betrouwbaarheid (bijv. Mc Cullough en Best 1979, Segal 1982, Leigh *et al.* 1984) en de validiteit (bijv. Green en Srinivasan 1978, Levin *et al.* 1982) van conjunct meten, vrijwel steeds met positief resultaat. Was het dan alleen maar rozegeur en maneschijn met conjunct meten? Nee, er werden ook problemen geconstateerd die de bredere toepassing van de methode in de weg stonden, vooral op het gebied van de praktische toepasbaarheid ervan. De methode werd als moeilijk ervaren, niet alleen door de onderzoekers zelf maar vooral ook door de respondenten. Voor velen was de gevraagde taak, het rangordenen van 9 tot soms zelfs 27 of meer profielen, te moeilijk. En ook waren de omschrijvingen van de profielen veelal onvoldoende duidelijk of relevant om een probleemloze interpretatie mogelijk te maken. Inmiddels zijn er oplossingen gezocht en gevonden voor deze problemen (bijv. Kroes 1986), die tevens mogelijk gemaakt werden door de ontwikkelingen die plaatsvonden in het computergestuurd enquêteren.

Tijdens diezelfde periode maakte de computer immers een sterke opmars in het marktonderzoek. Eerst werden de grotere computersystemen in het telefonisch onderzoekwerk (CATI) betrokken, vervolgens kwam de PC erbij. En sinds enkele jaren maakt ook het gebruik van PC's in face-to-face enquêtes (CAPI) een sterke groei door (bijv. Baden *et al.* 1987). Deze ontwikkeling leidde ertoe, dat de PC werd ingeschakeld als hulpmiddel bij het uitvoeren van conjuncte metingen. In eerste instantie alleen om de enquê-

teurs te helpen (zie bijv. Bradley 1988), maar spoedig ook als compleet instrument, waarbij de volledige enquêtetaak werd overgenomen. Door de ingebouwde rekenkracht van de PC werden nieuwe manieren van enquêteren mogelijk, sterk op de individuele respondent afgestemd, en met een interactief karakter. Hiermee bedoelen wij dat de computer tijdens de enquête de door de respondent gegeven informatie analyseert en op basis hiervan tot nieuwe, meer informatieve en relevante vragen komt.

In dit artikel bespreken wij de mogelijkheden die het moderne, computergestuurde conjunct meten aan de marktonderzoeker te bieden heeft. Daartoe geven wij eerst een kort overzicht van de ontwikkelingen die in het computergestuurd enquêteren zijn opgetreden. Vervolgens bespreken wij twee systemen die zijn ontwikkeld om computergestuurde conjuncte metingen uit te voeren: ACA en MINT. Beide systemen worden vergeleken, en vervolgens bespreken wij als voorbeeld een van de studies die wij met het MINT systeem hebben uitgevoerd. Wij ronden dit artikel af met een korte schets van de voor de nabije toekomst verwachte ontwikkelingen op het gebied van het computergestuurd enquêteren en conjunct meten in het bijzonder.

2. ONTWIKKELINGEN IN COMPUTERGESTUURD ENQUÊTEREN

In de jaren zeventig werd CATI (Computer-Assisted-Telephone-Interviewing) ontwikkeld, beproefd en toegepast in de Verenigde Staten. Pas in 1980 werd dit systeem in Nederland geïntroduceerd. Aanvankelijk voor grotere computersystemen, maar met de opkomst van de PC verscheen aangepaste programmatuur op de markt. De beginperiode verliep niet zonder kinderziektes. De programma's liepen in het begin te traag en het ontwerp van de vragenlijsten stelde bijzondere eisen waaraan te weinig aandacht was besteed.

De ontwikkeling van CAPI (Computer-Assisted-Personal-Interviewing) was de volgende stap in het computergestuurd enquêteren. De eerste toepassing hiervan was zelfs vóór de doorbraak van de micro-computers. In 1979 deden Bronner en de Hoog al experimenten met een interactieve computergestuurde vragenlijst. In het begin van de tachtiger jaren resulteerde dit in de eerste Nederlandse CAPI-toepassingen, zij het op kleine schaal en niet met behulp van portables. Van deze onderzoeken die Veldkamp in samenwerking met de Universiteit van Amsterdam uitvoerde, werd reeds in mei 1982 verslag gedaan op het Sociometrie Congres, Vrije Universiteit te Amsterdam. Bronner hield daar een lezing over CAPI-experimenten en de Hond en Groves over hun eerste CATI-ervaringen.

De opkomst van de portables maakte een toepassing van CAPI op grotere schaal mogelijk, aanvankelijk in Engeland (1983). Een jaar later begon het CBS met een CAPI-experiment, waarbij een groot aantal respondenten betrokken waren. Op basis hiervan werd in 1987 de Enquête Beroepsbevolking definitief gestart. In hetzelfde jaar kwamen onderzoeksbureaus waaronder Nipo en Interview met hun eigen programmatuur. Er kwamen ook standaardpakketten op de markt, zoals Ci2 van Sawtooth Software.

Een verdere ontwikkeling in de lijn van CATI en CAPI is het computergestuurd groepsinterview. Het bij Veldkamp (AGB) ontwikkelde DELPHISIE programma (Vogels, 1987) is

in die zin een CAGI-systeem (Computer-Assisted-Group-Interviewing) te noemen. Dit programma heeft als belangrijkste eigenschappen:
– Snelle inventarisatie – anoniem – van de individuele opinies van maximaal 10 respondenten door middel van druktoetseenheden, die via speciaal ontwikkelde randapparatuur aan een (portable) PC kunnen worden gekoppeld;
– Directe terugkoppeling – visueel – van de zojuist gegevens individuele opinies, waardoor ter plekke iedere respondent met de mening van de overige respondenten wordt geconfronteerd;
– Open communicatie – gestructureerd – met als uitgangspunt de verschillen in mening ten aanzien van het onderwerp dat aan de orde is en waar ieder individueel oordeel in doorklinkt;
– Flexibel instrumentarium – ad hoc – met betrekking tot zowel de vraagstelling als de presentatie;
– Overal inzetbaar – portable – waardoor het instrument op iedere gewenste lokatie gehanteerd kan worden.

Dit programma is de eerste Nederlandse toepassing van een computergestuurd groepsinterview in een onderzoekssituatie. Het interactieve element dat bij CATI en CAPI zo een grote rol speelt, is ook bij instrumenten als DELPHISIE van groot belang.
Uiteraard zijn ook de andere bekende eigenschappen zoals nauwkeurigheid, snelheid en toegesneden vraagstelling bij CAGI van toepassing. Door middel van computergestuurd enquêteren met behulp van CATI/CAPI/CAGI-systemen gaat de gegevensverzameling sneller en nauwkeuriger, en dat is in het belang van zowel onderzoeker als respondent.

3. COMPUTERGESTUURD CONJUNCT METEN

Voor zover ons bekend zijn er momenteel in Nederland twee systemen beschikbaar die het mogelijk maken om computergestuurde conjuncte metingen uit te voeren: ACA en MINT. Beide programma's zijn ongeveer gelijktijdig (en onafhankelijk van elkaar) ontwikkeld, ACA in de Verenigde Staten en MINT in Nederland. Wij gaan hieronder kort op de twee programma's in.

ACA staat voor Adaptive Conjoint Analysis, en is ontwikkeld in de Verenigde Staten door Sawtooth Inc (Johnson 1985). Het is een op zichzelfstaand 'standaard' programma, dat aansluit op andere door hetzelfde bedrijf ontwikkelde programmatuur voor toegepast marktonderzoek. De conjuncte meting verloopt via een standaardprotocol, waarbij door de respondent een vaste reeks stappen moet worden doorlopen:
1. Selectie van de attributen
2. Selectie van de niveaus van de attributen
3. Beoordeling van een reeks produktprofielen
4. Aangeven van de koopintentie voor enkele produktprofielen

In principe worden door de onderzoeker alle relevante attributen van het onderzochte produkt gedefiniëerd, maar elke respondent kiest daaruit een beperkte deelgroep, bijvoorbeeld de vier attributen die voor hem/haar het belangrijkst zijn. Vervolgens geeft de respondent voor elk van de attributen aan wat het acceptabele bereik van de niveaus van elk attribuut is, en wat de geprefereerde volgorde van deze niveaus is. Op basis van

de geselecteerde attributen en de als acceptabel aangemerkte niveaus genereert ACA daarna een aantal profielen die de respondent ter beoordeling worden voorgelegd. Dit gebeurt door telkens twee profielen gelijktijdig op het beeldscherm van de computer te laten verschijnen. De respondent dient op een preferentieschaal, veelal een 9-puntsschaal, aan te geven naar welk alternatief zijn/haar voorkeur uitgaat, en hoe sterk deze is. Uit de gegeven antwoorden leidt het programma door middel van een kleinste kwadraten schatting (vergelijk regressie) af wat de utiliteiten zijn die bij de diverse attribuutniveaus horen. Vervolgens kijkt het programma welke aanvullende informatie nodig is om de schatting van de utiliteiten te verbeteren, en genereert twee nieuwe profielen die weer beoordeeld moeten worden. Deze procedure herhaalt zich tot alle utiliteiten met een redelijke nauwkeurigheid kunnen worden geschat. In de laatste stap worden aan de respondenten drie profielen opnieuw ter beoordeling voorgelegd, waarbij nu echter de respondent een koopkans aan moet geven, op een schaal van 0 tot 100%. Vervolgens leidt het programma, weer door middel van regressie, een relatie af tussen de relatieve utiliteiten en de absolute koopkansen.

Het MINT programma is ontwikkeld door Mark Bradley van Hague Consulting Group. Het is, in tegenstelling tot ACA geen 'kant-en-klaar' standaardprogramma, maar een programma-omhulsel (shell) waarmee verschillende geavanceerde vormen van conjunct meten kunnen worden uitgevoerd, en waarmee ook de lay-out van het scherm vrij kan worden gekozen. Basis van het programma is een algemeen CAPI-enquêtesysteem, waarmee vragenlijsten 'geprogrammeerd' kunnen worden voor gebruik op gewone of portable PC's. Naast de gebruikelijke vormen van vraagstelling kent dit systeem tevens een reeks van 'modules', waarmee verschillende varianten van conjuncte metingen (en andere meer gecompliceerde vormen van vraagstelling) kunnen worden uitgevoerd. Deze modules zijn geheel geïntegreerd in het systeem, zodat bij het formuleren van de profielen voor de conjuncte meting gebruik gemaakt kan worden van alle informatie die voordien in het traditionele deel van de enquête is verzameld. Daardoor wordt het mogelijk om de profielen te formuleren in letterlijk exact dezelfde bewoordingen die door de respondent zijn gehanteerd bij het beantwoorden van eerdere vragen. Doel hiervan is het creëren van een context voor de conjuncte meting die voor elke respondent zo realistisch mogelijk is, opdat bias in de antwoorden zoveel mogelijk wordt vermeden.

De eerste module die voor MINT werd ontwikkeld voor het verrichten van conjuncte metingen vertoonde veel overeenkomst met de ACA opzet. Allereerst dienden door de respondent de meest relevante attributen te worden geselecteerd, en daarna diende per attribuut het hoogste en laagste acceptabele niveau te worden aangeduid. Vervolgens genereerde het programma verschillende paren van profielen, waarbij door de respondent voor één van beide een voorkeur moest worden geuit (beiden even goed ook toegestaan). Nadat enkele beoordelingen worden uitgevoerd rekende het programma door middel van regressie de geschatte utiliteiten uit, een bepaalde welke nieuwe paren profielen vervolgens het best ter beoordeling aangeboden konden worden. Dit herhaalde zich totdat alle utiliteiten voldoende nauwkeurig geschat waren, of tot het ingestelde maximale aantal beoordelingen was bereikt.

Op basis van de opgedane ervaringen met de respondenten en de wensen van de opdrachtgevers zijn inmiddels nieuwe modules voor MINT ontwikkeld, die eveneens con-

juncte metingen uitvoeren, maar volgens een iets andere procedure. Zo werd een module toegevoegd die werkt met een vast aantal attributen, een vast aantal niveaus en een standaard experimentele proefopzet (zij het wel sterk gerandomiseerd, zowel in de specificatie als in de presentatie van de profielen). Dit om te voldoen aan de behoefte van veel opdrachtgevers om een aantal vaste (nieuwe) produktconcepten ter beoordeling aan te bieden, en daarvoor statistisch zo rigide mogelijke resultaten te verkrijgen. Voorts werd een module ontwikkeld die gedurende de enquête uit de paarsgewijze vergelijkingen een volledige voorkeursrangordening afleidt, zodat per respondent bekend is welk produkt van alle beoordeelde produktprofielen als het beste wordt gezien. Dit beste produkt wordt later in de vragenlijst aan de respondent voorgelegd om te vragen naar diens koopintentie. Op deze wijze kan het maximaal marktpotentiëel worden gemeten.

Naast de conjuncte meetmodules zijn ook andere MINT-modules beschikbaar, die bijvoorbeeld Prijsmeterachtige vragen naar subjectief als 'hoog' en 'laag' ervaren niveaus van verschillende attributen vragen. Deze worden vervolgens tijdens de enquête gerelateerd aan de opgegeven koopintentie, en hieruit wordt een individuele koopkanskromme afgeleid. Desgewenst kan deze tijdens de enquête nog door de respondent worden gevalideerd (en desnoods bijgesteld). Op deze wijze kan de produktverkoop worden geschat als functie van verschillende attributen, waaronder de prijs van het produkt.

Tenslotte een korte vergelijking tussen ACA en MINT. ACA is een standaardprodukt, waarmee een betrekkelijk eenvoudige conjuncte meting kan worden uitgevoerd die tot op zekere hoogte is afgestemd op de specifieke omstandigheden en overtuigingen van de respondent. MINT wordt niet als los programma verkocht, maar uitsluitend aangeboden samen met adviesdienstverlening, als hulpmiddel bij meer geavanceerd onderzoek. ACA werkt volgens één vast concept, MINT biedt verschillende alternatieven, met de mogelijkheid om nieuwe varianten van conjunct meten 'op maat' te laten ontwikkelen en toe te passen. Beide programma's voeren on-line, tijdens de enquête, een statistische analyse uit waarmee zij de verdere vraagstelling sturen. ACA bevat een zogenaamde marktsimulator, waarmee voor de onderzochte respondenten de reacties op produktverandering voorspeld kunnen worden. Bij onderzoeken met MINT stellen wij indien de opdrachtgever dat wenst een PC simulatieprogramma ter beschikking dat kan variëren, afhankelijk van de behoefte van de opdrachtgever, van een eenvoudige marktsimulator tot een geïntegreerd prognosesysteem, compleet met databases van bevolkingsgegevens etcetera, waarmee opgehoogde prognoses voor marktsegmenten en totale populatie kunnen worden gemaakt.

De kracht van MINT is dat het alles kan wat ACA kan, en nog veel meer. Doordat het een eigen ontwikkeling is kan alles in de enquêteprocedure en de presentatie precies zo worden veranderd als nodig is om een zo goed mogelijk onderzoek uit te voeren. In de praktijk is gebleken dat deze mogelijkheid van MINT ertoe leidt dat niet één onderzoek met conjuncte meting gelijk is aan een voorgaande. Steeds blijken er specifieke wensen of eisen te zijn waaraan met een 'standaardprogramma' niet optimaal kan worden tegemoet gekomen.

4. PRAKTIJKERVARINGEN MET MINT

Het MINT-programma is ontwikkeld door Hague Consulting Group, maar wordt in Nederland regelmatig toegepast in samenwerking met Veldkamp (AGB), vooral als het om grotere aantallen respondenten gaat of om respondenten die moeilijk te traceren zijn. Inmiddels is een reeks onderzoeken met MINT uitgevoerd, waarbij op diverse heel verschillende manieren het veldwerk en de conjuncte metingen werden uitgevoerd:

1. face-to-face bij respondenten thuis; dit is de meest voor de hand liggende methode
2. face-to-face op de werklocatie; dit is met name interessant bij executive interviews
3. telefonisch maar wel computergestuurd; dit is met name interessant wanneer de adressen van de respondenten sterk verspreid liggen
4. per post, via toezending van een floppie-disk met het enquêteprogramma.

De eerste twee varianten werken in de praktijk het best. Hierbij zijn de momenteel beschikbare kwalitatief hoogwaardige draagbare PC's een onontbeerlijk instrument. Wij gebruiken daarvoor Zenith Z-181 machines, die compleet met oplaadapparatuur in een normale attachecase passen. Zowel bij 'gewone' huisenquêtes als bij interviews van topmanagers blijkt de enquêtemethode als zeer positief te worden ervaren. Overigens is onze ervaring dat het beter is om bij de huisenquête niet van tevoren aan te kondigen dat met behulp van een computer zal worden geënquêteerd: met name bij oudere respondenten kan daardoor onnodige argwaan ontstaan, die helemaal niet optreedt als de respondent direct met de PC wordt geconfronteerd. De managers waarderen veelal de efficiëncy van de computergestuurde enquête.

Maar ook de telefonische conjuncte meting is in de praktijk een goed toepasbaar alternatief. Hierbij wordt aan alle respondenten van tevoren een boekje opgestuurd, waarin alle profielen tweemaal naast elkaar zijn opgenomen. Elk profiel is aangeduid met een lettercode, en door het boekje door te bladeren kan de respondent iedere mogelijke combinatie van twee profielen voor zich leggen. Hij/zij ziet dan nagenoeg hetzelfde als wat op het computerscherm verschijnt. Tijdens de enquête genereert de computer steeds twee lettercodes, die de enquêteur aan de respondent doorgeeft. Deze draait in het profielenboekje de bijbehorende twee profielen voor, vergelijkt ze en noemt een voorkeur. De enquêteur toest deze dan in, en geeft vervolgens door welke lettercodes de computer daarna geëvalueerd wil zien. De verder verwerking en analyse gebeuren op de gebruikelijke manier.

Met de laatstgenoemde manier, toezending van de enquête per post op een floppie-disk, hebben wij nog weinig ervaring opgedaan. Wij passen deze vorm momenteel toe in een onderzoek voor AT & T, in de Verenigde Staten.

Ter illustratie geven wij hieronder een beknopte beschrijving van een onderzoek dat is uitgevoerd met behulp van MINT. Het betreft hier een opdracht van de Nederlandse Spoorwegen, die ons vroegen om voor de reisrelaties tussen de middelgrote steden in het oosten van het land naar de Randstad te onderzoeken hoe de vervoersmarkt is opgebouwd, hoe de concurrentiepositie van de trein is en hoe gevoelig deze markt is voor veranderingen in het treinprodukt. Daarbij ging het om zowel het vaststellen van het potentieël om automobilisten over te halen tot treingebruik, als om het risico om treinreizigers te verliezen aan de auto.

Voor het onderzoek was het noodzakelijk om zo'n 250 respondenten te identificeren die recent vanuit het oosten naar de Randstad waren gereisd en die konden kiezen tussen auto en trein. Daartoe is in het najaar van 1987 een telefonische screeningsenquête verricht onder huishoudens in Nijmegen. De respondenten die voldeden aan de selectiecriteria zijn vervolgens thuis geënquêteerd, met behulp van draagbare PC's en het MINT-enquêtesysteem. De keuze van deze wijze van uitvoering voor het veldwerk leverde ons twee belangrijke voordelen op:

1. hierdoor waren wij in staat te profiteren van alle standaard voordelen van CAPI, inclusief toegespitste vragen en antwoordcategoriën, permanente controle op waardebereik en consistentie van de antwoorden, automatische doorverwijzing, geautomatiseerde verwerking en opslag van de gegevens; tesamen leiden deze tot kwalitatief hoogwaardigere gegevens dan met traditionele onderzoeksvormen kan worden verkregen, die bovendien sneller voor analyse beschikbaar zijn.
2. maar ook waren wij hierdoor in staat om twee conjuncte metingen volledig in de enquête te integreren; bovendien konden deze voor elke respondent volledig 'op maat' worden gemaakt door zijn eigen reisgegevens in de ter beoordeling aangeboden profielen te verwerken, en kon de specificatie en de presentatie van de te beoordelen profielen in verregaande mate worden gerandomiseerd.

De conjuncte metingen zijn uitgevoerd om de gevoeligheden van de (potentiële) consumenten voor mogelijke veranderingen in het treinprodukt te kunnen kwantificeren, en om hieruit elasticiteiten te kunnen afleiden. Aan alle respondenten werd eerst een reeks omschrijvingen van de treinreis van Nijmegen naar Amsterdam, Rotterdam of Den Haag ter beoordeling voorgelegd. De respondent kreeg steeds twee varianten van zijn treinreis op het computerscherm gepresenteerd, met veranderingen in één of meer van de volgende variabelen:

- de prijs van het kaartje (op basis van de door de respondent opgegeven of geschatte prijs)
- de totale reistijd per trein (op basis van de opgegeven of geschatte reisduur, inclusief eventueel overstappen)
- het treincomfort (volgens door NS geleverde specificatie)
- het vereiste aantal overstappen (op basis van het door de respondent opgegeven aantal overstappen)

Daarbij werd hij/zij uitgenodigd om een voorkeur uit te spreken voor één van beide profielen. De opgegeven voorkeur werd in de PC ingevoerd, en direct door MINT geanalyseerd. Uit deze analyse werd afgeleid welke van alle mogelijke resterende paarsgewijze vergelijkingen naar verwachting een maximale additionele informatie zou opleveren, en deze werd vervolgens door de PC als volgende combinatie van twee treinreisprofielen ter beoordeling aangeboden. Deze procedure herhaalde zich totdat een volledige voorkeursvolgorde voor negen alternatieven was bepaald.

Daarna werd een tweede conjuncte meting uitgevoerd, dit keer in de vorm van een vergelijking van de trein met de auto. Hierbij werd de respondent gevraagd te doen alsof hij zich in de door de computer omschreven omstandigheden bevond, en aan te geven welke van de twee vervoermiddelen hij/zij voor zijn/haar reis zou kiezen. De reis per auto werd beschreven aan de hand van de autokosten en de totale reistijd per auto, de reis per trein aan de hand van eerdergenoemde vier attributen. Zowel de attributen van

de trein als die van de auto werden gevariëerd, en de PC zocht voor elke respondent systematisch naar het omslagpunt: dit was het punt waarop de betreffende reiziger zijn feitelijk gekozen vervoermiddel ruilde voor de alternatieve vervoerwijze.

De in de conjuncte metingen verzamelde gegevens zijn geanalyseerd met behulp van Maximum Likelihood Logit, en de resultaten zijn ondergebracht in een marktaandeelsimulatiemodel. Met behulp van dit model kunnen de reacties van de beschouwde markt op veranderingen van het treinprodukt per segement worden gesimuleerd en voorspeld. Op basis van uitgevoerde simulaties zijn bijvoorbeeld de in tabel 1 en 2 gegeven elasticiteitswaarden voor het treingebruik afgeleid. Op de resultaten en de conclusies van het besproken onderzoek gaan wij hier niet verder in. De geïnteresseerde lezer wordt verwezen naar Bradley *et al.* (1988) voor een meer gedetailleerde beschrijving van deze studie.

Tabel 1. Elasticiteiten voor treingebruik door automobilisten, bij verbetering van het treinprodukt.

ATTRIBUUT	*MARKTSEGMENT*		
	Zakelijk	Ontspanning	Sociaal/bezoek
Tarief	- 0.09	- 1.18	- 0.15
Reistijd	- 0.41	- 0.45	- 0.07
Comfort	0.05	0.10	0.02
Overstap	- 0.02	- 0.06	- 0.02

Tabel 2. Elasticiteiten voor treingebruik door treinreizigers, bij verslechtering van het treinprodukt.

ATTRIBUUT	*MARKTSEGMENT*		
	Zakelijk	Ontspanning	Sociaal/bezoek
Tarief	- 0.16	- 0.77	- 0.03
Reistijd	- 0.47	- 0.31	- 0.19
Comfort	0.16	0.25	0.07
Overstap	- 0.00	- 0.00	- 0.03

5. TOEKOMSTVERWACHTING

Wie in 1980 voorspeld zou hebben hoe CATI en CAPI systemen 8 jaar later in ons vak ingeburgerd zouden zijn, zou ongelovig aangestaard zijn. Uit de psychologie is bekend dat bij dit soort snelle ontwikkelingen het verwachtingspatroon voor de dan volgende toekomst snel opgeschroefd wordt. Meer apparatuur, nog snellere apparatuur, goedkopere hardware, meer flexibele programma's en meer gebruiksvriendelijke programma's. Een groot deel hiervan zal ongetwijfeld bewaarheid worden, maar de toekomst ligt waarschijnlijk niet alleen in deze verbetering van apparatuur en programmatuur. Onze verwachting is dat dit soort enquêtesystemen steeds meer gekoppeld zullen gaan worden aan andere systemen.

Wanneer een beoordeling of trade-off gemaakt moet worden zoals in de beschreven conjunctmeet programma's, kan de respondent bijvoorbeeld eerst een systeem raadplegen waarin kennis is opgeslagen. De respondent kan dan meer te weten komen over de produkten, voordat de beoordeling plaatsvind. Zo kunnen de CAPI/CATI/CAGI ontwikkelingen inhaken op voortgang op het gebied van artificiële intelligentie en expertsystemen.

Ook de ontwikkelingen rond Minitel en andere videotekst systemen bieden aanknopingspunten. Als dergelijke apparatuur ook in Nederland meer ingeburgerd is, wordt het mogelijk om op grotere schaal en meer frequent computergestuurd onderzoek te gaan doen. Hierbij is ook de aansluiting met panels waarin reeds gedragsregistratie plaats vindt niet uitgesloten. We kunnen dan denken aan panels die uitgerust zijn met 'people-meters' om hun kijkgedrag te registreren en ook aan panels waarvan het koopgedrag via 'bar-codes' wordt gemeten.

Bij de CATI/CAPI/CAGI systemen zelf zal steeds minder de nadruk komen te liggen op voordelen als snelheid en prijs en steeds meer op een betere kwaliteit van de data. Niet als verbetering 'l'art pour l'art', maar om betere schattingen en predicties te kunnen doen en als zodanig meer beleidsrelevante informatie voor beslissers te kunnen leveren.

LITERATUUR

Baden, R.W.D., E.M. Boerema en E.P.M. Bon (1987): *Computer Assisted Interviewing in a Face-to-Face Situation*, ESOMAR-Congress, Montreux, pp. 829-845.

Bradley, M. (1988): Realism and Adaption in Designing Hypothetical Travel Choice Concepts, *Journal of Transport Economics and Policy*, Volume XXII, no. 1.

Bradley, M., T. Grosvenor en A. Bouma (1988): *An Application of Computer-Based Stated Preference to Study Mode-Switching in The Netherlands*, Paper to be presented to the PTRC 16th Summer Annual Meeting, Bath, September 1988.

Bronner, A.E. (1982): Computer Assisted Interviewing (CAI), Lezing Tweede Sociometrie Congres, Vrije Universiteit Amsterdam. In *Sociometric Research* 1982, pp. 192-195.

Bronner, A.E. en R. de Hoog (1982): Enquêteren met behulp van de computer. In: *Jaarboek van de Nederlandse Vereniging van Marktonderzoekers*, pp. 109-131.

Cattin, P. en D.R. Wittink (1982): Commercial Use of Conjoint Analysis: A Survey, *Journal of Marketing*, Vol. 46, pp. 44-53.

Cattin, P. en D.R. Wittink (1986): *Commercial Use of Conjoint Analysis: An Update*, Paper presented to the ORSA/TIMS Marketing Science Meetings, Richardson.

Green P.E. en V. Srinivasan (1978): Conjoint Analysis in Consumer Research: Issues and Outlook, *Journal of Consumer Research*, Volume 5, pp. 103-121.

Johnson, R.M. (1985): *Adaptive Conjoint Analysis*, Sawtooth Software Inc, Idaho.

Kroes, E.P. (1986): Conjunct Keuze-onderzoek. In: *Recente Ontwikkelingen in het Marktonderzoek*, Jaarboek '86-87, Nederlandse Vereniging van Marktonderzoekers.

Leigh, T.W., D.B. McKay en J.O. Summers: Reliability and Validity of Conjoint Analysis and Self-Explicated Weights: A Comparison, *Journal of Marketing Research*, Vol. 21, pp. 456-462.

Levin, I.P., J.J. Louviere, A.A. Schepanski en K.L. Norman (1983): External Validity Tests of Laboratory Studies of Information Integration, *Organizational Behavior and Human Performance*, Vol. 31, pp. 173-193.

Mc Cullough, J. en R. Best (1979): Conjoint Measurement: Temporal Stability and Structural Reliability, *Journal of Marketing Research*, Vol. 16, pp. 26-31.

Raaij, W.F. van (1978): Het gebruik van de conjuncte meettechniek, Marktonderzoek en consumentengedrag. In: *Jaarboek van de Nederlandse Vereniging van Marktonderzoekers.*
Segal, M.N. (1982): Reliability of Conjoint Analysis: Contrasting Data Collection Procedures, *Journal of Marketing Research*, Vol. 19, pp. 139-143.
Sikkel, D. (1988): Computergestuurd enquêteren en data kwaliteit, Lezing NVvM-congres. In: *Congresboek 1988 van de Nederlandse Vereniging van Marktonderzoekers*, pp. 109-131.
Vogels, B.J. (1987): De kentering verkend. In: *Recente ontwikkelingen in het marktonderzoek.* Jaarboek '87-88 van de Nederlandse Vereniging van Marktonderzoekers, pp. 23-34.

3. Scanning: een tussenstand

P.S.H. LEEFLANG, L.A. VAN HAASTRECHT

SAMENVATTING

In dit artikel in de rubriek statistische gegevens en economische kengetallen zullen we de gegevens die door middel van scanning verkregen worden aan een nader onderzoek onderwerpen. We schetsen allereerst enige recente ontwikkelingen in de wijze van verkrijgen van scanning-gegevens.
Vervolgens zullen de scanning-gegevens vergeleken worden met 'corresponderende' gegevens die verkregen worden met behulp van een traditioneel winkelpanel. Daartoe worden enkele kengetallen van een aantal merken, afkomstig van twee markten, met elkaar vergeleken. Ten slotte zullen de prijs- en distributie-elasticiteiten voor de merken van één van de markten worden bepaald met behulp van traditionele gegevens en scanning-gegevens. Ook op deze wijze kan inzicht worden verkregen in de stand van zaken rond de ontwikkeling van scanning in Nederland: een tussenstand.

1. INLEIDING

In het Jaarboek van de Nederlandse Vereniging van Marktonderzoekers is in de afgelopen jaren een aantal malen aandacht besteed aan de gevolgen van de invoering van de uniforme artikel codering (UAC) voor marketing en marktonderzoek. Zo heeft Olivier (1981) de gevolgen van de invoering van UAC voor de structuur van het Nederlandse Marktonderzoek geschetst[1]. Plat en Bol (1985) hebben de ontwikkeling van 'scanning' in binnen- en buitenland beschreven. Daarbij wordt onder 'scanning' het inlezen van de streepjescode en de vertaling daarvan in een voor elk artikel unieke numerieke code verstaan. Plat en Bol hebben verder bij hun beschrijving van de ontwikkeling van scanning aandacht besteed aan de mogelijkheden die er zijn om scanning-gegevens te analyseren. Tevens is een illustratie van een dergelijke analyse gegeven, waarbij de effectiviteit van de marktinstrumenten op de marktaandelen van een aantal door een detaillist gevoerde merken uit een bepaalde produktcategorie bepaald is[2].
We zijn nu enkele jaren verder en de vraag dient zich aan hoe de stand van zaken is met betrekking tot de ontwikkeling van scanning. In paragraaf 2 zullen we een beknopte schets van de stand van zaken geven. Hieruit zal blijken dat 'scanning' nog volop in ontwikkeling is en dat we van een 'tussenstand' moeten spreken. Dit blijkt onder meer ook uit hetgeen in paragraaf 3 van dit artikel beschreven wordt. In paragraaf 3 zullen we de waarden die de marktaandelen, prijzen en de gewogen distributie van een aantal merken in de tijd aannemen bestuderen. Deze merken worden op twee verschillende markten afgezet. Beide markten zijn markten van frequent aangekochte niet-duurzame consumptiegoederen. Het produkt dat op markt 1 wordt afgezet is niet-eetbaar; het produkt van markt 2 is een voedingsmiddel. We zullen de waarden die voor de hiervoor

genoemde variabelen door middel van scanning verkregen worden vergelijken met de hiermee corresponderende waarden die gevonden worden met behulp van het, inmiddels traditionele, Nielsen winkelpanel.
In paragraaf 4 zullen we de prijs- en distributie-elasticiteiten voor de merken van markt 1 bepalen met behulp van zowel de traditionele Nielsen-gegevens als met behulp van de scanning-gegevens[3]. Op deze wijze kunnen we enig inzicht in de kwaliteit van de scanning-gegevens verkrijgen. Ten slotte zullen in paragraaf 5 onze conclusies worden samengevat.

2. ENKELE RECENTE ONTWIKKELINGEN

In zijn opsomming van de 'nieuwe mogelijkheden voor industrie en handel' als gevolg van de invoering van de UAC stelde Olivier in 1981 dat de toekomstige mogelijkheden van het gebruik van UAC uitgebuit moeten worden door te streven naar:
- meer informatie;
- betrouwbare informatie;
- meer verfijnd uitgesplitste informatie;
- snellere beschikbaarheid van informatie.

Met de onderstaande schets van enkele ontwikkelingen zullen we nagaan of en in welke mate deze doelstellingen inmiddels gerealiseerd zijn. Daarbij concentreren we ons op scanning in de levensmiddelendetailhandel en de door A.C. Nielsen (Nederland) B.V. te verstrekken gegevens[4]. Daarbij zij aangetekend dat Nielsen zorgdraagt voor de technische uitvoering van de uitwisseling van marktgegevens. De uitwisseling wordt gecoördineerd door het bureau Centrale Uitwisseling Marktgegevens (CUM). Dit bureau is eigendom van de Stichting UAC.

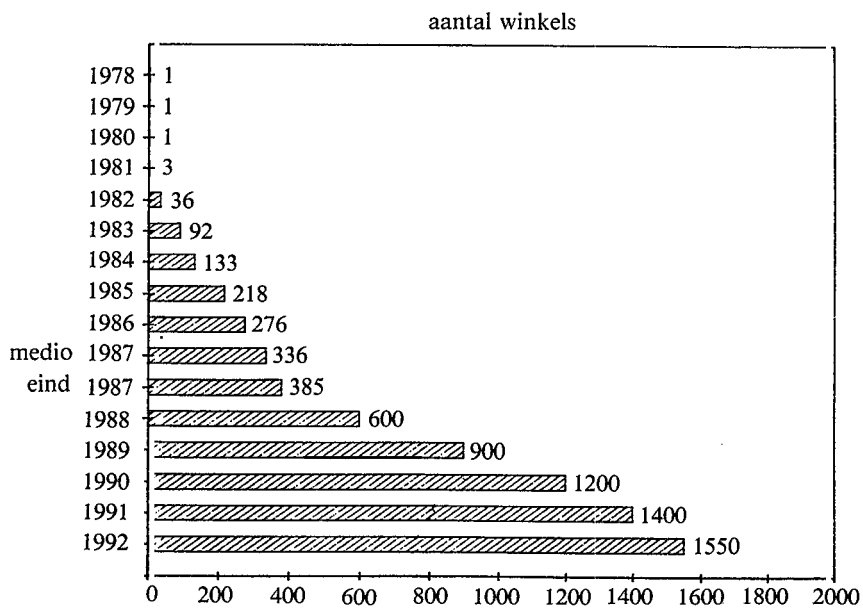
Sinds 1985 wordt door middel van '*Scantrack*' scanning-informatie die ontleend is aan gegevens afkomstig van scannende winkels verstrekt. Deze informatie was in 1985 gebaseerd op gegevens afkomstig van 30 scannende winkels. In 1988 is de rapportage van scanning-gegevens van de scannende winkels in Nederland gebaseerd op een steekproef van circa 140 winkels. Via celexpansie worden de gegevens afkomstig van dit winkelpanel verwerkt tot representatieve gegevens voor de *scannende winkels* in Nederland. De Scantrack-studies bevatten informatie over (onder meer) de volgende variabelen:
1. verkopen in stuks/kilogrammen/liters;
2. verkopen in geld;
3. consumentenprijzen;
4. (gewogen en numerieke) verkoopdistributie;
5. folder en dagbladadvertenties (overzichten);
6. een kalender met verkoopbeïnvloedende variabelen zoals bijzondere gebeurtenissen, weersgesteldheid, enz.;

Deze variabelen worden per merk en per produktgroep, per week gerapporteerd.
Door middel van de '*Scantrade Services*' wordt 'actuele medewerkersinformatie' voor scannende winkels en winkelorganisaties verstrekt. In de Scantrade rapporten worden onder meer de totaal omzetten en de omzetten van 50 produktgroepen van een scannen-

de winkel vergeleken met de hiermee corresponderende omzetten en prijsniveaus van winkels in dezelfde oppervlaktecategorie en jaaromzetklasse.
Ten slotte is het mogelijk om door middel van '*Scantest*' informatie over de uitkomsten van diverse experimenten te verkrijgen. Bij deze experimenten wordt van scanninggegevens gebruik gemaakt. Voorbeelden van deze experimenten zijn:
- het uittesten van verschillende displays;
- prijsgevoeligheidstesten;
- het testen van nieuwe produkten;
- het testen van nieuwe verpakkingen.

Vergeleken met de traditionele Nielsenrapporten is op basis van Scantrack en Scantrade de registratiefrequentie sterk toegenomen (van tweemaandelijks naar wekelijks) en is de rapportage veel sneller beschikbaar. Werd een tweemaandelijks rapport steeds na ongeveer een maand uitgebracht; nu worden de weekgegevens uit scanning elke vier weken ter beschikking gesteld, steeds twee weken na de betreffende periode.

Zowel het aantal produkten dat met een barcode is uitgerust als het aantal scannende winkels groeit sterk. De eerste ontwikkeling is niet alleen van belang voor de scannende winkels en het hierboven genoemde onderzoek wat met behulp van scannende winkels plaats vindt, doch ook voor het marktonderzoek waarin met behulp van scannende *consumenten* gegevens worden gegenereerd, zoals in het AGB-Consumerscan-project[5]. Voor een overzicht van de recente en verwachte ontwikkeling van het aantal scannende (levensmiddelen)winkels in Nederland verwijzen we naar figuur 1.



Figuur 1 Ontwikkeling aantal scannende levensmiddelenwinkels in Nederland[a]
Bron: Lemkes (1987)

a Tot en met 1987 werkelijk aantal scannende winkels, voor latere jaren zijn dit prognoses.

Wanneer we de gegevens die door middel van scanning verkregen worden vergelijken met de meer traditionele winkelpanel-gegevens dan kunnen we concluderen dat met behulp van scanning *meer*, en *meer verfijnd uitgesplitste*, informatie verkregen kan worden die bovendien *sneller beschikbaar* komt. Indicaties omtrent de betrouwbaarheid van de op basis van het scanningpanel verkregen gegevens zullen in de hierna volgende paragrafen besproken worden. Uit studies die met name in de Verenigde Staten verricht zijn, blijkt dat men zeer goed in staat is om betrouwbare schattingen te verkrijgen van de effecten van de hantering van marktinstrumenten op responsmaatstaven zoals het marktaandeel of de verkopen van een merk met modellen die geparameteriseerd zijn met behulp van scanning-gegevens[6]. We mogen al met al concluderen dat de door Olivier (1981) geformuleerde doelstellingen voor een belangrijk deel gerealiseerd (gaan) worden.

3. SCANNING-GEGEVENS VERSUS TRADITIONELE WINKELPANELGEGEVENS

In deze paragraaf zullen we het verloop van de waarden die de variabelen marktaandeel, prijs (per standaard eenheid) en gewogen distributie in de tijd aannemen bestuderen. Daarbij zijn we geïnteresseerd in de verschillen die we kunnen waarnemen tussen de registratie van de waarden der variabelen met behulp van het traditionele Nielsen-winkelpanel en de registratie op basis van scanning-gegevens. Daartoe zijn de scanning-gegevens geaggregeerd van wekelijkse waarnemingen tot twee-maandelijkse waarnemingen. Dit is gedaan voor zes merken die samen circa 60 procent van de markt van een niet-eetbaar levensmiddel voor hun rekening nemen (markt 1) en voor vijf merken die met elkaar een marktaandeel van ruim 70 procent bezitten op de markt van een voedingsmiddel (markt 2).

Tabel 1. Procentuele verschillen tussen op scanning gebaseerde marktaandelen en prijzen en op traditionele wijze gemeten gegevens, markt 1.

procentuele verschillen in marktaandelen

merk	perioden				
	mrt./april '87	mei/juni '87	juli/aug. '87	sept./okt. '87	nov./dec. '87
A	- 2.9	+ 1.0	0	- 13.5	+ 6.3
B	- 31.6	- 22.0	- 22.1	- 5.7	- 0,9
C	+ 4.9	- 5.1	- 2.8	+ 3.9	- 4.5
D	+ 6.8	- 7.0	+ 4.1	- 2.9	+ 4.8
E	+ 21.4	+ 1.6	+ 1.7	+ 15.1	+ 17.0
F	+ 10.7	+ 28.1	+ 18.3	+ 0.9	+ 26.1

procentuele verschillen in prijzen

merk	perioden				
	mrt./april '87	mei/juni '87	juli/aug. '87	sept./okt. '87	nov./dec. '87
A	+ 7.5	+ 5.2	+ 2.4	- 3.8	+ 3.0
B	+ 2.8	+ 3.3	+ 1.6	+ 1.4	+ 3.0
C	+ 0.7	- 2.5	+ 1.2	- 2.2	- 2.7
D	- 2.0	- 3.8	- 4.8	- 0.4	- 0.4
E	+ 5.9	+ 4.9	+ 5.9	+ 0.7	- 0.7
F	+ 4.1	+ 1.9	+ 2.1	+ 2.4	+ 2.4

In tabel 1 zijn de procentuele verschillen in marktaandelen en prijzen voor de merken van markt 1 in kaart gebracht. Daarbij zij aangetekend dat de merken B en D de goedkopere merken zijn en E en F de duurdere merken. We zien dat de verschillen in de marktaandelen zoals die gemeten worden met behulp van het traditionele winkelpanel en zoals gemeten met behulp van scanning-gegevens tamelijk groot kunnen zijn. Het goedkope merk B is in de eerste drie perioden slecht vertegenwoordigd in de scannende winkels. De dure merken E en F zijn steeds sterk vertegenwoordigd in het scannende panel. De verschillen in prijzen tussen beide panels zijn klein te noemen.
Voor deze markt (1) is het niet goed mogelijk om de op scanning gebaseerde distributiegegevens en de op traditionele wijze verkregen gegevens goed met elkaar te kunnen vergelijken. Elk van de merken bestaat uit een assortiment van artikelen met ieder hun eigen distributiegraad die zich op (geaggregeerd) merk-niveau moeilijk laat vergelijken. Wel is er een indicatie dat merken met een hoge distributiespreiding in de scannende winkels een hoger marktaandeel bezitten in het panel scannende winkels dan in het traditionele Nielsen-winkelpanel.

Tabel 2. Procentuele verschillen tussen op scanning gebaseerde marktaandelen, prijzen en gewogen distributie-gegevens en op traditionele wijze gemeten gegevens, markt 2.

procentuele verschillen in marktaandelen

merk	perioden			
	febr./maart '87	april/mei '87	juni/juli '87	aug./sept. '87
A	+ 7.5	+ 8.6	+ 14.5	+ 10.4
B	+ 44.6	+ 26.2	+ 29.2	+ 38.1
C	+ 28.6	+ 37.5	+ 20.9	+ 4.3
D	– 0.1	+ 1.8	+ 5.1	+ 11.1
E	+ 25.0	+ 7.7	+ 21.4	+ 35.7

procentuele verschillen in prijzen

merk	perioden			
	febr./maart '87	april/mei '87	juni/juli '87	aug./sept. '87
A	– 3.0	– 6.3	– 1.6	– 1.6
B	– 2.4	– 4.5	– 0.7	– 0.4
C	– 4.1	– 12.7	– 2.0	0.0
D	– 3.8	– 6.0	– 0.7	– 0.7
E	– 1.3	– 2.5	+ 0.3	+ 0.3

procentuele verschillen in gewogen distributie

merk	perioden			
	febr./maart '87	april/mei '87	juni/juli '87	aug./sept. '87
A	+ 0.2	+ 4.2	+ 4.2	+ 4.2
B	+ 11.2	+ 6.5	+ 8.7	+ 8.7
C	+ 34.8	+ 18.2	+ 9.2	+ 2.4
D	+ 5.3	– 2.1	– 1.1	– 4.3
E	+ 26.6	+ 26.6	+ 26.6	+ 25.0

In tabel 2 zijn de procentuele verschillen in marktaandelen, prijzen én gewogen distributie-gegevens voor vijf merken van markt 2 in kaart gebracht. Deze merken zijn A-merken. De merken die niet in beschouwing zijn genomen, zijn distribuantenmerken. Voor de A-merken zijn in praktisch alle perioden de marktaandelen hoger in het panel van de scannende winkels dan de marktaandelen in het traditionele panel. Deze positieve verschillen kunnen voor een belangrijk deel worden verklaard uit de positieve verschillen in de gewogen distributie-gegevens; zie het onderste gedeelte van tabel 2. De verschillen in prijzen die bestaan tussen de twee verschillende panels zijn gering te noemen.

Op basis van hetgeen in de tabellen 1 en 2 is weergegeven kunnen we concluderen dat er met name voor de variabelen marktaandeel/afzet en gewogen distributie aanzienlijke verschillen bestaan tussen gegevens afkomstig van een winkelpanel en gegevens die gebaseerd zijn op een panel van scannende winkels.

4. BEPALING PRIJS- EN DISTRIBUTIE-ELASTICITEITEN

In dit deel van onze studie naar de ontwikkeling van scanning-gegevens zullen we de mogelijkheden om de vraag naar een aantal merken door middel van numeriek gespecificeerde modellen te beschrijven, onderzoeken. Daartoe parameteriseren we voor de merken van markt 1 een aantal marktaandeelmodellen. Parameterisering vindt allereerst plaats met behulp van scanning-gegevens en daarna met traditionele winkelpanel-gegevens.
Aangezien de traditionele (Nielsen-) winkelpanel-gegevens betrekking hebben op een periode van twee maanden en de scanning-gegevens per week beschikbaar zijn, zullen we voor parameterisering van de modellen met behulp van de traditionele winkelpanel-gegevens een langere periode beschouwen dan wanneer parameterisering op basis van scanning-gegevens plaats vindt. Voor de parameterisering van de marktaandeelmodellen op basis van de Nielsen-winkelpanel-gegevens gebruiken we gegevens die betrekking hebben op de periode maart-april 1984 t/m maart-april 1987. Dit zijn gegevens die betrekking hebben op 19 twee-maandelijkse perioden. Voor de parameterisering van de modellen, gebaseerd op scanning-gegevens, gebruiken we 24 paren wekelijkse waarnemingen. Deze hebben betrekking op de periode oktober 1986-maart 1987.
We parameteriseren een relatief eenvoudig marktaandeelmodel waarmee, althans in principe, prijs- en distributie-elasticiteiten kunnen worden bepaald. Dit model heeft de volgende vorm:

$$m_{jt} = e^{\beta_{oj}} \left[\frac{p_{jt}}{\sum_{r=1}^{n} p_{rt}} \right]^{\beta_{1j}} \left[\frac{d_{jt}}{\sum_{r=1}^{n} d_{rt}} \right]^{\beta_{2j}} e^{u_{jt}}, \; j=1,\ldots,n, \; t=1,\ldots,T \qquad (1)$$

waarbij:

m_{jt} = marktaandeel merk j in periode t,
e = grondtal van de natuurlijke logarithme,
p_{jt} = prijs van merk j in periode t per eenheid (bijv. per kg),
n = aantal beschouwde merken, n=6,
d_{jt} = gewogen distributie van merk j in periode t,
u_{jt} = storingsterm in de relatie,
T = aantal waarnemingen, T=19 voor de traditionele Nielsen-winkelpanel-gegevens, T=24 voor de scanning-gegevens.

De parameters β_{1j}, β_{2j} representeren respectievelijk de (relatieve) prijs- en de distributie-elasticiteiten.
Relatie (1) is voor elk van de zes merken en met behulp van de twee verschillende gegevensbronnen geschat door middel van de methode der kleinste kwadraten. De op basis van de traditionele gegevens verkregen parameterschattingen, de bij deze schattingen behorende standaardfouten, de determinatiecoëfficiënten (R^2) en de waarden van de Durbin-Watson-statistic (D.W.) staan in tabel 3 vermeld.

Tabel 3. Parameterschattingen en statistische validatiecriteria relatie (1) – traditionele gegevens.

merk	relatieve prijs		relatieve distributie			
	β_{1j}	$\beta_{1j} / \sigma_{\beta_{1j}}$ [a]	β_{2j}	$\beta_{2j} / \sigma_{\beta_{2j}}$	R^2	D.W.
A	- 1.63	– 2.35[b]	4.69	2.78[b]	.72	1.37
B	- .13	- .14	2.07	.86	.02	1.75
C	- 1.26	- .54	2.23	.46	.32	2.06
D	- .68	- .73	- .09	- .06	.40	2.12
E	- 3.09	- 2.84[b]	- .70	- .20	.44	1.50
F	- 2.52	- 2.07[b]	- .11	- .05	.65	1.82

a = t-waarden
b = parameterschatting significant op het 5%-niveau (eenzijdige toets).

De resultaten die in tabel 3 vermeld staan leiden tot de volgende conclusies:

1. Met behulp van dit model en deze gegevens kunnen de fluctuaties in de marktaandelen van met name de merken B, C, D en E tamelijk slecht worden verklaard.
2. Slechts 4 van de 12 geschatte parameters zijn statistisch significant (op 5%-niveau).
3. Er is geen sifnificante autocorrelatie voor de merken B, C, D en F. Voor de merken A en E kunnen geen conclusies omtrent de mate van autocorrelatie worden getrokken.

De parameterschattingen en de hierbij behorende statistische validatiecriteria die met behulp van scanninggevens worden verkregen voor relatie (1) zijn weergegeven in tabel 4.

Tabel 4. Parameterschattingen en statistische validatiecriteria relatie (1) – scanninggegevens.

merk	relatieve prijs		relatieve distributie			
	β_{1j}	$\beta_{1j} / \sigma_{\beta_{1j}}$ [a]	β_{2j}	$\beta_{2j} / \sigma_{\beta_{2j}}$	R^2	D.W.
A	- 4.84	- 6.37[b]	.79	1.40	.50	1.59
B	- 3.46	- 2.43[b]	0.96	1.87[b]	.53	1.83
C	- 2.48	- 1.88[b]	- .23	- .77	.04	1.95
D	- 5.51	- 2.66[b]	1.01	3.46[b]	.89	1.73
E	- 6.63	- 5.60[b]	1.72	2.08[b]	.81	1.69
F	- 4.17	- 4.04[b]	.58	.89	.35	1.96

a = t-waarden
b = parameterschatting significant op het 5%-niveau (eenzijdige toets).

Aan de hand van hetgeen in tabel 4 vermeld staat kunnen de volgende conclusies worden getrokken:

1. Ook ditmaal kunnen de fluctuaties in de marktaandelen van slechts een beperkt aantal merken redelijk worden verklaard. Dit zijn de merken D en E. Dit is niet zo verwonderlijk omdat model (1) een tamelijk eenvoudig model is waarin variabelen zoals de (relatieve) reclame-uitgaven en promoties ontbreken.
2. De met behulp van scanning-gegevens verkregen prijselasticiteiten zijn allen significant. Ook 3 van de 6 distributie-elasticiteiten blijken significant te zijn.
3. Vergeleken met tabel 3 zijn in tabel 4 de waarden van de prijselasticiteiten relatief hoog. Dit is niet verwonderlijk gezien de periodiciteit van de (wekelijkse) scanning-gegevens. De traditionele winkelpanel-gegevens zijn 'gesmoothde' gegevens waarin de twee-maandelijkse fluctuaties als het ware 'uitgemiddeld' zijn.[7]
4. Er is geen autocorrelatie voor elk van de merken A t/m F.

Vergelijken we de resultaten die in tabel 4 vermeld staan met de resultaten uit tabel 3 dan kunnen we concluderen dat met behulp van scanning-gegevens meer significante parameters gevonden kunnen worden dan met de traditionele Nielsen-winkelpanel-gegevens. We dienen hierbij echter aan te tekenen dat de resultaten op verschillende perioden betrekking hebben. Teneinde een beter inzicht in de genoemde verschillen te verkrijgen kunnen we proberen om prijs- en distributie-elasticiteiten te schatten met behulp van de twee gegevensbronnen die op dezelfde periode betrekking hebben. Daartoe specificeren we model (2):

$$m_{jt} = e^{\beta_{oj}} \left[\frac{p_{jt}}{\sum_{r=1}^{n} p_{rt}} \right]^{\beta_1} \left[\frac{d_{jt}}{\sum_{r=1}^{n} d_{rt}} \right]^{\beta_2} e^{u_{jt}}, \; j = 1, \ldots, n, \; t = 1, \ldots, T \qquad (2)$$

Dit model heeft een structuur die analoog is aan de structuur van (1).
In tegenstelling tot hetgeen bij (1) het geval is, zijn de elasticiteiten in (2) niet merk-afhankelijk (heterogeen) doch merk-onafhankelijk (homogeen). De homogene responsparameters zijn a.h.w. gemiddeld over alle merken. Relatie (2) wordt nu geschat met behulp van waarnemingen die betrekking hebben op de periode oktober 1986-maart 1987. Daartoe worden de gegevens over de 6 merken 'gepooled'[8] en wordt relatie (2) geschat met behulp van:

a. 6 x 3 (twee-maandelijkse) gegevens = 18 waarnemingsparen;
b. 6 x 24 (wekelijkse) gegevens = 144 waarnemingsparen.

De pooling van de twee-maandelijkse gegevens over merken is noodzakelijk omdat de lengte van de tijdreeks (3 waarnemingen) te klein is om relatie (1) voor elk merk afzonderlijk te schatten.
De parameterschattingen van (2) zijn weergegeven in tabel 5. De resultaten in tabel 5 bevestigen de hiervoor geformuleerde conclusies. De absolute waarden van de prijselasticiteiten en de bij deze elasticiteiten behorende t-waarden die met behulp van scanning-gegevens verkregen worden zijn hoog. Ook worden met scanning-gegevens significante schattingen van distributie-elasticiteiten verkregen.

Tabel 5. Parameterschattingen en statistische validatiecriteria relatie (2).

bron	relatieve prijs		relatieve distributie		
	β_1	β_1/σ_{β_1} [a]	β_2	β_2/σ_{β_2}	R^2
traditioneel winkelpanel	-2.18	- 2.72[b]	3.27	.74	.69
scanning	-5.09	-10.10[b]	.66	4.05[b]	.79

a = t-waarden
b = parameterschatting significant op het 5%-niveau (eenzijdige toets).

5. SAMENVATTING EN CONCLUSIES

Deze tussenstand van zaken met betrekking tot scanning leidt tot het formuleren van de volgende conclusies:

1. Met behulp van scanning kan *meer, meer verfijnd uitgesplitste*, informatie worden verkregen die *sneller beschikbaar komt.*
2. Er zijn aanzienlijke verschillen in de marktaandelen zoals die met behulp van het traditionele winkelpanel worden gemeten en de marktaandelen gebaseerd op scanning-gegevens. Alhoewel er ook aanzienlijke verschillen kunnen bestaan tussen de 'werkelijke' afzetgegevens en de gegevens zoals die met een winkelpanel gemeten worden[9] zijn de verschillen tussen de afzetgegevens afkomstig van het traditionele winkelpanel en de afzetgegevens gebaseerd op het panel scannende winkels aanzienlijk groter. Deze verschillen kunnen voor een belangrijk deel worden toegeschreven aan verschillen tussen de distributiegraad van merken in scannende winkels en de distributiegraad van merken in de populatie van 'alle' (levensmiddelen-) winkels. De verschillen tussen werkelijke en gemeten prijzen zijn gering te noemen.
3. Gegeven de samenstelling van het panel scannende winkels is de *kwaliteit van de scanning-gegevens* goed te noemen. Het is mogelijk om relaties te leggen tussen marktaandeelfluctuaties en fluctuaties in prijzen en gewogen distributie-gegevens. Daarbij is het opmerkelijk, doch verklaarbaar, dat korte-termijn prijselasticiteiten zo hoog zijn. Met behulp van scanning-gegevens kunnen ook significante distributie-elasticiteiten gevonden worden.

NOTEN

1. Zie ook Olivier (1987, blz. 269-292).
2. Zie ook Plat (1987, 1988, blz. 130-134).
3. Deze analyse is ook weergegeven in Leeflang, Plat (1988).
4. De navolgende informatie is ontleend aan Boonman (1987).
5. Zie bijv. Luijten, Valstar (1987).
6. Zie bijv. Guadagni, Little (1983); Neslin, Henderson, Quelch (1985); Blattberg, Levin (1987); Wittink, Addona, Hawkes, Porter (1988).
7. Zie ook Plat, Bol (1985), Leeflang (m.m.v. Beukenkamp) (1987, blz. 299-300).

8. Zie voor een dergelijke aanpak ook bijv. Leeflang, Broeksteeg (1979).
9. Zie bijv. Leeflang, Olivier (1985); Olivier (1987).

LITERATUUR

Blattberg, R.C. en A. Levin (1987), 'Modelling the Effectiveness and Profitability of Trade Promotions', *Marketing Science*, vol. 6, blz. 124-125.

Boonman, A.M. (1987), 'Scanning Services in Nederland voor Handel en Industrie', lezing gehouden op 30 november 1987 voor de Marketing Associatie Groningen (MARUG), te Groningen.

Guadagni, P.M. en J.D.C. Little (1983), 'A Logit Model of Brand Choice Calibrated on Scanner Data', *Marketing Science*, vol. 3, blz. 203-238.

Leeflang, P.S.H. m.m.v. P.A. Beukenkamp (1987), *Probleemgebied Marketing, een Mangementbenadering*, deel Ia, 2e druk, H.E. Stenfert Kroese b.v., Leiden/Antwerpen.

Leeflang, P.S.H. en A.G.M. Broeksteeg (1979), 'Segmentatie en Aggregatie, Enkele Illustraties', *Marktonderzoek en Consumentengedrag, Jaarboek Nederlandse Vereniging van Marktonderzoekers*, vol. 5, blz. 161-179.

Leeflang, P.S.H. en A.J. Olivier (1985), 'Bias in Consumer Panel and Store Audit Data', *International Journal of Research in Marketing*, vol. 2, blz. 27-41.

Leeflang, P.S.H. en F.W. Plat (1988), '*Scanning Scanning Opportunities*', Proceedings 41th ESOMAR Marketing Research Congress, Lisbon, Portugal, blz. 471-484.

Lemkes, F.J. (1987), 'De Ontwikkeling van Scanning in de Levensmiddelendetailhandel', lezing gehouden op 30 november 1987 voor de Marketing Associatie Groningen (MARUG), te Groningen.

Luijten, A.L.J.M. en C.D. Valstar (1987), 'Consumentenpanelonderzoek in Beweging', in *Recente Ontwikkelingen in het Marktonderzoek, Jaarboek '87-'88 van de Nederlandse Vereniging van Marktonderzoekers*, blz. 51-64.

Neslin, S.A., C. Henderson en J. Quelch (1985), 'Consumer Promotions and the Acceleration of Product Purchases', *Marketing Science*, vol. 4, blz. 147-165.

Olivier, A.J. (1981), 'Uniforme Artikel Codering in de Tachtiger Jaren: Gevolgen van Invoering voor de Structuur van het Nederlandse Marktonderzoek', *Marktonderzoek en Consumentengedrag, Jaarboek van de Nederlandse Vereniging van Marktonderzoekers*, blz. 43-52.

Olivier, A.J. (1987), *Het Samenstellen en Beheer van Gegevens als Onderdeel van een Beslissingsondersteunend Systeem ten behoeve van het Marketing Management*, proefschrift ter verkrijging van het Doctoraat in de Economische Wetenschappen aan de Rijksuniversiteit te Groningen.

Plat, F.W. (1987), 'Het Gebruik van Scanning-gegevens in Marketing-modellen: Enkele Voorbeelden', lezing gehouden op 30 november 1987 voor de Marketing Associatie Groningen (MARUG), te Groningen.

Plat, F.W. (1988), *Modelling for Markets, Applications of Advanced Models and Methods for Data Analysis*, Proefschrift ter verkrijging van het Doctoraat in de Economische Wetenschappen aan de Rijksuniversiteit te Groningen, Druk Universiteitsdrukkerij.

Plat, F.W. en J.W. Bol (1985), 'Het gebruik van Scanning bij de Distributie van Levensmiddelen: Implicaties voor Marketing en Marktonderzoek', *Marktonderzoek en Consumentengedrag, Jaarboek van de Nederlandse Vereniging van Marktonderzoekers*, blz. 45-62.

Wittink, D.R., M.J. Addona, W.J. Hawkes en J.C. Porter (1988), 'SCAN*PRO^{R}: The Estimation, Validation and Use of Promotional Effects Based on Scanner Data', Internal Paper, Cornell University.

4. AGB Gain-Loss-Analyse

Een beschrijvend model ten behoeve van concurrentie-analyse op basis van consumentenpanelgegevens

A.L.J.M. LUIJTEN

SAMENVATTING

In dit artikel wordt een model behandeld waarmee het mogelijk is de ontwikkeling van merken/soorten binnen markten van 'fast moving consumer goods' te beschrijven. Het model is ontworpen ten behoeve van de analyse van aankoophistories die verzameld worden met behulp van het huishoudpanel van AGB Attwood bv.
Binnen het model worden de aankoophistories van de individuele panelhuishoudens afzonderlijk geanalyseerd, waarna de individuele analyse-resultaten geaggregeerd worden over de respondenten, teneinde de dynamiek in markten diepgaand te kunnen beschrijven.
De structuur van het model wordt uitvoering beschreven en geïllustreerd aan de hand van een rekenvoorbeeld.

1. INLEIDING

De panelmethode kan worden gekenschetst als een dataverzamelingsmethode die zich in de praktijk van andere onderscheidt door het gebruik van verfijnde registratiemethoden en het gebruik van overlappende steekproeven. Doordat met behulp van deze methode individuele gedragsgeschiedenissen geregistreerd worden, is het mogelijk om bijvoorbeeld veranderingen in markten van fast moving consumergoods te beschrijven op basis van veranderingen op het individuele niveau.[1]
In de literatuur zijn vele stochastische koopgedragsmodellen ontwikkeld die voor een belangrijk gedeelte gebaseerd zijn op gegevens, afkomstig uit consumentenpanelonderzoek. Belangrijke bijdragen op laatstgenoemd gebied zijn onder meer geleverd door Chatfield, Ehrenberg en Goodhardt (1966), Chatfield en Goodhardt (1970, 1973), Wierenga (1974), Goodhardt, Ehrenberg en Chatfield (1984), Herniter (1971), Jones en Zufryden (1980, 1982), Leeflang en Boonstra (1982) en Zufryden (1978). Voor de Nederlandse situatie is in Rotterdam nader onderzoek naar aankoopstructuren verricht door Wijnberger (1980) en Verhulp, Wijnberger (1982, 1983). Binnen laatstgenoemde studies werd o.a. expliciet aandacht besteed aan de houdbaarheid van de veronderstellingen van het Negatief Binomiaal Model van Ehrenberg. In dit artikel wordt het AGB Gain-Loss-Analysemodel beschreven. Binnen dit model, dat geen stochastische koopgedragsmodel is, wordt gebruik gemaakt van gedachten die door Ehrenberg c.s. ont-

4. AGB Gain-Loss-Analyse

Een beschrijvend model ten behoeve van concurrentie-analyse op basis van consumentenpanelgegevens

A.L.J.M. LUIJTEN

SAMENVATTING

In dit artikel wordt een model behandeld waarmee het mogelijk is de ontwikkeling van merken/soorten binnen markten van 'fast moving consumer goods' te beschrijven. Het model is ontworpen ten behoeve van de analyse van aankoophistories die verzameld worden met behulp van het huishoudpanel van AGB Attwood bv.
Binnen het model worden de aankoophistories van de individuele panelhuishoudens afzonderlijk geanalyseerd, waarna de individuele analyse-resultaten geaggregeerd worden over de respondenten, teneinde de dynamiek in markten diepgaand te kunnen beschrijven.
De structuur van het model wordt uitvoering beschreven en geïllustreerd aan de hand van een rekenvoorbeeld.

1. INLEIDING

De panelmethode kan worden gekenschetst als een dataverzamelingsmethode die zich in de praktijk van andere onderscheidt door het gebruik van verfijnde registratiemethoden en het gebruik van overlappende steekproeven. Doordat met behulp van deze methode individuele gedragsgeschiedenissen geregistreerd worden, is het mogelijk om bijvoorbeeld veranderingen in markten van fast moving consumergoods te beschrijven op basis van veranderingen op het individuele niveau.[1]
In de literatuur zijn vele stochastische koopgedragsmodellen ontwikkeld die voor een belangrijk gedeelte gebaseerd zijn op gegevens, afkomstig uit consumentenpanelonderzoek. Belangrijke bijdragen op laatstgenoemd gebied zijn onder meer geleverd door Chatfield, Ehrenberg en Goodhardt (1966), Chatfield en Goodhardt (1970, 1973), Wierenga (1974), Goodhardt, Ehrenberg en Chatfield (1984), Herniter (1971), Jones en Zufryden (1980, 1982), Leeflang en Boonstra (1982) en Zufryden (1978). Voor de Nederlandse situatie is in Rotterdam nader onderzoek naar aankoopstructuren verricht door Wijnberger (1980) en Verhulp, Wijnberger (1982, 1983). Binnen laatstgenoemde studies werd o.a. expliciet aandacht besteed aan de houdbaarheid van de veronderstellingen van het Negatief Binomiaal Model van Ehrenberg. In dit artikel wordt het AGB Gain-Loss-Analysemodel beschreven. Binnen dit model, dat geen stochastische koopgedragsmodel is, wordt gebruik gemaakt van gedachten die door Ehrenberg c.s. ont-

wikkeld zijn bij de analyse van het koopgedrag in twee opeenvolgende perioden. Het hier gepresenteerde model is evenwel eenvoudiger van structuur, waardoor het toegankelijk wordt voor een groter publiek van marktonderzoekers en marketingmanagers, die geïnteresseerd zullen zijn in de effecten van gepleegde marktetinginspanningen op het gedrag van consumenten.

2. HET AGB GAIN-LOSS-MODEL

2.1. Inleiding

Binnen het ontwikkelde model wordt het patroon van merkentrouw (brandloyalty) en merkenwissel (brandswitching) tussen twee, niet noodzakelijk aaneensluitende perioden van een gelijke lengte aan een nader onderzoek onderworpen. Bij het vergelijken van marktgegevens over twee perioden kan veelal geconstateerd worden dat bepaalde merken hoeveelheid winnen of verliezen.
Deze geregistreerde winsten of verliezen zijn het gevolg van opgetreden veranderingen in het aankoopgedrag van individuele huishoudens. Zelfs in situaties waarbij gesproken kan worden van een constant volume voor een bepaald merk in de loop van de tijd, zullen er veranderingen op het individuele niveau kunnen optreden, die op geaggregeerd niveau tegen elkaar wegvallen. Er is derhalve in dit geval sprake van een ogenschijnlijk statische situatie voor het betrokken merk.

Om het totale proces van merkentrouw en merkenwissel, zoals dit plaatsvindt binnen markten (produktveld), adequaat te kunnen representeren, dient het op te stellen model derhalve gedefinieerd te worden op het niveau van het individuele huishouden.
Aggregatie van de resultaten op het individuele niveau leidt tot de standaardtabellen die bij de gain-loss-analyse worden afgedrukt.
De interpretatie van de resultaten van de analyse zullen moeten geschieden in het licht van het gehanteerde model en de daarbij gehanteerde vooronderstellingen.

In het vervolg van deze paragraaf zal uitvoering ingegaan worden op het gehanteerde rekenmodel en de daarbij behorende veronderstellingen. Het model wordt toegelicht aan de hand van een eenvoudig rekenvoorbeeld.

2.2. Het marktgedrag van huishoudens in twee perioden

Binnen het gain-loss-model vindt het onderzoek naar merkentrouw en merkenwissel plaats op het niveau van de afzonderlijke huishoudens. Alvorens het gedrag van huishoudens met betrekking tot één of meer merken te bestuderen, is het zinvol het *marktgedrag* van de huishoudens in de twee onderscheiden analyseperioden te typeren.
Dit *marktgedrag* wordt gedefinieerd als de totale gekochte hoeveelheid van het *produkt* in de onderscheiden perioden. Op basis van het geregistreerde gedrag in twee perioden, kunnen 5 soorten kopende huishoudens onderscheiden worden, te weten:
- huishoudens met een constant marktgedrag (hoeveelheid produkt in periode 1 = hoeveelheid produkt in periode 2);
- huishoudens die meer zijn gaan kopen (hoeveelheid produkt in periode 1 < hoeveelheid produkt in periode 2);

- huishoudens die minder zijn gaan kopen (hoeveelheid produkt in periode 1 > hoeveelheid produkt in periode 2);
- huishoudens die uitsluitend in periode 2 kopen. Deze huishoudens worden *nieuwe kopers* genoemd ongeacht hun gedrag in overige niet beschouwde perioden;
- huishoudens die uitsluitend in periode 1 kopen. Deze huishoudens worden *verloren kopers* genoemd, ongeacht hun gedrag in overige niet beschouwde perioden.

Daarnaast zullen er huishoudens kunnen voorkomen die in beide perioden het produkt niet aangekocht hebben. Voor ieder kopend huishouden, behorend tot één van de vijf bovengenoemde kopersgroepen, wordt vervolgens het patroon van merkentrouw en -wissel nader bestudeerd.

2.3. Het model voor huishoudens met een constant marktgedrag

Als voorbeeld van een huishouden met een constant marktgedrag, nemen wij 'huishouden 1'. Dit huishouden koopt in zowel periode I als periode II 25 produkteenheden, dat wil zeggen er is sprake van een constant marktgedrag. De verdeling van de gekochte hoeveelheden over de merken wijzigt zich echter in de loop van de tijd. Op basis van de panelgegevens, kan het volgens overzicht worden samengesteld:

Aankopen huishouden 1

Merk	*Periode I*	*Periode II*
A	8	0
B	5	2
C	2	5
D	4	8
E	6	8
F	0	2
Totaal produkt	25	25

Op basis van het vergelijken van bovenstaande twee cijferreeksen, moeten nu uitspraken gedaan worden over de loyaliteit van huishoudens 1 met betrekking tot de onderscheiden merken en zullen winsten en verliezen van de diverse merken toegerekend moeten worden naar herkomst en bestemming. Teneinde dit mogelijk te maken, dienen definities en vooronderstellingen te worden geïntroduceerd.
De loyaliteit van een huishouden aan een merk wordt tot uitdrukking gebracht met de term *herhalingsaankopen.*
De herhalingsaankopen van een merk door een individueel huishouden, worden gelijkgesteld aan de kleinste waarde van de aangekochte hoeveelheden over de twee analyseperioden.
Een merk maakt *winst* bij een huishouden indien in de tweede periode meer van het merk gekocht wordt dan in de eerste periode en *verlies* indien de omgekeerde situatie zich voordoet.
Voor het rekenvoorbeeld betekent dit:

Merk	*Periode I*	*Periode II*	*Herhaling*	*Winst*	*Verlies*
A	8	0	0	0	8
B	5	2	2	0	3
C	2	5	2	3	0
D	4	8	4	4	0
E	6	8	6	2	0
F	0	2	0	2	0
Totaal produkt	25	25	14	11	11

Naar aanleiding hiervan merken wij op dat in de beschouwde situatie geldt:
1. totale winst = totaal verlies
2. aankopen periode I = herhalingsaankopen + verlies
3. aankopen periode II = herhalingsaankopen + winst

De relatie onder 2 en 3 vermeld, gelden zowel op merkniveau als op produktniveau. Het model voor huishouden 1 is echter dan volledig indien de winsten en verliezen zijn toegerekend naar herkomst en bestemming.
In het rekenvoorbeeld wil dit zeggen dat bijvoorbeeld aangegeven moet worden van welke merken de winst van D ter grootte van 4 afkomstig is.
Om de winst- en verliestoerekening op een eenduidige wijze te laten plaatsvinden, wordt een vooronderstelling geïntroduceerd die in de literatuur bekend staat onder de naam '*onafhankelijkheid*'. Deze vooronderstelling impliceert dat de verliezende merken bijdragen aan de winst van de winnende merken naar rato van hun verliezen. De resultaten van de winst- en verliestoerekening worden bewaard in de gain-loss-matrix.
Voor huishouden 1 betekent dit:

Gain-loss-matrix-huishouden 1

		winnende merken				
		C	D	E	F	Totaal
verliezende merken	A	24/11	32/11	16/11	16/11	8
	B	9/11	12/11	6/11	6/11	3
		3	4	2	2	11

De herhalingsaankopen per merk en de gain-loss-resultaten voor huishouden 1 leveren hun bijdrage aan de tabellen die het eindprodukt zijn van de analyse over de te onderzoeken groep van huishoudens.

2.4. Het model voor huishoudens die meer gaan kopen van het produkt.

Ter illustratie van de werking van het model in deze situatie, gaan wij uit van 'huishouden 2'. Op basis van de panelgegevens kan het volgende aankopenoverzicht worden samengesteld:

Aankopen huishouden 2

Merk	*Periode I*	*Periode II*
A	4	2
B	0	2
C	2	0
D	0	6
Totaal produkt	6	10

In deze situatie constateren wij dat er sprake is van meer aankopen van het produkt in periode II ten opzichte van periode I ter grootte van 4. De situatie wordt ook wel aangeduid met *meerkoop*.
In de meerkoopsituatie worden de herhalingsaankopen op dezelfde wijze bepaald als in paragraaf 2.3. Bij de verlies- en winsttoerekening ontstaat echter een probleem, aangezien de totale winst nu niet meer gelijk is aan het totale verlies.
Om dit probleem op te lossen, introduceren wij in het model het *imaginaire merk X* dat 4 eenheden (de omvang van de meerkoop) aan de markt verliest. Wij kunnen derhalve het volgende gemodificeerde aankopenoverzicht voor huishoudens 2 opstellen.

Aankopen huishouden 2

Merk	*Periode I*	*Periode II*
A	4	2
B	0	2
C	2	0
D	0	6
'X'	4	0
Totaal produkt	10	10

Op basis van dit aangepaste aankopenoverzicht, kan de analyse voor deze consumptieve eenheid op soortgelijke wijze worden uitgevoerd als voor huishoudens met een constant marktgedrag. Voor huishouden 2 betekent dit:

Huishouden 2

Merk	*Periode I*	*Periode II*	*Herhaling*	*Winst*	*Verlies*
A	4	2	2	0	2
B	0	2	0	2	0
C	2	0	0	0	2
D	0	6	0	6	0
'X'	4	0	0	0	4
Totaal produkt	10	10	2	8	8

Gemodificeerde gain-loss-matrix huishouden 2

		winnende merken		
		B	D	Totaal
verliezende merken	A	4/8	12/8	2
	B	4/8	12/82	2
	'X'	8/8	24/8	4
		2	6	8

De gemodificeerde matrix wordt nu opgesplitst in twee delen, namelijk:
- de gain-loss-matrix voor huishouden 2 die betrekking heeft op de winst- en verliestoerekening voor merken behorende tot het produktveld;
- het meerkoopoverzicht (winst ten koste van het buiten de markt gedefinieerde merk 'X').

Dit betekent dat de winst van de winnende merken gesplitst wordt in twee componenten, namelijk:
- de winst van andere merken;
- de winst door meerkoop.

De uitsplitsing van de gemodificeerde gain-loss-matrix leidt tot de volgende overzichten:

Gain-loss-matrix huishouden 2

		winnende merken		
		B	D	Totaal
verliezende merken	A	4/8	12/8	2
	C	4/8	12/8	2
		1	3	4

Winst door meerkoop huishouden 2

winnende merken		Totale winst door meerkoop
B	D	
8/8	24/8	4

Tot slot merken wij op dat in de meerkoop-situatie de volgende relaties gelden:
1. aankopen periode I = herhalingsaankopen + verlies
2. aankopen periode II = herhalingsaankopen + winst
   = herhalingsaankopen + winst van andere merken + winst door meerkoop

De herhalingsaankopen, de gain-loss-resultaten en de meerkoopgegevens leveren hun bijdrage aan de tabellen die het eindprodukt zijn van de analyse over de te onderzoeken groep.

2.5. Het model voor huishoudens die minder gaan kopen van het produkt

Voor huishouden 3 is op basis van het panelaankopenbestand het volgende samengesteld:

Aankopen huishouden 3

Merk	*Periode I*	*Periode II*
A	2	4
B	2	0
C	0	2
D	6	0
Totaal produkt	10	6

Deze situatie wordt aangeduid met de term *minderkoop*. In het concrete geval van huishouden 3 zien wij dat 4 eenheden van het produkt minder zijn gekocht.

Wij introduceren nu een imgainair merk 'Y', waaraan de markt 4 eenheden verliest, teneinde op basis van de onafhankelijkheidshypothese de winst- en verliestoerekening te kunnen uitvoeren.

Het gemodifceerde aankopenoverzicht voor huishouden 3 met de daaraan gekoppelde herhaling, winst- en verliesrekening ziet er nu als volgt uit:

Aankopen huishouden 3

Merk	*Periode I*	*Periode II*	*Herhaling*	*Winst*	*Verlies*
A	2	4	2	2	0
B	2	0	0	0	2
C	0	2	0	2	0
D	6	0	0	0	6
'Y'	0	4	0	4	0
Totaal produkt	10	10	2	8	8

Op basis van de laatste twee kolommen uit het bovenstaande overzicht, kan de gemodificeerde gain-loss-matrix worden opgesteld:

Gemodificeerde gain-loss-matrix-huishouden 3

		winnende merken			
		A	C	'Y'	Totaal
verliezende	B	4/8	4/8	8/8	2
merken	D	12/8	12/8	24/8	6
		2	2	4	8

Ook nu moet de gemodificeerde matrix worden opgesplitst in twee delen, namelijk:
- de gain-loss-matrix voor huishouden 3, welk betrekking heeft op de winst- en verliestoerekening voor de merken behorende tot het produktveld;
- het minderkoopoverzicht (verlies aan het buiten de markt gedefinieerde merk 'Y').

Dit betekent dat het verlies van de verliezende merken gesplitst wordt in twee componenten, namelijk:
- het verlies aan andere merken;
- het verlies door minderkoop.

De uitsplitsing van de gemodificeerde matrix leidt tot de volgende twee overzichten:

Gain-loss-matrix-huishouden 3

		winnende merken		
		A	C	Totaal
verliezende	B	4/8	4/8	1
merken	D	12/8	12/8	3
		2	2	4

Verlies door minderkoop huishouden 3

verliezende merken	B	8/8
	D	24/8
Totaal verlies door minderkoop		4

In de minderkoopsituatie gelden de volgende relaties:
1. aankopen periode I = herhalingsaankopen + verlies
   = herhalingsaankopen + verlies aan andere merken + verlies door minderkoop
2. aankopen periode II = herhalingsaankopen + winst.

De resultaten met betrekking tot de herhalingsaankopen, gain-loss en minderkoop leveren hun bijdrage aan de eindrapportage van de analyse over de te onderzoeken groep.

2.6. Het model voor nieuwe kopers

In de gain-loss-analyse worden met de term nieuwe kopers huishoudens die in periode I niet en in periode II wel het produkt kopen, aangeduid. Stel dat voor huishouden 4 het volgende aankopenoverzicht kan worden opgesteld:

Aankopen huishouden 4

Merk	*Periode I*	*Periode II*
A	0	2
B	0	3
Totaal produkt	0	5

In dergelijke situaties worden de aankopen die in de tweede periode gerealiseerd zijn, aangezien er geen herhalingsaankopen plaatsvinden, bestempeld als winst voor de aangekochte merken en afzonderlijk genoteerd als winst door nieuwe kopers.

Winst door nieuwe kopers		
winnende merken		
A	B	Totale winst door nieuwe kopers
2	3	5

2.7 Het model voor verloren kopers

Voor als volledigheid beschouwen we een huishouden welke uitsluitend in periode I aankopen binnen het produktveld verrichtte.
Voor het model wordt dit huishouden getypeerd als een verloren koper ongeacht het feit of na periode II al dan niet aankopen binnen het produktveld gedaan zijn.
Als voorbeeld gaan we uit van:

Aankopen huishouden 5

Merk	*Periode I*	*Periode II*
C	1	0
E	1	0
F	1	0
Totaal produkt	3	0

In deze situatie worden de aankopen in periode I, aangezien uitsluitend in deze periode gekocht is, tevens bestempeld als verlies voor de betrokken merken veroorzaakt door verloren kopers.

Verlies door verloren kopers		
verliezende merken	C	1
	E	1
	F	1
Totaal verlies door verloren kopers		3

2.8. Het model op geaggregeerd niveau

Binnen het gain-loss-programma wordt voor ieder huishouden bepaald tot welke van de 5 onderscheiden groepen kopende huishoudens het behoort. Vervolgens wordt per huishouden het bij de groep behorend rekenmodel doorgerekend, waarna de resultaten tezamen gebracht worden in een schema. De resultaten van de analyse zijn in de voorgaande paragrafen omkaderd en hebben betrekking op:
- herhalingsaankopen;
- gain-/loss-matrices;
- winsten door meerkoop;
- verliezen door minderkoop;
- winsten door nieuwe kopers;
- verliezen door verloren kopers.

De definitie die op het individuele niveau per merk en voor het produktveld gelden zijn ook geldig voor het geaggregeerde niveau. Zij luiden:

1. Aankopen in periode I = herhalingsaankopen + verlies, waarbij verlies onderverdeeld wordt naar oorzaak.
   Verlies = verlies aan andere merken + verlies minderkoop + verlies verloren kopers.
2. Aankopen in periode II = herhalingsaankopen + winst, waarbij winst onderverdeeld wordt naar oorzaak.
   Winst = winst van andere merken + winst meerkoop + winst nieuwe kopers.

Uitgaande van de geaggregeerde gain-loss-matrix waarin de winsten van en verliezen aan andere merken per merk gespecificeerd zijn, kan het totaal schema als volgt voorgesteld worden. De getalswaarden in het schema ontstaan hier als de som van de getalswaarden van de 5 huishoudens, die in de vorige paragrafen geïntroduceerd zijn.

		Winsten door						Totaal verlies t.g.v.				
		A	B	C	D	E	F	Andere merken	Minder-koop	Verloren kopers	Herhl. aank	Aank. per I
Verliezen door:	A	*	.5	2.18	4.41	1.45	1.45	10	0	0	4	14
	B	.5	*	1.32	1.09	.55	.55	4	1	0	2	7
	C	0	.5	*	1.5	0	0	2	0	1	2	5
	D	1.5	0	1.5	*	0	0	3	3	0	4	10
	E	0	0	0	0	*	0	0	0	1	6	7
	F	0	0	0	0	0	*	0	0	1	0	1
Totale winst t.g.v	Andere merken	2	1	5	7	2	2	19	4	3	18	44
	Meer-koop	0	1	0	3	0	0	4				
	Nieuwe kopers	2	3	0	0	0	0	5				
Herhalings aankopen		4	2	2	4	6	0	18				
Aankopen 8 periode II		7	7	14	8	2	46					

Aangezien in de praktijksituaties het aantal onderscheiden merken erg groot (bijv. 15) kan zijn is de computeroutput anders opgesteld en opgedeeld in twee delen.
In de situatie van het voorbeeld zou de computeroutput de volgende structuur te zien geven.

Blad 1
Gain-loss-analyse AGB Periode I:..... Periode II:.....

				VERLIES			WINST		
Merk	Per. I	Per. II	Herhl.	And. merk	Minder-koop	Verl. kopers	And. merk	Meer-koop	Nieuwe kopers
A	14	8	4	10	0	0	2	0	2
B	7	7	2	4	1	0	1	1	3
C	5	7	2	2	0	1	5	0	0
D	10	14	4	3	3	0	7	3	0
E	7	8	6	0	0	1	2	0	0
F	1	2	0	0	0	1	2	0	0
Tot.	44	46	18	19	4	3	19	4	5

Blad 2
Gain-loss-analyse AGB Periode I:..... Periode II:.....

Merk	A	B	C	D	E	F	Totaal
A	*	.50	2.18	4.41	1.45	1.45	10
B	.50	*	1.32	1.09	.55	.55	4
C	0	.50	0	1.50	0	0	2
D	1.50	0	1.50	*	0	0	3
E	0	0	0	0	0	*	0
F	0	0	0	0	0	0	0
Totaal	2	1	5	7	2	2	19

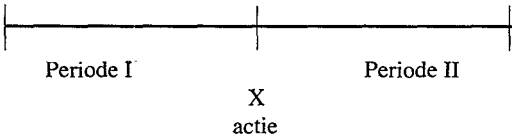
Op het eerste blad worden de basisgegevens voor de markt en het model weergegeven. Centraal staan hierbij de twee formules die aan het begin van paragraaf 2.8 gegeven zijn.
Op het tweede blad wordt de specificatie van de winst- en verliesrekening per merk gegeven. De randtotalen van deze winst- en verliestoerekening corresponderen als volgt met gegevens uit het eerste blad:
- rijtotalen blad 2 = verliezen aan andere merken blad 1;
- kolomtotalen blad 2 = winsten van andere merken blad 1.

Essentieel bij de interpretatie van de gevens zijn de gehanteerde modelvooronderstellingen.

3. HET GEBRUIK VAN DE AGB GAIN-LOSS-ANALYSE

De meeste markten voor 'fast-moving packaged consumer goods' worden gekenmerkt door intensieve marketing-inspanningen gericht op het beïnvloeden van consumenten in het koop- en keuzeproces. Ten gevolge van veranderingen in koopgewoonten zullen merken binnen markten enerzijds winst en anderzijds verlies incasseren.
Het analysemodel is derhalve geschikt om na te gaan wat de effecten op het koopgedrag zijn van bijvoorbeeld:
- introductie van nieuwe merken;
- herpositionering (relaunch) van merken;
- assortimentsuitbreidingen;
- speciale promotiecampagnes
- wijzigingen in distributiepolitiek ten behoeve van een bepaald merk.

De analyse toont aan, conform het model, wat het effect van de gepleegde marketing-inspanningen is, waarbij tevens rekening gehouden is met de effecten van ontwikkelde acties van concurrenten binnen de beschouwde markt. De keuze van de analyseperioden (even lang) hangt derhalve samen met het startsein van de geleverde marketing-inspanning. Schematisch weergegeven:



Bij het uitvoeren van de analyse zal de onderzoeker zichzelf beperkingen op moeten leggen in termen van:
- minimale lengte van de perioden, afhankelijk van de koopfrequentie in het produktveld;
- merkuitsplitsing, afhankelijk van het aantal te analyseren produktaankopen per periode, derhalve mede afhankelijk van de koopfrequentie en gehanteerde periodelengte.

De output is dusdanig weergegeven dat het wel en wee van merken overzichtelijk gepresenteerd kan worden in termen van:
- nieuwe/verloren *produkt*veld kopers;
- meer en minder kopen in het *produkt*veld;
- merkwisseling in termen van winsten en verliezen zoals dat per merk in totaliteit optreedt en ten opzichte van belangrijke concurrerende merken;
- herhalingsaankopen.

Ter afsluiting van dit artikel wordt een voorbeeld gegeven van de besproken analysetechniek, toegepast om de effecten van een gevoerde promotionele campagne voor melk en melkprodukten bij een bepaald afzetkanaal te bepalen. In plaats van het analyseren van het koopgedrag naar merken vindt derhalve nu de analyse van de aankopen naar de afzetkanalen plaats. De afzetkanalen worden aangegeven met behulp van de codes A tot en met G.
De perioden waarop de analyse betrekking heeft zijn twee niet aansluitende 12-wekelijkse perioden in 1986. In de tussenliggende periode werd door organisatie A een campagne gevoerd voor het genoemde produktveld. De resultaten, welke betrekking hebben op een steekproef van 1.700 huishoudens, zijn weergegevens in het hierna volgende overzicht[2].

** GAIN-LOSS-ANALYSE ** AGB ** PERIODE I : 8613-8624 PERIODE II: 8637-8648 OMVANG STEEKPROEF: 1700
MELK EN MELKPRDUKTEN, HOEVEELHEDEN IN LITERS

						VERLIES			WINST		
Kanaal	Per. I	MA %	Per. II	MA %	Herh.	Andere kanalen	Minder-koop	Verloren kopers	Andere kanalen	Meer-koop	Nieuwe kopers
A	20.600	19	24.158	22	17.775	1.245	1.576	4	3.537	2.829	17
B	14.451	13	13.080	12	11.045	1.979	1.419	8	840	1.195	1
C	9.482	9	9.218	8	7.291	1.212	979		925	994	8
D	18.372	17	18.448	17	14.832	1.921	1.619		1.826	1.771	19
E	18.500	17	18.488	17	15.134	1.815	1.539	12	1.877	1.473	4
F	17.239	16	17.074	15	14.183	1.673	1.383		1.431	1.460	1
G	12.044	11	11.045	10	9.667	1.417	960		827	551	
Totaal	110.689	100	111.511	100	89.926	11.263	9.476	25	11.263	10.272	49

GAIN-LOSS MATRIX

Kanaal	A	B	C	D	E	F	G	Verlies andere kanalen
A		201	90	291	348	197	118	1.245
B	661		277	416	409	138	78	1.979
C	310	85		285	149	182	200	1.212
D	922	154	152		387	182	124	1.921
E	579	214	189	263		444	126	1.815
F	742	82	95	271	303		180	1.673
G	322	103	123	300	281	289		1.417
Winst andere kanalen	3.537	840	925	1.826	1.877	1.431	827	11.263

Hieruit blijkt dat de gevoerde campagne als succesvol bestempeld kan worden. Het marktaandeel voor A is, binnen het beschouwde produktveld, toegenomen tot 22% terwijl de afzetkanelen B en G en belangrijke deel van hun omzet in de loop van de tijd zijn kwijtgeraakt.
De oorzaken voor verschuiving in de marktpositie van A binnen het produkt moeten duidelijk niet gezocht worden in het vertoonde gedrag van verloren en nieuwe kopers. Een relatief gering verlies bij de minderkoopsituaties en een anderzijds relatief hoge winst bij meerkoop, verklaren voor een belangrijk deel de verstevigde positie van A binnen het produktveld.
Voor een nog groter deel moet de positieverbetering gezocht worden in een groei ten koste van andere afzetkanalen.

Uit de Gain-loss-matrix blijkt voor A het volgende:

Verlies aan en winst van andere afzetkanelen door A

Kanaal	*Verlies*	*Winst*	*Saldo (W-V)*
B	201	661	+ 460
C	90	310	+ 220
D	291	922	+ 631
E	348	579	+ 231
F	197	174	+ 545
G	118	322	+ 204
Totaal	1.245	3.537	+ 2.292

Hieruit volgt dat per saldo van ieder concurrerend afzetkanaal terrein gewonnen is waarbij vooral B, D en F een belangrijke bijdrage geleverd hebben.
De analyse toont voorts overduidelijk aan dat bijvoorbeeld in de situatie van D, waarbij op basis van het marktaandeel in beide perioden gesproken kan worden over een stabiele situatie, er toch op basis van de analyse van het gedrag van afzonderlijke huishoudens de nodige dynamiek in de opbouw van het marktaandeel blijkt.

Het unieke karakter van het model bestaat uit het beschrijven van veranderingen in markten op basis van de analyse van gedragsveranderingen op het individuele niveau. De gain-loss analyse is derhalve een geschikt instrument om op bezit van consumenten panelgegevens de effecten van marketinginspanningen te volgen.

NOTEN

1. Zie bijvoorbeeld Van der Zwan en Verhulp (1980, pag. 510 e.v.)
2. De cijfers in de tabel zijn afgeronde grootheden. Derhalve behoeft de som van de gepresenteerde cijfers in een kolom niet exact overeen te komen met het afgeronde kolomtotaal.

LITERATUUR

Chatfield C., G.J. Goodhart (1970), The Beta-binominal model for consumer purchasing behaviour, *Applied Statistics*, Vol. 19, pp. 240-250.

Chatfield C., A.S.C. Ehrenberg, G.J. Goodhardt (1966), Progress on a simplified model of stationary purchasing behaviour, *Journal of the Royal Statistical Association*, Vol. 129A, pp. 317-367.

Chatfield C., G.J. Goodhardt (1973), A Consumer purchasing model with erlang interpurchase times, *Journal of the American Statistical Association*, Vol. 68, pp. 828-835.

Ehrenberg A.S.C. (1972), *Repeat buying: Theory and applications* North-Holland.

Goodhardt G.J., A.S.C. Ehrenberg, C. Chatfield (1984), The Dirichlet: A Comprehensive Model of Buying Behaviour, *Journal of the Royal Statistical Association*, Vol. 147A, pp. 621-655.

Herniter J. (1971), A probalistic market model of purchase timing and brand selection, *Management Science*, Vol. 18, pp. 102-103.

Jones J.M., F. Zufryden (1980), Adding explanatory variables to a consumer purchase behaviour model: An exploratory study, *Journal of Marketing Research*, Vol. 17, pp. 323-334.

Jones J.M., F. Zufryden (1982), An approach for assessing demographic and price influences on brand purchase behaviour, *Journal of Marketing*, Vol. 46, pp. 36-46.

Leeflang P.S.H., A. Boonstra (1982), Some comments on the development and application of linear learning models, *Management Science*, Vol. 28, pp. 1233-1246.

Verhulp J., J.P. Wijnberger (1982), *Aankoopstructuren: Verslag van een panelstudie*, Rapport 8203/M, Centrum voor Bedrijfseconomisch Onderzoek, EUR.

Verhulp J., J.P. Wijnberger (1983), *Aankoopstructuren (2), nader getoetst*, Rapport 8327/M, Centrum voor Bedrijfseconomisch Onderzoek, EUR.

Wijnberger, J.P. (1980), *Statistische analyse van het aankoopgedrag in consumentenpanels: verslag van een onderzoek bij een drietal produkten.* Discussienota 8015/M, Centrum voor Bedrijfseconomisch Onderzoek, EUR.

Wierenga B. (1974), *An Investigation of Brand Choice Processes*, Rotterdam, University Press.

Zufryden F. (1978), An empirical evaluation of a composite heterogeneneous model of brand choice and purchase timing behaviour, *Management Science*, Vol. 24, pp. 761-773.

Zwan, A. van der, J. Verhulp (1980), *Grondslagen en techniek van de Marktanalyse*, Stenfert Kroese.

5. Een meetsysteem voor gegevens ten behoeve van de constructie van een produktselectie model

B.J.J.A.M. BRONNENBERG

SAMENVATTING

Aan de Faculteit der Bedrijfskunde van de Universiteit Twente wordt door de auteur een empirisch onderzoek uitgevoerd naar de voorspelbaarheid van succes en falen van nieuwe produkten. Doel hiervan is te komen tot een model waarmee de selectie van nieuwe produkten ondersteund kan worden door een voorspelling omtrent het commerciële resultaat van een nieuwe produkt. Dit artikel gaat in op het meten van gegevens in empirisch onderzoek en op de invloed die de gehanteerde methodologie en de herkomst van die gegevens op de meting ervan heeft.
In een conceptueel model wordt het succes van falen van nieuwe produkten afhankelijk gesteld van 'bedrijfsmiddelen', 'marktomstandigheden' en 'produkteigenschappen'. De variabelen die binnen deze drie entiteiten te onderscheiden zijn, alsmede het succes en falen zelf zullen, ter calibratie van het model gemeten moeten worden binnen ondernemingen. Hiervan en van de calibratie wordt een voorbeeld gepresenteerd. De meting van deze, in principe subjectieve, gegevens vindt in eerste instantie op a posteriori basis plaats. Hierna zal een cyclisch proces van meting, calibratie en validatie op a priori basis worden onderhouden. Het te presenteren meetsysteem heeft t.o.v. eerdere studies voordelen op het gebied van met name validatie en onafhankelijkheid van gegevens. Het artikel behandelt tevens vijf probleemgebieden m.b.t. gegevensverwerving in deze specifieke studie.

1. INLEIDING

In dit artikel staat de modelbenadering bij de selectie van nieuwe produkten centraal. Hiermee wordt in dit verband bedoeld dat beslissingen van het management omtrent de selectie van produkten ondersteund worden met een model.
Een van de gevolgen van de voortschrijdende informatie-technologie is dat informatie steeds toegankelijker wordt. De hoeveelheid informatie die bij besluitvorming ter beschikking staat wordt navenant groter. Het mag dan ook verwacht worden dat de trend naar grotere beschikbaarheid van gegevens zal resulteren in nog grotere complexiteit van beslissingen. Dientengevolge ontstaat de behoefte aan een bruikbaar instrument om de grote hoeveelheid gegevens 'in te dikken' tot gestructureerde en doelgerichte informatie.
De mogelijkheden om een bruikbaar managementinstrument te construeren zijn historisch gezien steeds groter geworden. Allereerst is er een duidelijke toename te constateren van de mogelijkheden die geboden worden door rekenapparatuur m.b.t. het construeren van modellen. Door deze ontwikkeling is het mogelijk steeds complexere modellen te construeren, die probleemgebieden op een minder vereenvoudigde, dus reëlere wijze kunnen weergeven.

Daarnaast komen Lilien en Kotler (1983) tot een vijftal trends, waarom bedrijven in toenemende mate gebruik gaan maken van een modelachtige aanpak bij (in dit geval) marketingbeslissingen:

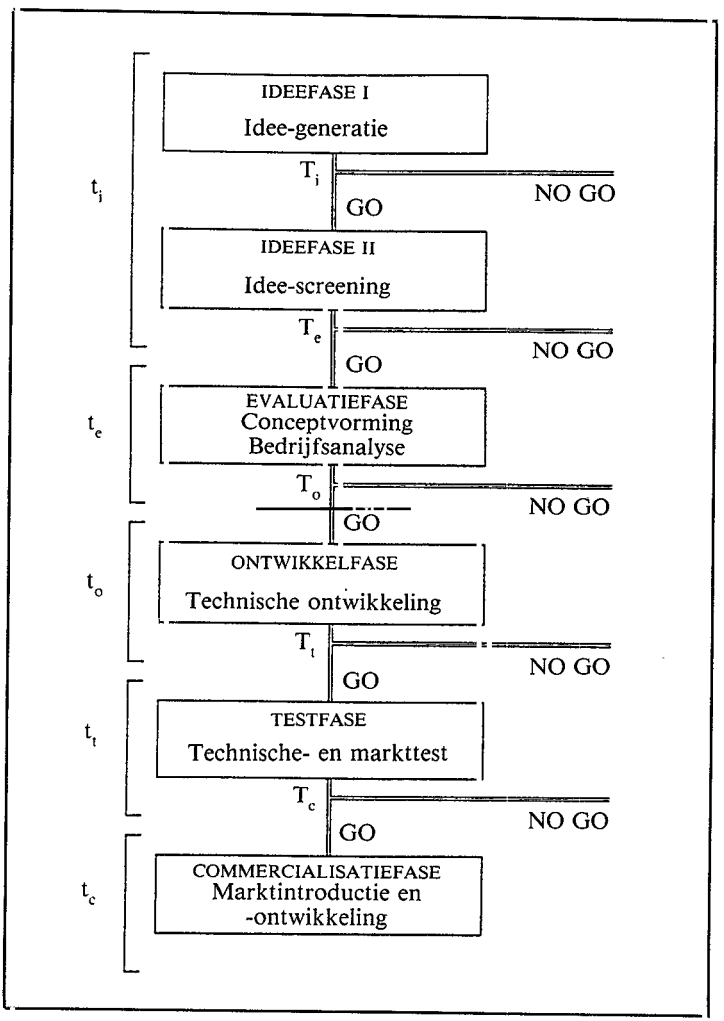
- Marketing decision support systemen worden ontwikkeld en raken in gebruik. De hoeveelheid en kwaliteit van gegevens nemen explosief toe en de eisen aan de rapportagetijden worden steeds hoger.
  Marketing modellen voorzien in instrumenten om gegevens samen te vatten, te interpreteren en te normaliseren.
- Kwantitatief opgeleide bedrijfskundigen en bedrijfseconomen krijgen langzamerhand leidinggevende posities binnen het bedrijfsleven. Deze nieuwe generatie van managers is opgeleid in en staat open voor decision support systemen.
- De theorievorming omtrent marketing is de laatste jaren aanmerkelijk verbeterd hetgeen leidt tot verbeterde modelvorming.
- De waarde van marketingmodellen wordt de laatste jaren gerapporteerd en vormt daardoor een stimulans voor bredere toepassing.
- Marketingbudgetten worden de laatste decennia steeds ruimer.
  Hierdoor wordt de behoefte van het management tot verhoogde produktiviteit steeds groter.

Het probleemgebied 'Nieuwe Produkten' is de laatste jaren in hevige ontwikkeling. Allereerst is te constateren dat de sterfte-ratio van produktideeën vanaf 1968 tot 1982 zeer sterk is gedaald. In een door Booz, Allen en Hamilton uitgevoerd onderzoek werd in 1968 een sterfte-ratio van 60:1 gevonden. Deze ratio wil zeggen dat 60 produktideeën uiteindelijk één succesvol produkt voortbrengen. Volgens Booz, Allen en Hamilton (1982) is de verhouding tussen produktideeën en uiteindelijk succesvolle produkten in 1982 over de jaren heen gedaald tot 7:1. Enige voorzichtigheid bij deze conclusies is geboden omdat de auteurs niet vermelden of deze verbetering daadwerkelijk gerealiseerd wordt door de introductie van meer succesvolle produkten. Volgens de onderzoeksgegevens van Booz, Allen & Hamilton is het aantal succesvolle produkten in de periode tussen 1968 en 1982 licht gedaald van 67% naar 65% van alle introducties. Het verloop van het aantal introducties tussen 1968 en 1982 wordt niet vermeld.
Nieuwe produkten zijn van vitaal belang voor de groei en continuïteit van de meeste producerende organisaties (Cooper en Kleinschmidt, 1987a). Nieuwe produkten zijn blijkens het onderzoek van Booz, Allen en Hamilton (1982) ook zeer risico-vol; ongeveer 46% van de middelen die worden toegewezen aan produktontwikkeling in brede zin, gaan naar falende produkten die ongeveer 35% van het aantal introducties betreffen (zie ook Crawford, 1979).

Samenvattend kan men stellen dat het 'klimaat' voor het construeren van, en het werken met modellen steeds gunstiger is geworden gedurende de laatste jaren en dat er op voorhand geen reden is om aan te nemen dat hierin voorlopig verandering komt.
Tevens is er t.a.v. nieuwe produkten een zeer gunstig klimaat voor de modelbenadering, voornamelijk omdat de behoefte aan een modelbenadering zo groot lijkt. De laatste 20 jaar heeft namelijk geen verbetering laten zien in het percentage succesvolle produkten. Het ligt dan ook voor de hand een instrument te ontwikkelen waarbij de selectiemethode van nieuwe produkten verbeterd kan worden door het commerciële resultaat van een produktidee te voorspellen.

2. DOEL VAN EEN MODEL TEN BEHOEVE VAN PRODUKTSELECTIE

Het model om nieuwe produkten te selecteren gaat uit van het beoordelen van een produktidee, waarbij men de keuze heeft dit al dan niet verder te ontwikkelen. Deze zgn. 'go-on go' beslissing zal bij een positief oordeel omtrent het produktidee nog een aantal malen genomen moeten worden. Het gaat in dit artikel echter voornamelijk om de vroege tot zeer vroege beslissingen. Bekijkt men het produktontwikkelingstraject, zoals dat wordt gehanteerd door Booz, Allan en Hamilton (1982), dan gaat de interesse voornamelijk uit naar de eerste twee à drie stadia van het produktontwikkelingstraject.



Figuur 1. Het Produktontwikkelingstraject.
t_i, enz.: periode tussen twee GO/NO GO beslissingen.
Ti, enz.: tijdstip van een GO/NO GO beslissing.

Het model om van nieuwe produktideeën te voorspellen of het commercieel aantrekkelijk danwel onaantrekkelijke ontwikkelingen zijn heeft de volgende doelstellingen:

Systematische stemmingsmeter
Bij de oordeelvorming omtrent de haalbaarheid van een nieuw produkt zal informatie omtrent de marktomstandigheden, bedrijfsmiddelen en de gewenste produkteigenschappen moeten worden verzameld, gecombineerd en beoordeeld.
In een onderzoek, dat in 1986/1987 werd uitgevoerd aan de Universiteit Twente werd een model getest van R.G. Cooper waarbij gegevens over nieuwe produkten d.m.v. een vragenlijst werden opgevraagd aan een team van bedrijfsmedewerkers. Een van de eisen aan dit team was dat het moest zijn samengesteld uit vertegenwoordigers van een aantal disciplines; bijv. een algemeen manager, een verkoopleider, een produktiechef, een marketing manager, iemand van R & D, etc. Deze personen leverden antwoorden op een 48-tal vragen, die op voorhand wezenlijk leken te zijn in het onderscheid tussen succes en falen van een nieuw produkt. Door middel van tabellering en vergelijking van deze vragen kan men zodoende hiaten in kennis of informatie opsporen of meningsverschillen detecteren. Speciaal indien men bedenkt dat een dergelijk team tevens het produkt binnen het bedrijf moet 'verkopen', is het vaststellen van *meningen* omtrent de ingeschatte levenskansen èn de discussie over de geconstateerde verschillen van meningen en belangrijke doelstelling die met behulp van een model gerealiseerd kan worden.

Correctie tijdens het produktontwikkelingstraject
Een van de functies van een goed beslissingsmodel is dat aan de hand van de gegevens de sterkten en zwakten van een bepaald produktidee kunnen worden onderkend. Aangezien het screenen van produkten in een vroeg stadium plaatsvindt, bestaat de mogelijkheid om bij gebleven zwakten, indien deze door het bedrijf te verbeteren zijn, de ontwikkeling bij te sturen.

Doelgerichte toewijzing van middelen
In een ideaal situatie zouden middelen t.b.v. ontwikkeling en commercialisering van nieuwe produkten slechts toegewezen worden aan commercieel succesvolle produkten. Zoals reeds in de inleiding werd opgemerkt is de praktijk anders en wordt 46% van de middelen toegewezen aan falende produkten. Een laatste doel van een model voor produktselectie is, om door middel van een voorspelling omtrent het commerciële resultaat, middelen meer gericht aan succesvolle produkten toe te wijzen.

3. KORTE SCHETS VAN DE ONDERZOEKSOPZET

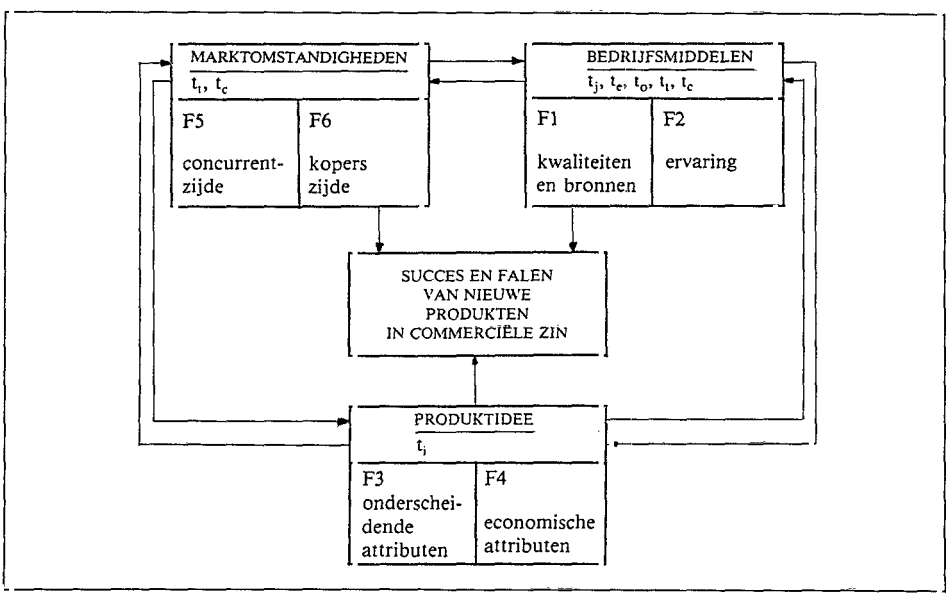
Het onderzoek zal op empirische wijze worden uitgevoerd. Dit betekent dat er sprake zal zijn van data-acquisitie. Allereerst zal een conceptueel model worden gebruikt om een lijst van indicatoren vast te stellen die het succes of falen van nieuwe produkten zou kunnen verklaren. Tevens worden de relaties uit dit model gebruikt voor hypothesevorming. Per deelgebied wordt een inventarisatie gemaakt van indicatoren die op basis van de theorie van belang zijn met betrekking tot het onderscheid van succes en falen.
Vervolgens wordt de lijst van indicatoren alsmede een maat voor succes en falen geme-

ten in de praktijk. Op de meting in de praktijk wordt in het onderstaande dieper ingegaan.
Tot slot zullen de gegevens worden geanalyseerd en zullen causale relaties tussen de indicatoren enerzijds en het succes of falen anderzijds worden vastgelegd. Tevens zullen hypothesen worden getest op hun juistheid.
Hieronder wordt een conceptueel model gepresenteerd waarin relaties worden verondersteld m.b.t. het succes en falen van nieuwe produkten. Dit concept wordt gebruikt om de vraag naar te meten gegevens in dit specifieke geval te formuleren.

4. EEN CONCEPTUEEL MODEL

4.1. Indeling en onafhankelijke variabelen

In dit conceptuele model wordt kwalitatief vastgelegd welke verbanden worden verondersteld tussen bepaalde onafhankelijke factoren enerzijds en het succes of falen van nieuwe produkten anderzijds. Het concept telt drie entiteiten: marktomstandigheden, bedrijfsmiddelen en produktidee, die het commerciële resultaat van een nieuw produkt bepalen.
In het conceptuele model dienen de marktomstandigheden en bedrijfsmiddelen in strikte relatie tot het produktidee gezien te worden.
Voor de marktomstandigheden geldt dat de relatie tot succes en falen van nieuwe produkten vooral tot uiting komt in de laatste fasen van het ontwikkelingstraject. De marktomstandigheden die gelden bij de commercialisatie, en niet die bij de ideefase, van een nieuw produkt zullen namelijk het slagen of mislukken mede bepalen.



Figuur 2. Het conceptuele model
***N.B.** De fasen die vermeld staan bij de drie hoofdgebieden hebben betrekking op de fasen waarin de, in deze studie, relevante gegevens over de corresponderende gebieden bekend worden.*

Tabel 1. Invulling van het conceptuele model[1]

DEELGEBIEDEN	VARIABELEN
Bedrijfsmiddelen	
F1: Kwaliteiten en Bronnen	1. financiële bronnen 2. R & D vaardigheden 3. ontwerp vaardigheden 4. management vaardigheden 5. marketing research vaardigheden 6. produktie vaardigheden 7. verkoop en distributie 8. reclame en promotie
F2: Ervaring	1. potentiële klanten nieuw 2. nieuwe produktklasse voor markt 3. nog nooit gemaakt/verkocht 4. nieuw produktieproces 5. nieuwe R & D 6. nieuw distributiesysteem 7. nieuwe promotie 8. nieuwe concurrenten
Produkteigenschappen	
F3: Onderscheidende Attributen	1. unieke kenmerken van produkt 2. produkt is superieur t.o.v. concurr. 3. prod. heeft unieke taak voor klant 4. hogere kwaliteit dan concurrentie 5. innovatief produkt
F4: Economische Attributen	1. lagere prijs per eenheid 2. kosten klant reducerend 3. hogere prijs dan concurrentie
Marktomstandigheden	
F5: Aanbodzijde	1. intensiteit concurrentie 2. intensiteit prijsconcurrentie 3. aantal concurrenten 4. één dominante concurrent? 5. toetredingsintensiteit concurrenten 6. introductiefrequentie van conc. prod.
F6: Vraagzijde	1. massa markt vs. één klant 2. grote behoefte van klanten 3. potentiële of reële vraag? 4. marktomvang in geld 5. produkttrouw (bij concurrenten) 6. tevredenheid (bij concurrenten) 7. is de markt communicat. bereikbaar? 8. marktgroei 9. dynamiek in de markt.

Voor de bedrijfsmiddelen geldt de relatie tot succes of falen gedurende het gehele ontwikkelingstraject en gedurende de totale commercialisatiefase. Wel is het zo dat binnen de groep van bedrijfsmiddelen accentverschuivingen zijn te constateren in belangrijkheid. Zo zal het accent in de ontwikkelingsfase meer gericht zijn op technische middelen dan in de commercialisatiefase, enz.
De drie entiteiten zijn onderverdeeld in telkens 2 deelgebieden. Bij 'marktomstandigheden' wordt, in het kader van succes of falen van nieuwe produkten, onderscheid gemaakt tussen variabelen die de aanbodzijde (concurrenten) van de markt beschrijven en variabelen die de vraagzijde (kopers) van de markt beschrijven.
Bij 'bedrijfsmiddelen' wordt onderscheid gemaakt tussen de kwaliteiten en bronnen enerzijds en de ervaring waarop een onderneming, ten aanzien van een project, kan bogen anderzijds. De eerste categorie meet in hoeverre de behoeften van een project en de vrij te maken middelen van een bedrijf op elkaar aansluiten. De tweede categorie meet de nieuwheid van het project voor het bedrijf.
Tot slot wordt de entiteit 'produktidee' gesplitst in economische attributen en onderscheidende attributen. De eerste groep attributen heeft betrekking op de kostenaspecten voor de klant, hetzij in gebruikerskosten, hetzij de aanschafkosten. De onderscheidende attributen beschrijven de uniekheid, kwaliteit, innovativiteit, etc. Technische attributen zijn in het conceptuele model als afzonderlijke groep buiten beschouwing gelaten om twee redenen. Allereerst is een aantal technische attributen vertegenwoordigd in de bovengenoemde onderscheidende groep van attributen. Ten tweede wees een eerder, aan de Universiteit Twente uitgevoerd, onderzoek uit dat deze groep niet tot de meest belangrijke groep variabelen kan worden gerekend in het *onderscheid* van succes of falen van nieuwe produkten (Bronnenberg, 1987).
In tabel 1 volgt een opsomming van mogelijke variabelen per deelgebied. De opsomming is niet uitputtend doch slechts als indicatie bedoeld. Uiteindelijk werden 48 verschillende onafhankelijke variabelen in de studie betrokken.

4.2. Afhankelijke variabelen

De afhankelijke variabele van het conceptuele model is het *succes* of *falen* van nieuwe produkten. Er is in principe een groot aantal mogelijkheden om het succes en falen van nieuwe produkten te omschrijven. Kotler (1967) onderscheidt bijvoorbeeld drie vormen van falen. Hij omschrijft een 'absolute failure' als een produkt waarvan de revenuen de vaste en variabele kosten niet dekken, een 'partial failure' als een produkt, waarvan de revenuen de vaste kosten ten dele dekken maar waar nog steeds verlies op wordt geleden (dus een positieve dekkingsbijdrage, doch op integrale kostprijsbasis verliesgevend) en een 'relative failure' als een produkt waarmee weliswaar winst wordt gerealiseerd maar waarbij een gestelde winstdoelstelling niet wordt bereikt.
In het reeds eerder aangehaalde onderzoek in Twente werden twee definities van succes en falen gehanteerd en vergeleken. Deze definities luidden:
- de mate van succes/falen is de mate waarin met dit produkt de minimaal acceptabele winstgevendheid van dit soort van produkt en investering wordt overschreden/achterblijft (nulpunt is relatief) (Cooper, 1981)
- de mate van succes is de mate waarin de inkomsten met dit produkt de uitgaven ervan overtreffen (nulpunt is gelijk aan het break-even punt)

Beide definities zijn in de praktijk gebruikt om de succesmaat te meten en zijn vergele-

ken op het citerium eenduidigheid. Van de beide definities bleek de tweede beter te functioneren in die zin, dat het een consistentere schatting gaf van (de mate van) succes en falen van nieuwe produkten[2].

5. EEN EERSTE AANZET TOT MODELVORMING; HET TWENTSE ONDERZOEKSPROJECT

Aan de Faculteit der Bedrijfskunde van de Universiteit Twente werd zoals reeds gesteld in 1987 een onderzoek uitgevoerd waarvan het hoofddoel was het NewProd model[3] te testen op betrouwbaarheid en validiteit. De tijdens deze testfase ingewonnen getalsmatige informatie over 9 succesvolle en 10 falende produkten, bood de mogelijkheid om een onafhankelijk onderzoek uit te voeren naar de voorspelbaarheid van succes en falen van nieuwe produkten. Het conceptueel model zoals hierboven beschreven werd hierbij als uitgangspunt genomen.

5.1. De meting van variabelen

Het meten van gegevens ten behoeve van het testen van het NewProd model (en dus tevens ten behoeve van de constructie van het Twentse model) vond plaats bij negen middelgrote en grote ondernemingen, opererend op zowel industriële- als consumentenmarkt. Deze ondernemingen waren werkzaam in de machinebouw (1), electronische industrie (1), chemische industrie (2), verpakkingsindustrie(1), zuivelindustrie (1), kunststof verwerkende industrie (2) en in de textielindustrie (1).
De (48) onafhankelijke variabelen van het NewProd model werden met behulp van een schriftelijke enquête op a posteriori basis gemeten. Elke variabele werd in stellingvorm geformuleerd en kon bevestigd of ontkend worden op een [0, ..., 10] Likert-scale. De gebruikte afhankelijke variabele, op basis van het break-even criterium[4], werd gemeten op een [-5, ..., +5] schaal, waarbij negatieve cijfers duidden op een falend produkt en positieve cijfers op een succesvol produkt.
Per produkt werd een team samengesteld met vertegenwoordigers van diverse disciplines binnen het bedrijf om zoveel als mogelijk de kennis binnen het bedrijf te benutten. Een 'screeningsteam' bestond in het algemeen uit een verkoopleider, produktiemanager, algemeen manager of projectmanager, een vertegenwoordiger van R & D en een vertegenwoordiger van de financiële afdeling van de betreffende onderneming. Deze vijf personen waardeerden *individueel* de variabelen in plenaire sessies, waarbij de onderzoekers aanwezig waren. Vanwege de verschillen in functies van de respondenten, werd het waarschijnlijk geacht dat bepaalde variabelen buiten de oordeelkundigheid van een individuele respondent zouden vallen. In de vragenlijst werd hiermee rekening gehouden, door de invoering van de 'mate van zekerheid' (ook op een [0, ..., 10] schaal), waarmee respondenten kunnen aangeven hoe zeker zij ervan zijn dat hun gegevens antwoord juist is. Deze maten van zekerheid dienden als weegfactoren bij het totaaloordeel van een team dat derhalve werd geconstrueerd als de gewogen som van de individuele beoordelingen.

5.2. Selectie van variabelen

Er werden per case 48 variabelen gemeten die benodigd zijn voor de toepassing van het

NewProd model. Tabel 1 wordt grotendeels uit deze variabelen samengesteld. Per gebied werd gezocht naar één maatstaf, die kon bestaan uit maximaal twee variabelen, waarmee dat gebied beschreven kon worden.
Er diende derhalve een beperkt aantal variabelen te worden geselecteerd, die bovendien evenwichtig zijn verdeeld over de zes gebieden van het conceptuele model. De selectie van variabelen vond plaats aan de hand van de volgende criteria (in volgorde van belangrijkheid):

1. De variabelen in het model dienen goed meetbaar en eenduidig van betekenis te zijn.
2. De geselecteerde variabelen dienen zo hoog mogelijk te correleren met zoveel mogelijk niet geselecteerde variabelen. Hierdoor wordt een kleine set variabelen geselecteerd, die een grotere set variabelen (grotendeels) beschrijft.
3. De geselecteerde variabelen dienen zo min mogelijk te correleren met andere geselecteerde variabelen. Het aantal correlaties met een absolute waarde groter dan 0.5 mag niet meer bedragen dan 20% van het totaal aantal correlaties. Dit is overigens een arbitrair criterium.
4. De geselecteerde variabelen dienen een significante relatie op te leveren. Na de toe-

Tabel 2: Definiëring en meetwijze van onafhankelijke variabelen uit het model.

	geheel mee oneens	*geheel mee eens*	*mate van zekerheid*
Bedrijfsmiddelen			
F1.2 De vaardigheden op het gebied van onderzoek en ontwikkeling zijn in ons bedrijf meer dan voldoende voor dit project	0 1 2 3 4 5 6 7 8 9 10		_____
F1.4 Het management van ons bedrijf heeft meer dan voldoende vaardigheden en middelen om het project te beheersen	0 1 2 3 4 5 6 7 8 9 10		_____
F2.3 We hebben nog nooit produkten gemaakt of verkocht om in deze behoefte van mogelijke klanten te voorzien	0 1 2 3 4 5 6 7 8 9 10		_____
Produktidee			
F3.1 Vergeleken met produkten van concurrenten zal ons produkt als enige een aantal unieke kenmerken en/of voordelen hebben voor de klant	0 1 2 3 4 5 6 7 8 9 10		_____
F4.3 Ons produkt zal significant duurder zijn dan produkten van concurrenten	0 1 2 3 4 5 6 7 8 9 10		_____
Marktomstandigheden			
F5.2 De markt van dit produkt wordt gekenmerkt door hevige prijsconcurrentie	0 1 2 3 4 5 6 7 8 9 10		_____
F6.4 De verwachte omvang van deze markt is in geld uitgedrukt zeer groot	0 1 2 3 4 5 6 7 8 9 10		_____
F6.9 De markt is dynamisch m.a.w. de behoefte van klanten verandert snel	0 1 2 3 4 5 6 7 8 9 10		_____

passing van de eerste drie criteria bleef een aantal variabelen per deelgebied over die werd gebruikt voor de calibratie van het conceptuele model. Met deze variabelen werd een roulerende procedure uitgevoerd waarin steeds één deelgebied centraal staat. Van dit deelgebied wordt de variabele bepaald die het meest significante model oplevert, gegeven de andere functie gebieden. Hierdoor wordt het 'startmodel' verbeterd. Bij herhaling van deze procedure treedt er uiteindelijk geen verbetering meer op.

Aan de hand van de bovenstaande criteria is een set van variabelen geselecteerd die staat weergegeven in tabel 2 (zie pag. 66). De gebieden F1 en F6 worden ieder door twee variabelen beschreven. Aan de hand van de correlatie van de variabelen met het succes van een produkt wordt bij 'Kwaliteiten en Bronnen van het bedrijf (F1)' de score op F1.2 en F1.4 opgeteld en bij 'Vraagzijde van de Markt (F6)' de score van F6.4 en F6.9 van elkaar afgetrokken. Deze samengestelde variabelen worden in het vervolg als één variabele per deelgebied beschouwd.

5.3. Data-analyse en Validatie

Met behulp van multiple regressie-analyse werden de parameter van het conceptueel model geschat. Omdat standaardisatie van variabelen de toepasbaarheid van dit model, zonder kennis omtrent de verdelingen van de variabelen, onmogelijk maakt worden hier ten behoeve van de lezer de parameters op basis van de ongestandaardiseerde variabelen gepresenteerd.
De resultaten van de regressie-analyse waren als volgt:

1. De regressieverlijking o.b.v de ongestandaardiseerde variabelen luidt:
   $PS_{break\text{-}even} = -0.1 + 0.3*F1 - 0.3*F2 + 0.5*F3 - 0.3*F4 - 0.7*F5 + 0.4*F6$,
   waarin $PS_{break\text{-}even}$ gelijk is aan de mate waarin de inkomsten met dit produkt de uitgaven ervan overtreffen [-5, ..., +5).
2. De verklaarde variantie van de regressievergelijking is 90%. De standard error of estimate is 1.2 op een schaal van –5 tot +5.
3. Het aantal hoge correlaties ($|r| \geq 0.5$) tussen variabelen in het model is gelijk aan 3. Bij 6 variabelen zijn 15 correlaties mogelijk zodat aan de randvoorwaarde van ten hoogste 20% hoge correlaties is voldaan. Hoge correlaties zijn geconstateerd tussen F1 en F6 (+), tussen F2 en F5 (–) en tussen F3 en F6 (+).
   De set van variabelen in het model correleert hoog met 15 variabelen van tabel 1 die niet geselecteerd werden. Hiermee werd aan het tweede selectiecriterium tegemoet gekomen.
4. De significantie van het totale model, onder de nulhypothese dat alle parameters gelijk zijn aan nul, is groter dan 99.9% (F-test met een ratio van 11.7)
5. De tekens van de parameters lijken intuïtief gezien allen juist.

De validatie van het model vond plaats m.b.v. een op de methode van Lachenbruch (van Frederikslust, 1976) gelijkende procedure. Deze procedure werkt op basis van een roulerend principe. Er wordt een regressievergelijking geschat voor elke steekproef, die wordt verkregen door van de oorspronkelijke steekproef steeds één case weg te laten. De gegevens die weg zijn gelaten worden vervolgens gesubstitueerd in de geschatte regressievergelijking en de voorpelling wordt vergeleken met de werkelijke commerciële uitkomst van het produkt. Zodoende kan op eenvoudige wijze een indruk worden ver-

kregen omtrent de juistheid van classificatie van het model. Men schat dus uiteindelijk evenveel regressievergelijkingen als er cases zijn. Omdat in tegenstelling tot bijv. cross-split-half validatie de databank hier niet in twee gelijke delen wordt gesplitst, op basis waarvan calibratie van parameters moet plaatsvinden, is deze validatiemethode bij uitstek geschikt voor validatie van het model op basis van kleine steekproeven.

6. De validatieresultaten wezen uit dat er van de 19 cases 17 juist werden geclassificeerd (91%). De fouten in classificatie betroffen beide falende produkten, die als 'succesvol' werden geclassificeerd. De uitspraak 'falend' lijkt op basis van dit materiaal derhalve betrouwbaarder te zijn dan de uitspraak 'succesvol'.
7. De 19 modellen die ten behoeve van de methode van Lachenbruch werden geschat, varieerden niet sterk. Zo werd geen enkele keer een parameter met, een van het totaalmodel, afwijkend teken aangetroffen. De verschillen in waarden van de parameters van de 19 modellen ten opzichte van totaalmodel bedroeg in geen enkel geval meer dan 0.3. Het lijkt derhalve zo te zijn dat zich tussen de 19 cases geen extreme uitschieters bevinden.

Deze resultaten suggereren, dat het numerieke model goed onderscheid maakt tussen succes en falen van nieuwe produkten. Het model dat op basis van slechts 6 (deels samengestelde) variabelen werkt, stelt echter wel eisen aan de betrouwbaarheid van de gegevens. Door aan de omvang van de respondentengroep behorende bij één case de eis van minimaal 5 personen te stellen, lijkt een voldoende betrouwbare meting van deze gegevens mogelijk.
Het lijkt tevens zo te zijn dat het conceptuele model een juiste basis is om een dergelijk model op te stoelen. De zes deelgebieden suggeren op basis van deze resultaten, als totaal, sterk onderscheidend te zijn voor succes en falen van nieuwe produkten.
Het casuïstische karakter van het onderzoek noodzaakt zeer voorzichtig te zijn met deze resultaten. Nogmaals zij vermeld dat dit model gebaseerd is op slechts 19 cases en dat verder onderzoek noodzakelijk is. Ons inziens zijn bovenstaande resultaten dan ook slechts als indicatief te beschouwen.

6. PROBLEMEN BIJ HET METEN VAN DE BENODIGDE GEGEVENS

De calibratie van bovenstaand model had feitelijk als doel om te testen of het conceptuele model een bruikbaar uitgangspunt vormt voor de voorgenomen constructie van een uitgebreider model dat aan de geformuleerde doelstellingen voldoet.
De meting van gegevens ten behoeve van modelconstructie in deze specifieke toepassing stelt ons voor enkele (vermeende) problemen, die deels op eenvoudige wijze kunnen worden omzeild. Onderstaand betoog geeft een inventarisatie van deze problemen en presenteert een meetsysteem waarin oplossingen zijn verwerkt.

6.1. De tijdsongelijkheid van het conceptuele model

Het conceptuele model legt, in kwalitatieve zin, de relaties vast tussen de onafhankelijke variabelen en de maat van succes en falen van nieuwe produkten, als afhankelijke variabele. De onafhankelijke variabelen zijn ingedeeld naar drie entiteiten te weten be-

drijfsmiddelen, produktidee en marktomstandigheden, die zijn opgesplitst in telkens 2 deelgebieden. Bij deze deelgebieden staan in figuur 2 tijdvakken vermeld die corresponderen met de fasen zoals die worden weergegeven in figuur 1.
In paragraaf 4.1. werd reeds uiteengezet dat de causale relaties van het conceptuele model op verschillende tijdstippen gelden. De implicatie van deze zgn. tijdsongelijkheid van het ontwikkelingstraject is dat, als we uitgaan van de juistheid van het gepresenteerde conceptuele model, objectieve en feitelijke meting van gegevens in het begin van het ontwikkelingstraject niet mogelijk is. Dit duidt op de noodzaak om een meetsysteem op te zetten dat eerder gekarakteriseerd wordt door een optimale benutting van de aanwezige gegevens en kennis dan door de volledig objectieve meting van die gegevens.

6.2. De variabelen en bedrijfstakafhankelijkheid

De toepassingsgebied van het te construeren model dient zo breed mogelijk te zijn. Een onderscheid tussen industriële en consumenten goederen is denkbaar, echter van een separate modelconstructie per bedrijfstak wordt vooralsnog afgezien. Dit is mede ingegeven door de goede resultaten van het onderzoek van R.G. Cooper[3] in Canada, waarbij een model werd geconstrueerd voor selectie van industriële goederen (zie ook de Brentani, 1986). Bij de meting van de variabelen zou echter toch een bedrijfstakafhankelijkheid een rol kunnen spelen. De vraag is echter of deze eventuele afhankelijkheid de meting *onjuist* beïnvloedt. De beïnvloeding van de aard van de bedrijfstak op de variabelen van de entiteiten '*bedrijfsmiddelen*' en '*produktidee*' lijkt niet groot. De variabelen zijn namelijk zo geformuleerd dat ze in feite relatief gesteld worden ten opzichte van concurrerende produkten en daarmee beïnvloeding door bedrijfstakkenmerken uitsluit. Er wordt slechts gevraagd naar zaken als 'zijn er voldoende mensen, middelen, etc. aanwezig om dit project uit te voeren' en 'biedt dit produkt een differentieel voordeel ten opzichte van de concurrentie voor de klant'. Deze variabelen zijn onafhankelijk van bedrijfstakken.
De entiteit 'marktomstandigheden' is wel bedrijfstakafhankelijk. Echter als men teruggaat naar het uitgangspunt dan is de vraag of deze afhankelijkheid die de meting beïnvloedt onjuist is. Cooper (1979) en Maidique en Zirger (1984) vonden in hun onderzoeken dat zowel de vraagzijde als de aanbodzijde van de markt het succes van een nieuw produkt beïnvloeden. Voor marktsituaties die gekenmerkt worden door grote groei, behoefte en omvang is de kans op succes groter dan in marktsituaties van kleine groei, behoefte en omvang. Daarmee lijkt de beïnvloeding door bedrijfstakafhankelijkheid m.b.t. marktgegevens juist.

6.3. A posteriori of a priori meting van gegevens?

In principe zijn er twee manieren om gegevens over nieuwe produkten in te winnen. Allereerst kan men gegevens opvragen van produkten die reeds geïntroduceerd zijn. Men kan dan bijvoorbeeld per bedrijf gegevens opvragen over één succesvol en één falend produkt. Het grote nadeel van deze zgn. a posteriori aanpak is het reeds bekend zijn met het commerciële resultaat van deze produkten. Hierdoor kan een vertekening in de gegevens ontstaan die deze niet langer valide maken. Tevens wordt een probleem gevormd door de tijdsspanne die ligt tussen het moment van (bijvoorbeeld) introductie en het moment waarop gegevens verkregen worden. Door de grenzen aan het menselijk

geheugen en selectieve perceptie zijn hierdoor vertekeningen in de gegevens te verwachten, die de betrouwbaarheid van deze gegevens beïnvloeden. Daarmee lijkt de a posteriori methode niet erg aantrekkelijk.
Een ander optie (a priori meting) is vanuit praktisch oogpunt en uitvoerbaarheid tevens verre van ideaal. Allereerst is er sprake van een aanzienlijke tijdsperiode, die men na het inwinnen van gegevens in acht moet nemen. Deze periode is opgebouwd uit de resterende duur van het ontwikkelingstraject binnen het bedrijf en de periode die verstrijkt alvorens men kan vaststellen of het desbetreffende produkt, conform de definitie(s) van het onderzoek, succesvol danwel falend is. Exacte cijfers over het tijdsverschil tussen ideestadium en een zinvolle evaluatie van het commercieel resultaat van een produkt zijn niet bekend, doch zullen in termen van jaren zijn uit te drukken.
Een ander nadeel van deze aanpak is dat de steekproef die benodigd is voor het empirische deel van het onderzoek veel groter zal moeten zijn dan bij de retrospectieve benadering. Gegeven een uitkomst van het reeds eerder gememoreerde onderzoek van Booz, Allen en Hamilton (1982) waaruit blijkt dat van de zeven produktideeën uiteindelijk één produkt succesvol zal zijn, zal gegeven een minimale hoeveelheid van bijv. 100 succesvolle en 100 falende produkten de steekproef 700 produktideeën dienen te omvatten. Het equivalent hiervoor bij de a posteriori benadering is 200 produkten.
De gedachte gaat op dit moment uit naar een combinatie van beide benaderingen, waarbij een eerste calibratie van het model wordt gebaseerd op a posteriori onderzoek. Tegelijkertijd met het inwinnen van a posteriori zullen a priori cases worden gemeten, die na het bekend zijn van hun commerciële resultaat gebruikt zullen worden voor nieuwe calibraties van het model. Tot slot kan vermeld worden dat het inwinnen van a priori cases een continu karakter zal krijgen. In het uiteindelijke meetsysteem dan tot slot van dit artikel wordt gepresenteerd wordt deze continue meting weergegeven in een cyclisch proces van meting, calibratie en validatie.

6.4. Betrouwbaarheid van gegevens

Zoals in het voorafgaande werd gesteld is betrouwbaarheid van de gegevens belangrijk. Men is geneigd te stellen dat bij het verzamelen van gegevens betrouwbaarheid een te maximaliseren grootheid is. Vanuit een oogpunt van pragmatiek en besluitvorming binnen bedrijven zijn echter andere standpunten in te nemen. Omdat het verhogen van betrouwbaarheid van de meetprocedure over het algemeen de kosten ervan hoger maakt en andere vormen van fouten in de hand kan werken (Tull en Hawkins, 1976) zal het gewenste niveau ook betrouwbaarheid afhankelijk zijn van het voorhanden zijnde problemen. Ook Naert en Leeflang (1978) voeren aan dat overwegingen omtrent kosten en baten een rol moeten spelen in de contructie van modellen.
Als we teruggrijpen naar de doelstelling omtrent het toepassingstijdstip van het te construeren model moeten we constateren dat een zekere mate van onbetrouwbaarheid in de gegevens welhaast onvermijdelijk is. Enerzijds kan er sprake zijn van hiaten in kennis, hoewel deze vorm van fouten als systematisch (en dus als validiteits-kwestie) bestempeld kan worden. Anderzijds blijkt in de praktijk een zekere mate van intuïtie onvermijdelijk bij het waarderen van de variabelen van tabel 1.

6.5. Onafhankelijkheid van cases

In strikte zin is het opvragen van meerdere cases per bedrijf een aanleiding tot afhanke-

lijkheid in deze cases. Speciaal de onafhankelijke variabelen op het gebied van de bedrijfsmiddelen zullen voor diverse cases van één bedrijf gerelateerd zijn. Dit heeft gevolgen voor de keuzevrijheid van de calibratietechnieken. In een situatie van afhankelijke cases is het niet toegestaan om bijvoorbeeld gebruik te maken van multiple regressie-analyse (Norusis, 1985; Madsen en Moeschberger, 1986; Hamburg, 1977). Wel is het mogelijk paarsgewijze technieken te gebruiken, echter deze lijken voor dit onderzoek niet bijster geschikt.
Als illustratie daarvoor kan gewezen worden op het SAPPHO onderzoek (Rothwell, 1972). Hier werden 43 paren van innovaties onderzocht. Elk tweetal van innovaties was in zekere mate vergelijkbaar qua produkteigenschappen en marktomstandigheden. Eén innovatie was echter succesvol en de ander falend. Enigszins chargerend kan men stellen dat per paar het produkt en de markt 'constant' werden gehouden. De resultaten van dit onderzoek lijken door deze aanpak zwaar beïnvloed, aangezien het onderscheid tussen succesvolle en falende innovaties voornamelijk gevormd wordt door organisatorische kenmerken en niet door marktoverwegingen. Omdat de innovaties paarsgewijs bovendien sterke gelijkenissen vertoonden komen 'produkteigenschappen' niet eens voor in de resultaten van het onderzoek. Maidique en Zirger (1984), die tot een vergelijkbare conclusie komen, merken bovendien op dat het onderzoek van Cooper[3] een zelfde soort vertekening vertoont. In tegenstelling tot het SAPPHO onderzoek worden hier niet het produkt en de markt constant gehouden maar het bedrijf. Cooper verzamelde van elk bedrijf één falend en één succesvol produkt. Zijn resultaten (o.a. op basis van multiple regressie-analyse) laten zien dat verreweg het belangrijkste onderscheid tussen commercieel succes en falen wordt gevormd door het produkt zelf. Bovendien moet hier nog aan worden toegevoegd dat zijn gebruik van multiple regressie-analyse, o.a. uit hoofde van de eis van onafhankelijke gegevens bij de toepassing van deze techniek, discutabel te noemen is.
Een oplossing voor dit probleem is simpel en voor de hand liggend. De methode die Cooper gebruikte kwam er op neer dat per bedrijf één succesvol en één falend produkt werd opgevraagd ten behoeve van modelcalibratie. Met betrekking tot het verzamelen van gegevens ten behoeve van deze studie bevinden we ons in een analoge situatie. De wijze waarop deze data gebruikt worden zal echter anders zijn, dan in het geval van Cooper. Er zal namelijk per bedrijf willekeurig slechts één succesvol of één falend produkt worden genomen ten behoeve van de calibratie van het model. De resterende groep produkten zal worden gebruikt voor validatie, waarmee de vaak bekritiseerde 'cross-split-half' validatie methode tevens omzeild kan worden. Hierdoor ontstaat de situatie dat de entiteiten 'bedrijfsmiddelen', 'marktomstandigheden' en 'produktidee' steeds in een onafhankelijke combinatie voorkomen, zodat beïnvloedingen uit hoofde van methodologische aspecten niet te verwachten zijn.

7. HET UITEINDELIJKE MEETSYSTEEM; CONCLUSIES

In dit uiteindelijke meetsysteem is zoveel mogelijk rekening gehouden met de meetproblemen die enerzijds ontstaan door de aard en diversiteit van de gegevens en anderzijds kunnen ontstaan uit hoofde van de later te hanteren methodologie. Uit bovenstaande formulering (ontstaan versus kunnen ontstaan) ligt reeds besloten dat niet alle problemen oplosbaar zijn.
Tevens is getracht lering te trekken uit voorgaande onderzoekservaringen en worden

voor onjuistheden ten gevolge van, met name, de methodologie suggesties gedaan voor een bruikbaar meetsysteem.

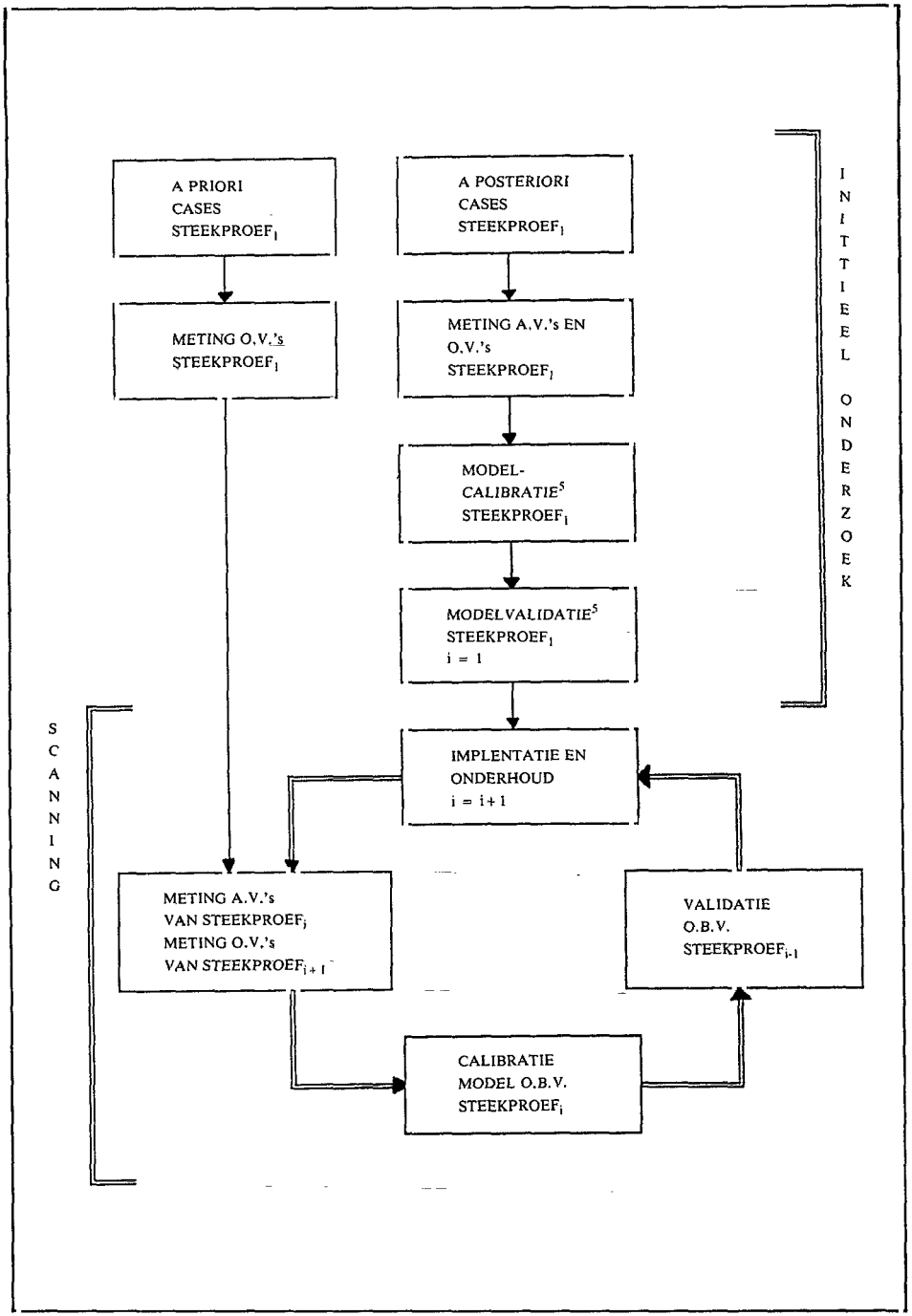
In figuur 3 wordt een schematische weergave gegeven van datgene wat in het bovenstaande werd voorgesteld. In feite worden twee metingen tegelijkertijd opgezet. Eén meting, die dienst zal doen bij een initiële callibratie van het model en één meting die de start inhoudt van een continu onderzoek. De initiële meting heeft tot doel een eerste model samen te stellen, waarbij de callibratie op a posteriori onderzoek gebaseerd is. Tevens zullen gegevens worden gemeten op een continue a priori basis en zal er na het opstellen van het eerste model sprake zijn van een cyclisch proces van meting, calibratie en validatie.
De beoordeling van dit meetsysteem kan plaatsvinden op basis van de doelstelling zoals die eerder werd vermeld. Betrouwbaarheid van validiteit waren daar de kernbegrippen en zullen bij de beoordeling als zodanig worden gehanteerd.

7.1. Betrouwbaarheid

De problematiek rond het a priori resp. a posteriori vraagstuk, in het kader van betrouwbaarheid, is gelegen in de tijdsperiode die verstreken is tussen moment van meting en het daadwerkelijke ideestadium van een betreffende case. Dit probleem is opgelost door een combinatie van beide benaderingen. In eerste aanleg zal een model worden geconstrueerd op basis van a posteriori meting, later zullen nieuwe calibraties van het model plaatsvinden op a priori basis. De praktische nadelen, die kleven aan de laatstgenoemde benadering worden hiermede omzeild aangezien er relatief snel sprake kan zijn van een eerste calibratie en de steekproef geleidelijk kan worden opgebouwd naar een, voor a priori onderzoek, gewenste omvang. Het opzetten van a priori onderzoek houdt tevens in dat het effect van vertekening van gegevens, gedurende het tijdvak tussen meting van variabelen en het ideestadium van produkten, te niet wordt gedaan. Dit tijdvak dat bij de a posteriori aanpak onvermijdelijk is wordt bij de a priori aanpak nihil.
De aard van de te meten gegevens is iets wat feitelijk buiten het bereik van een meetsysteem valt. Echter deze gegevens zijn in hun aard subjectief en dus in zekere mate onbetrouwbaar. Dit probleem zal voor het selecteren van nieuwe produkten blijven bestaan, aangezien het tijdstip van meting van variabelen of gebruik van het model dusdanig vroeg is dat onzekerheden onvermijdelijk zijn. Ook uit hoofde van de tijdsongelijkheid van het ontwikkelingstraject van nieuwe produkten is onzekerheid onvermijdelijk. Eerdere onderzoekservaringen leerden echter dat dit geen belemmering hoeft te zijn m.b.t. de constructie van een bruikbaar model.

7.2. Validiteit

De validatie van het model kan in het cyclische gedeelte van het voorgestelde meetsysteem steeds worden uitgevoerd op basis van nieuwe gegevens. Als ervan uit wordt gegaan, dat elke calibratie van het model geschiedt op basis van een nieuwe steekproef (dus niet op basis van een cumulatieve steekproef) kan validatie van het model geschieden op basis van de steekproef die aan de meest recente vooraf gaat. Zodoende kunnen de bezwaren die kleven aan de 'cross-split-half' methode[6] worden omzeild.
Per bedrijf worden uit hoofde hiervan drie cases gevraagd, te weten één succesvol, één



Figuur 3. Het Meetsysteem.
O.V. = onafhankelijke variabelen
A.V. = afhankelijke variabelen

falend en één nieuwe produkt. Voor de a posteriori calibratie van het model zal per bedrijf slechts één willekeurige (a posteriori) case gebruikt worden. Het gebruik van beide cases zou, door de afhankelijkheid ervan, de validiteit van het model kunnen schaden. Tevens kan dan in het niet-cyclische deel van het meetsysteem de bovengenoemde validatietechniek worden vermeden.

7.3. Slotwoord

Het meten van gegevens ten behoeve van de constructie van een kwantitatief marketing model is niet vrij van imperfecties. In dit specifieke geval blijft de aard van de gegevens en het vroege tijdstip van meting een punt van zorg. De doelstelling van dit meetsysteem is dan ook niet het optimaliseren van de betrouwbaarheid van de gegevens, maar eerder om ondanks de beperkingen een optimale basis voor model-calibratie te verkrijgen. Met name op het gebied van validering, alsmede rondom de problematiek met betrekking tot het a posteriori respectievelijk a priori karakter van gegevens heeft de voorgestelde meetwijze voordelen.
Hierdoor zou een verbetering in onderzoeksresultaten ten opzichte van voorgaande studies mogelijk kunnen zijn.

NOTEN

1. Grotendeels naar Cooper (1979).
2. Cooper en Kleinschmidt (1987a) maken in dezen een inventarisatie waarbij zij komen tot een driedeling in maatstaven te weten: 'financial performance standards, marketshare standards en opportunity windows.'
3. Zie bijvoorbeeld Cooper (1981).
4. Zie paragraaf 4.2.
5. Één case per bedrijf, zoals gesteld in 6.5.
6. Zie bijv. Segers (1983).

LITERATUUR

Booz, Allen & Hamilton (1968), '*Management of new products*', Booz, Allen & Hamilton Inc., New York.
Booz, Allen & Hamilton(1982), '*New product management for the 1980's*', Booz, Allen & Hamilton Inc. New York.
Brentani U. de (1986), 'Do firms need a custom designed new product screening model', *Journal of product innovation management*, 3, (June), blz. 108-119.
Bronnenberg B.J.J.A.M. (1987), '*Het NewProd Project: Het testen en beoordelen van een Canadees model dat het succes of falen van nieuwe produkten op voorhand voorspelt*', Interne Uitgave, Universiteit Twente.
Cooper R.G. (1979), 'Identifying Industrial New Product Success: Project NewProd', *Industrial Marketing Management*, 8 (May) blz. 124-135.
Cooper R.G. (1981), 'An empirically derived new product project selection model', *IEEE, transactions on engineering management*, vol EM 28, no. 3, blz. 54-61.
Cooper R.G. (1985), 'Selecting winning new product project', *Journal of product innovation management*, 2 (March), blz. 34-44.
Cooper R.G. en Kleinschmidt E.J. (1987a), 'New Products: What seperates winners from losers?', *Journal of product innovation management*, 4 (September), blz. 169-184.
Cooper R.G. en Kleinschmidt E.J. (1987b), 'What makes a new product a winner: success factors at the project level', *R & D management*, 3 (September), blz. 175-189.

Crawford C.M. (1979), 'New Product failure-rates, facts and fallacies' *Research Management*, 22 (September), blz. 9-13.
Frederikslust van (1976), 'Voorspelbaarheid van insolventie', *M.A.B.*, 50 (april), blz. 177-213.
Hamburg M. (1985), '*Statistical analysis for decision making*', second edition, Harcourt Brace Jovanovich Inc, New York.
Kotler Ph. (1967), '*Marketing Management Analysis, Planning and Control*', Fourth edition, Prentice Hall, Englewood Cliffs.
Lilien G.L. en Kotler Ph. (1983), '*Marketing decision making; a modelbuilding approch*', Harper & Row, Cambridge.
Madsen R.W. en Moeschberger M.L. (1986), '*Statistical concepts with applications to business and economics*', second edition, Prentice Hall, Englewood Cliffs.
Maidique M.A. en Zirger B.J. (1984), 'A study of success and failure in product innovation: the case of the US electronics industry', *IEEE transactions on engineering management*, vol EEM-31, No. 4., blz. 192-203.
Naert Ph. en Leeflang P.S.H. (1978), '*Building Implementable Marketing models*', Martinus Nijhoff, Leiden.
Norusis M.J. (1985), '*SPSS*^{x}, *Advanced Statistics Guide*', McGraw Hill, New York.
Pride W.M. en Ferrell O.C. (1985), '*Marketing, basic concepts and decisions*', fourth edition, Houghton Mifflin, Boston.
Rothwell R. (1972), 'Factors for success in industrial innovation', uit '*Project SAPPHO – A comparative study of success and failure in industrial innovation*', Science Policy Research Unit, University of Sussex, Brighton, U.K.
Segers J.H. (1983), '*Sociologische Onderzoeksmethoden*', *deel I*, Van Gorcum, Assen.
Tull D.S. en Hawkins D.I. (1976), '*Meaning, measurement, and method of marketing research*', MacMillan Publishing Co., Inc., New York.

6. Energiebesparing op verwarming van woningen als gevolg van electronische dagelijkse feedback

J.H. VAN HOUWELINGEN EN W.F. VAN RAAIJ

SAMENVATTING

Voorlichtingscampagnes om energiezuinig gedrag te propageren hebben in het algemeen een beperkt resultaat. In aanvulling op voorlichting is feedback een doeltreffend middel om energiebesparing van huishoudens te bevorderen. In dit onderzoek is nagegaan, wat het effect is van informatieverstrekking door middel van electronische dagelijkse feedback op het gasverbruik in woningen. In nieuwbouwwoningen te Nieuwegein zijn een jaar lang gasverbruiksindicatoren in gebruik geweest, waardoor de bewoners dagelijks geconfronteerd werden met hun eigen gasverbruik. Onderzocht is, of deze indicator een effectievere vorm van gedragsbeïnvloeding is dan maandelijkse, schriftelijke informatie en zelfregistratie van gasmeterstanden. Een indicator leidt tot een grotere besparing dan de andere vormen van gedragsbeïnvloeding. Na verwijdering van de indicatoren uit de woningen verdween het besparingseffect grotendeels. Het verdient derhalve aanbeveling om de indicatoren permanent in de woningen te installeren.

1. INLEIDING

Ten gevolge van de stijging in energieprijzen en de afhankelijkheid van het buitenland voor de invoer van olie is het efficiënt gebruik van energie sinds 1974 één van de hoofddoelen van het energiebeleid van het Ministerie van Economische Zaken. Zo zou het efficiënt gebruik van energie in het jaar 2000, ten opzichte van 1977, moeten leiden tot een daling van 45 procent van het energieverbruik voor verwarmingsdoeleinden in de huishoudelijke sector.

Er zijn verscheidene mogelijkheden om energie te besparen in huishoudens. Installatie van thermische isolatie en energiezuinige ketels zijn voorbeelden van energiebesparing bij de aanschaf van duurzame goederen. Een andere vorm van energiebesparing is gerelateerd aan gedragsverandering (Van Raaij en Verhallen, 1983). Stern en Gardner (1980) maken een soortgelijk onderscheid tussen aankoop- en verbruiksgerelateerd besparingsgedrag.
Het energiegedrag van bewoners heeft nog steeds een aanzienlijk besparingspotentieel in zich. Sonderegger (1978) vond in het onderzoek in Twin Rivers, New Jersey (V.S.) dat 54% van het verbruik in de winter verklaard kon worden uit de bouwwijze van de woningen. Van de resterende 46% werd 71% veroorzaakt door specifiek verbruiksgedrag van bewoners en 29% door verschillen in kwaliteit van de woningen. 33% van de totale variantie van het energieverbruik was dus toe te schrijven aan het gedrag van be-

woners. Van Raaij en Verhallen (1980) konden in hun onderzoek in Vlaardingen 26% van de totale variantie toeschrijven aan het gedrag van personen in huishoudens.

Bij veel mensen is echter de feitelijke kennis over het eigen energieverbruik klein. Energienota's zijn gebaseerd op het jaarverbruik en worden in vaste maandelijkse of tweemaandelijkse termijnen betaald. Hierdoor krijgen consumenten geen informatie over hun dagelijkse energieverbruik; de informatie die ze krijgen, komt zeer vertraagd. Ook geeft een betaling in vaste termijnen, niet de indruk dat het verbruik in werkelijkheid sterk van dag tot dag varieert.
Er is de laatste jaren geprobeerd om deze situatie te veranderen. Van overheidswege zijn voorlichtingscampagnes gehouden, de ontwikkeling van nieuwe energiebesparende produkten is bevorderd en onderzoeksprogramma's op het gebied van energiebesparing zijn op gang gebracht.

Van Raaij en Verhallen (1983; zie ook Van de Maele et al, 1980) onderscheiden vijf mogelijkheden om energiebesparing te bevorderen:
1. Algemene informatie verstrekken over energieverbruik en energiebesparing.
2. Specifieke informatie verstrekken over de energiekosten van specifieke gedragingen.
3. Eisen stellen aan de woningbouw met betrekking tot isolatie.
4. Subsidies en prijsstelling van energie.
5. Feedback verstrekken over het eigen verbruik.

In vergelijking met de andere mogelijkheden tot energiebesparing is feedback of terugkoppeling een veelbelovende methode om energiezuinig gedrag van consumenten te bevorderen. De resultaten van feedback worden echter in hoge mate bepaald door de wijze, waarop feedback wordt toegepast: Het soort informatie dat men geeft en de duur van het interval tussen verbruik en informatie over het verbruik (Seligman en Darley, 1977; Becker, 1978). Men dient voorzichtig te zijn met algemene bevindingen over feedback. Feedback is vooral effectief, wanneer er sprake is van een te bereiken doel.

Deze voorwaarde is niet los te zien van de functies van feedback. Van Raaij en Verhallen (1983) beschrijven drie functies van feedback. In de eerste plaats is er een *leerfunctie*, d.w.z. dat nieuwe gedragspatronen worden aangeleerd, terwijl andere gedragingen worden afgeleerd. Maar de informatie die gegeven wordt, zal begrijpelijk moeten zijn. In het bijzonder dient het aan te sluiten bij het niveau van de reeds aanwezige kennis. Terugkoppelingsinformatie over de energiekosten en de gebruikte hoeveelheden energie is effectiever voor het verminderen van energieverbruik dan algemene informatie over energiebesparing en aansporingsinformatie (prompts). Specifieke informatie betreft onder meer informatie over de energiekosten die gemaakt zijn in een bepaalde periode, alsmede informatie over het aantal verbruikte kWh elektriciteit of m^3 gas en informatie over het verbruik van energie door huishoudens in vergelijkbare omstandigheden.
Een tweede functie is het *aanleren van gewoontegedrag*, dat wil zeggen gedragingen, die routines zijn geworden zonder dat men erbij nadenkt. Gewoonten kunnen door middel van feedback beloond worden met als gevolg dat beloond gedrag telkens wordt bevestigd en daarmee wordt versterkt.
Een derde functie van feedback is de *internalisatie van gedrag*, wat betekent dat energiebewuste attitudes zullen ontstaan ten gevolge van energiebesparende gedragingen.

Deze laatste functie zou verklaard kunnen worden door de theorie van de zelfperceptie (Bem, 1967). Door middel van feedback worden nieuwe energiezuinige gedragingen aan het licht gebracht. Na enige tijd maken mensen zich attitudes eigen, die overeenkomen met hun eigen gedrag. Deze nieuwe energiebewuste attitudes blijven bestaan, ook nadat de experimentele condities verdwenen zijn.
Een vierde functie van feedback is de *beloningsfunctie* (Mauser en Hutton, 1981). Als een individu of huishouden zich vooraf voor zichzelf een energiebesparingsdoel stelt en de feedback toont dat het doel bereikt kan worden, wordt dit opgevat als een beloning. Het bereiken van een doel kan een gevoel van voldoening geven. Hier komen de bereikte financiële besparingen dan nog bij.

Craig en McCann (1978) zijn van mening dat het gestelde *doel* bereikbaar en geloofwaardig dient te zijn. Becker (1978) concludeert dat het gestelde doel niet te eenvoudig moet zijn. Atkinson (1964) vond dat een doel een gemiddelde moeilijkheidsgraad moet hebben. Een te gemakkelijk of te moeilijk doel leidt tot de situatie dat men bij het slagen of bij het mislukken van het te bereiken doel, dit aan externe factoren gaat toeschrijven. Het bereiken van een doel met een gemiddelde moeilijkheidsgraad kan aan de eigen inspanning worden toegeschreven. In deze omstandigheden kan feedback motiverend werken, omdat men er door leert, of men al dichter tot het doel is genaderd.
Locke, Saari, Shaw en Latham (1981) bespreken de rol van het stellen van doelen, waar gedragingen op gericht moeten zijn. Feedback toont de mate van slagen of mislukken aan van nieuw of veranderd gedrag voor het bereiken van het doel. Succes stimuleert mensen om gedrag te continueren, het te verbeteren en langer vol te houden.

Financiële beloningen die ter aanvulling zijn bedoeld, verhogen het effect van feedback (Battalio, Kagel, Winkler en Winett, 1979; Hayes en Cone, 1981; Kohlenberg, Phillips en Proctor, 1976; Winett, Kagel, Battalio en Winkler, 1978; Winett, Neale en Grier, 1979). Seaver en Patterson (1976) bevelen ook sociale waardering aan om de effectiviteit van feedback te verhogen.

Een ander belangrijk aspect is de *tijdsperiode* tussen het feitelijk gedrag en de feedback over de gevolgen voor het energieverbruik. Hoe korter deze periode is, des te effectiever zal de feedback over het algemeen zijn. Ester (1985) heeft groepen bewoners met twee-wekelijkse en maandelijkse feedback over het eigen energieverbruik vergeleken en komt tot de conclusie dat de kortere feedbackperiode effectiever is om energiebesparing te bereiken.

De huidige ontwikkelingen in (elektronische) feedbackmonitoring hebben een stimulerende werking op verder onderzoek naar feedback. Nu is het mogelijk geworden om de verstrekking van feedback in produkten in te bouwen. In het bijzonder geldt dit voor het tijdsinterval, dat tot een minimum kan worden teruggebracht, daar de informatie nu vrijwel onmiddellijk verstrekt kan worden.

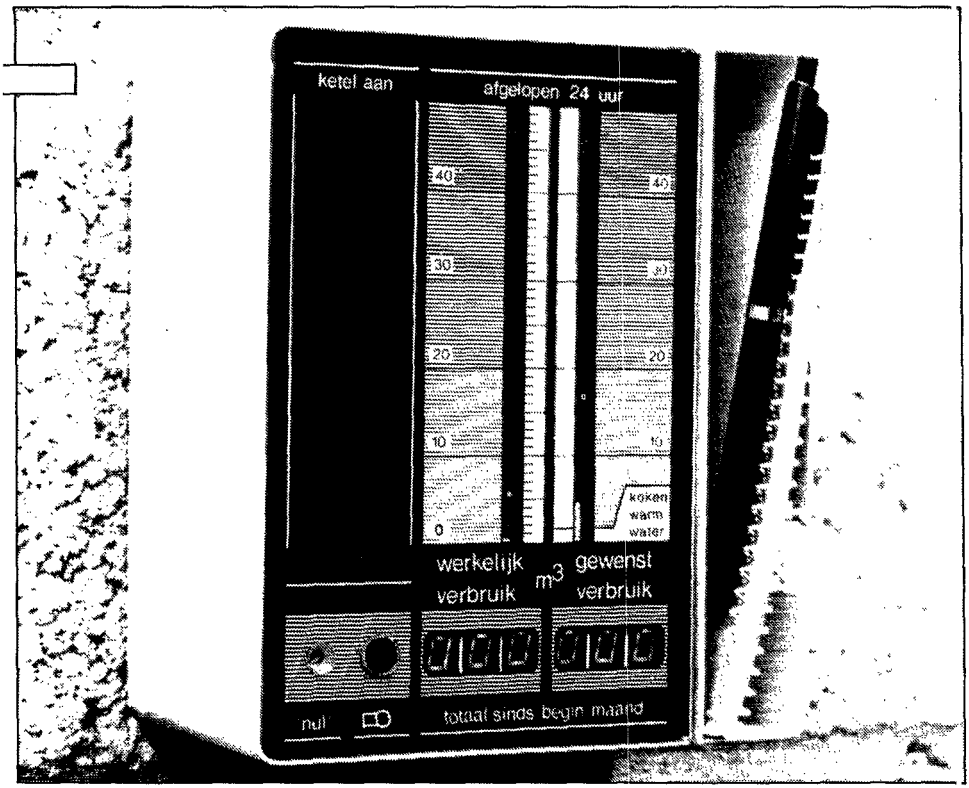
2. FEEDBACK EN DE ENERGIEMONITOR

In de Verenigde Staten en Canada, is de toepassing van elektronische feedback op energiebesparing onderzocht. Het betrof hier een Energy Cost Indicator (ECI), een appa-

raat, dat in huis informatie verschaft over de kosten van het totale energieverbruik van het huishouden (zowel gas als elektriciteit). De eerste experimenten met de ECI vonden in 1978 plaats in het kader van het reeds eerder vermelde Twin Rivers' onderzoek (Becker, Darley en Seligman, 1979). Deze experimenten leverden echter geen duidelijke resultaten op, als gevolg van extreem lage temperaturen, die ten tijde van het onderzoek zijn gemeten. In het gehele gebied, werd namelijk de energielevering voor alle gezinshuishoudens gerantsoeneerd, inclusief de huishoudens die aan het onderzoek meewerkten.
Het gebruik van de ECI is onderzocht in experimenten in vier steden in de V.S. en Canada (Hutton, Mauser, Filiatrault en Ahtola, 1986). De onderzoekers komen tot de conclusie dat de ECI een gunstig effect heeft op de bewustwording en kennis van energieverbruik van consumenten en dat het de bewoners helpt bij het verminderen van hun energieverbruik. Tevens geven consumenten blijk van positieve attitudes ten opzichte van de ECI. Er bleek een verschil in nationaliteit te bestaan: Canadezen besparen meer met een indicator dan Amerikanen. In de Verenigde Staten vond men zelfs geen significant verschil tussen de experimentele en controlegroepen.

In Nederland is een apparaat ontwikkeld dat overeenkomt met de ECI. Het is een meter (huiskamerindicator; energiemonitor) voor de registratie van het gasverbruik, die in de woonkamer op een zichtbare plaats geïnstalleerd dient te worden (Van Beurden, 1982). Het ontwerp is gedeeltelijk gebaseerd op een studie van Becker (1978) over de ontwikkeling van elektronische apparaten ten behoeve van energiebesparing. De Nederlandse indicator geeft de dagelijkse verbruikte hoeveelheid en een dagelijkse genormeerde hoeveelheid gas aan. De verbruikers kunnen hiermee hun eigen verbruik vergelijken met een referentieverbruik, zichtbaar gemaakt door displays met staafdiagrammen en met digitale cijfers. De staaf aan de linkerkant geeft het dagelijks verbruik aan en de staaf aan de rechterkant het normverbruik. Het hoogteverschil tussen de beide staafdiagrammen is de afwijking van het werkelijke verbruik ten opzichte van het genormeerde of gewenste verbruik. De hoeveelheid die gebaseerd is op het gemiddelde verbruik van het afgelopen etmaal, verandert ieder uur. De digitale cijferdisplays laten het dagelijks verbruik zien in m^3 gas op het linkerdisplay en het referentieverbruik op het rechterdisplay. Het dagelijks verbruik wordt gesommeerd met het verbruik van de voorafgaande dagen. Deze cumulatie kan worden stopgezet, met behulp van de reset-knop, waarmee de getallen naar nul verspringen. Ten slotte heeft deze indicator ook een signaleringsfunctie. Aan de voorzijde van het apparaat bevindt zich een controlelampje dat aangeeft, wanneer de CV-ketel aanslaat en daardoor veel gas wordt verbruikt (zie afbeelding).

Het referentie- of gewenst verbruik is gebaseerd op het gasverbruik van het huishouden in het afgelopen jaar, verminderd met een vooraf vastgesteld besparingspercentage (in het onderzoek 10%). Het totaalverbruik in een jaar wordt omgerekend naar dagelijkse hoeveelheden gewenst verbruik. Deze berekening is als volgt: het jaartotaal minus 10% wordt gedeeld door het aantal graaddagen in het jaar[1]. De uitkomst is het referentieverbruik per graaddag. Het bepalen van het aantal graaddagen op een dag is mogelijk door de verbinding met een buitentemperatuurvoeler. Door middel van deze voeler kan een indicator het referentieverbruik berekenen. De graaddagencorrectie is nodig voor de aanpassing van het verbruik aan de buitentemperatuur. Zonder deze correctie is het voor bewoners moeilijk om de wisselende hoeveelheden verbruik te verklaren, vanwe-



ge de wisselvalligheid van het weer. Met de graaddagencorrectie wordt de invloed van het weer op het gasverbruik geminimaliseerd, terwijl het gedragsgerelateerd verbruik duidelijker wordt gemaakt. Concluderend kan gezegd worden dat in vergelijking met de ECI de indicator, die in Nederland is ontwikkeld, dagelijks informatie geeft over het werkelijke en het referentieverbruik, een signaleringsfunctie heeft bij hoog verbruik, maar dat het geen informatie geeft over het toekomstig verbruik en over de financiële gevolgen van het verbruik.

3. PROBLEEMSTELLING EN HYPOTHESEN

Het doel van het onderzoek is om het effect van electronische feedback, die dagelijks gegeven wordt, vast te stellen. Daartoe zijn 50 indicatoren in woningen geïnstalleerd en gedurende een jaar in gebruik geweest. Het onderzoek is zodanig opgezet dat de werking en het effect van de indicatoren is vergeleken met andere vormen van feedback die in dezelfde periode in identieke woningen is verstrekt.

De eerste hypothese heeft betrekking op de werking van de indicatoren: Het gebruik van een huiskamerindicator in combinatie met voorlichting leidt tot een vermindering van het gasverbruik. Er wordt verondersteld dat deze vorm van informatieverstrekking

de aandacht vestigt op het energieverbruik; dat het discussies tussen bewoners op gang brengt over het eigen verbruik; dat het gedragsveranderingen met zich meebrengt met als gevolg een daling van het energieverbruik.
De tweede hypothese heeft betrekking op de verschillen in gasverbruik als gevolg van de verschillende vormen van feedback. Naast het plaatsen van een indicator in woningen zijn de effecten van twee andere vormen van feedback in het onderzoek bestudeerd, te weten maandelijkse feedback en zelfregistratie door de bewoners op een verbruikskaart. Deze hypothese veronderstelt dat de daling van het gasbruik in de indicator-conditie groter is dan in de andere experimentele condities.
De derde hypothese gaat over de vergelijking tussen de experimentele en controlegroepen: Het gebruik van feedback in combinatie met voorlichting leidt tot een vermindering van het gasverbruik. Als hypothese 3 wordt bevestigd, is de vermindering van het gasverbruik in de experimentele condities groter dan in de controlecondities.
De vierde hypothese betreft lange-termijn effekten van feedback door middel van een indicator. Hypothese 4 is dan ook dat de vermindering van het gasverbruik als gevolg van het gebruik van een indicator blijft bestaan na de experimentele periode van een jaar (dus zonder de indicator).

Tenslotte is in het onderzoek nagegaan, hoe de bewoners de huiskamerindicator beoordelen en welke ideeën zij naar voren brengen in verband met verdere ontwikkelingen van het apparaat.

4. ONDERZOEKSOPZET EN STEEKPROEFTREKKING

In totaal zijn zes experimentele groepen samengesteld, overeenkomend met de experimentele en controlecondities:
1. Huishoudens met een huiskamerindicator en voorlichting (Indicatorgroep).
2. Huishoudens, die maandelijks schriftelijke feedback en voorlichting ontvangen (Maandelijkse feedbackgroep).
3. Huishoudens, die een verbruikskaart hebben ontvangen om zelf in te vullen en voorlichting (Zelfregistratiegroep).
4. Huishoudens, die alleen voorlichting ontvangen (Eerste controlegroep of voorlichtingsgroep).
5. Huishoudens, die niet op de hoogte zijn gebracht van deelname aan het onderzoek (Tweede of 'blinde' controlegroep).
6. Huishoudens, die niet aan het onderzoek willen deelnemen (Derde controlegroep, de weigeraars).

Bij de keuze van een geschikte locatie voor het onderzoek is gezocht naar een plaats met veel recente nieuwbouw van identieke woningen. Nieuwegein in de provincie Utrecht voldeed aan deze vereisten. Uit het totaal van 690 eengezinswoningen in Nieuwegein zijn 475 woningen aselect gekozen. Hierbij bleken 95 woningen/huishoudens niet geschikt voor deelname aan het onderzoek om uiteenlopende redenen, en weigerden 80 bewoners na het eerste interview verdere medewerking.
Het gasverbruik van deze 80 huishoudens verschilde niet significant van de 300 huishoudens die wel wilden deelnemen aan het onderzoek. Het gemiddelde gasverbruik komt op 0,72 m^3 gas per graaddag van zowel de non-respons als de responsgroep (de

standaarddeviaties zijn respectievelijk, 0,24 en 0,25 m^3). Het verbruik is gebaseerd op het jaarverbruik 1982; de eerste selectie van huishoudens heeft in de herfst in 1983 plaatsgevonden.
Het gemiddeld energieverbruik in de overgebleven woningen is 1747 m^3 (standaarddeviatie 588). De variantie in energieverbruik lijkt hoog, maar in vergelijking met andere onderzoeken is deze standaarddeviatie niet uitzonderlijk hoog te noemen (Naber en Zwetsloot, 1979).
Huishoudens die kort voor de aanvang van het onderzoek in de woning waren getrokken of waarbij veranderingen in het verwarmingssysteem waren aangebracht, werden alsnog uit de steekproef verwijderd. De resterende 300 huishoudens zijn zodanig over de experimentele groepen verdeeld, dat er een matching plaatsvond ten aanzien van het aantal gasapparaten in de woning, het woningtype, en het gasverbruik in het jaar 1983. Huishoudens, die zijn ingedeeld in de eerste vijf groepen hebben alle aangegeven bereid te zijn tot een energiebesparing van minimaal 10%. Een aantal huishoudens die niet wilden besparen, is in groep 6 ondergebracht.

De indicatoren zijn geplaatst in 50 woningen die behoren tot groep 1. De bewoners is verzocht elke maand het werkelijk en gewenst gasverbruik, zoals aangegeven op het apparaat, te noteren en daarna de indicator op nul te zetten. De huishoudens in groep 2 hebben maandelijks schriftelijke informatie ontvangen over hun eigen gasverbruik. De huishoudens uit groep 3 ontvingen een zelfregistratiekaart met het verzoek om maandelijks hun gasverbruik te noteren. De huishoudens uit de groepen 1, 2, 3 en 4 ontvingen voorlichtingsmateriaal over energiebesparing in de woning. Huishoudens uit de laatste twee groepen, de condities 5 en 6, kregen geen informatie toegestuurd en waren niet op de hoogte van hun deelname aan het onderzoek.

De experimentele periode begon op 16 februari 1984 en eindigde een jaar later. De standen van de gasmeters in de woningen in de groepen 1, 2, 3 en 4 zijn regelmatig opgenomen tijdens het experimentele jaar. In aanvulling op de registratie door de bewoners zijn gegevens over het gasverbruik vanaf 1979 door het gasbedrijf Centraal Nederland verstrekt. De medewerking van huishoudens aan het onderzoek was tijdens de gehele onderzoekperiode goed: De hoogste non-respons was 17% in conditie 4.

5. RESULTATEN

Het gemiddelde gasverbruik in 1983, het jaar voorafgaande aan het onderzoek, is weergegeven in Tabel 1. In Totaal gaat het om 285 huishoudens, die de gehele experimentele periode (februari 1984 tot februari 1985) aan het onderzoek hebben meegewerkt. Het gemiddelde gasverbruik is berekend, nadat een standaardhoeveelheid voor het verbruik voor koken en warm water van het totaal is afgetrokken, te weten, respectievelijk 100 en 450 m^3 gas per jaar (Naber en Zwetsloot, 1979). In 1983 zijn er tussen de condities geen significante verschillen gevonden in het gemiddelde gasverbruik.
De vermindering van het gasverbruik in de experimentele periode in vergelijking met 1983 is in de indicatorgroep 12,3%; in de maandelijkse schriftelijke feedbackgroep 7,7%; in de zelfregistratiegroep 5,1%; en in de voorlichtingsgroep 4,3%. Dit zijn gemiddelde percentages van de vier groepen. De 12,3% besparing in de indicatorgroep is hoger dan het gestelde doel van 10% besparing.

Tabel 1. Gemiddeld gasverbruik in 1983 in m³; gemiddeld verbruik per graaddag en aantallen huishoudens.

Conditie	Gemiddeld gasverbruik³ (1983)	Gemiddeld gasverbruik per graad-dag³ (1983)	Aantal huishoudens begin (1984)	midden (1985)	einde (1986)
1. Indicator	1885 (510)	.76 (.22)	50[1]	45	40
2. Maandelijkse feedback	1812 (587)	.74 (.24)	55	48	44
3. Zelfregistratie	1793 (567)	.72 (.26)	52[2]	48	43
4. Voorlichting	1785 (681)	.72 (.30)	55	46	42
5. Controle 2	1808 (584)	.74 (.28)	55	48	46
6. Controle 3	1789 (436)	.75 (.22)	55	49	43

1. Het aantal geplaatste indicatoren was 50.
2. Het aantal kaarten voor zelfregistratie was 52.
3. De tussen haakjes staande cijfers geven de standaardafwijking aan.

In de eerste hypothese wordt gesteld dat het gebruik van een indicator plus voorlichting (conditie 1) leidt tot een vermindering van het gasverbruik per graaddag. Er is een vermindering van het gasverbruik te constateren in het experimentele jaar en in 1984, beide jaren vergeleken met 1983. Een logaritmische transformatie is noodzakelijk om te corrigeren voor de scheve verdeling van het energieverbruik. In de eerste vier condities is een significante vermindering van het gasverbruik gevonden. In beide perioden is de daling van het gasverbruik in de indicatorconditie groter dan in de overige condities.
Bij een vergelijking tussen de condities is een significante vermindering van het gasverbruik geconstateerd in conditie 1 ten opzichte van de overige feedbackcondities 2 en 3 ($F = 3{,}60^{*}$), experimenteel jaar versus 1983[2], en ten opzichte van alle overige condities ($F = 5{,}97^{**}$), 1984 versus 1983. Op basis van dit resultaat wordt hypothese 1 niet verworpen.

Hypothese 2 heeft betrekking op de effectiviteit van de verschillende toepassingsmogelijkheden van feedback. De indicatorgroep is vergeleken met de maandelijkse, schriftelijke feedbackgroep en met de zelfregistratiegroep. De uitkomsten staan in Tabel 2. De besparingen in conditie 1 (Indicator) zijn significant hoger dan in de condities 2 en 3 afzonderlijk en de condities 2 en 3 samen, zowel voor het experimentele jaar als voor het kalenderjaar 1984. Ook hypothese 2 wordt niet verworpen.

In de derde hypothese worden de experimentele en controlegroepen vergeleken. De vermindering van het gasverbruik in de groepen 1, 2 en 3 is groter dan in groepen 4, 5 en 6 in het kalenderjaar 1984 ($F = 2{,}89^{*}$). Ook deze hypothese wordt niet verworpen.

Voorts is op basis van het gemiddeld verbruik van 0,72 m³ per graaddag een indeling gemaakt in type verbruikers. Het verbruik in de experimentele periode van de hoge en

Tabel 2. Besparingsverschillen tussen de experimentele condities (t-toetsen), experimenteel jaar en 1984 vergeleken met 1983.

	Experimenteel jaar	*1984*
Periode:	16 febr. 1984, 15 febr. 1985.	1 jan. 1984, 31 dec. 1984.
1 vs. 2	1,35*	1,23*
1 vs. 3	2,06**	2,39**
2 vs. 3	0,76	1,18*
1 vs. (2+3)	1,90**	2,07**

Conditie 1: Indicator.
Conditie 2: Maandelijkse feedback.
Conditie 3: Zelfregistratie.

** $p < .01$
* $p < .10$

de lage verbruikers is vergeleken. De hoge verbruikers bleken een groot besparingspotentieel te hebben en bespaarden naar verhouding meer dan de lage verbruikers (T = 1,86*). Bij de groep lage verbruikers nam het gemiddeld verbruik zelfs toe. Mogelijk is er sprake van een 'regressie naar het midden' (Tversky en Kahneman, 1974), d.w.z. dat hoge verbruikers minder gaan gebruiken en lage verbruikers meer. De groepen hoge en lage verbruikers zijn niet significant verschillend over de condities verdeeld.

6. MENINGEN OVER DE INDICATOR

Bewoners waarbij een indicator in de woning is geïnstalleerd, zijn aan het eind van de onderzoeksperiode geïnterviewd, waarbij naar hun mening is gevraagd over het ontwerp en het nut van de indicator.

De informatie is op de indicator in digitale cijfers en in staafdiagrammen af te lezen. 56% van de gebruikers geeft de voorkeur aan de aflezing in cijfers omdat deze informatie exact is, terwijl de staafdiagrammen indicatief zijn. 29% geeft de voorkeur aan de staafdiagrammen, terwijl 15% geen duidelijk voorkeur heeft. Een verklaring voor deze verschillen zou kunnen zijn dat de diagrammen niet optimaal functioneerden. De afmetingen van de schaal waren gebaseerd op het (gemiddeld veel hogere) gasverbruik van enkele jaren daarvoor. De vergelijking tussen het werkelijke en het gewenste verbruik vond men zinvol.

Het controlelampje blijkt tot handelen aan te zetten. Een te frequente signalering, leidt tot een verminderde aandacht. Dit betekent dat de hoeveelheid gasverbruik, waarop het lampje reageert, zorgvuldig moet worden bepaald.

De algemene opinie over de indicator is gunstig. De meeste bewoners verklaren dat de indicator hen bewuster had gemaakt van het gasverbruik. Ook vonden zij dat ze de resultaten konden zien van hun pogingen om te besparen. Men kan uit de antwoorden

concluderen dat een indicator beschouwd kan worden als een voortdurende controle op het gasverbruik. De indicator geeft echter geen inzicht in het energieverbruik van afzonderlijke apparatuur (verwarming, koken en warm water). Met betrekking tot de tijdsinterval tussen verbruik en informatie geeft tweederde van de respondenten de voorkeur aan dagelijkse informatie, terwijl 22% er de voorkeur aan geeft frequentere informatie te krijgen, zelfs informatie, die onmiddellijk op elke handeling volgt. In totaal is 83% van de respondenten een voorstander van een grootschalige verspreiding van de indicator.

Op de vraag naar de gewenste duur van de aanwezigheid van de indicator in de woning wil 22% de indicator voor slechts één jaar in huis hebben, terwijl 57% de voorkeur geeft aan een permanente installatie in de woning. Dit wijst er op dat er een markt is voor zowel verhuur als verkoop van een indicator. Indien blijkt dat een korte periode voldoende is om energiebesparing aan te leren, is verhuur een geschikte mogelijkheid. Indien een permanente feedback noodzakelijk is, zou verkoop en permanente installatie (in bij voorbeeld nieuwe woningen) de voorkeur verdienen. In het gesprek met de bewoners is de mogelijkheid van aankoop of huur ter sprake gebracht. 61% van de bewoners geeft de voorkeur aan het huren van een indicator.

7. LANGE-TERMIJN EFFECTEN VAN DE INDICATOREN

De uitkomsten van de na-meting zijn niet alleen van belang voor de keuze tussen verkoop en verhuur van indicatoren, maar geven ook inzicht in het lange-termijn effect van de indicator op het gasverbruik.

In dit onderzoek is, nadat het experiment is beëindigd, het gasverbruik een jaar later opnieuw gemeten. Lange-termijneffecten zijn over twee perioden gemeten. Voor de groepen 1, 2, 3 en 4 is het jaar na het experimentele jaar (februari 1985 tot februari

Tabel 3. Verschillen tussen de experimentele condities (t-toetsen), na-experimenteel jaar en 1985 vergeleken met 1983.

	Na-experimenteel jaar	*1985*
Periode:	16 febr. 1985, 15 febr. 1986.	1 jan. 1985, 31 dec. 1985.
1 vs. 2	-0.12	0.41
1 vs. 3	0.40	1.23*
2 vs. 3	0.59	0.82
1 vs. (2+3)	0.14	0.93

Conditie 1: Indicator.
Conditie 2: Maandelijkse feedback.
Conditie 3: Zelfregistratie.

* $p < .10$

1986) genomen als periode voor de nameting, alsmede het kalenderjaar 1985. Over het kalenderjaar 1985 is ook het gasverbruik van de andere groepen gemeten, inclusief de groep bewoners, die zich niet bewust was van deelname aan het onderzoek en de groep weigeraars (Tabel 3).

Er is sprake van een vermindering van het gasverbruik van het na-experimentele jaar en het jaar 1985 vergeleken met het basisjaar 1983. In de indicatorgroep en de schriftelijke, maandelijkse feedbackgroep is de vermindering van het verbruik in beide perioden significant. De zelfregistratie en de voorlichtingsgroep hebben alleen een significante vermindering in de periode 1985. De controle-groepen hebben een significante daling in 1985 ten opzichte van 1983. Het verloop van het verbruik van de groepen 5 en 6 is interessant omdat het een indicatie is voor de algemene trend in het gasverbruik tussen de jaren 1983 en 1985.

Behalve de meting van verschillen binnen de groepen is ook gekeken naar verschillen tussen de groepen, waarbij de in het bijzonder de verschillen in gasverbruik tussen de indicatorgroep en de overige groepen van belang zijn. Met betrekking tot de na-experimentele periode kan worden geconcludeerd dat de indicatorgroep geen significant lager verbruik heeft in vergelijking met beide andere feedbackgroepen. Dit geldt zowel voor veranderingen in verbruik ten opzichte van het basisjaar 1983 (Tabel 3) als ten opzichte van het experimtenele jaar en het basis(kalender)jaar 1984 (Tabel 4). De uitkomsten weerleggen hypothese 4 over het lange-termijnbesparingseffect van de indicator.

Het gasverbruik in de eerste vier groepen in de na-experimentele periode is significant toegenomen in vergelijking met de experimentele periode. Maar in de tweede periode, de periode van de vergelijking tussen de kalenderjaren is deze toename niet te constateren. De tegenstrijdige uitkomsten in de tweede periode in vergelijking met de eerste periode laten zich niet gemakkelijk verklaren. Denkbaar is dat het verschil in aantallen graaddagen een belangrijke rol heeft gehad op de uiteindelijke resultaten.

Tabel 4. Verschillen tussen de experimentele condities (t-toetsen).

	Na-experimenteel jaar vergeleken met experimenteel jaar	*1985 vergeleken met 1984.*
Periode:	16 febr. 1985, 15 febr. 1986.	1 jan. 1985, 31 dec. 1985.
1 vs. 2	– 1.18*	– 1.05
1 vs. 3	– 0.77*	– 0.66
2 vs. 3	0.43	– 0.54
1 vs. (2+3)	0.72	1.00

Conditie 1: Indicator.
Conditie 2: Maandelijkse feedback.
Conditie 3: Zelfregistratie.

* $p < .10$

Zo was het aantal graaddagen in het experimentele jaar 2474; dat is 275 meer dan in het hele kalenderjaar 1984, toen er 2199 graaddagen waren. Beide jaren zijn in de berekening als basisjaar genomen. In het na-experimentele jaar was het aantal minder dan in 1985, te weten 2403 versus 2516 graaddagen; een verschil van 113. Deze extra graaddagen vielen overigens vooral in de periode van 1 januari tot 15 februari 1984.
Ondanks de tegenstrijdige uitkomsten kan geconcludeerd worden dat de lagere niveaus van gasverbruik van bewoners uit conditie 1, gemeten in de experimentele periode en in 1984, in 1985 zijn verdwenen. In de na-experimentele periode en in 1985 vinden we een toename in gasverbruik.

Dezelfde ontwikkelingen in gasverbruik is gevonden in conditie 2. Het lagere gasverbruik in de experimentele periode en in 1984 in vergelijking met de zelfregistratieconditie is in de na-experimentele periode verdwenen. Met betrekking tot conditie 3 is een conclusie over het lange-termijneffect moeilijk te trekken.

8. EVALUATIE VAN DE INDICATOR EN BELEIDSAANBEVELINGEN

Feedback over het energieverbruik met behulp van de indicator is veelbelovend. Informatie over het gasverbruik laat zien dat dagelijkse feedback met behulp van een indicator meer energie bespaart dan maandelijkse feedback of zelfregistratie van energieverbruik. In deze drie experimentele condities wordt gedurende de experimentele periode minder energie verbruikt dan in de drie controlecondities. Onderzoek naar het lange-termijn effect indiceert dat een tijdelijke installatie niet leidt tot het bereiken van lange-termijn effecten.
Men zou een zuiverder schatting van het lange-termijneffect hebben verkregen, indien slechts de helft van het aantal indicatoren zou zijn weggehaald uit de woningen en de andere helft nog langer in gebruik zou zijn gebleven. Een vergelijking zou dan kunnen plaatsvinden tussen beide groepen huishoudens. Vanwege praktische problemen, zoals de financiering van het onderzoek, het beschikbare aantal indicatoren en de noodzakelijke technische begeleiding, is van deze onderzoeksopzet voor een nameting afgezien.

De uitkomsten van de nameting veronderstellen dat het belang van de feedbackfuncties verschilt. Zoals gezegd worden vier functies aan feedback toegeschreven. De eerste functie is een leerfunctie; de tweede is het aan- en afleren van gewoontes; de derde betreft de internalisatie van gedrag; en de vierde een beloningsfunctie. Het ziet er niet naar uit dat de indicator bij bewoners heeft geleid tot een energiegedrag dat zich weerspiegelt in een laag verbruik op de lange termijn, ook als de indicator niet meer aanwezig is. Dit resultaat wijst erop dat de beloningsfunctie van grote invloed is, maar dat zich weinig blijvende gedragsveranderingen hebben voorgedaan. De indicator is blijvend nodig om het energiezuinige gedrag mogelijk te maken en te ondersteunen.
Feedback stelt mensen in staat de resultaten te zien van hun pogingen om te besparen, wat door 85% van de respondenten werd bevestigd. Mensen voelen zich beloond door de informatie, die zij krijgen over hun zuiniger verbruik, als hun pogingen tot energiebesparing slagen. Door een indicator worden zij voortdurend *geattendeerd* op het succes of falen van hun pogingen. Aan de andere kant wordt energieverspillend gedrag snel geregistreerd. De indicator wordt dan beschouwd als een controleapparaat over ei-

gen gedrag en het gedrag van andere leden van de huishouding.
Verdere ontwikkeling van de indicator is gezien de uitkomsten van het onderzoek zeker aan te bevelen. De indicator heeft een positieve invloed op het gedrag van consumenten ten aanzien van zuinig energieverbruik en vergroot de kennis over eigen verbruik, wat energiebesparing vergemakkelijkt. De informatie in cijfers over het dagelijks verbruik gesommeerd over een langere periode en de vergelijking met gewenste hoeveelheden verbruik wordt als zinvol ervaren.

De uitkomsten uit de nameting wijzen erop dat verkoop en permanente installatie in woningen te prefereren is boven een tijdelijke installatie. De bewoners zullen bij een permanente installatie zelf de hoogte van het referentiegebruik kunnen instellen. Een voortdurende besparing, jaar na jaar, lijkt niet aannemelijk, maar stabilisatie van het verbruik is wel mogelijk. Gecorrigeerd voor de weersgesteldheid is het dan voor de verbruikers mogelijk om het gasverbruik op een bepaald niveau te handhaven.

Met betrekking tot een verdere ontwikkeling van het apparaat valt te denken aan een functie, waarin het toekomstige verbruik, bijvoorbeeld voor het hele kalenderjaar wordt berekend, aan de hand van het verbruik zoals dat op een bepaald moment plaatsvindt. Dit zou de consument meer inzicht geven in de hoogte van het verbruik als men bepaalde gewoonten handhaaft. Aan de hand van de graaddagenmethode is de ontwikkeling van een functie in deze richting denkbaar. Bovendien kan een indicator zelf één van de functies worden en een zogenaamd 'household control unit', een elektronisch apparaat ter controle van allerlei huishoudelijke activiteiten (energie, wassen, koken en schoonmaken). Het kan gemakkelijk ingepast worden in een thermostaat en gegevens zouden kunnen worden geregistreerd door een computer speciaal geïnstalleerd ten behoeve van het huishouden (Toffler, 1981).

NOTEN

1. Onder een graaddag wordt verstaan een afwijking van de gemiddelde etmaaltemperatuur van de buitenlucht van een graad Celsius beneden de stookgrens. De stookgrens is de buitentemperatuur, waarboven in het algemeen geen en waaronder in het algemeen wel gas voor ruimteverwarming wordt gebruikt. Bij deze berekening worden er vaste hoeveelheden gas van het jaarverbruik afgetrokken, te weten: 100 m^3 gas voor het koken en 450 m^3 gas voor het gebruik van warm water. Een eventuele besparing op het gebruik van warm water is in dit onderzoek niet te scheiden van een besparing op de woningverwarming.
2. De F-toets heeft de volgende significantie-niveaus: **: $p < .10$; *: $p < .01$.

LITERATUUR

J.W. Atkinson, 1964, *An Introduction to Motivation*, New York: Van Nostrand.
R.C. Battalio, J.H. Kagel, R.C. Winkler en R.A. Winett, 1979, Residential electricity demand: An experimental study, *Review of Economics and Statistics, 61*, blz. 180-189.
L.J. Becker, 1978, Joint effect of feedback and goal setting on performance: A field study of residential energy conservation, *Journal of Applied Psychology, 63*, blz. 428-433.
L.J. Becker, J.M. Darley en C. Seligman, 1979, '*Psychological strategies to reduce energy consumption.* A project summary report', Princeton University, Center for Energy and Environmental Studies. Report 41.
Bem, D.J., 1967, Self-perception: An alternative interpretation of cognitive dissonance phenomena, *Psychological Review, 74*, blz. 183-200.
K.M.M. van Beurden, 1982, 'Indicator voor huishoudelijk gasverbruik', *GAS*, 102, blz. 58-64.

C.S. Craig en J.M. McCann, 1978, Assessing communication effects on energy conservation, *Journal of Consumer Research, 5*, blz. 82-88.
P. Ester, 1985, *Consumer Behavior and Energy Conservation*, Dordrecht: Reidel.
S.C. Hayes en J.D. Cone, 1981. Reduction of residential consumption of electricity through simple monthly feedback, *Journal of Applied Behavior Analysis, 14*, blz. 81-88.
J.H. van Houwelingen en W.F. van Raaij, 1986, Energiebesparing met een huiskamer indicator, *GAS*, 106, blz. 340-343.
R.B. Hutton, G.A. Mauser, P. Filiatrault en O.T. Ahtola, 1986, Effects of cost-related feedback on consumer knowledge and consumption behavior: A field experimental approach, *Journal of Consumer Research, 13*, blz. 327-336.
R. Kohlenberg, T. Phillips en W.A. Proctor, 1976, A Behavioral analysis of peaking in residential electrical-energy consumers, *Journal of Applied Behavioral Analysis, 9*, blz. 13-18.
E.A. Locke, L.M. Saari, K.N. Shaw en G.P. Latham, 1981, Goal setting and task performance 1969-1980, *Psychological Bulletin, 90*, blz. 125-152.
M.C.L. van de Maele-Vaernewijck, W.F. van Raaij en Th.M.M. Verhallen, Energiegedrag in de woning: Literatuuroverzicht en gedragsmodel, *Jaarboek van de Nederlandse Vereniging van Marktonderzoekers*, 1980, blz. 169-189.
G.A. Mauser en R.B. Hutton, 1981, *Feedback: A tool for energy conservation*, Halifax, NS, Canada: Dalhousie University.
Ministerie van Economische Zaken, *Nota Energiebeleid*, Den Haag: Staatsuitgeverij, 1979.
W.F.A.M. Naber en M.G.F. Zwetsloot, Basisonderzoek aardgas kleinverbruik, *Gas, 99*(12), 1979.
W.F. van Raaij en T.M.M. Verhallen, 1980, Huishoudelijk gedrag en stookgasverbruik, *Economisch Statistische Berichten, 65*, (3236), blz. 8-13.
W.F. van Raaij en T.M.M. Verhallen, 1983, A behavioral model of residential energy use, *Journal of Economic Psychology, 3*, blz. 39-63.
W.B. Seaver en A.H. Patterson, 1976, Decreasing fuel-oil consumption through feedback and social commendation, Journal of Applied Behavior Analysis, 9, blz. 147-152.
C. Seligman en J.M. Darley, 1977, Feedback as a means of decreasing residential energy consumption, *Journal of Applied Psychology, 62*, blz. 363-368.
R.C. Sonderegger, 1978, 'Movers and stayers: The resident's contribution to variation across houses in energy consumption for space heating', in: R.H. Socolow (red.), *Saving Energy in the Home, Princeton's Experiments at Twin Rivers*, Cambridge, MA: Ballinger, blz. 207-230.
P.C. Stern en G.T. Gardner, 1980, A review and critique of energy research in psychology, *Social Science Energy Review, 3*, (New Haven, CT: Yale University)
A. Toffler, 1981, *De Derde Golf*, Utrecht: Veen.
A. Tversky en D. Kahneman, 1974, Judgment under uncertainty: Heuristics and biases, *Science, 185*, blz. 1123-1131.
R.A. Winett, J.H. Kagel, R.C. Battalio en R.C. Winkler, 1978, Effects of monetary rebates, feedback, and information on residential electricity conservation, *Journal of Applied Psychology, 63*, blz. 73-80.
R.A. Winett, M.S. Neale en M.S. Grier, 1979, Effects on self-monitoring and feedback on residential electricity consumption, *Journal of Applied Behaviour Analysis, 12*, blz. 173-184.

7. Het marktonderzoek als bewijsmiddel in het mededingings- en merkenrecht[1]

P.J. KAUFMANN

SAMENVATTING

In dit artikel wordt aan de orde gesteld, dat de Nederlandse rechter weinig aandacht besteed aan het marktonderzoek als bewijsmiddel om de centrale begrippen misleiding en verwarringsgevaar in het reclamerecht en het merkenrecht te bepalen. Dit in tegenstelling tot belangrijke industrielanden als West-Duitsland en de Verenigde Staten waar de rechterlijke macht zich, bij de beantwoording van de vraag of er sprake is van misleidende reclame of van merkinbreuk, laat leiden door de ervaring van de consumenten. Dit heeft dan ook geleid tot een goede samenwerking tussen rechters en marktonderzoekers.
Aan de hand van veel voorbeelden wordt de vraag in positieve zin beantwoord of dit ook niet een voor Nederland wenselijke situatie is.

1. INLEIDING

In Nederland wordt aan het bewijsmiddel van het marktonderzoek weinig aandacht besteed. Van Nieuwenhoven Helbach, die zelf het onderwerp twee maal aan de orde heeft gesteld,[2] spreekt daar zijn verbazing over uit, temeer daar het marktonderzoek in andere landen met een hoog ontwikkeld mededingingsrecht, zoals West-Duitsland en de Verenigde Staten van Amerika, al tientallen jaren is geaccepteerd.
Dat de Nederlandse rechter dit bewijsmiddel met scepsis, zo niet met reserve beziet, blijkt uit de dikwijls summiere wijze waarop de methode of de resultaten van het marktonderzoek worden beoordeeld. Een recent voorbeeld hiervan is de uitspraak van het Hof te Amsterdam in de Q/Q8-zaak[3]. In deze stond Kuwait Petroleum met zijn merk Q8 tegenover het merk Q van het Engelse Duckham, een bijna volledige dochteronderneming van BP. Nadat het Hof uitgebreid zijn eigen opvattingen over de onderscheidende kracht van het Q-merk en het mogelijke verwarringsgevaar tussen het merk en het teken heeft uiteengezet, wordt het door gedaagde geproduceerde marktonderzoek afgewezen met de niet erg verhelderende overweging dat 'in die enquête geen rekening is gehouden met een aantal hiervoor aan de orde gekomen, voor het onderzoek naar het gestelde merkinbreuk relevant geachte bijzonderheden'. Wat die bijzonderheden precies zijn, wordt niet verder aangegeven. Blijkbaar vertrouwt het Hof meer op zijn eigen inzicht in en ervaring met de markt, dan op de resultaten van een empirisch onderzoek naar de mening van de consumenten.
Anders is de benadering van de West-Duitse rechters. In een uitgebreide jurisprudentie zijn onder leiding van het Bundesgerichtshof, het hoogste West-Duitse rechtscollege, sinds de vijftiger jaren richtlijnen opgesteld hoe het marktonderzoek gebruikt moet

worden in de verschillende situaties. De waarde hiervan wordt in de literatuur benadrukt omdat de rechter gedwongen wordt zijn beslissingen te nemen op basis van een feitelijke vaststelling van het weten of de perceptie van de consumenten[4].
In de Verenigde Staten is het marktonderzoek dermate vanzelfsprekend als bewijsinstrument, dat het de partijen, die geen marktonderzoek hebben doen uitvoeren, negatief wordt aangerekend in de procedure[5].

De vraag die zich opdringt, waarom in Nederland zo terughoudend geoordeeld wordt over het gebruik van het marktonderzoek als bewijsmiddel, laat zich beantwoorden door te kijken hoe de begrippen voorkoming van verwarringsgevaar en misleiding in het merken- en het mededingingsrecht behandeld worden. Deze begrippen staan centraal, omdat het in die rechtsgebieden gaat om de handhaving van het communicatieproces tussen producenten en consumenten, waarbij misleiding van de consument één van de belangrijkste verstorende factoren is. In het Benelux-merkenrecht is nu voor de inbreuksvraag niet bepalend of het publiek werkelijk in verwarring kan geraken. Verwarringsgevaar wordt 'in abstracto' door de rechter vastgesteld[6]. Ook bij misleiding in het mededingingsrecht lijkt mij het accent meer te liggen op de bepaling door de rechter hoe het gemiddelde publiek op de onwaarheid reageert, dan op een empirische meting van het misleidingseffect op de consumenten[7].

Tegen de toepassing van het marktonderzoek wordt verder ingebracht, dat het recht niet beoogt te beschermen tegen bestaande verwarring, maar reeds tegen het gevaar ervan. Het doen van een onderzoek naar bestaande verwarring bij het publiek is dan geen goede graadmeter, omdat het ongewenst is te wachten totdat de verwarring realiteit geworden is.
Helbach concludeert op grond hiervan dat de rechter zelfstandig verwarringsgevaar moet vaststellen en dat het marktonderzoek hier geen hulp kan bieden. De vraag van het verwarringsgevaar vergt bovendien een normatieve beoordeling en bij gebrek aan voorhanden feiten kan de rechter dat het beste beoordelen.
Ook de Duitse rechter is van mening dat alleen als er feitelijke waarnemingen te bewijzen zijn, het nuttig en geboden is een marktonderzoek te doen uitvoeren. Dit is bijvoorbeeld het geval bij de meting van de onderscheidende kracht van het merk, voor hoeveel produkten en diensten en hoe sterk het als identificeringssymbool op de markt dient of kan dienen. Bij de vraag of een niet onderscheidende aanduiding merkbetekenis heeft gekregen, kan marktonderzoek ook een goede dienst bewijzen. Dit zijn allemaal feitelijk vast te stellen zaken, hetgeen het constateren van verwarrings*gevaar* niet is. In die laatste situatie zouden de consumenten om hun mening maar niet naar feiten gevraagd worden[8].

Uit onderzoek naar consumentengedrag blijkt echter dat deze onderscheiding niet verdedigbaar is. In de consumententheorie wordt een onderscheid gemaakt tussen de voorstellingen, die de consumenten reeds hebben over produkten vóórdat zij met deze in de werkelijkheid geconfronteerd worden en de spontane reacties bij de eerste aanblik van de produkten zelf. Beide elementen zijn onlosmakelijk met elkaar verbonden in de percepties van de consumenten en bepalend voor hun waarneming. Het is niet duidelijk waarom hun spontane reacties en feitelijk bestaande situaties wel te meten zouden zijn, maar niet wanneer deze in een laboratoriumsituatie gesimuleerd worden[9].
Ook door ondernemingen wordt de levensvatbaarheid van nieuwe merken en produkten

dikwijls via marktonderzoek getest. Deze 'test-marketing' zoekt niet naar de mening of het oordeel van het publiek, zoals in een verkiezingsonderzoek, maar probeert door de realiteit zo goed mogelijk te imiteren te zien hoe het merk het bij introductie op de markt zou doen.[10].
Ditzelfde is nu mogelijk bij het meten van gevaar voor verwarring, dat door een teken wordt veroorzaakt, ook als die verwarring zich nog niet op de markt heeft voorgedaan. Eerst nadat verwarring of verwarringsgevaar bij de consumenten feitelijk of in de test-situatie is vastgesteld, zal de rechter de rechtsregel normatief kunnen toepassen door te bepalen of de opgetreden verwarring of het dreigend verwarringsgevaar voldoende omvangrijk is gezien alle omstandigheden van het geval, om tot verwarringsgevaar in de zin van het recht te concluderen.

Wat de techniek en de methode van gebruik van het marktonderzoek betreft, zijn er enige duidelijke kritische kanttekeningen te maken bij de huidige gang van zaken. In de eerste plaats bestaat er de neiging bij de rechter om de voorkeur te geven aan open vragen boven gesloten vragen, omdat deze laatste de teneur zouden hebben de ondervraagden te sturen. Hoewel 'leading questions' natuurlijk vermeden moeten worden, lijkt het onjuist gesloten vragen uit een goed opgezette enquête te weren. Het zal uiteindelijk van de probleemstelling afhangen met welk soort vragen de geënquêteerden het best benaderd kunnen worden.
Wel is belangrijk dat de formulering van de vragen op geen enkel wijze suggereert aan de respondenten wat van hen als antwoord verwacht wordt[11]. Tevens wordt in de Verenigde Staten de eis gesteld dat het samenstellen van de steekproef en de opstelling van de vragenlijsten en de interviews volgens objectieve statistische normen uitgevoerd worden en dat daarom de advocaten geen enkele bemoeienis daarmee moeten hebben. De bewijskracht van de resultaten wordt daardoor verhoogd. Ook zou het de objectiviteit ten goede komen als de partij, die het marktonderzoek wenst uit te voeren, de opzet en de uitvoeringsmethode aan de andere partij voorlegt vóórdat het onderzoek uitgevoerd wordt. Daardoor zullen de testresultaten grotere overtuigingskracht hebben dan wanneer één of beide partijen met een ook qua methode geheel eigen onderzoek komen[12].
Deze internationale trend tot verhoging van de objectiviteit van het marktonderzoek zou dan uiteindelijk kunnen uitmonden in een ideale situatie, waarbij in plaats van op initiatief van de partijen, de rechter de opdracht tot een marktonderzoek geeft in de vorm van een deskundigenbericht.

Nu uit het bovenstaande blijkt dat marktonderzoek als methode geen belemmering behoeft te vormen voor een ruimer gebruik als bewijsinstrument dan nu in Nederland in het mededingingsrecht gebruikelijk is, zal aan de hand van een analyse van de toepassingsmogelijkheden in het mededingings- en het merkenrecht in Duitsland en de Verenigde Staten bezien worden of de terughoudendheid in ons land gewettigd is.

2. HET MARKTONDERZOEK IN HET MEDEDINGINGSRECHT

West-Duitsland

Het marktonderzoek wordt voornamelijk toegepast bij de vaststelling van de vraag of

een mededeling in het economisch verkeer, in de vorm van reclame, misleidend is voor de consumenten (Art. 3 van het Gesetz gegen den unlauteren Wettbewerb, UWG). De Lübecker Marzipan-uitspraak van het Bundesgerichtshof[13] geeft aan hoe intensief dit hoogste rechtscollege zich bezighoudt met de sturing en de correctie van het gepresenteerde marktonderzoek. De beide lagere gerechten hadden vastgesteld dat de reclamemededeling dat echte Lübecker marsepein alleen in Lübeck gemaakt wordt, onjuist was omdat slechts een gering deel van de ondervraagden (13,7%) blijkens het onderzoek er waarde aan hechte of deze marsepein ook inderdaad uit Lübeck kwam. De rechter was tot deze conclusie gekomen door naast de vragen omtrent de herkomst een andere vraag beslissend te achten, namelijk of het publiek deze kennis omtrent de herkomst relevant achtte voor de aankoopbeslissing.
Deze vraag: 'Legen Sie persönlich Wert darauf, dass es sich bei "Lübecker Marzipan" auch wirklich immer um Marzipan handelt, das in Lübeck hergestellt ist, oder ist Ihnen das gleichgültig?' gaf echter volgens het Bundesgerichtshof de ondervraagden niet de mogelijkheid om aan te geven, dat zij weliswaar de herkomst waardeerden, maar dat dit slechts één van de eigenschappen van het produkt was.
Daarom stelde het Bundesgerichtshof een nieuwe neutrale vraagstelling voor, namelijk of de consumenten bij gelijke prijs en uiterlijk van het produkt aan een in Lübeck gemaakt marsepein de voorkeur gaven boven een in een andere stad geproduceerde.
Opvallend is dat na terugverwijzing het onderzoeksinstituut, dat vervolgens van het Hof in Keulen de opdracht kreeg, deze aanwijzingen volgde omdat ook de lagere rechter de beantwoording van deze vraag door het publiek doorslaggevend achtte voor de constatering of Lübecker marsepein een herkomst- dan wel een kwaliteitsaanduiding is.
Het Bundesgerichtshof, zo blijkt uit deze beslissing, geeft duidelijke richtlijnen aan de lagere rechters, hoe enquêtes opgezet en op welke wijze de testresultaten geïnterpreteerd moeten worden.

Dat wil niet zeggen dat ten alle tijde automatisch naar het marktonderzoeksmiddel wordt gegrepen. Wanneer de uitleg van een reclamemededeling overduidelijk is, oordeelt de rechterlijke macht op grond van eigen ervaring[14]. Echter, in geval van twijfel corrigeert het Bundesgerichtshof de lagere rechter en eist een marktonderzoek. Zo moest het Hof te Keulen zijn werk overdoen, nadat het meende op grond van eigen 'algemene levenservaring' te kunnen beoordelen of de mededeling 'naturrot' uitsluitend stond voor rode bakstenen, of dat deze benaming ook in reclame voor betonnen stenen, die rood gekleurd waren, gebruikt mocht worden. De zaak werd naar het Gerechtshof terugverwezen omdat 'eine Beurteilung allein auf grund eigener Sachkunde des Berufungsgerichts Bedenken begegnet'.

Een laatste punt, dat in de jurisprudentie veel aandacht krijgt, is dat de enquête zodanig is opgesteld dat rekening wordt gehouden met alle consumenten, tot wie de mededeling gericht is. Het gaat daarbij niet alleen om hen die reeds enige kennis of ervaring van het produkt waarover de mededeling gaat hebben, maar ook om hen die potentieel tot de koperskring kunnen behoren en tot op het moment van de enquête geen actieve interesse hebben getoond[15]
Dit toont nog eens aan dat de hiervoor behandelde eis aan het marktonderzoek, dat het alleen aanwezige feitelijke kennis zou kunnen meten, niet houdbaar is. Ook de reactie van de potentiële doelgroep is meetbaar en relevant voor het testresultaat.

Kennismaking van de jurisprudentie leert dat de rechterlijke macht zich met een sinds de Tweede Wereldoorlog opgebouwde ervaring actief bedient van het marktonderzoeksinstrument.
Dit gebeurt bovendien voornamelijk in kort geding procedures. Hiermee wordt de veel gehoorde stelling gelogenstraft dat het marktonderzoek een te lange opzet vergt om in een kort geding gebruikt te kunnen worden. Het is algemeen aanvaard in de Verenigde Staten dat 'because injury can be so swift and devastating, the courts must handle these cases swiftly and expertly'[16]. Uit de hierna te behandelen Philip Morris- en Reynolds-uitspraken blijkt dat het mogelijk is voor één van de partijen een voor de rechter acceptabele controletest te houden in één weekend[17].
In de rechterlijke uitspraken worden de verschillende testvormen gewogen en de resultaten met elkaar vergeleken. Thomson geeft een overzicht hoe de rechters met deze materie omgaan en soms duidelijk voorkeur geven aan de ene boven de andere testmethode[18].
Een goed voorbeeld hiervan vormen genoemde sigarettenuitspraken. Gedaagde Lorillard had een nationale smaaktest gehouden van zijn sigaret Triumph met laag teergehalte in vergelijking met topmerken van Philip Morris en Reynolds. Door het percentage ondervraagden, dat de vergeleken merken gelijk achtte, op te tellen bij de groep die Triumph beter vond, kwam Lorillard tot de stelling dat 'an amazing 60% said 3 mg. Triumph tastes as good or better than 8 mg. Merit' (van Philip Morris).
Philip Morris bewees van zijn kant op grond van een weekend-enquête volgens de rechter overtuigend dat de consumenten deze advertentie van Lorillard als een superioriteitsclaim begrepen en hij oordeelde deze daarom misleidend. Naar aanleiding van de vraag of een reclamemededeling misleidend is, merkt de rechter op dat 'professionally conducted consumer perception surveys are best evidence of effect of deceptive advertisement on viewers'. Vervolgens worden de door partijen geïntroduceerde enquêtes – vijf in totaal – geanalyseerd en met behulp van deskundigen gewogen.
Deze twee voorbeelden, die met vele andere aangevuld kunnen worden, rechtvaardigen de conclusie dat de Amerikaanse rechter geen oordeel zal uitspreken over het misleidende karakter van een reclamemededeling zonder via een enquête de feitelijke waarnemingen van het publiek vast te doen stellen. Hij zal dit doen door òf zelf een onderzoek te gelasten òf het door partijen aangereikte onderzoek te toetsen.

Wanneer deze rechtsvergelijkende analyse met de *Nederlandse* praktijk vergeleken wordt, is het opvallend dat daar de rechter eigener beweging vaststelt of het publiek door een reclamemededeling misleid wordt. De behoefte om de waarnemingen van het publiek te testen, blijkt niet.
In het Canderel-arrest meende het Hof te Amsterdam zelfstandig te kunnen beoordelen wat de consument verstaat onder de mededeling dat een zoetstof de natuurlijke smaak van suiker heeft en in smaak niet van suiker zou zijn te onderscheiden[19]. Nu het Hof zelf opmerkt dat de uitdrukking 'natuurlijke smaak van suiker' niet duidelijk is, is het toch verre te verkiezen om de wijze waarop het publiek deze reclame waarneemt via onderzoek vast te stellen, dan om dit voor de consument te doen. Het vergelijkbare hiervoor behandelde 'Naturrot'-arrest toont aan dat de West-Duitse rechter terughoudender is in de veronderstelling het eigen oordeel over wat misleidend is over de hoofden van de betrokken consumenten heen te kunnen vaststellen.

Tot wat voor merkwaardige situaties deze zelfstandige rechterlijke beoordeling van misleiding kan leiden, tonen de uitspraken van de President Rechtbank Amsterdam en de Reclame Code Commissie inzake de betekenis van de slagzin: ''t Smaakt beter omdat er lemon en lime in zit' voor de frisdrank Sprite van Coca-Cola[20]. Ziet de President in deze reclame een onschuldige superlatief reclame, zo beslist de Code Commissie dat wel degelijk van misleiding sprake is, omdat de suggestie gedaan wordt dat Sprite ingrediënten bevat die andere dranken, die ook lemon en lime bevatten, missen.
Onbeantwoord blijft op deze wijze welke interpretatie werkelijk in overeenstemming met de consumetenwaarneming is. Door te onderzoeken hoe de consument de reclame opvat, zou hierop antwoord gegeven kunnen worden.
In ieder geval is duidelijk dat de Nederlandse rechter de vraag van misleiding in het mededingingsrecht langs deductieve c.q. normatieve weg beoordeelt, zonder zich empirisch van de consumentenopvattingen, die door marktonderzoek verkregen worden, te vergewissen.

3. MARKTONDERZOEK IN HET MERKENRECHT

In het merkenrecht spelen de begrippen onderscheidingskracht van het merk en de soortgelijke waren, waarop gelijkende tekens gebruikt worden en het verwarringsgevaar dat hieruit kan voortvloeien een centrale rol. Een merk heeft onderscheidingskracht, wanneer het in staat is als herkenningssymbool voor een producent of een dienst te kunnen dienen. Hoe intensiever het merk gebruikt wordt en hoe meer reclame voor dit merk gemaakt wordt, hoe groter de onderscheidende kracht zal worden.
Ook hier zal nagegaan worden in hoeverre het marktonderzoek een bijdrage kan leveren aan de vaststelling van bovenstaande begrippen. Daarbij zal dezelfde weg gevolgd worden als bij het mededingingsrecht. Eerst wordt de West-Duitse en de Amerikaanse standpunten bezien, waarmee dan het Nederlandse standpunt vergeleken wordt.

West-Duitsland

In de Duitse rechtsopvatting wordt veel waarde gehecht aan de dynamische wisselwerking tussen de onderscheidingskracht van een merk en de waren waarvoor het als kenteken dient. Dit leidt tot de geaccepteerde stelregel, dat hoe breder het merk zich in het verkeer heeft gevestigd, hoe breder de beschermingsomvang is[21].
De vaststelling van de zogenaamde 'Verkehrsgeltung' is bij de bepaling van de reikwijdte van de onderscheidingskracht van het merk van groot belang. Hierbij is de consumentopvatting bepalend en er wordt dan ook regelmatig naar het middel van het marktonderzoek gegrepen om deze verkeersopvatting vast te stellen.
Zo besliste het Hof te München dat Alcantara bescherming verdiende, zelfs tegen een weinig gelijkend teken als Amara voor gebruik op uit synthetisch op wildleer gelijkend materiaal gemaakte kledingstukken, omdat van de 17% van de ondervraagden die weleens die type kleren kochten, ruim 60% Alcantara noemde als merk. Op grond van deze uitkomst werd aan dit merk een brede beschermingsomvang gegeven[22].
Bij de vraag of een teken ingeschreven kan worden als merk, zag het Bundesgerichtshof aanleiding om de vraag van mogelijk verwarringsgevaar te doen bepalen via een onderzoek. Omdat slechts 5% van de ondervraagden San Remo op kledingstukken ook begreep als geografische herkomstaanduiding, werd de inschrijving als merk

als niet verwarringswekkend beschouwd[23].
Wanneer het om een beschermingsclaim voor een beroemd merk gaat, wordt voor de vaststelling van de beroemdheid dikwijls een marktonderzoek voorgeschreven, omdat de beschermingsomvang van het beroemde merk bij uitzondering zich ook uitstrekt tot de bescherming van de goodwill[24]. Dat wil zeggen dat het merk ook beschermd wordt tegen gebruik op andere dan gelijk(soortige) producten.
In de enquêtes worden ook meestal vragen opgenomen om vast te stellen voor welke waren of diensten het publiek het merk herkent, omdat dit noodzakelijk geacht wordt om daarvan afgeleid te kunnen bepalen wat soortgelijke waren zijn[25].

Deze voorbeelden geven aan dat er in het Duitse merkenrecht een sterke tendens bestaat om het marktonderzoek te benutten als bewijsmiddel. Hoe actief het Bundesgerichtshof hierbij leiding geeft, wordt verduidelijkt door twee kort na elkaar gedane uitspraken over het uiterlijk en de vormen van waren, de zogenaamde 'Ausstattung' in de zin van art. 25 Warenzeichengesetz[26].
In de eerste zaak ging het om de vraag of Maggi een langjarig gebruik van de kleurcombinatie geel-rood als merkteken kon handhaven tegenover één van Knorr's produktverpakkingen die dezelfde kleuren voerde. Volgens het Bundesgerichtshof moest om dit te beantwoorden vastgesteld worden of de kleurcombinatie als zelfstandig merkteken apart van het totale uiterlijk van het produkt en zijn verpakking onderscheidende kracht had verworven. Hiervoor achtte het Hof een marktonderzoek de aangewezen weg, waarbij twee methoden bruikbaar waren.
In het eerste onderzoek wordt een fles van eiseres (Maggi), geneutraliseerd zonder merk en verdere aanduidingen, alleen met de kleuren getoond aan de ondervraagden. Getest wordt dan of het publiek de kleuren met eiseres verbindt. Dit wordt door het Bundesgerichtshof een acceptabele methode genoemd, omdat bij gecombineerde tekens het publiek in feite zich ook laat leiden door abstracties in zijn herinnering. Wel werd er tegelijkertijd op gewezen dat bij de interpretatie van de uitkomsten van een dergelijk onderzoek rekening gehouden zal moeten worden met het feit dat de ondervraagden een van de werkelijkheid afwijkend beeld wordt getoond.
Als sterkere methode werd aangemerkt een onderzoek waarbij een verpakking van de gedaagde getoond wordt met weglating van zijn firmanaam en vervolgens getest wordt of het publiek deze verpakking zonder merk dan met de eisende partij en zijn produktverpakking associeert.
Het Bundesgerichtshof wees ook in de tweede beslissing op de noodzaak om de verkeersopvatting feitelijk te toetsen, dit ter vaststelling van de onderscheidingskracht en het verwarringsgevaar. De verschillende testmethoden werden uitgebreid besproken met ook hier weer een voorkeur voor de neutralisering van de fles van gedaagde tot uitsluitend de omstreden kenmerken, de rode stippen op de fles, om te onderzoeken of deze verwarring veroorzaakten.

Hoewel de vaststelling van verwarringsgevaar via marktonderzoek in West-Duitsland in de literatuur met enige terughoudendheid beoordeeld wordt, omdat men zich ten onrechte de vraag stelt of gevaar voor verwarring in plaats van verwarring zelf wel via een onderzoek bewezen zou kunnen worden, blijkt uit deze beide arresten dat het Bundesgerichtshof dit niet voor ongewenst of onmogelijk houdt. Het wordt als een begin van bewijs gezien voor de door de rechter uiteindelijk te nemen beslissing als uit het onderzoek gebleken is dat de merktekens verward kunnen worden[27].

In de Verenigde Staten is het marktonderzoek niet meer weg te denken uit de praktijk van het merkenrecht. De rechterlijke macht beschouwt het marktonderzoek als een vanzelfsprekend instrument om verwarring en verwarringsgevaar te constateren en gelast zeker zelf een onderzoek als partijen niet reeds het initiatief hiertoe genomen hebben.
Het bezwaar tegen het gebruik van marktonderzoek ter vaststelling van verwarringsgevaar, omdat het om toekomstige verwarring zou gaan die niet feitelijk aanwezig en daarom niet te meten zou zijn, wordt noch in de jurisprudentie noch in de literatuur gehoord[28]. Blijkbaar is men zich er goed van bewust dat de potentiële marktsituatie en het gevaar van verwarring dat daar kan optreden, even goed te meten is in een nagebootst model als reeds opgetreden verwarring. Hoewel het een steeds weerkerende overweging is, dat 'there can be no more positive proof of likelihood of confusion than evidence of acutal confusion',[29] wordt deze opvatting nergens ter discussie gesteld.

Een tweede opvallend punt in de jurisprudentie is dat een strak onderscheid gemaakt wordt tussen het meten van onderscheidingskracht en van gevaar voor verwarring[30].
Onderscheidingskracht, pregnant omschreven als 'secondary meaning', is de associatie die een merk oproept ten opzichte van een produkt of dienst. Dit betekent dat het onderzoek zich moet richten op '... the non-infringing party's mark and product and the connection the consumer makes between the product bearing the mark and its source'.
Het gaat er in dit geval om de onderscheidingskracht van het merk, waarop inbreuk gemaakt wordt via marktonderzoek, vast te stellen.
Aan de andere kant richt het marktonderzoek inzake verwarringsgevaar zich op 'the product bearing the infringing mark and whether the consumer believes the product bearing that mark emanates from or is endorsed by the plaintiff'. Daarbij moet het inbreukmakende merk zo getrouw mogelijk aan de consumenten voorgelegd worden om vast te stellen of het geassocieerd wordt met het nagebootste merk. Om verwarringsgevaar te bepalen, is het onjuist om het inbreukmakende merk in een beklede vorm aan het publiek te tonen. Als dat gedaan wordt kan hooguit indirect de 'secondary meaning' van het vergeleken onderdeel van het merk van de niet-inbreukmakende partij gemeten worden.
Hoe scherp deze beide onderzoeksmethoden uit elkaar worden gehouden, blijkt uit de 'Zippo-aansterker'-beslissing[31]. De vraag was of het publiek verward zou worden door de wijze waarop het uiterlijk van gedaagdes Rogers-aansteker leek op de 'Zippo' van eiser.
In de eerste plaats verklaarde de rechter uitdrukkelijk dat hij de resultaten van marktonderzoek aanzienlijk hoger achtte ter beantwoording van deze vraag, dan ander bewijsmateriaal, zoals verklaringen van getuigen of deskundigen. Deze laatsten zijn meestal zorgvuldig geselecteerd en geïnstrueerd door partijen, terwijl het marktonderzoek berust op een enquête bij een door middel van een a-selecte steekproef gevormde groep.
De personen uit de steekproef werden met twee onderzoeken geconfronteerd. In de eerste werd een 'Zippo' getoond zonder het woordmerk, om vast te stellen of het uiterlijk van de aansteker zelf voldoende onderscheidingskracht had om als merk te dienen. Daarna werd het publiek een 'Rogers' getoond, niet ontdaan van één van zijn merktekens, maar precies zoals hij in werkelijkheid verkocht werd. Op grond van dit laatste onderzoek achtte de rechter het pas mogelijk vast te kunnen stellen of de 'Roger-aansteker' inderdaad zo verwarringwekkend op de 'Zippo' leek dat de consument direct of

indirect een verband zou leggen naar de maker van deze aansteker.

De scheiding van beide criteria en het niet-gebruiken van 'secondary meaning'-onderzoeksmethoden om verwarringsgevaar te meten, is naast de vanzelfsprekendheid waarmee het marktonderzoek door de rechterlijke macht als onontbeerlijk bewijsmiddel geaccepteerd wordt, het meest opvallende in de praktijk van het juridisch gebruik van het marktonderzoek in de Verenigde Staten.

Nederland

Sinds de beslissing van het Hof te Amsterdam inzake de Silex- en Biotex-merken in 1966 is er in de acceptatie van het marktonderzoek het een en ander veranderd. In dit arrest werd nog overwogen dat het niet om de opvatting van het publiek omtrent de mogelijkheid van verwarring gaat. Bovendien zou het marktonderzoek de werkelijkheid niet voldoende benaderen[32].
Daarentegen oordeelde de President te Haarlem twintig jaar later naar aanleiding van kritiek van één van de partijen dat een bepaalde methode van steekproefsgewijze selectie van voorbijgangers de realiteit onvoldoende zou benaderen, dat een dergelijke methode van onderzoek zeer gebruikelijk en daarom aanvaardbaar is[33].
Blijkbaar ziet de rechter tegenwoordig het marktonderzoek meer als betrouwbare analyse van de werkelijkheid dan voorheen, dit ondanks de statistische formaliteiten.

De kritiek op het gebruik van marktonderzoek ter vaststelling van verwarringsgevaar richt zich voornamelijk op het onjuiste gebruik in enige belangrijke uitspraken[34]. Deze kritiek lijkt mij in de gegeven gevallen juist.
Als gevraagd wordt naar de associatie bij het woord Up en een groot percentage noemt Seven-Up, dan bewijst dat alleen de onderscheidende kracht van Up als merkteken verwijzend naar het produkt van de Seven-Up-onderneming. Dergelijke vragen testen immers associatievermogen, echter geen verwarringsgevaar. Om dat laatste te bewijzen moet het onderzoek zich richten op het jongere teken in zijn werkelijke verschijningsvorm, zoals het zich aan het publiek voordoet. Dan kan onderzocht worden of de respondenten dit in verband brengen met het oudere merk en zijn bron.
In de Cola-beslissing lijkt mij de vraag naar verwarringsgevaar en die naar het beschrijvende karakter van een deel van de merktekens niet voldoende uit elkaar gehouden te zijn door het gelijktijdig stellen van vragen naar beide begrippen.
Als in plaats daarvan eerst vastgesteld was of Cola als soortnaam ervaren werd, had daarna in het onderzoek het verwarringsgevaar tussen de merken Royal Crown Cola en Coca Cola aan de orde kunnen komen. De soortnaamvraagstelling heeft betrekking op het oudere merk, Coca-Cola, terwijl de vraag naar verwarringsgevaar uitgaat van het inbreukmakend karakter van het jonge teken. Door beide vragen tegelijk te stellen, kreeg men nu de merkwaardige situatie dat volgens de deskundigen 40% van de ondervraagden omtrent de herkomst in verwarring verkeerde, terwijl volgens Helbach tevens uit de enquête bleek dat 77% Cola als soortnaam voor een dergelijke drank kende; een resultaat dat vermoedelijk meer door de gevolgde verwarrende onderzoeksmethode veroorzaakt is, dan door het feit dat verwarringsgevaar niet via onderzoek gemeten kan worden.

Dat het van grote betekenis is de vraag naar de onderscheidingskracht te scheiden van

het verwarringsgevaaronderzoek, wordt duidelijk uit de Fenjal-beslissing[35]. Of gedaagdes douche-flacon verwarringsgevaar kon opleveren met de als merk ingeschreven flacon Fenjal, trachtte de producent van Fenjal te bewijzen door een onderzoek waarbij aangetoond werd hoe een geheel wit gekleurde Fenjal-flacon door 78 onderscheidenlijk 69 van de 100 respondenten toch als een Fenjal-flacon werd geïdentificeerd. Het Hof legde echter het resultaat van dit onderzoek terzijde, omdat hier niet de vraag naar de onderscheidende kracht van deze flacon aan de orde was, maar of gedaagdes flacon met opschriften en in zijn werkelijke uitvoering bij het publiek associaties met de Fenjal-flacon wekte.

Toch bepaalt helaas nog maar al te vaak de rechter zelf, niet gesteund door verder onderzoek, het verwarringsgevaar. Dit blijkt bijvoorbeeld uit de Aqua fresh-beslissing. Een opvallende zaak, omdat hier wel onderzoek gedaan was naar de onderscheidingskracht van het oudere merk Prodent[36].
Als 52% van de respondenten een sliert blauw-wit gestreepte tandpasta als Aqua fresh herkent en daaruit de conclusie getrokken wordt dat 'een belangrijk deel van het kopend publiek blauw-wit gestreepte tandpaste associeert met Aqua fresh', dan is het merkwaardig dat de president de verwarringsgevaar veroorzakende elementen van het jongere teken en zijn uiterlijk, zonder verder onderzoek te gelasten, zelf beoordeelt. Was het niet juister geweest om de visuele indruk van het jongere Prodent-teken in zijn verschijningsvorm ook met blauw-witte ondergrond en de mogelijkheid van verwarringsgevaar te toetsen aan de waarneming hiervan door het publiek?

Uit deze recente uitspraken blijkt evenwel dat het marktonderzoek meer en meer ingang vindt. Als ook de rechter meer aandacht zal gaan besteden aan dit middel om verwarringsgevaar empirisch vast te stellen, zou dit de bruikbaarheid voor partijen ongetwijfeld verhogen. In dit verband moet opgemerkt worden dat de Nederlandse rechter zich afwachtender opstelt dan zijn collega's in West-Duitsland en de Verenigde Staten. In die landen wordt regelmatig door de rechter zelf een onderzoek gelast. Daarmee wordt voorkomen dat tegenstrijdige onderzoeken en resultaten eerst tijdens de zitting over en weer bekritiseerd worden. Als er daarentegen wel twee marktonderzoeken geproduceerd worden, is het van belang dat de rechter ook in die situatie het initiatief neemt om de gebruikte methoden en de resultaten te toetsen en daarvan in de motivering te doen blijken.
Zo werden in de Schenkstroop-zaak twee tegengestelde onderzoeksrapporten door partijen voorgelegd. Het onderzoek van gedaagde had als resultaat dat Schenkstroop een sterk beschrijvend karakter had, terwijl eisers onderzoek in de richting van een niet gering associatievermogen wees, namelijk dat 82,8% van de ondervraagden de benaming kende, waarvan 36,2% een verband legde tussen de benaming en eiser.
De overwegingen van het Hof inhoudende, dat geenszins aannemelijk gemaakt was dat de benaming Schenkstroop zodanig was ingeburgerd dat zij voor merkbescherming in aanmerking kwam, geven niet duidelijk aan in hoeverre en op grond van welke resultaten uit de beide onderzoeken tot deze slotsom gekomen is[37].

3. CONCLUSIE

Uit de ontwikkeling die het marktonderzoek als bewijsmiddel in de vergeleken jurisdicties heeft doorgemaakt, blijkt dat de Amerikaanse en de West-Duitse rechterlijke macht zich oorspronkelijk terughoudend opstelden tegenover een onderzoeksinstrument dat voortkomt uit een andere discipline. De empirische vaststelling van de waarnemingen van de consumenten werd als hulpmiddel voor het rechterlijk oordeel minder hoog aangeslagen dan rechters eigen levenservaring en het oordeel van deskundigen. Zo werd het marktonderzoek in de Verenigde Staten op bewijs-juridische gronden tegengehouden als 'hearsay-evidence', dat is bewijs dat niet uit eerste directe waarneming voortkomt.
Eerst langzaam drong het bewustzijn door dat de resultaten van marktonderzoek een objectieve basis vormen om verwarring of verwarringsgevaar te meten.
Het aanvaarden van dit bewijsmiddel ontneemt echter de rechter niet zijn eigen bevoegdheid om de rechtsregel uiteindelijk normatief op de omstandigheden van het specifieke geval toe te passen. Het constateren van een bepaald percentage verwarring(sgevaar) houdt immers nog geen verwarring(sgevaar) in de zin van het recht in.

Opvallend in de Nederlandse ontwikkeling is dat in de recent gepubliceerde jurisprudentie, waarin marktonderzoek een rol speelt, de rechter in zijn overwegingen méér geneigd is aandacht te besteden aan de methode van het onderzoek en de totstandkoming van de resultaten dan voorheen, zodat hierin misschien een begin van aanvaarding gezien kan worden. Echter, de uitgebreide wijze waarop in de Duitse en Amerikaanse rechtspraak de methoden van het onderzoek en de resultaten geanalyseerd worden, is bij ons nog toekomstmuziek. Daardoor is het moeilijk te beoordelen waarom het marktonderzoek wel of niet door de Nederlandse rechter bij zijn uitspraak is betrokken.

Het scherpe onderscheid tussen een onderzoek naar verwarringsgevaar en een onderzoek naar onderscheidingskracht en de fouten die gemaakt worden als het laatste gebruikt wordt om verwarringsgevaar te onderzoeken, tonen aan dat uitgebreidere analyse van de onderzoeksresultaten in de uitspraak, tot een betere beoordeling van de bruikbaarheid en daardoor tot grotere aanvaarding zal leiden.

Nu aangetoond is dat het marktonderzoek als bewijstechnisch instrument geen belemmering vormt voor het meten van verwarring, noch van verwarringsgevaar, is het wenselijk dat het marktonderzoek in het mededingingsrecht inclusief het merkenrecht meer toegepast zal worden. Zodoende zullen de waarnemingen van de consumenten, van de marktdeelnemers zelf, de grondslag vormen van het rechterlijk oordeel.

NOTEN

1. Dit artikel is een bewerking van een inleiding voor de Vereniging voor Mededingingsrecht in 1987 en voor de Vereniging van Marktonderzoek in maart 1988.
2. E.A. van Nieuwenhoven Helbach, 'Het opinieonderzoek als bewijsmiddel in het mededingingsrecht', in: *Een goede procesorde*, Kluwer 1983, p. 287 e.v.; en in *BIE* 1948, p. 167 e.v.
3. Hof Amsterdam 11 december 1986, IER 1986, p. 121.
4. Winfried Tilmann, *Die Verkehrsauffassung in Wettbewerbs- und Warenzeichenrecht*, GRUR 1984, Heft 10, p. 717.

5. Joseph G. Smith, S. Walance, e.a., Legal Standards for Consumer Survey Research, *Journal of Advertising Research*, vol. 23, nr. 5, 1983.
6. W.H. Drucker en G.H.C. Bodenhausen, *Kort begrip van het recht betreffende de intellectuele eigendom*, bewerkt door L. Wichers Hoeth, 1984, p. 95-96.
7. D.W.F. Verkade, *Ongeoorloofde mededinging*, 1986, p. 123 e.v.
8. Roland Knaak, *Demoskopische Umfragen in der Praxis des Wettbewerbs- und Warenzeichenrechts*, 1986, p. 83 e.v.
9. W. Kroeber-Riel, *Konsumentenverhalten*, 1980, p. 55 e.v. en p. 226 e.v.
10. B. Verhage en W.H. Cunningham, *Grondslagen van het marketing management*, 1984, p. 168-171.
    A. van der Zwan en J. Verhulp, *Grondslagen en techniek van de marktanalyse*, 1980, p. 61-64.
11. Robert L. Sorensen, *Survey Research Execution in Trademark Litigation: Does Practice Make Perfection:*, 73 TMR 1983, p. 349.
12. Fred W. Morgan, 'The Admissability of Consumer Surveys as Legal Evidence in Courts', *Journal of Marketing*, Fall 1979, p. 33.
13. BGH, GRUR 1981, p. 71 (Lübecker Marzipan).
14. Tilmann, *op.cit.*, noot 3, p. 719.
15. Gerhard Schricker, *Zum Begriff der 'beteiligten Verkehrskreise' im Ausstattungsrecht*, GRUR 1980, p. 462-470.
16. David I.C. Thomson, *Problems of Proof in False Comparative Product Advertising: How Gullible is the Consumer?*, 72 TMR, p. 385.
17. Philip Morris Inc., 511 F. Supp. 855 (1980).
    R.J. Reynolds Tobacco Co., 511 F. Supp. 867 (1980).
18. Thomson, *op.cit.*, noot 15. p. 393-398.
19. Hof Amsterdam, 5 februari 1987, IER 1987, p. 39 (Canderel).
20. President Rechtbank Amsterdam, 10 september 1987 (Lemon en Lime I), IER 1987, p. 118.
    Reclame Codecommissie, 9 oktober 1987 (Lemon en Lime II), IER 1987, p. 119.
21. A. Baumbach en W. Hefermehl, *Warenzeichenrecht*, 1985, par. 31, nr. 15.
22. OLG München, 5 augustus 1982, GRUR 1983, 322 (Alcantara).
23. BGH, 18 februari 1972, GRUR, 361 (San Remo).
24. OLG Düsseldorf, 17 september 1981, GRUR 1983, p. 389 (Rosenthal).
25. Knaak, *op cit.*, noot 7, p. 57-60.
26. BGH 23 juni 1967, GRUR 1968, p. 371-376 (Maggi).
    BGH 12 juli 1967, GRUR 1968, p. 581-583 (Blunazit).
27. E. Noelle-Neumann en C. Schramm, *Testen der Verwechlungsgefahr*, GRUR 1976, p. 51-63.
28. R. Bradlee Boal, *Techniques for Ascertaining Likelihood of Confusion and the Meaning of Advertising Communications*, 73 TMR 1983, p. 405 e.v.
29. Boal, *idem*, p. 410-411.
30. Vincent N. Palladino, *Techniques for Ascertaining If There Is Secondary Meaning*, 73 TMR 1983, p. 391.
31. Zippo Manufacturing Co., 216 F. Suppl. 670 (1963).
32. Hof Amsterdam, 24 november 1966, BIE 1967, p. 271 (Biotex-Silex).
33. President Rechtbank Haarlem, 10 april 1987, BIE 1987, P. 219 (Homburg/Homeburger).
34. Van Nieuwenhoeven Helbach, *op.cit.*, noot 1, p. 294-295.
35. Hof Amsterdam, 25 februari 1986, BIE 1987, p. 36 (Fenjal-flacon/Karnemelk-flacon).
36. President Rechtbank Haarlem, 3 februari 1987, IER 1987, p. 36 (Aqua fresh/Prodent).
37. Hof Amsterdam, 27 november 1986, IER 1987, p. 30 (Schenkstroop-Schenkhoning).

8. Maakt de gelegenheid de dief?

Onderzoek naar de psychologische achtergronden van het verschijnsel winkeldiefstal

T. POIESZ, D. BAKKEREN EN A. PRUYN

SAMENVATTING

In deze bijdrage wordt een drietal onderzoeken gerapporteerd over de betekenis van psychologische variabelen bij de verklaring van winkeldiefstal door jongeren. De eerste studie toont aan dat jongeren een lage kans op aanhouding percipiëren, en dat ze geen zeer negatieve gevolgen bij aanhouding verwachten. Deze studie kon echter geen duidelijkheid verschaffen over de aard van de motieven tot winkeldiefstal. Studie 2 richtte zich met name op de motieven, gebruikmakend van scenario-beschrijvingen waarbij de jongeren de kans op winkeldiefstal aangaven. Eén van de gevolgtrekkingen uit de tweede studie was dat het vóórkomen van individuele verschillen in winkeldiefstalgedrag toe te schrijven is aan verschillen ten aanzien van de inschatting van de gevolgen van dit gedrag. Studie 3, eveneens met jongeren als onderzoeksdeelnemers, toonde aan dat de kansen op achtereenvolgens detectie, aanhouding, aangifte, toekennen van straf en ondergaan van de straf na het plegen van winkeldiefstal vaak laag worden ingeschat.

1. INLEIDING

Het centrale thema in de marketing is de aankoopbeslissing van de consument. De ene marktpartij biedt een product aan tegen een bepaalde prijs, en de andere marktpartij besluit om het product wel of niet te nemen. Een consument kan 'het nemen van een product' op twee manieren opvatten en in gedrag tot uiting brengen. In de (vooralsnog) meeste gevallen wordt eronder verstaan het meenemen van het artikel uit de winkel nadat er de aangegeven prijs voor is betaald. In een toenemend aantal gevallen wordt er echter alleen het eerste onder verstaan: het meenemen van een artikel zonder dat ervoor betaald wordt, ofwel winkeldiefstal.
Marketing en de marketingliteratuur besteden de meeste aandacht aan 'gewenst' consumentengedrag, het gedrag waarmee de consument zich aan de impliciete wederzijdse afspraak tussen de marktpartijen houdt. De hoeveelheid aandacht die besteed wordt aan deviant gedrag als winkeldiefstal staat in geen verhouding tot de hoeveelheid geld die ermee gemoeid is, een observatie die niet alleen te maken is voor de Nederlandse, maar ook voor bijvoorbeeld de Amerikaanse situatie (Klemke, 1982).
Winkeldiefstal is een fenomeen dat de laatste jaren in omvang is toegenomen en daardoor een groeiende bron van zorg is voor de detailhandel, politie, het justitiële apparaat en de maatschappij in haar geheel (zie Van der Valk, 1986a, 1986b; Moerland & Roder-

mond, 1980). De vaak toch al niet ruime marges van de detailhandel verkleinen als gevolg van diefstal omdat met het gestolene niet alleen de winst op het betreffende item uit de winkel verdwijnt, maar tevens de daarvoor betaalde inkoopprijs en de gemaakte kosten.
Ook voor politie en justitie is winkeldiefstal een probleem. De omvang van het verschijnsel is zodanig dat adequate preventie-, detectie- en correctiemaatregelen niet meer mogelijk lijken. De belasting in mensuren die het met zich meebrengt is aanzienlijk. 'Het oordeel van de minister van justitie nadert het punt waarop het gat tussen het aantal geconstateerde gevallen van winkeldiefstal en het aantal geregistreerde strafrechtelijke correcties zo groot wordt dat de geloofwaardigheid van de justitie in haar geheel in het geding kan komen.' (Minister van Justitie, 2e Kamer, Vergaderjaar 1986-1987, 18995, nrs. 19-20, blz. 21). Voor de maatschappij als geheel is winkeldiefstal een probleem omdat zij uiteindelijk opdraait voor de kosten. Dit betreft de kosten van het gestolene zelf, de kosten van justitie, en de bijkomende kosten van preventie, detectie, vervolging, publiciteit, etc. Er zijn echter ook kosten die meer psychologisch dan strikt financieel van aard zijn. Winkeldiefstal maakt onderdeel uit van de zogenaamde 'kleine criminaliteit', het geheel van activiteiten dat in staat is het leven van de burgers in een aantal opzichten onaangenamer te maken. De burger voelt zich in toenemende mate onveilig, en winkeldiefstal behoort tot het complex van verschijnselen dat het gevoel van onveiligheid veroorzaakt.

De schattingen omtrent de financiële omvang van het verschijnsel winkeldiefstal lopen uiteen, maar worden alle met 9-cijferige bedragen aangeduid. Het zou gaan om enkele honderden miljoenen guldens per jaar. Hier wordt niet ingegaan op de vraag welke schatting de juiste is, maar wordt volstaan met de constatering dat het probleem omvangrijk is en dat het in toenemende mate groeit. Deze constatering wordt ook onderschreven door buitenlandse auteurs, zie bijvoorbeeld Velocci (1978); Moschis, Saliagas & Kellaris (1986). Er is daarom aanleiding tot een nadere beschouwing van de oorzaken van het verschijnsel omdat, in het algemeen, kennis van oorzaken essentieel is voor de vaststelling van de meest gewenste en effectieve tegenmaatregelen.
In deze bijdrage wordt getracht inzicht te krijgen in de oorzaken van winkeldiefstal gepleegd door jeugdigen. Winkeldiefstal wordt dikwijls door jongeren gepleegd en blijkt vaak de poort te zijn tot de criminaliteit (Angenent, 1986, 1980). Nagegaan is wat in de gedragswetenschappelijke literatuur en bij deskundigen bekend is over winkeldiefstal. Daarnaast zijn een drietal studies verricht onder 'echte' deskundigen: (jeugdige) winkeldieven zelf, en jongeren die vaak rechtstreeks met winkeldieven in contact komen.

Bevindingen in de gedragswetenschappelijke literatuur

In de literatuur wordt een beschrijving geboden van een aantal aspecten van het verschijnsel winkeldiefstal: welke mensen plegen winkeldiefstal, wat stelen ze, wat kost het gestolene, waar wordt het gestolen, op welke tijdstippen, en welke vormen van winkeldiefstal zijn er? Voor dit onderzoek werd echter met name gezocht naar mogelijke verklaringen van het verschijnsel. De in de literatuur genoemde oorzaken zijn onder te brengen in een beperkt aantal categorieën: omgevingskenmerken, kenmerken verbonden aan producten en persoonskenmerken.

– *Omgevingskenmerken*
In de traditie waarin omgevingskenmerken verantwoordelijk geacht worden voor het optreden van winkeldiefstal gaat men ervan uit dat de gelegenheid de dief maakt (Cameron, 1964; Edwards, 1970).
De consumenten worden door de omgeving tot winkeldiefstal verleid en het zijn ook de maatregelen in de omgeving die winkeldiefstal moeten terugdringen. Tot de omgevingskenmerken behoren bouwkundige aspecten zoals de toegang tot artikelen en de situering van kassa's, en electronische maatregelen.

– *Productkenmerken*
Het spreekt vanzelf dat de aantrekkelijkheid van het product, zoals die geldt voor een consument/winkeldief een belangrijke oorzaak vormt van winkeldiefstal. De meer luxe artikelen die een duidelijke voldoening geven bij de consumptie zijn meer in trek dan produkten die tot de dagelijkse 'necessities' behoren (Cameron, 1964). Kleine artikelen lopen uiteraard een grotere kans tot de buit van een winkeldief te gaan behoren dan grotere producten (Brady & Mitchell, 1971).

– *Persoonskenmerken*
Tot de persoonskenmerken worden die kenmerken gerekend waarmee individuen zijn te beschrijven. Het betreft naast socio-demografische en -economische variabelen ook psychologische variabelen. Sociodemografische en -economische factoren zijn dikwijls in verband gebracht met winkeldiefstallen, en tevens zijn de meer duurzame persoonlijkheidskenmerken van winkeldieven onderwerp van studie geweest (Moschis et al, 1986; Klemke, 1982; Beck & McIntyre, 1977; Cameron, 1964; Robin, 1963).

Geen van de bovengenoemde soorten factoren of kenmerken bleek behulpzaam bij het vinden van een afdoende en praktisch betekenisvolle verklaring voor het onderhavige verschijnsel. Het verwijzen naar de omstandigheid als de belangrijkste verklaringsgrond maakt niet duidelijk *waarom* er in dezelfde omstandigheden mensen zijn die wel en die niet tot winkeldiefstal overgaan. Ook in praktisch opzicht bieden ze een beperkt houvast; het veranderen van de omstandigheden in winkel A heeft wellicht alleen tot gevolg dat winkeldiefstal naar winkel B wordt geëxporteerd. Het verhuizen van een probleem is niet hetzelfde als het oplossen van dat probleem.
Zo kan ook de sociale groep niet *verklaren* waarom winkeldiefstal optreedt. Omgevingskenmerken en sociale kenmerken kunnen naar onze mening dan ook hoogstens worden opgevat als remmende of stimulerende factoren en niet als de werkelijke oorzaken van winkeldiefstal. Het zijn met name de persoonlijke factoren waarin deze oorzaken moeten worden gezocht. Uiteindelijk zijn het de individuele personen die een beslissing nemen wel of niet tot diefstal over te gaan, hier niet toe gedwongen (maar eventueel wel gestimuleerd) door fysieke of sociale omstandigheden.
In een eerste studie, eind 1986, hebben we daarom getracht nader inzicht te verkrijgen in de aard van de persoonlijke, psychologische factoren die van belang zouden kunnen zijn bij de verklaring van winkeldiefstal.

2. STUDIE 1: PSYCHOLOGISCHE DETERMINANTEN VAN WINKELDIEFSTAL

Vanwege het ontbreken van een theorie omtrent winkeldiefstalgedrag waarin psychologische factoren en hun onderlinge relaties nader zijn uitgewerkt is de eerste studie inventariserend van aard en wordt alleen een globaal onderscheid gemaakt tussen antecedente, psychologische en met winkeldiefstal geassocieerde afhankelijke variabelen.

Tot de antecedente variabelen worden gerekend persoonsgebonden, maar niet psychologische kenmerken. Hieronder vallen leeftijd, geslacht, opleidingsniveau, woonplaats, aantal vrienden (volgens Klemke, 1982, positief gecorreleerd met winkeldiefstal), en hoeveelheid vrij te besteden geld.
De psychologische variabelen die in deze studie zijn meegenomen betreffen percepties van de gevolgen van winkeldiefstal, de evaluatie/acceptatie van winkeldiefstal, de subjectieve norm ten aanzien van winkeldiefstal (in de traditie van Fishbein, 1967) en de mogelijke motieven tot winkeldiefstal.
De afhankelijke variabelen, tot slot, betreffen de frequentie waarmee winkeldiefstal is gepleegd, en de intentie tot het in de toekomst plegen van winkeldiefstal.

Methode

In totaal namen 177 jongeren aan deze studie deel, 106 jongens en 71 meisjes. Hun leeftijd varieerde van 13 tot 19 jaar. De respondenten waren afkomstig van drie verschillende scholen, een middenschool (n = 88) en twee scholen voor lager beroepsonderwijs. Deze scholen waren gesitueerd in het noorden van het land. Onder andere vanwege de locatie en de verdeling over opleidingsniveaus kan deze steekproef niet als representatief worden beschouwd voor de populatie van Nederlandse jongeren in deze leeftijdsklasse.
Vragenlijsten werden afgenomen in de klaslokalen. In enkele klassen kostte het de onderzoeker moeite om de klas onder controle te houden, maar het onderzoek werd niet gesaboteerd.
In een korte inleiding werd de anonimiteit van de respondenten en de vertrouwelijkheid van de gegevens gegarandeerd, en kregen respondenten de uitleg over de procedure die dit moest bewerkstelligen. De procedure was zodanig dat de respondenten de vragenlijst konden invullen zonder dat anderen hun persoonlijke antwoorden konden achterhalen.
De vragen werden zodanig geformuleerd dat de respondenten niet een suggestie kregen ten aanzien van sociaal gewenst gedrag. Zo werd winkeldiefstal in de vragenlijst omschreven als 'het wegnemen van iets uit een winkel zonder ervoor te betalen'.
Geladen woorden als 'stelen' en 'diefstal' werden niet gebruikt.

Resultaten

Van de respondenten geeft 54.2 procent aan wel eens iets weggenomen te hebben zonder te betalen. Het aantal gestolen artikelen in een jaar varieert sterk. Het gemiddelde is 51 artikelen, de mediaan is echter slechts 3, en de modus is 1. De scheve verdeling maakt duidelijk dat het gemiddelde hier geen goede indicatie vormt van de centrale tendentie in de verdeling. De gemiddelde prijs van het gestolene laat eveneens een scheve

verdeling zien, met als gemiddelde bedrag 15.53 gulden, mediaan 3 gulden en modus 1 gulden. De jongeren die in deze steekproef winkeldiefstal plegen stelen dus in het algemeen laaggeprijsde items, met enkele uitschieters naar dure artikelen.
De helft van de respondenten (50.4 procent) geeft aan in de toekomst misschien (ook) iets weg te nemen zonder te betalen. De gemiddelde score op een 5-punts schaal, 5 is 'zeker wel' en 1 is 'zeker niet', is 1.89 met een standaard deviatie van 1.06.

Van elke 100 keer dat respondenten behorend tot de laatste 50.4 procent iets wegnemen uit een winkel zonder te betalen verwacht men gemiddeld 14.5 gepakt te zullen worden. Standaard deviatie: 19.6. De mediaan van deze variabele is 10 keer. Bijna de helft van de respondenten (47.3 procent) zou het 'een beetje onplezierig' tot 'helemaal niet onplezierig' vinden om gepakt te worden. Op de 4-punts schaal (1 = 'helemaal niet onplezierig') is het gemiddelde 2.7 en de standaard deviatie .98. De modus is 2: 'een beetje onplezierig'. Elk van beide laatste resultaten, en met name hun combinatie, is interessant bij de beoordeling hoe afschrikwekkend, in termen van kansen en gevolgen, de risico's van winkeldiefstal zijn.
Respondenten konden aangeven in hoeverre men verschillende sociaal onacceptabele feiten, waaronder winkeldiefstal, als ernstig beschouwde. De winkeldiefstalvarianten, alle betrekking hebbend op een artikel met een winkelwaarde van 25 gulden, werd ter beoordeling voorgelegd. De gemiddelde score op een 10-punts schaal (10 = zeer ernstig) varieerde tussen 6.25 en 6.81, en standaarddeviaties van rond 2.30. Het stelen van iets (grammofoonplaat, sigaretten, snoepgoed, in supermarkt, kleine buurtwinkel of warenhuis) wordt binnen deze steekproef blijkbaar niet als ernstig beschouwd. Bovendien blijken de oordelen weinig te verschillen over verschillende situaties en over verschillende producten.
Een factoranalyse op de 23 items suggereerde 3 factoren: een factor 'kleine criminaliteit', een factor 'criminele feiten' en als derde factor 'overtredingen'. Tot de eerste factor behoorden alle winkeldiefstal-items, en ook items als het stelen van 25 gulden uit een kassa (behoorde dus niet tot factor 2), en het zwartrijden in openbaar vervoer, eveneens ter waarde van 25 gulden. De tweede factor betrof het handelen in drugs, alcoholmisbruik in het verkeer, het stelen van 25 gulden van ouders, en het stelen van hetzelfde bedrag uit het huis van vreemden (wat blijkbaar anders is dan het stelen van 25 gulden uit een kassa).

Met betrekking tot de relatie met de antecedente variabelen zijn de volgende verbanden geobserveerd:
Onderwijs: De opleiding van respondenten werd in twee niveaus verdeeld, hoger en lager. Een tweede onderscheid werd gemaakt tussen respondenten die wel en niet winkeldiefstal hebben gepleegd. De chi-kwadraat van de geconstrueerde 2 x 2 matrix is gelijk aan 10.48. Met één vrijheidsgraad is dit significant op het 1 procentsniveau. Opleiding is negatief geassocieerd met winkeldiefstal. Dezelfde significante relatie treedt niet op ten aanzien van de intentie tot het plegen van winkeldiefstal in de toekomst.
Geslacht: Jongens vertonen iets vaker winkeldiefstalgedrag dan meisjes (chi-kwadraat = 4.47 df = 1, $p < .05$). Ook zijn ze vaker van plan in de toekomst winkeldiefstal te plegen dan meisjes (chi-kwadraat = 8.17, df = 1, $p<.01$).
Leeftijd: Leeftijd is zwak positief gecorreleerd met winkeldiefstal ($r = .14$, $p < .05$). Leeftijd en intentie tot winkeldiefstal in de toekomst zijn niet significant gecorreleerd ($r = .04$, $p = .32$).

Economische/financiële situatie: De frequentie van winkeldiefstal neemt toe met de hoeveelheid vrij te besteden zakgeld ($r = .33$, $p < .001$), de mate waarin dit bedrag als onvoldoende wordt ervaren ($r = .28$, $p < .001$), en de omvang van het extra bedrag dat men nodig acht om net genoeg te hebben ($r = .39$, $p < .001$).
Winkeldiefstal neemt toe met de mate dat men geïrriteerd is over de constatering dat anderen meer te besteden hebben ($r = .25$, $p < .001$). Winkeldiefstal is echter niet geassocieerd met verwachtingen ten aanzien van economische ontwikkelingen en werkgelegenheid (respectievelijk $r = .02$ en $r = .07$).

Voor de *intentie* tot het in de toekomst plegen van winkeldiefstal ziet het patroon van resultaten er ongeveer hetzelfde uit, met dit verschil dat de gevonden correlaties lager zijn. Hoeveelheid vrij te besteden zakgeld: $r = .17$; mate waarin men dit bedrag als onvoldoende ervaart: $r = .14$; omvang benodigd extra bedrag: $r = .13$; hoeveelheid geld relatief ten opzichte van anderen: $r = .06$ (niet significant). De intentie tot winkeldiefstal is zwak negatief (!) gecorreleerd met de gepercipieerde kans op betaalde arbeid in de toekomst: $r = -.17$ ($p < .05$).

Bij de chi-kwadraat matrices dient erop te worden gewezen dat de randtotalen van deze matrices niet overeen hoeven te komen met de totalen zoals die in werkelijkheid gelden, en dat de associaties geen causaliteiten aangeven.

Regressie-analyses werden uitgevoerd om na te gaan hoe de attitude ten opzichte van winkeldiefstal, de sociale norm en de intentie tot winkeldiefstal bijdragen aan de verklaring van winkeldiefstal. De R-kwadraat van de regressieoplossing is .27, met beta-coefficiënten van .40, .21, en –.15 voor sociale norm, intentie en attitude ten opzichte van winkeldiefstal. Indien de intentie tot winkeldiefstal als afhankelijke wordt genomen dan verschijnt alleen de sociale norm als predictor, en wel met een beta-coefficiënt van .34.

Een voor de hand liggend onderdeel in de vragenlijst werd gevormd door de motieven die jongeren zouden kunnen hebben om wel of niet tot winkeldiefstal over te gaan. Verschillende motieven werden in de vragenlijst opgenomen en hun belangrijkheid werd beoordeeld door diegenen die aangaven wel eens iets weggenomen te hebben ($n = 88$). De belangrijkheidsoordelen van de afzonderlijke motieven correleerden niet significant met de frequentie van winkeldiefstal. Het betrof de volgende motieven, met tussen haakjes de bijbehorende correlaties: heb het nodig maar heb het geld niet ($r = .12$); gewoon meedoen met de anderen ($r = .00$); vind het leuk om te hebben maar kan het niet betalen ($r = .12$); voor de opwinding/de 'kick' ($r = .17$); anderen halen me over het te doen ($r = -.09$); om van een slecht humeur af te komen ($r = -.03$); om bij de groep te horen ($r = .08$); en: omdat ik een bepaalde winkeleigenaar niet aardig vindt ($r = .13$).
Twee van deze motieven toonden wel een significante positieve correlatie met de intentie tot winkeldiefstal: om van een slecht humeur af te komen ($r = .26$, $p < .01$), en omdat ik een bepaalde winkeleigenaar niet aardig vindt ($r = .23$, $p < .05$).

Twee discriminant analyses met als afhankelijke variabelen frequentie van winkeldiefstal ('nooit' versus rest) en intentie tot winkeldiefstal ('nooit' versus rest) leverden niet meer duidelijkheid op met betrekking tot motieven tot winkeldiefstal. Zie Tabel 1 voor

Tabel 1. correlaties met de discriminantfunctie.

I: variabelen met betrekking tot winkeldiefstal*frequentie* ('nooit' versus rest);
II: variabelen met betrekking tot winkeldiefstal*intentie* ('nooit' versus rest).

	I	II
Factorscore 'kleine criminaliteit'	–.38	–.41
Factorscore 'criminele feiten'	–.07	
Leeftijd	.21	.31
Aantal vriend(inn)en	.37	.32
Voldoende zakgeld	.41	
Winkeldiefstal door vrienden	.66	.26
Afkeuring winkeld. door gezin	.44	.56
Afkeuring winkeld. door vrienden	.53	.55
Opleidingsniveau	–.26	
'Winkeldiefstal is *altijd* fout'	–.25	–.27
Zich vervelen		.31
'Als je maar regelmatig wel betaalt mag je ook wel eens niet betalen'		.37
'Winsten zijn toch hoog genoeg'		.47
Percentage correcte classificaties*	74.58	70.86
A priori waarschijnlijkheid	50.04	50.00

* n.b. bij de berekening van de discriminantfunctie en bij de classificatie is gebruik gemaakt van dezelfde respondenten.
Analytisch correct(er) is het beide analyses op verschillende groepen respondenten uit te voeren. In dit geval was de omvang van de steekproef hiervoor echter te gering.

de correlaties tussen de discriminerende variabelen en de discriminantfunctie, voor beide criteriumvariabelen.

In deze discriminantanalyses waren onder andere factorscores opgenomen van de factoren die uit de beoordelingen van een groot aantal sociaal onacceptabel geachte gedragingen waren getrokken: 'kleine criminaliteit' en 'criminele feiten'.

In de eerste kolom van Tabel 1 is te zien dat de grootste bijdrage aan de verklaring van het verschil tussen respondenten die nooit en die ooit winkeldiefstal hebben gepleegd wordt geleverd door de variabelen winkeldiefstal door vrienden, afkeuring door vrienden, afkeuring door het gezin, de mate waarin het vrij te besteden zakgeld als voldoende wordt ervaren, de beoordeling van kleine criminaliteit en het aantal vrienden. Deze variabelen zijn over het algemeen ook van belang voor het onderscheid tussen respondenten die zeker niet of (misschien) wel in de toekomst tot winkeldiefstal overgaan. Het verschil met de tweede kolom (intentie tot winkeldiefstal) lijkt met name gelegen in de rationalisaties die voor winkeldiefstal kunnen worden gegeven. Diegenen die eerder geneigd zijn tot winkeldiefstal zijn het meer eens met opmerkingen als:
- 'Als je maar regelmatig betaalt is het niet erg als je eens iets wegneemt zonder te betalen' en
- 'Iets wegnemen zonder te betalen is niet erg want winkels maken toch winst genoeg'.

Deze eerste studie suggereerde enkele psychologische variabelen die voor de verklaring van winkeldiefstal mogelijk relevant zouden kunnen zijn, zoals de lage gepercipieerde kans op aanhouding en de evaluatie van de gevolgen van aanhouding.
Duidelijke motieven konden echter niet worden gevonden. Hiervoor zijn een paar mogelijke redenen aan te geven. Eén van de redenen is dat de motieven die in de vragenlijst ter beoordeling werden voorgelegd niet relevant waren. Ook bestaat de mogelijkheid dat vragen naar achterliggende, abstracte redenen voor respondenten met lagere opleidingsniveaus moeilijk of zelfs onmogelijk is.
Een derde mogelijke reden is dat winkeldiefstal door de onderzoekers teveel als een 'pre-planned' activiteit is opgevat terwijl het in werkelijkheid wellicht meer een impulsgedrag is.
Naar aanleiding van de uitkomsten van dit eerste onderzoek en gelet op deze nabeschouwing is besloten een tweede onderzoek te verrichten. Dit onderzoek bestond uit twee delen: een kwalitatief onderzoek onder deskundigen met als specifiek doel in te gaan op de mogelijke winkeldiefstal-motieven, en een kwantitatief onderzoek onder jongeren waarin de in deze nabeschouwing besproken punten zou worden verdisconteerd.

Ter nadere oriëntatie zijn een aantal gesprekken gevoerd met deskundigen, personen die vanuit hun beroep of functie regelmatig geconfronteerd worden met het verschijnsel, en ook beleidsbeslissingen dienen te nemen ten aanzien van het fenomeen; Gesproken is met vertegenwoordigers van politie, justitie, detailhandel, en opbouwwerk (jongerenwerk) en met een beveiligingsdeskundige. De conclusie die wij over de inhoud van de gevoerde gesprekken moesten trekken is dat men bij de vraag naar de *verklaring* van het verschijnsel geneigd is een nadere *beschrijving* te bieden, of dat men verklaringen geeft van een zeer algemeen niveau (zoals bijvoorbeeld: 'de oorzaak is gelegen in een gebrek aan verantwoordelijkheidsgevoel'). Daar waar motieven door verschillende deskundigen werden genoemd bleken deze niet dezelfde te zijn. Het blijkt daarom niet mogelijk een eenduidige samenvatting te bieden van de inhoud van de gevoerde gesprekken, die informatief waren maar niet op het punt van de motieven tot winkeldiefstal.

In het tweede onderzoek wordt getracht de mogelijke beperkingen van het eerste onderzoek te ondervangen. Uitgangspunt voor de tweede studie is dat motieven zich niet rechtstreeks laten onderzoeken maar dat motieven gaan gelden in een situatie waarin de persoon zich bevindt. In dit onderzoek wordt daarom gekozen voor een scenario-benadering: situaties, die met een bepaald motief zijn geassocieerd worden ter beoordeling aan respondenten voorgelegd. Hiertoe zijn een aantal mogelijke motieven geïnventariseerd. Ze hebben betrekking op:
Klant-onvriendelijk winkelpersoneel, klant-onvriendelijke winkelsituatie, financiële correctie van eerdere kassa-fouten, frustratie, stimulering door anderen, politieke/maatschappelijke factoren, materiële behoeften, relatieve financiële achterstand, gelegenheid tot winkeldiefstal, wraakgevoelens ten opzichte van de winkeleigenaar, marketingtechnieken, verwachte straf bij aanhouding, anonimiteit, en het kopiëren van diefstalgedrag van anderen (modelling).

3. STUDIE 2: BEOORDELING WINKELSCENARIO'S OP KANS OP WINKELDIEFSTAL

Methode

Respondenten:
In totaal namen 131 respondenten aan dit onderzoek deel, 57 jongens (43.5 procent) en 74 meisjes (56.5 procent). Hun leeftijd varieerde van 16 tot 21 jaar. De gemiddelde leeftijd was ongeveer 17 jaar. Opleidingsniveau: kort middelbaar beroepsonderwijs. Het onderzoek vond plaats in de regio Groningen/Friesland. Mede in verband met de geringe spreiding over opleidingsniveaus en de geringe regionale spreiding gaat dit onderzoek niet uit van representativiteit van de steekproef. Resultaten zijn dan ook op te vatten als indicatief.

Procedure:
Het onderzoek vond plaats via vragenlijsten die in anonimiteit konden worden ingevuld en ingeleverd. Het onderzoek vond collectief plaats in een klaslokaal. De leerkrachten van de betreffende scholen werden voor het onderzoek geïnstrueerd door de onderzoekers. In de instructie stond dat het een onderzoek betrof naar 'het wegnemen van iets uit een winkel zonder ervoor te betalen'.

Resultaten

a. Algemeen
Meer dan de helft (58%) van de jongeren gaf aan vroeger wel eens iets weggenomen te hebben uit een winkel zonder te betalen. Dit percentage daalde echter bij de vraag of men de laatste tijd wel eens iets weggenomen had. Slechts 10 procent gaf hierop een positief antwoord, zijnde 13 van de 131 responses. Dat dit een onderschatting zou kunnen zijn van het werkelijke percentage wordt gesuggereerd door het totaal van de antwoorden op de vraag of het idee om iets weg te nemen pas in de winkel of van te voren opkomt. In totaal gaven 27 respondenten op deze vraag een antwoord, hetgeen neerkomt op ongeveer 20%. Het totaal neemt nog toe als gekeken wordt naar de resultaten met betrekking tot de vraag of men achteraf spijt had na het wegnemen van iets zonder betaling. 15 Respondenten geven aan er 'heel erg' of 'nogal spijt' van te hebben en 21 'een beetje' of 'niet'. Het totaal komt hiermee op 36 personen of 27 procent, aannemende dat respondenten alleen reageren op de vraag of ze spijt hadden als ze ook daadwerkelijk winkeldiefstal hebben gepleegd. Overigens is dit percentage beduidend lager dan het percentage dat in het vorige onderzoek is gevonden. In de eerste studie gaf een kleine *meer*derheid aan wel eens iets weggenomen te hebben uit winkels zonder ervoor te betalen. Het verschil is wellicht te verklaren door het verschil in het aantal mannelijke en vrouwelijke respondenten. In het eerste onderzoek waren mannelijke respondenten in de meerderheid, in het tweede vrouwelijke. Jongens zijn vaker betrokken bij winkeldiefstal dan meisjes (zie ook Studie 1). Ook waren de respondenten in de tweede studie gemiddeld een paar jaar ouder dan die in de eerste studie. Daarnaast lijkt een ander onderzoekskenmerk relevant voor de verklaring van het verschil: in het eerste onderzoek is geen onderscheid gemaakt tussen recent en vroeger gepleegde winkeldiefstallen, in het laatste onderzoek werd expliciet gevraagd naar recent gepleegde diefstallen.

b. Beoordeling scenario's
Het gaat om een totaal van 30 scenario's, die dienden te worden beoordeeld op de kans dat de respondent iets zou wegnemen zonder te betalen. Op de beoordelingen over alle scenario's werd een factoranalyse uitgevoerd.
De eerste factor kan geïnterpreteerd worden als betrekking hebbend op het hebben van een gelegenheid tot winkeldiefstal bij afwezigheid van externe controle.
Belangrijke items binnen deze factor zijn (achter elk item wordt de factorlading weergegeven):
- De winkelier stalt zijn spullen zo uit dat ze voor het grijpen liggen (.74)
- Je bent in een winkel in een vreemde stad waar je waarschijnlijk nooit terug zult komen (.73)
- Het is koopavond, het is donker en terwijl je in een winkel loopt vallen de lichten uit (.74)
- Je hebt haast om op tijd te komen op school. Je moet in de winkel nog snel een pen kopen. Er is niemand bij de kassa (.58)
- In een winkel zie je dat een andere klant ongezien een artikel kan wegnemen zonder te betalen (.61)
- In een rek staan doosjes met aardbeien. Hoe groot is de kans dat je een paar aardbeien van een ander doosje bij je eigen doosje doet? (.58)

Factor 2 wordt geïnterpreteerd als reactie op frustratie.
Belangrijke items:
- Op school verveel je je rot. Tijdens de lunchpauze ga je naar een winkel (.67)
- Je hoort in een winkel de winkelier klagen over de jeugd van tegenwoordig (.77)
- Terwijl je in de winkel loopt merk je dat je beurs gestolen is (.59)

Factor 3 duidt op sociale druk:
- Je vriend(innen) vragen je mee te gaan naar een winkel om te proberen iets zonder betaling mee te nemen (.60)
- Je komt erachter dat andere jongeren van jouw leeftijd meer kunnen besteden dan jij (.71)
- Je vriend(innen) nodigen je uit om samen met hen iets weg te nemen uit een winkel (.66)

Factor 4, tenslotte, weerspiegelt gevoelens van wraak:
- Je bent in een winkel waar het personeel heel erg onvriendelijk is (.72)
- Toen je de vorige keer iets kocht in een winkel heeft de cassière je teveel laten betalen. Je bent nu opnieuw in die winkel (.74)
- In de winkel zie je een artikel dat je heel graag wilt hebben maar je vindt de prijs ervoor belachelijk hoog (.58)

Bij deze uitkomsten valt op dat er in feite een grote factor naar voren komt, die bestaat uit een relatief groot aantal items, die onderling vrij hoog correleren. Deze factor (factor 1) verklaart 50% van de variantie. De overige factoren bestaan uit een kleiner aantal items, en verklaren ieder minder dan 5% van de variantie. Als we mogen aannemen dat de item-set waaruit de factoren zijn gedestilleerd een set was waarin alle potentieel relevante motief-aspecten voorkwamen, dan kunnen we stellen dat 4 motieven een rol

spelen bij de verklaring van het verschijnsel. Het eerste motief leent zich echter mogelijk voor een nadere uitsplitsing.

Een kernachtige omschrijving van Factor 1 is: 'de gelegenheid maakt de dief'. Het voorkomen van dit motief lijkt op het eerste gezicht nogal evident. Als er geen gelegenheid is zal er immers geen winkeldiefstal kúnnen voorkomen. De uitkomst is echter ook op een andere manier te bezien. Het zich voordoen van een gelegenheid houdt niet in dat deze gelegenheid ook zal worden benut. Toch blijken zich met name ten aanzien van dit punt de verschillen tussen de antwoordpatronen van de respondenten voor te doen. De gelegenheid, zoals omschreven in de afzonderlijke scenario's, is gelijk voor alle respondenten,winkeldief of niet; toch maakt de eerste er gebruik van en de laatste niet. Dit impliceert dat de gelegenheid de dief maakt, zij het niet alleen.
De gelegenheid vormt een noodzakelijke maar niet een voldoende voorwaarde. Het impliceert tevens dat de oorzaak van en de oplossing voor winkeldiefstal niet uitsluitend gezocht moeten worden in aanpassingen van situaties en producten, maar dat ook gekeken dient te worden naar meer algemene, over situaties heen geldende psychologische determinanten van winkeldiefstal.

Alvorens hier verder op in te gaan is het interessant eerst nader aandacht te besteden aan de inhoud van de eerste factor. Gelet op de bevinding dat deze factor vele items bevat en het overgrote deel van de variantie verklaart is besloten om exclusief op de items van de *eerste* factor een aparte factoranalyse uit te voeren. De vraag die hiermee diende te worden beantwoord was: zijn er in het totaal van de items die de 'gelegenheids'-factor vormen zelf nog patronen te ontdekken die meer duidelijkheid over de aard van de gelegenheden zouden kunnen geven?
Een drie- (in feite twee-, zie later) factoren oplossing leek hiertoe in staat, met 33 procent, 30 procent en 4 procent van de verklaarde variantie voor de respectievelijke factoren (totaal 67 procent). We zullen deze factoren 1a, 1b en 1c noemen om aan te geven dat het hier gaat om een uitsplitsing van de eerste factor uit de oorspronkelijke factoranalyse.

Op *Factor 1a* hebben de volgende items de hoogste lading:
- De winkelier stalt zijn spullen zo uit dat ze voor het grijpen liggen (.79)
- Je bent in een winkel in een vreemde stad waar je waarschijnlijk nooit terug zult komen (.81)
- Het is koopavond, het is donker en terwijl je in een winkel loopt vallen de lichten uit (.68)
- In een winkel zie je dat een andere klant ongezien een artikel kan wegnemen zonder te betalen (.72)

Op *Factor 1b* laden de volgende items het hoogst:
- In een supermarkt zie je een al geopende verpakking van chocoladerepen. Je hebt zo'n trek dat je een reep *in de winkel* zou willen opeten (.55)
- In een rek staan doosjes met aardbeien. Elk doosje bevat 200 gram en kost 2 gulden. Hoe groot is de kans dat je een paar aardbeien van een ander doosje bij de eigen doosje doet (zonder dat je er extra voor betaalt)? (.84)
- Je moet in een supermarkt zelf het fruit verpakken en wegen. De weegschaal maakt dan een bonnetje waarop de prijs staat. Hoe groot is de kans dat je nog wat extra fruit in de zak doet nadat de prijs al is vastgesteld? (.49)

Factor 1c, tenslotte, kent één hoogladend item:
- Je wordt opeens ontzettend verkouden, je hebt geen zakdoek en ook geen geld bij je. Toch moet je iets hebben om je neus te snuiten. In een winkel zie je een pakje papieren zakdoekjes. (.91)

De eigenwaarden van deze factoren zijn 4.54, 4.22 en .54, waaruit blijkt dat er in feite twee factoren zijn. De interpretatie van de inhoud van de factoren lijkt gezocht te moeten worden in de ernst van de diefstal. Factor 1a lijkt te duiden op 'echte' diefstal – het misbruik maken van de situatie ten nadele van de winkelier, waarbij het nadeel niet onbeduidend is. Factor 1b suggereert dat er ook een andere categorie winkeldiefstal is: diefstal tussen aanhalingstekens. Dit soort diefstal lijkt in de ogen van de respondenten niet tot de categorie criminele feiten te behoren maar tot de categorie 'fout maar slim'. De daad wordt onfatsoenlijk geacht maar te onbetekenend om er werk van te maken. Strikt genomen wordt er wel iemand benadeeld, maar het nadeel blijft in de ogen van de respondent beperkt en zal waarschijnlijk niet eens worden opgemerkt door het slachtoffer. De derde soort winkeldiefstal lijkt tot een nog lichtere categorie te behoren. Hoewel de betreffende factor nauwelijks als factor kan worden geïnterpreteerd, ligt de interpretatie ervan in het logisch verlengde van die van de eerste twee factoren: het gaat hier om het wegnemen van iets wat de consument in de genoemde omstandigheid bij wijze van spreken zelfs zonder betaling van de winkelier zou kunnen krijgen als hij erom zou vragen.

De laatste drie factoren maken onderdeel uit van een grote factor die 'gelegenheid' is genoemd. Gegeven de gelegenheid doen zich dus naar het schijnt gradaties voor in de ernst van de winkeldiefstal. Deze gradaties hebben op zich geen duidelijke associatie met bepaalde motieven. Blijft de vraag hoe het komt dat verschillende respondenten verschillend op een gegeven gelegenheid reageren.
In zijn algemeenheid is te stellen dat winkeldiefstal niet mag worden opgevat als een geïsoleerd verschijnsel. Winkeldiefstal en andere vormen van kleine criminaliteit maken onderdeel uit van een verschijnsel waarvan het veranderende normbesef één van de achtergronden vormt. In dit verband moet ook worden gewezen op het belang van de variabele sociale norm bij de verklaring van winkeldiefstal en intentie tot winkeldiefstal in de eerste studie. Veranderend normbesef is echter niet meer dan een beschrijving, het kan niet dienen tot de verklaring van het verschijnsel. Een verklaring lijkt eerder gezocht te moeten worden in de fundering van normen. Voor het (voort)bestaan van normen is het noodzakelijk dat deze gefundeerd zijn in het verband dat mensen zien tussen (hun) gedrag en de gevolgen van dat gedrag. Gedrag dat niet binnen de norm valt moet tot negatieve gevolgen leiden, en gedrag dat wel binnen de norm past tot positieve. Als de (perceptie van de) relatie tussen gedrag en gevolgen teniet gedaan wordt valt tevens het fundament weg onder de normen. De vraag is dan: welke (gepercipieerde) gedrag-gevolg relaties spelen mogelijk een rol bij het winkeldiefstalgedrag? In het geval van winkeldiefstal zijn gevolgen uiteen te rafelen in een aantal componenten die elk een noodzakelijke schakel vormen in de relatie tussen gedrag en haar consequenties. Elk van deze componenten of schakels is met een kansschatting te verbinden:
- Wat is de kans dat winkeldiefstal leidt tot ontdekking?
- Wat is de kans dat ontdekking leidt tot aanhouding door het winkelpersoneel?
- Wat is de kans dat aanhouding resulteert in aangifte bij de politie?
- Wat is de kans dat dit leidt tot de toekenning van een straf?

- Als een straf wordt toegekend, wat is dan de kans dat deze straf ook moet worden ondergaan? en
- Als de straf is ondergaan door de winkeldief, hoe groot is dan de kans dat hij/zij weer tot winkeldiefstal zal overgaan?

Aan elk van deze relaties is een objectieve en een subjectieve kans te verbinden. Voorbeelden van objectieve kansen zijn:
- In 1986 werden bij rijks- en gemeentepolitie bijna 50.000 gevallen van winkeldiefstal geregistreerd. Uit enquêtes bij bedrijven blijkt dat per jaar meer dan een kwart miljoen gevallen van winkeldiefstal worden geconstateerd, en dat het werkelijke aantal gepleegde winkeldiefstallen vele malen hoger is (F. Korthals Altes, Tweede Kamer der Staten Generaal, 18995, nrs. 19-20, blz. 20).
  De objectieve kans om als winkeldief geregistreerd te worden is dus klein.
- Volgens het CBS (1985) en het EIM (1986) wordt in één op de vier à vijf ontdekte gevallen van winkeldiefstal daadwerkelijk aangifte gedaan bij de politie.

Voor degene die wel of niet tot het plegen van winkeldiefstal zal besluiten zijn niet zozeer de objectief vast te stellen kansen als wel de subjectieve inschattingen van deze kansen relevant.
Het is op basis van deze subjectieve kansen dat de afweging wordt gemaakt of de beslissing wordt genomen. Nu de eerste twee studies naar winkeldiefstal-motieven niet de gewenste duidelijkheid hebben geboden, en de sociale norm een relevante variabele is gebleken (zie Studie 1) is het van belang vast te stellen hoe jongeren de genoemde relaties tussen gedrag en mogelijke consequenties van dat gedrag inschatten. Hiertoe is een derde studie onder jongeren uitgevoerd.

4. STUDIE 3: DE BEOORDELING VAN GEDRAG - GEVOLG RELATIES

Het specifieke doel van dit onderzoek was inzicht te verkrijgen in de validiteit van de voorlopige conclusie dat de oorzaak van (kleine) criminaliteit mede te zoeken is in de subjectieve waarneming van relaties tussen oorzaken en gevolgen in het detectie- en vervolgingstraject door jongeren. Daarnaast had het onderzoek tot doel nadere gegevens te verkrijgen over de beoordeling van bestaande in-store preventiemaatregelen.

Methode

Respondenten:
In totaal namen 163 jongeren aan dit onderzoek deel, 66 jongens en 97 meisjes. De gemiddelde leeftijd was 16.4 jaar, met een standaarddeviatie van .86 jaar. Drie opleidingsniveaus waren in de deelnemersgroep vertegenwoordigd: Lager Huishoudelijk en Nijverheids Onderwijs (LHNO), Middelbaar Algemeen Vormend Onderwijs (MAVO) en Voorbereidend Wetenschappelijk Onderwijs (VWO). Leeftijdsgroepen waren min of meer gelijk over de schooltypen gespreid (respectievelijk 53, 55 en 55), maar mannen en vrouwen niet omdat de LHNO-groep uitsluitend vrouwen bevatte (mannen op respectievelijk MAVO en VWO 26 en 40; vrouwen op LHNO, MAVO en VWO respectievelijk 53, 29, 15). In de analyses wordt met deze ongelijke verdeling van mannen en

vrouwen over de opleidingsniveaus rekening gehouden. De betreffende scholen bevinden zich in Brabant. Op elke school hebben twee volledige klassen aan het onderzoek deelgenomen.
Gelet op onder andere de verdeling van respondenten over de beide geslachten en de beperkte geografische spreiding van de respondenten mag niet uitgegaan worden van representativiteit van de steekproef voor de populatie van Nederlandse jongeren binnen deze drie opleidingsniveaus.
Bij sommige resultaten werd het van belang geacht deze te kunnen contrasteren met de dito resultaten afkomstig van een kleine steekproef uit de andere partij in de markt, i.c. financiële directeuren van middelgrote tot grote kledingzaken, alle met filialen. In totaal werden 16 personen in een mondeling interview verzocht te reageren op een aantal items zoals die ook aan de jongeren werden voorgelegd. Waar relevant worden de betreffende resultaten vergeleken.

Materiaal:
Respondenten kregen een vragenlijst met open en gesloten vragen. In de vragenlijst werd elke suggestie van goed- of afkeuring van winkeldiefstal vermeden. Naar winkeldiefstal werd weer verwezen met de aanduiding: 'het wegnemen van spullen uit winkels zonder ervoor te betalen'. De volgorde van de vragen was zodanig dat eerst vragen over winkeldiefstal in het algemeen werden gesteld en daarna vragen over het eventuele eigen winkeldiefstal-gedrag.

Procedure:
Het onderzoek werd bij de scholieren aangekondigd als een onderzoek naar het kopen van dingen. De docent van de betreffende klas introduceerde de onderzoeker die vervolgens de procedure verbaal en visueel communiceerde. Nadruk werd gelegd op de anonimiteit en de maatregelen ter garantie hiervan. De respondenten werden verzocht de ingevulde vragenlijsten in een blanco gesloten envelop te doen en deze ter plekke aan de onderzoeker te retourneren. De deelnemers kregen de toezegging dat elke deelnemende klas na de analyses en de rapportage een samenvatting van de resultaten zou ontvangen.

Resultaten

Is deze resultatensectie zullen achtereenvolgens aan bod komen:
1. het eigen winkeldiefstal-gedrag van respondenten;
2. de inschattingen van de oorzaak-gevolg relaties; en
3. de evaluaties van de maatregelen ter preventie van winkeldiefstal
(Deze volgorde correspondeert niet met de volgorde van de items in de vragenlijst).

ad 1. *Het eigen winkeldiefstal-gedrag van de respondenten*
Op de vraag of de jongeren 'vroeger wel eens iets mee hadden genomen uit een winkel zonder te betalen' antwoordde bijna de helft (45.4 procent) positief. ('Vroeger' werd niet gedefinieerd). Dit percentage daalde aanzienlijk bij de voorzichtig geformuleerde vraag of dit de laatste tijd het geval is. Een positief antwoord gaf hier slechts 8 procent van de respondenten. Deze respondenten hadden ook allen 'vroeger' winkeldiefstal gepleegd. Als spontaan in een open vraag opgegeven reden voor winkeldiefstal werd opgegeven: besparing (54 procent), spanning (27 procent) en 'proletarisch winkelen' (9

procent). Deze percentages hebben betrekking op 45 procent van de totale groep, dus op ruim zeventig respondenten (hetgeen de percentages enigszins relativeert).
Respondenten dienden een schatting te geven van 'hoeveel van de 100 personen wel eens dingen wegnemen uit winkels zonder ervoor te betalen'. Het gemiddelde aantal (percentage) was 44 mensen, met als standaarddeviatie 23. Het opgegeven aantal houdt geen verband met het eigen winkeldiefstal-gedrag in het verleden. Tussen respondenten die wel en niet aangaven vroeger winkeldiefstal te hebben gepleegd is geen significant verschil geconstateerd tussen de inschatting van de frequentie van winkeldiefstal door anderen (gemiddelde, respectievelijk, 46 en 43 procent).
De gemiddelde percentages van de opleidingen verschillen wel significant van elkaar: 57, 36 en 40 voor LHNO, MAVO en VWO, respectievelijk ($F\ 2,\ 161 = 13.18$, p. = 0.000). Respondenten van de beide hogere opleidingen schatten het percentage mensen dat winkeldiefstal pleegt significant lager in dan respondenten van de lagere opleiding. Dit is geen effect van de confounding van opleiding en geslacht, omdat ook het effect bij meisjes alléén significant is: 57, 41 en 41 voor LHNO, MAVO en VWO ($F\ 2,\ 94 = 6.68$, $p = .02$).

ad 2. *De inschattingen van de oorzaak – gevolg relaties*
De zes verschillende oorzaak-gevolg relaties zijn ter beoordeling aan de respondenten aangeboden. Het betrof de volgende relaties:
- het door het personeel ontdekt worden van een geval van winkeldiefstal (plegen-ontdekt worden);
- het door het personeel aanhouden van een gedetecteerde winkeldief (ontdekking--aanhouding);
- het door het winkelpersoneel aangeven van de aangehouden winkeldief aan de politie (aanhouding-aangifte);
- het verkrijgen van straf na aangifte bij de politie (aangifte-toekennen straf);
- het daadwerkelijk moeten ondergaan van de straf na de toekenning van de straf (toekennen straf-straf); en
- het na het ondergaan van de straf wederom overgaan tot winkeldiefstal (straf-winkeldiefstal).

Respondenten werden verzocht per relatie een indruk weer te geven van het aantal mensen op de 100 die van de ene situatie in de andere zouden geraken. Een voorbeeld van de vraagstelling zoals die bij de zes relaties is gehanteerd: 'Stel dat 100 mensen zijn aangehouden door het winkelpersoneel. Hoeveel mensen, denk jij, zullen door het personeel worden aangegeven bij de politie?'
Tabel 2 bevat de percentages zoals die door de respondenten van de verschillende opleidingsniveaus zijn ingeschat. Tevens is in deze tabel opgenomen het gemiddelde percentage over alle respondenten, en het gemiddelde percentage zoals dat ingeschat werd door de financiële directeuren.

In Tabel 2 is te zien dat de geobserveerde percentages de verwachting ondersteunen dat consumenten/jongeren een betrekkelijk lage inschatting hebben van de kans dat in het opsporings- en vervolgingstraject een bepaalde gebeurtenis consequent tot het (door de wet bedoelde) gevolg leidt. Strikt theoretisch is de maximum score telkens 100 procent, hoewel dit percentage uiteraard niet als reëel referentiepunt kan dienen.
Voorbijgaand aan de vraag wat het haalbare percentage is, is te constateren dat de con-

Tabel 2. Gemiddelde inschatting van oorzaak-gevolg relaties onderscheiden naar opleidingniveau, voor het totaal van de jongeren en voor financiële directeuren (B).

Relatie	*LHNO*	*MAVO*	*VWO*	*TOTAAL*	*(B)*
diefstal-detectie	28	28	21	26	27
detectie-aanhouding	54	69	72	65	28
aanhouding-aangifte	48	65	59	57	39
aangifte-straf	37	54	48	46	09
straf-ondergaan	42	58	62	54	25
.....					
ondergaan-diefstal	50	42	55	49	65

sument de eerste stap in het traject als de minst riskante ervaart. Ook is te constateren dat zich verschillen tussen de opleidingsniveaus voordoen (die mogelijk te wijten zijn aan verschillen tussen mannen en vrouwen, hierover later).
De ondervraagde financiële directeuren nemen de oorzaak-gevolg relaties als nog zwakker waar dan de ondervraagde jongeren. Hierbij dient in aanmerking te worden genomen dat het aantal ondervraagde directeuren beperkt is en dat hierdoor de standaardfout van de gevonden percentages hoog is. Toch kunnen we constateren dat de bij deze personen gevonden percentages over de gehele linie lager liggen dan voor de ondervraagde jongeren. Omdat de relaties geacht worden met elkaar in een logisch volgtijdelijk verband te staan en omdat de relaties in kansen zijn uitgedrukt, is het, althans volgens mathematische principes, mogelijk de kans te berekenen die betrekking heeft op de overgang van een eerder naar een later punt in het traject. Hiertoe dienen de kansen van de betreffende stappen te worden vermenigvuldigd. De overall kansen, verkregen door de combinatie (vermenigvuldiging) van de kansen, staan weergegeven in onderstaand schema, voor zowel jongeren als financiële directeuren.

	jongeren	*fin. direct.*
winkeldiefstal-detectie	26.0	27.0
detectie-aanhouding	16.9	7.6
aanhouding-aangifte	9.6	3.0
aangifte-straf	4.4	0.3
straf-ondergaan	2.4	0.1

In deze wijze van benadering 'zien' jongeren een 2.4 procent kans dat winkeldiefstal uiteindelijk leidt tot het ondergaan van een straf.
Het is evident dat een dergelijke berekeningswijze in zoverre riskant is dat het product van de kansen afhankelijk is van het aantal relaties dat de onderzoeker aan de respondent ter beoordeling voorlegt. Het is zeer waarschijnlijk dat in geen enkele relatie de kans gelijk is aan 100 procent. Het toevoegen van meerdere relaties leidt daarom tot het (kunstmatig?) verlagen van de overall kans, wat vereist dat de relaties en de bijbehorende kansen met de nodige reserve worden geïnterpreteerd.
In ieder geval wijzen de relaties op de meer kritieke fasen in het opsporings- en vervolgingstraject, en geven ze aan waar de perceptie van de consument het meest afwijkt van de feitelijke of gewenste toestand.

Vanwege de niet-onafhankelijkheid van de factoren opleiding en geslacht was het niet

mogelijk een two-way variantie-analyse op deze factoren uit te voeren. Daarom zullen significantietoetsen binnen de groep mannen en binnen de groep vrouwen afzonderlijk plaatsvinden en niet over de totale groep. Zie de Tabellen 3.2. en 3.3.

Tabel 3. Gemiddelde inschatting van oorzaak-gevolgrelaties voor vrouwen voor de drie opleidingsniveaus.

Relatie	*LHNO*	*MAVO*	*VWO*	*F 2, 94*	*sign.*
diefstal-detectie	28	34	23	2.10	n.s.
detectie-aanhouding	54	68	74	4.55	<.01
aanhouding-aangifte	48	62	59	2.69	n.s.
aangifte-straf	37	54	46	4.01	<.05
straf-ondergaan	42	56	68	5.24	<.01
.....					
ondergaan-diefstal	50	41	69	7.88	<.001

Tabel 3 laat zien dat er bij vrouwen duidelijke verschillen zijn tussen de opleidingsniveaus voor wat betreft de inschattingen van de oorzaak-gevolg relaties. In het algemeen kan worden gesteld dat de kansschattingen stijgen met het opleidingsniveau, hoewel niet consequent lineair. De scheidslijn is duidelijk gelegen tussen LHNO enerzijds en MAVO en VWO anderzijds. Een soortgelijke vergelijking is te maken voor mannen, voor wie in dit onderzoek resultaten van twee opleidingsniveaus beschikbaar zijn. Zie Tabel 4.

Tabel 4. Gemiddelde inschatting van oorzaak-gevolg relaties voor mannen voor twee opleidingsniveaus.

Relatie	*MAVO*	*VWO*	*F 1, 63*	*sign.*
diefstal-detectie	23	22	.04	n.s.
detectie-aanhouding	70	73	.24	n.s.
aanhouding-aangifte	67	60	.25	n.s.
aangifte-straf	54	50	.33	n.s.
straf-ondergaan	61	62	.02	n.s.
.....				
ondergaan-diefstal	44	50	.69	n.s.

In Tabel 4 valt op dat de percentages die voor mannen gevonden worden bij de verschillende relaties in het algemeen hoger zijn dan die voor vrouwen. Tevens valt op dat ook hier het onderscheid tussen MAVO en VWO, net als bij vrouwen, niet significant is.

Naast een onderscheid naar geslacht en opleidingsniveau zijn de verschillende relaties ook onderscheiden naar het eerdere winkeldiefstalgedrag van de respondent. De bij de relaties opgegeven percentages staan vermeld in Tabel 5 samen met de bijbehorende aanduiding van de significantie van het verschil tussen de groepen. De groepen 'laatste tijd gestolen' en 'vroeger gestolen' zijn in deze uitwerking verder niet gescheiden.

Tabel 5. Beoordeling van de oorzaak-gevolg relaties, onderscheiden naar het wel of niet gepleegd hebben van winkeldiefstal.

	Winkeldiefstal gepleegd		
Relatie	*ja*	*nee*	*sign.*
diefstal-detectie	22	31	0.003[1]
detectie-aanhouding	69	62	n.s.
aanhouding-aangifte	57	57	n.s.
aangifte-straf	48	45	n.s.
straf-ondergaan	57	52	n.s.
ondergaan-diefstal	49	49	n.s.

1. F 1, 161 = 9.10

Opvallend is dat in Tabel 5 alleen een verschil wordt geobserveerd tussen wel- en niet winkeldieven bij de relatie diefstal-detectie. Bij de overige relaties zijn de eventuele kleine verschillen niet significant (hetgeen overigens niet impliceert dat deze relaties irrelevant zijn). Voormalige winkeldieven schatten de kans op ontdekking bij het plegen van winkeldiefstal lager in dan de niet-winkeldieven. Mogelijk zijn winkeldieven (uit ervaring?) optimistischer over het ondekt worden. Uiteraard is uit deze resultaten geen indicatie te verkrijgen over het bestaan en de richting van een eventueel causaal verband.

De eerste van de zes relaties blijkt de minst 'harde' te zijn. Respondenten verwachten slechts in één op vier of vijf gevallen van feitelijke winkeldiefstal ontdekt te zullen worden. Dit *lijkt* erop te wijzen dat maatregelen ter voorkoming of ontdekking in de winkel relatief belangrijk zijn ten opzichte van maatregelen gericht op de versteviging van de andere relaties. Dit houdt *niet* in dat maatregelen gericht op de andere relaties daarom onbelangrijk zouden zijn. Integendeel: maatregelen gericht op deze andere relaties zijn wellicht effectiever, omdat ze aansluiten op situaties waarin de dief reeds gedetecteerd is. Voordat hierover concrete uitspraken gedaan mogen worden dient echter eerst te worden vastgesteld hoe de relaties worden gerpercipieerd in een representatieve steekproef van Nederlandse jongeren. Het is in dit verband van belang erop te wijzen dat de verschillende steekproeven verschillende indicaties hebben gegeven over de frequentie waarmee winkeldiefstal wordt gepleegd en over de gepercipieerde kans tot aanhouding. Het is niet uit te sluiten dat ook de percepties van de kansen in andere oorzaak-relaties afhankelijk zijn van de aard van de steekproef.

De kritieke vraag ten aanzien van de oorzaak-gevolg relaties is of (potentiële) winkeldieven ze inderdaad als zodanig ervaren.
Duidelijk is gemaakt dat de berekening van de overall-kans op het uiteindelijk moeten ondergaan van straf na het plegen van winkeldiefstal afhankelijk is van het aantal relaties dat de onderzoeker door de respondent laat scoren. Het realiteitsgehalte van de afzonderlijke relaties is dus een belangrijk thema dat in vervolgonderzoek nader aan de orde dient te komen. Dit thema is tevens nauw verwant aan de vraag of winkeldiefstal opgevat moet worden als een 'pre-planned' of als een 'impulse' gedrag. Het is niet

houdbaar dat een consument, voor het schap staande, eerst een inschatting maakt van de verschillende kansen zoals die bij de relaties horen om dan vervolgens op basis van de overall kansschatting te bepalen of diefstal een acceptabel risico inhoudt. Het is wellicht meer aannemelijk dat de kansschattingen behorend bij de relaties die met het optreden van politie en justitie te maken hebben verband houden met de bereidheid om überhaupt winkeldiefstal te overwegen, dus *voordat* men met een concrete gelegenheid wordt geconfronteerd. De kans op detectie en aanhouding in de winkel, ofwel de gelegenheid (Factor 1 in Studie 2), vormt dan een 'go- no go'-factor. In deze opvatting zou het beter zijn de relaties die hierboven als eerste in de reeks werden vermeld achteraan in de serie te plaatsen. De eerste afweging betreft winkeldiefstal in het algemeen; de tweede afweging heeft betrekking op diefstal in een specifieke winkelsituatie.
Afhankelijk van de uitkomst van de eerste afweging komt men aan de tweede afweging wel of niet toe. Op zich doet de volgorde van de relaties echter niets af aan hun plausibiliteit.

Als we mogen aannemen dat de door respondenten opgegeven kansschattingen binnen de gegeven steekproef validiteit bezitten, hoe moeten deze dan worden opgevat? De kans van 60 procent bij bijvoorbeeld de relatie aanhouding-aangifte kan men opvatten als een zeer positief teken, als een mooi resultaat voor de winkeldiefstalpreventie. Anderzijds is een meer realistische visie te hanteren met de volgende overwegingen:

a. een percentage van 60 wijkt sterk af van het ideale percentage;
b. deze relatie vormt slechts een kleine stap in een serie van stappen, en deze serie is als totaal van stappen nog zeer zwak;
c. winkeldiefstal is nog steeds een groeiend verschijnsel ondanks de genomen maatregelen;
d. respondenten zijn wellicht pessimistischer ten aanzien van de kansen zoals *anderen* die hebben dan ten aanzien van hun *eigen* kansen (zoals ook uit Studie 1 bleek); de gevonden percentages zijn dus eerder over- dan onderschattingen van de percepties die betrekking hebben op de eigen persoon; en
e. een geobserveerd percentage is alleen te interpreteren in verhouding tot het percentage zoals dat in een nulmeting is geconstateerd. Een nulmeting is niet beschikbaar, dus mag de mogelijkheid niet uitgesloten worden geacht dat de gevonden percentages gelijk zijn aan of misschien zelfs lager dan die van een aantal jaren geleden. Met andere woorden, er dient voor gewaakt te worden dat dit soort percentages via selectieve waarneming *a priori* opgevat worden als een ondersteuning voor reeds gevoerd beleid. Wel is het mogelijk de relaties, na hun nadere validering, onderdeel te laten uitmaken van een instrument waarmee vorderingen op het terrein van winkeldiefstal-preventiemaatregelen kunnen worden geregistreerd, en wel per fase van het detectie- en vervolgingstraject. De opgegeven kansen zijn immers tevens op te vatten en te gebruiken als indicatoren voor de mate waarin het anti-winkeldiefstalbeleid succesvol is.

ad 3. *Evaluatie preventiemaatregelen*
Vooruitlopend op de validering van de afzonderlijk relaties kan worden gesteld, dat van deze relaties die tussen winkeldiefstal en detectie in ieder geval aandacht behoeft. Dit is de relatie met de laagst waargenomen kans. In reëele winkelsituaties worden een aantal maatregelen gehanteerd die met name op deze relatie betrekking hebben. Daarom is de respondenten verzocht een oordeel te geven over de maatregelen die dienen ter voorko-

ming of ter detectie van winkeldiefstal. In twee open vragen werd nagegaan welke maatregelen respondenten spontaan noemen en welke hiervan volgens hen 'het beste werken om ervoor te zorgen dat er geen dingen worden weggenomen zonder te betalen'. De in de spontane antwoorden sterkst naar voren komende maatregelen zijn: electronische systemen (68%), camera's (59%), inktsystemen (36%), spiegels (30%), personele bewaking (30%) en het winkelpersoneel (22%). De hiervan als best beoordeelde maatregelen zijn: electronische systemen (44%), inktsystemen (14%) en camera's (11%).
Bij deze resultaten dient in aanmerking te worden genomen dat elders in de vragenlijst een uitleg werd gegeven van de wellicht onbekende termen 'electronische systemen' en 'inktsystemen'. Deze uitleg was niet zichtbaar tijdens de beantwoording van de open vraag, maar een eventuele bias is niet uit te sluiten.

De beoordeling van de maatregelen is tevens gevraagd in een gesloten vraag waarin de verschillende soorten maatregelen aan de respondent werden voorgelegd voor een meer expliciete vergelijking. De resultaten staan vermeld in Tabel 6.

Tabel 6. Beoordeling op een 10-punts schaal (10 is meest effectief) van in-store preventiemaatregelen.

	gemiddelde	*stand. dev.*
max. drie scholieren in winkel	5.46	2.22
max. drie kledingstukken in paskamer	5.10	2.53
paskamerbewaking	6.09	2.66
electronisch systeem	8.56	1.24
goede verlichting	4.31	2.34
camera's	7.42	1.87
spiegels	6.00	2.19
kettingen met sloten	6.91	2.48
afsluitbare kleerhangers	6.63	2.46
inktsysteem	7.25	2.23
waarschuwingsborden	3.57	2.33
bewaking in uniform	6.03	2.64

Electronische systemen, camera's en inktsystemen worden door de respondenten als het meest effectief ervaren, een goede verlichting en waarschuwingsborden als relatief het minst. Hierbij is wel in aanmerking te nemen dat het effect van sommige maatregelen door de respondent tijdens de beoordeling beter te visueliseren is dan andere. Goede verlichting bijvoorbeeld kan door een respondent het best worden beoordeeld in een concrete situatie. Toch zou de maatregel effectief kunnen werken. De in Tabel 6 gegeven oordelen zijn dus mogelijk beïnvloed door de mate waarin de respondent zich de betreffende maatregelen concreet kan voorstellen in winkelsituaties.

De beoordelingen van de drie als meest effectief beoordeelde maatregelen verschillen significant van elkaar. Electronische systemen scoren positiever dan camera's ($t = 6.9$, $df = 162$, $p = .000$), en positiever dan inktsystemen ($t = 7.18$, $df = 160$, $p = .001$). De gemiddelde beoordelingen van camera's en inktsystemen verschillen echter niet significant ($t = .80$, $df = 160$, $p = .43$). Camera's worden overigens significant effectiever be-

oordeeld dan bewakers in uniform ($t = 5.90$, $df = 162$, $p = .000$) en dan spiegels ($t = 7.79$, $df = 162$, $p = .000$), hoewel elk van deze maatregelen (minstens) gericht is op visuele registratie. Er is in zoverre een interessant verschil tussen electronische systemen en inktsystemen dat de eerste tracht de actieradius van de winkeldief te beperken. De grens van de actieradius wordt gevormd door een relatief klein gebied rondom de poortjes bij de winkeluitgang. De beperking door inktsystemen is gelegen in de veronderstelde vernietiging van een kledingstuk bij pogingen de inkthouder te verwijderen. Daarom is gevraagd hoeveel tijd de respondenten dachten nodig te hebben voor het thuis verwijderen van de klemmen van beide systemen. Het gemiddelde aantal minuten voor inktsystemen is 71 minuten (st. dev. 80 minuten) en 48 (st. dev. 64 minuten) voor de electronische systemen. Dit verschil is significant ($t (130) = 5.11$, $p = .000$).
Opvallend was dat (voormalig) winkeldieven en niet-winkeldieven nauwelijks verschilden in hun beoordeling van de effectiviteit van de verschillende maatregelen. Personen die de laatste tijd hadden gestolen zagen, in vergelijking met personen die de laatste tijd niet stalen, camera's, spiegels en waarschuwingsborden als significant minder effectief (resp. gemiddelden op een 10-punts effectiviteitsschaal, 10 is meer effectief: camera's 6.1. vs. 7.5; spiegels 4.5 vs 6.1; waarschuwingsborden 2.2 vs. 3.7; resp. $t (161) = 2.61$, $p = 0.01$; $t (161) = 2.57$, $p = 0.01$; en $t (161) = 2.19$, $p = .03$). Door deze respondenten werd effectiviteit van kettingen met sloten echter positiever beoordeeld dan door de niet-winkeldieven (resp. 8.3 vs. 6.8; $t (159) = 2.07$, $p = .04$).

5. ALGEMENE NABESCHOUWING

In de gerapporteerde studies is getracht om vanuit de invalshoek van het consumentengedrag een nader inzicht te bieden op de psychologische achtergronden van het verschijnsel winkeldiefstal. Met nadruk dient te worden gesteld dat de studies methodische (steekproef)-beperkingen hadden. Anderzijds zijn indicaties verkregen die aangeven in welke richting de aanpak van winkeldiefstallen zinvol kan zijn. Dit type onderzoek is daarom in twee aspecten functioneel. Het geeft aan hoe consumenten zelf tegen het thema winkeldiefstal aankijken en suggereert daarmee waar de zwakke plekken zitten die voor beleid toegankelijk zijn. In de tweede plaats worden concrete suggesties gedaan tot het ontwikkelen van een instrument dat in staat moet zijn de veranderingen in de genoemde zwakke plekken te registreren. In de derde plaats wordt aangegeven dat verschillende groepen consumenten de afschrikwekkende werking van de reeds gehanteerde maatregelen op een verschillende manier percipiëren. Dit impliceert dat een technische vergelijking van anti-winkeldiefstal maatregelen aangevuld zou moeten worden met een vergelijking zoals die door consumenten(groepen) wordt gemaakt.
Gelet op de omvang en de groei van het verschijnsel winkeldiefstal, en rekening houdend met de problemen die het voor de verschillende partijen in de markt en daarbuiten met zich meebrengt is het noodzakelijk het verschijnsel vanuit meerdere invalshoeken tegelijk te benaderen: de juridische, de (bedrijfs)economische, de technische, de sociologische en de psychologische invalshoek. In deze bijdrage is nader ingegaan op de psychologische invalshoek. Duidelijk is dat een psychologische benadering complementair is aan de andere benaderingen omdat hierin nagegaan wordt hoe de maatregelen die door de andere disciplines worden gesuggereerd hun uitwerking hebben op de individuele consument. Duidelijk is ook dat het onderzoek op dit terrein beperkt is en dat de hier gerapporteerde studies de start dienen te vormen van een theoretisch en methodologisch uitgebreider onderzoeksprogramma.

LITERATUUR

Angenent, H. *Winkeldiefstal.* 's Gravenhage, VUGA, 1980.

Angenent, H. *Over lange vingers en ladenlichters; alles over winkeldiefstal.* Groningen, Wolters Noordhoff, 1986.

Beck, E.A. & McIntyre, S.C. MMPI Patterns of Shoplifters within a College Population. *Psychological Reports*, 1977, 41, blz. 1035-1040.

Brady, J.F. & Mitchell, J.G. Shoplifting in Melbourne. *Australian and New Zealand Journal of Criminology*, 1971, 4, blz. 154-162.

Cameron, M.O. *The Booster and the Snitch: Department Store Shoplifting.* New York, Free Press of Glencoe, 1964.

Centraal Bureau voor de Statistiek, 1985. Zie *Jaarverslag Openbaar Ministerie 1986.* 's Gravenhage, Staatsuitgeverij 1987.

Edwards, L. *Shoplifting and Shrinkage Procedures for Stores.* Springfield, Illinois, Thomas, 1970.

EIM, Economisch Instituut voor Midden- en Kleinbedrijf, 1986. Zie *Jaarverslag Openbaar Ministerie 1986.*

Fishbein, M. Attitude and the Prediction of Behavior. In: Fishbein, M. (ed.) *Readings in Attitude Theory and Measurement*, New York, Wiley, 1967.

Jaarverslag Openbaar Ministerie 1986. 's Gravenhage, Staatsuitgeverij, 1987.

Klemke, L.W. Exploring Juvenile Shoplifting. *Sociology and Social Research*, 1982, 67, blz. 59-75.

Korthals Altes, F. *Actieplan voor de aanpak van veel voorkomende criminaliteit.* Tweede Kamer der Staten Generaal, vergaderjaar 1986-1987, 18955, nrs. 19-20.

Moerland, H. & Rodermond, J.G. *Winkeldiefstal? Een analyse van ongewenst klantengedrag* Niet gepubliceerd rapport. Mededelingen van het Juridisch Instituut, 12, Erasmus Universiteit Rotterdam, 1980.

Moschis, G.P. Cox, D. & Kellaris, J.J. *An Exploratory Study of Adolescent Shoplifting Behavior.* Paper Annual Conference Association for Consumer Research, Toronto, 1986.

Robin, G.D. Patterns of Department Store Shoplifting. *Crime and Delinquency*, 1963, blz. 163-172.

Van der Valk, W.D.M. Meer personeel als wapen tegen winkeldiefstal? *EIM, Mededelingen*, 1, 1986a.

Van der Valk, W.D.M. Winkelcriminaliteit fors toegenomen. *EIM, Mededelingen*, 5, 1986b.

Velocci, T. Can Business Collar White Collar Crime? *Nations Business*, 1978, 66, blz. 35-44.

9. Afdanken van een duurzaam consumptiegoed

G. ANTONIDES

SAMENVATTING

De betrouwbaarheid en de overlevingskans van wasmachines is onderzocht m.b.v. gegevens uit een consumenten-survey in 1983. De overlevingskans van de wasmachine hangt af van de faalkans van het produkt en van de afdankbeslissing van de consument, gegeven een defect. Determinanten van de faalkans en van de afdankbeslissing zijn onderzocht en conclusies voor de marketing van duurzame goederen en het beleid van consumenten-organisaties worden getrokken.

1. INLEIDING

Marktonderzoek is doorgaans gericht op het aankoopgedrag van de consument. De gebruiks- en de afdankfase van consumptie krijgen veel minder aandacht in het marktonderzoek. Toch zijn gebeurtenissen in de latere fase van consumptie van groot belang voor de vervanging van duurzame goederen, herhalingsaankopen en merktrouw van consumenten. Determinanten van de afdankbeslissing m.b.t. een duurzaam goed behoren derhalve evenzeer tot het terrein van het marktonderzoek als de determinanten van de aanschafbeslissing.

Onderzoek naar de beslissing tot afdanken betreft meestal de *reden* van afdanken, zie bijvoorbeeld Jacoby e.a. (1977) en Hanson (1980). Afdankredenen, genoemd door consumenten, zijn bijvoorbeeld het falen of slijtage van een produkt, verhuizing, en incidentele inkomsten. Uit Tabel 1 blijkt dat een defect of slijtage de voornaamste reden vormt voor vervanging van grote huishoudelijke toestellen, vooral bij de wasmachine. In Conn e.a. (1978) wordt bij kleine huishoudelijke apparaten in 50% van de gevallen een defect als reden van afdanken opgegeven. De mate waarin tijdens de levensduur defecten optreden wordt aangeduid met de betrouwbaarheid van het produkt.

Tabel 1. Defect/slijtage als percentage van het aantal redenen voor vervanging.
Bron: Statistisch Bulletin, Centraal Bureau voor de Statistiek (1982).

koelkast	66.2%
diepvriezer	61.4%
wasmachine	90.1%
kleurentelevisie	80.2%
zwart-wit televisie	74.3%
vaatwasmachine	67.1%

Het falen van een produkt (hier beschouwd als synoniem van een defect) betreft een

storing van de functievervulling tijdens de levensduur van een produkt. Hoewel falen de oorzaak van afdanken kan zijn, kunnen zeer veel produkten echter ook worden gerepareerd. Voor het marktonderzoek naar reparatiediensten zijn, naast de betrouwbaarheid van een produkt, ook de determinanten van een reparatiebeslissing van belang.

De beslissing tot afdanken impliceert meestal het einde van de levensduur van een produkt, tenzij het produkt nog tweedehands in de oorspronkelijke functie of voor een andere dan de oorspronkelijke functie wordt gebruikt. Voor het onderzoek naar tweedehands markten en de vervangingsmarkt is de levensduur van een produkt van belang. De levensduur van duurzame consumptiegoederen wordt enerzijds bepaald door de ingebouwde betrouwbaarheid en duurzaamheid, anderzijds door een afdankbeslissing van de consument. In de termen 'technische levensduur', 'economische levensduur' en 'psychologische veroudering' (Box, 1981; Laurentius, 1987) komt het complementaire karakter van technische en sociale determinanten van de levensduur echter niet tot uitdrukking. Deze termen zullen hier dan ook niet wordt gebruikt.

De overlevingskans van een produkt betreft de kans dat een goed op een bepaalde leeftijd nog niet is afgedankt. Onderzoek naar de voornaamste reden van afdanken, het defect, resulteert in een faalanalyse van het produkt. In deze faalanalyse wordt er meestal geen rekening mee gehouden dat vorige defecten beslissingen tot reparatie (dus niet-afdanken) impliceren. Dit leidt waarschijnlijk tot vertekende schattingen van de faalkans, de overlevingskans en de levensduur van produkten.

Volgens Tabel 1 wordt van de grote huishoudelijke toestellen de wasmachine het vaakste afgedankt om reden van een defect. Tevens is de penetratiegraad van wasmachines in Nederland zeer hoog: 87% (CBS, 1982). Vanwege de onderzoekbaarheid en het belang als consumptie-artikel is de wasmachine geselecteerd voor een studie naar het afdanken van een duurzaam consumptiegoed (Antonides, 1988). De studie betreft een analyse van de betrouwbaarheid en de levensduur van een duurzaam goed, c.q. de wasmachine, en de determinanten van een afdank- of reparatiebeslissing, gegeven een defect.

2. DE LEVENSDUUR VAN EEN DUURZAAM CONSUMPTIEGOED

De kans op beëindiging van de levensduur van een produkt op of vóór de leeftijd t wordt voorgesteld door de functie G(t):

(1) $$G(t) = \text{Prob}\ (T \leq t)$$

G(t) is een cumulatieve verdelingsfunctie, corresponderend met de dichtheidsfunctie g(t). De waarschijnlijkheid dat een produkt precies op tijdstip t wordt afgedankt is g(t). Voor een directe schatting van g(t) is het nodig produkten tijdens de gehele levensduur waar te nemen. In praktijkonderzoek kan echter meestal niet zo lang worden geobserveerd en beschikt men veelal slechts over cross-sectionele gegevens, d.w.z. over waarnemingen van produkten die reeds enige tijd hebben overleefd. De kans op afdanken in het leeftijdsinterval (t,t+h), gegeven het overleven van t, is de conditionele waarschijnlijkheid:

$$(2) \qquad (G|\,t) = \frac{G(t+h) - G(t)}{1 - G(t)}$$

Als h nadert tot 0 kan de hazard, of afdankintensiteit, $\theta(t)$, worden gedefinieerd als:

$$(3) \qquad \theta(t) = \lim_{h \to 0} \frac{1}{h} \frac{G(t+h) - G(t)}{1 - G(t)} = \frac{g(t)}{1 - G(t)}$$

met g(t) de kansdichtheids-functie van afdanken. De afdankintensiteit kan worden beschouwd als de kans op afdanken in een oneindig klein interval (t,t+h), gegeven de leeftijd t. De overlevingsfunctie kan nu worden uitgedrukt als een functie van de hazard door beide zijden van (3) te integreren:

$$(4) \qquad 1 - G(t) = \exp \{ - {}_0\int^t \theta\,(u)\,du \}$$

Zie o.a. Barlow en Proschan (1967), Kalbfleisch en Prentice (1980) en Cox en Lewis (1966).
Vergelijkingen (2)-(4) kunnen direct worden toegepast op gegevens uit een consumenten-survey waarin het afdanken van goederen gedurende enige tijd wordt geobserveerd.

De parameter θ hangt niet alleen af van de leeftijd van een duurzaam goed, maar ook van de kwaliteit en van het gebruik ervan. In het onderzoek is de prijs, P, beschouwd als proxy voor de kwaliteit, en de grootte van het huishouden, H, als proxy voor de gebruiksfrequentie van het goed. Een parametrisering van θ is:

$$(5) \qquad \theta(t) = \exp\,(\alpha_0 + \alpha_1 \ln P + \alpha_2 \ln H + \beta \ln t)$$

Zoals blijkt uit vgl. (3) kan θ (t) worden opgevat als de kans op beëindiging van de levensduur in een oneindig klein interval (t,t+h), gegeven de leeftijd t. Bij wasmachines wordt de levensduur meestal beëindigd door een afdankbeslissing van de consument, als gevolg van een defect. Zowel de faalkans van het produkt als de afdankkans, gegeven een defect, dienen te worden beschouwd. Dit leidt tot een samengestelde afdankintensiteit. Onder verwaarlozing van andere oorzaken van afdanken wordt deze nl. ontbonden in een faalintensiteit, ϕ, en een afdankkans, gegeven een defect, δ (de conditionele afdankkans). De samengestelde afdankintensiteit is dan $\phi\delta$. Bij een constante conditionele afdankkans δ en een parametrisering van ϕ, analoog met (5), luidt de ontbinding:

$$(6) \qquad \delta\,\phi\,(t) = \delta \exp\,(\alpha_0' + \alpha_1' \ln P + \alpha_2' \text{in } H + \beta' \ln t)$$

De faalintensiteit en de conditionele afdankkans kunnen worden geschat m.b.v. gegevens uit een consumenten-survey waarin defecten, reparaties en afdanken gedurende een bepaalde periode worden waargenomen. De uitkomsten zijn hier niet vertekend, zoals in de klassieke faalanalyse, omdat rekening is gehouden met de waarschijnlijkheid dat een goed wordt waargenomen in het survey.

3. AFDANKEN GEGEVEN EEN DEFECT

De economische literatuur m.b.t. afdanken van duurzame goederen laat de reden van afdanken buiten beschouwing. Wel worden soms psychologische variabelen als economische verwachtingen en depreciatie van goederen gemodelleerd, naast economische variabelen als inkomen, leeftijd en kosten van goederen, en prijsindices (Westin, 1975; Stone en Rowe, 1957, 1958, 1960; Jorgenson, 1974).

In het model van paragraaf 2 is de afdankreden als het ware ontbonden in een faalcomponent en een (conditionele) afdankcomponent. De afdankcomponent is beschouwd als een constante omdat informatie over deze component corresponderend met defecten in het verleden zonder een longitudinaal onderzoek nauwelijks is te verkrijgen. In praktijk is de conditionele afdankkans natuurlijk niet constant.

De voorwaardelijke afdankkans kan nader worden beschouwd in een survey waarin defecten, afdank- en reparatie-beslissingen en hun determinanten worden waargenomen. De conditionele afdankkans kan worden gespecificeerd m.b.v. een economisch-psychologisch model van de determinanten van de beslissing. In het navolgende wordt dit economisch-psychologisch afdankmodel beschouwd.

De beslissing over afdanken of repareren, gegeven een defect, kan worden beschouwd als een discrete keuze uit twee alternatieven. De kans op een bepaald keuze kan worden gespecificeerd als een normale verdeling, N, van een functie van determinanten van de beslissing, f(Z):

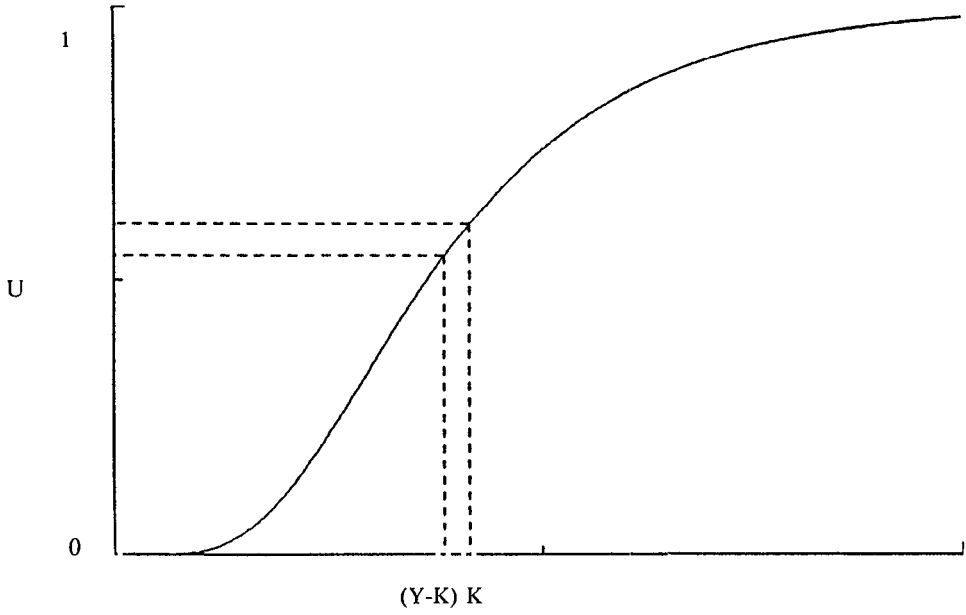
(7)
$$\text{Prob (repareren)} = N\{ f(Z) ; 0,1 \}$$
$$\text{Prob (afdanken)} = 1 - N\{ f(Z) ; 0,1 \}$$

Met behulp van een Probit-analyse (Maddala, 1983) kan de functie f(Z) worden geschat uit waargenomen beslissingen, gegeven een defect. Eerst zullen nu de determinanten van de afdankbeslissing worden beschouwd.

Aangenomen wordt dat consumenten rationeel beslissen in de zin dat gekozen wordt voor het alternatief met het grootste nut. Het nut bestaat uit de som van het nut, ontleend aan de consumptie van het produkt in kwestie, U, en van het nut, ontleend aan de consumptie van alle andere goederen en diensten, V (assumptie van additief separabel nut). Het is duidelijk dat uitgaven aan het goed in kwestie leiden tot minder consumptie van andere goederen en diensten, en dus tot een een lagere waarde van V. Voor de alternatieven van reparatie, 1, en afdanken, 1′, is het nut U(1)+V(1) resp. U(1′)+V(1′). De functie f(Z) kan nu worden geschreven als:

(8)
$$\begin{aligned} f(Z) &= f[\{U(1)+V(1)\} - \{U(1')+V(1')\}] \\ &= \xi_0 + \xi_1 \{V(1) - V(1')\} + \xi_2 \{U(1)-U(1')\} \end{aligned}$$

Het nut van alle uitgaven aan goederen en diensten wordt door Van Praag (1968) voorgesteld als een lognormale functie van het huishoudbudget, Y, met locatie-parameters μ en σ, Λ (Y; μ, σ). Een voorbeeld van deze functie, de individuele Welvaartsfunctie van het Inkomen (WFI) genoemd, is weergegeven in Figuur 1. Uit Figuur 1 kan het nut



Figuur 1. Een voorbeeld van een individuele Welvaartsfunctie van het Inkomen.

worden afgelezen dat correspondeert met het netto huishoudinkomen, Y. Stelt men de uitgaven aan het duurzame goed op K, dan zal de evaluatie van de uitgaven aan andere goederen en diensten, Y-K, overeenkomen met Λ (Y-K; μ, σ). De evaluatie van de uitgaven Y-K wordt gesubstitueerd voor het nut, ontleend aan andere goederen en diensten, V, in (8).

De WFI wordt gemeten m.b.v. enquête-vragen van het volgende type:
'Welke netto gezinsinkomen zou u, in uw omstandigheden, beschouwen als zeer slecht?'
De vraag wordt herhaald voor inkomens die worden beschouwd als resp. slecht, onvoldoende, voldoende, goed en zeer goed. De gerapporteerde inkomens worden uitgezet op de horizontale as van Figuur 1, de kwalificaties zeer slecht t/m zeer goed krijgen ieder een numerieke waarde, u_i, op de verticale as, zodanig dat de intervallen tussen deze kwalificaties gelijk zijn. De lognormale functie wordt geschat m.b.v. de regressie:

$$Y_i = \mu + \sigma N^{-1} (u_i) + \epsilon_i \tag{9}$$

met N^{-1} de inverse van de standaard normale verdeling.
De uitgaven aan het duurzame goed, K, zullen i.h.a. verschillen naargelang de keuze van het alternatief. Bij afdanken van het oude produkt en vervanging door een nieuw moet direct een uitgave voor de aanschaf worden gedaan. Bij reparatie moet een uitgave voor de reparatiekosten worden gedaan (de mogelijkheid van garantie blijft hier buiten beschouwing) en, na verloop van tijd, voor een nieuw produkt als het gerepareerde produkt aan vervanging toe is. Om de eenmalige uitgaven bij reparatie en vervanging te

kunnen vergelijken worden de kosten per tijdseenheid berekend. Het uitgangspunt hiervoor is de contante waarde van (toekomstige) uitgaven, gedeeld door de tijdsspanne waarin de uitgaven worden gedaan. Toekomstige uitgaven worden gedisconteerd m.b.v. een discontovoet, ρ.
De contante waarde van de uitgaven in het reparatie-alternatief, E(1), wordt berekend uit de eenmalige reparatiekosten, R′, de reparatiekosten per tijdseenheid, R, de aanschafprijs van een nieuw produkt, P, de resterende levensduur van het oude produkt, T_0, en de levensduur van het nieuwe produkt, T_1. De eventuele opbrengt van het oude produkt (resp. de verwijderkosten) blijven hier buiten beschouwing.

$$E(1) = R' + P \exp(\rho T_0) + \int_0^{T_0+T_1} R \exp(-\rho t)\, dt \tag{10}$$

De contante waarde van de uitgaven in het afdank-alternatief, E(1′), wordt berekend uit de reparatiekosten per tijdseenheid, R, de aanschafprijs van een nieuw produkt, P, en de levensduur van het nieuwe produkt, T_1.

$$E(1') = P + \int_0^{T_1} R \exp(-\rho t) dt \tag{11}$$

Zij K de gemiddelde uitgaven aan het duurzame goed per tijdseenheid, gedurende een periode T, dan geldt (Su, 1975):

$$\begin{aligned} E &= \int_0^{T} K \exp(-\rho t)\, dt \\ K &= \rho E / \{ 1 - \exp(-\rho T) \} \end{aligned} \tag{12}$$

Volgens (12) kunnen de gemiddelde uitgaven per tijdseenheid in de beide alternatieven, K(1) resp. K(1′), worden berekend. Vgl. (10) en (11) zijn restrictief om complicaties in de modelvorming en in de schattingen te vermijden. Zo kunnen reparatiekosten toenemen met de leeftijd van een goed of verschillen naargelang de kwaliteit ervan, terwijl deze hier constant zijn genomen.

In (8) komt de term V(1)-V(1′) voor. Deze term zal nu eerst worden uitgewerkt door subsitutie van de WFI en de gemiddelde uitgaven per tijdseenheid, K(1) en K(1′).

$$\begin{aligned} V(1)-V(1') &= \Lambda \{Y-K(1); \mu,\sigma\} - \Lambda(Y-K(1'); \mu,\sigma\} \\ &\approx \lambda (Y; \mu,\sigma) \{K(1')-K(1)\} \\ &= \lambda (Y; \mu,\sigma) [\{T_0/(T_0+T_1)\}\{P(1+\rho T_1)/T_1-R'/T_0\}] \end{aligned} \tag{13}$$

met $\lambda(.)$ de eerste afgeleide van $\Lambda(.)$ en enige algebra. Vervolgens worden de termen U(1) en U(1′) in (8) beschouwd.

Het nut van het duurzame goed in het reparatie-alternatief wordt opgevat als een multi-attribuut attitude t.a.v. de kenmerken (attributen) van het produkt.

$$U(1) = \sum_{j=1}^{J} b_j w_j / \sum_{j=1}^{J} w_j \tag{14}$$

met b_j de opinie t.a.v. attribuut j en w_j het gewicht dat aan attribuut j wordt gegeven.

Het nut van het duurzame goed in het afdank-alternatief wordt opgevat als de partiële welvaart genoten door consumptie van het nieuwe produkt, Kapteyn e.a. (1979) beschouwen de partiële welvaart als een lognormale functie van de uitgaven die t.b.v. het produkt worden gedaan. De partiële welvaartsfunctie wordt bepaald m.b.v. enquête-vragen van het type:
'Er is een grote kans dat een wasmachine mij zeer goed zou bevallen als ik minstens *f*... zou uitgeven'.
De vraag volgt na een korte inleiding over prijzen en kwaliteiten van wasmachines en wordt een aantal keren herhaald met (achtereenvolgens) de kwalificaties goed, redelijk, matig, slecht en zeer slecht (Kapteyn e.a., 1978). De schatting van de partiële welvaartsfunctie verloopt analoog met de regressie in (9). Gegeven de uitgave aan een nieuwe wasmachine wordt het nut van het produkt in het afdank-alternatief, $U(1')$, berekend.

De reden dat het nut van het produkt in de beide alternatieven op verschillende wijzen wordt bepaald is dat tijdens de keuze van het alternatief de informatie over het produkt verschillend is. Van het oude produkt heeft de consument een tamelijk gedetailleerd beeld, zodat de multi-attribuut attitude kan worden bepaald. Van het nieuwe produkt zal waarschijnlijk nog geen gedetailleerd beeld bestaan, maar de orde van grootte van de aanschafprijs is vaak wel bekend. Men zou in het laatste geval kunnen zeggen dat de generieke keuze (vervanging) is gemaakt, maar de specifieke keuze (van het produkttype) nog niet (Van Raaij en Wandwossen, 1978).

4. HET CONSUMENTEN-SURVEY

In 1983 is een onderzoek gehouden onder 1253 bezitters van een wasmachine van vijf jaar of ouder (Schelbergen en Serail, 1987). De leeftijdsbeperking is aangebracht omdat uit een leeftijdsverdeling van wasmachines (CBS, 1982) bleek dat vóór het vijfde levensjaar nauwelijks wordt afgedankt. Bij de start van het survey werd de respondenten (telefonisch) gevraagd een toekomstige defect aan de machine te melden. Tevens werd gevraagd naar de grootte van het huishouden, het inkomen (in klassen) en de financiële situatie van het huishouden, de aanschafprijs van de oude wasmachine, de datum van het meest recente defect, opinies m.b.t. produkt-attributen, en de verwachte resterende levensduur.

Na de melding, maar vóór een beslissing over afdanken of repareren was genomen, werd een telefonisch vraaggesprek gevoerd om nadere informatie over de determinanten van de afdankbeslissing te verkrijgen. Gevraagd werd naar de ernst van het defect en de verwachte reparatieprijs, de moeite om informatie over de reparatie of vervanging te verkrijgen, de verwachte resterende levensduur van de oude wasmachine, opinies t.a.v. attributen en attribuut-gewichten, de verwachte aanschafprijs en de verwachte levensduur van een nieuwe wasmachine, en de financiële situatie van het huishouden.

Zes weken na het defect werd een mondeling interview afgenomen om achtergrond-informatie te verkrijgen. Dit betrof de beslissing over afdanken of repareren, het exacte inkomen van het huishouden, de welvaartsfunctie van het inkomen, de partiële wel-

vaartsfunctie, de reparatieprijs resp. de aanschafprijs en het merk van de wasmachine, en het gebruik en de defect-geschiedenis van het oude produkt. Met dit interview werd het survey voor de betreffende consument beëindigd.

Tenslotte werden, ongeveer tien maanden na de start van het survey, alle consumenten in het survey, die geen defect hadden gemeld, telefonisch ondervraagd om ontbrekende meldingen op te sporen. Hiermee werd ook voor deze consumenten het survey beëindigd.

5. RESULTATEN

Van de 1253 consumenten in het survey hebben er 216 een defect aan hun wasmachine gemeld. Uit de slot-enquête bleek dat 140 defecten niet waren gemeld, deze waren i.h.a. van minder ernstige aard. Het percentage afdankbeslissingen in de beide groepen was 31%, resp. 15%. Vergeleken met een overeenkomstige steekproef van het CBS (1982) bevat het survey 10% meer vierpersoons huishoudens en wat minder eenpersoons huishoudens. Op een aantal andere kenmerken, de leeftijd van de hoofdkostwinner, het inkomen van het huishouden, de grotte van de woonplaats, en de leeftijdsverdeling van de wasmachine, verschillen de huishoudens in het survey niet of nauwelijks van die in de CBS steekproef.
De schatting van de afdankintensiteit (5) is uitgevoerd op 950 waarnemingen met bekende leeftijd en aanschafprijs van de wasmachine (excl. tweedehands exemplaren). De aanschafprijs is gecorrigeerd m.b.v. een CBS-reeks van prijsindices voor wasautomaten. Hoewel in de specificatie van de afdankintensiteit (5) voor een belangrijk deel rekening is gehouden met heterogeniteit van de waarnemingen (proxy's voor kwaliteit en gebruiksfrequentie), is het mogelijk dat er nog niet-geobserveerde heterogeniteit van de waarnemingen aanwezig is. Hiermee is rekening gehouden door toevoeging van een storingsterm aan (5), waarvan de variatie, σ_ε kan worden geschat (Lancaster, 1979). In Tabel 2 worden de schattingen van (5) gepresenteerd.

Tabel 2. Schatting van de afdankintensiteit.
(standaardfouten tussen haakjes)

α_0 (constante)	2.94 (3.67)
α_1 (aanschafprijs)	-1.66 (.70)
α_2 (gezinsgrootte)	1.65 (.50)
β_2 (leeftijd)	2.56 (.85)
σ_ε^2 (storing)	.30 (.47)

Het effect van de aanschafprijs op de afdankintensiteit is negatief. Dit resultaat is volgens verwachting, omdat een hogere kwaliteit van het produkt tot minder afdanken zal leiden. De parameters kunnen worden beschouwd als elasticiteiten, zodat een procentuele verandering van de aanschafprijs in praktijk overeenkomt met een vermindering van de afdankintensiteit van 1.66%.
Het effect van de gezinsgrootte op de afdankintensiteit is positief. Ook dit resultaat is volgens verwachting, omdat de afdankkans zal toenemen naarmate een produkt vaker wordt gebruikt. De coëfficiënten voor de aanschafprijs en de gezinsgrootte zijn ongeveer even groot, maar hebben een tegengesteld effect. Om een toename van de afdank-

intensiteit door toename van de gezinsgrootte (bijv. van 20% bij uitbreiding van het gezin van vier naar vijf personen) te compenseren is dus een even grote procentuele toename van de aanschafprijs van een wasmachine nodig (bijv. van *f* 1000,– naar *f* 1200,–).
Het effect van de leeftijd van het produkt op de afdankintensiteit is volgens verwachting positief. Hoe ouder het produkt, hoe groter de afdankintensiteit.
De absolute waarde van de bovengenoemde parameters is groter dan één, d.w.z. dat de procentuele verandering van de afdankintensiteit meer dan evenredig is met de procentuele verandering van de variabele.

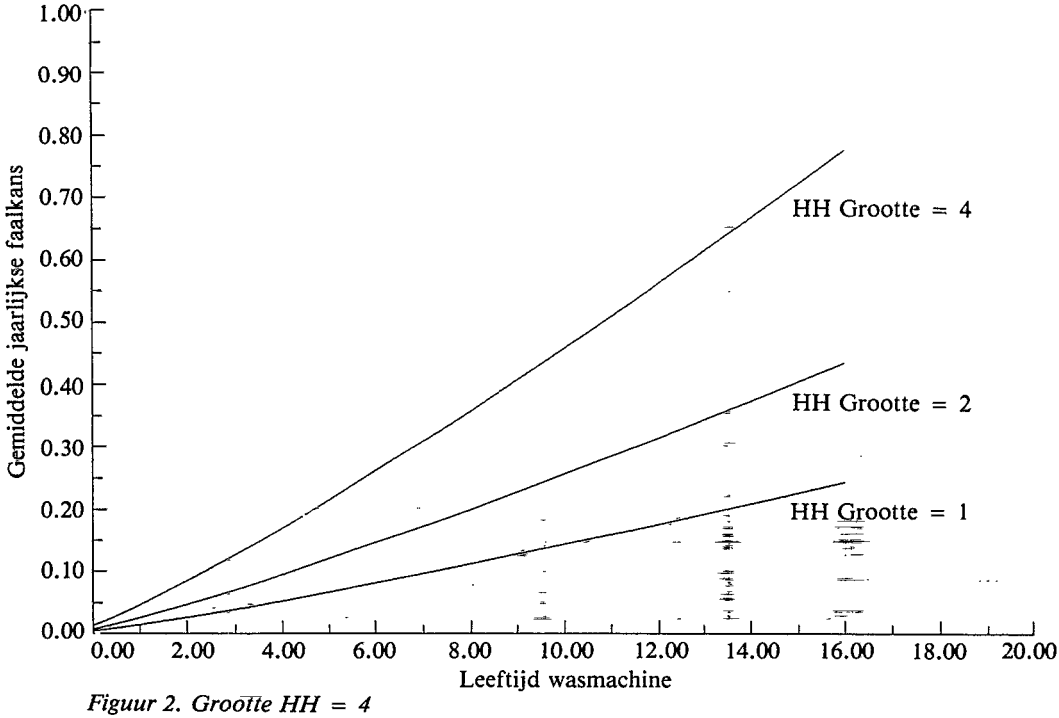
De storingsvariantie is klein en niet significant, zodat de niet-geobserveerde heterogeniteit van de waarnemingen als verwaarloosbaar kan worden beschouwd. Een pseudo-R^2 (Maddala, 1983), gebaseerd op een vergelijking met een model met een constante afdankintensiteit, bedraagt 8.5%.
Met behulp van de afdankintensiteit kan de verwachting van de levensduur van de wasmachines in het survey worden berekend (voor het model met de storingsterm voor heterogeniteit wordt deze verwachting gegeven in Lancaster, 1979). Deze verwachte levensduur bedraagt gemiddeld 13.7 jaar. Dit is hoger dan op grond van eerder onderzoek (Butlin, 1976) werd verwacht, temeer daar tweedehands wasmachines in de berekeningen buiten beschouwing zijn gelaten. In eerder onderzoek is echter geen rekening gehouden met vertekeningen in de gegevens en met een veranderende afdankkans tijdens de levensduur van het produkt.

Een schatting van de samengestelde afdankintensiteit (6) is gemaakt met behulp van de waargenomen defecten en beslissingen in het survey. In feite is deze informatie nog uitgebreid met het tijdstip van het meest recente defect vóór de start van het survey (zie Antonides, 1988, pp. 152-155). Een schatting van de parameters van de faalintensiteit en van de conditionele afdankkans is weergegeven in Tabel 3.
De parameters corresponderend met de aanschafprijs, de gezinsgrootte en de leeftijd van de wasmachine in Tabel 3 hebben dezelfde richting als die in Tabel 2. Voor de faalintensiteit geldt dus ook dat een hogere aanschafprijs de kans op defecten vermindert, en dat een groter huishouden en een hogere leeftijd van het produkt de kans op defecten vergroot.

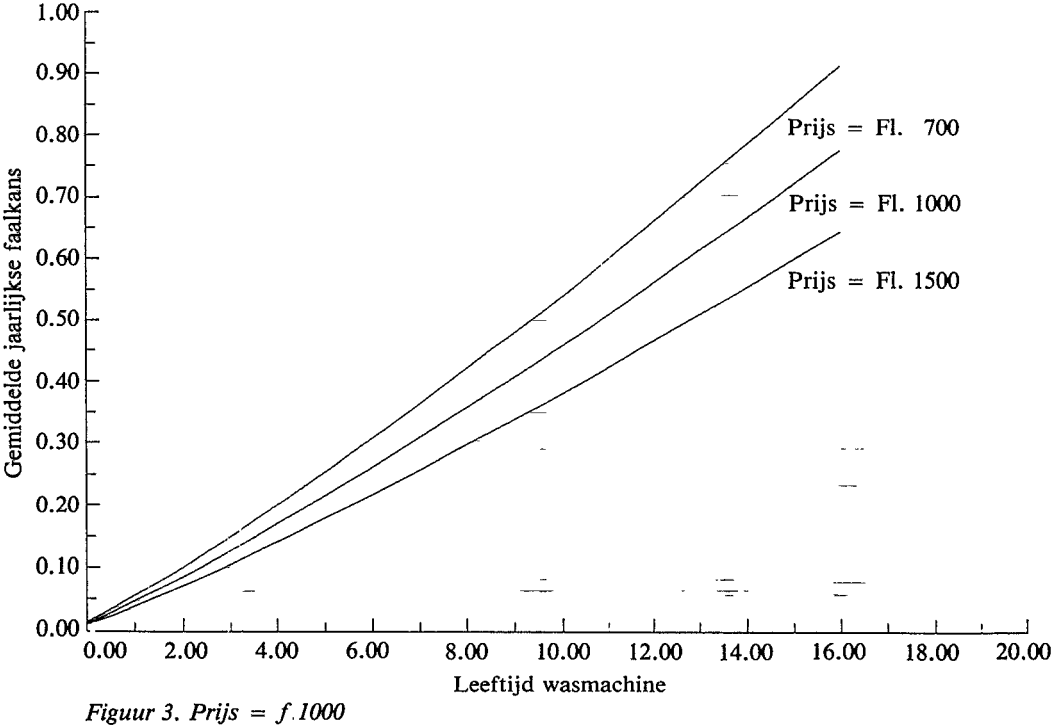
Tabel 3. Schatting van de samengestelde afdankintensiteit.
(standaardfouten tussen haakjes)

α_0'	(constante)	-1.53 (.57)
α_1'	(aanschafprijs)	-.46 (.09)
α_2'	(gezinsgrootte)	.84 (.11)
β'	leeftijd)	1.17 (.09)
δ	(afdanken/defect)	.31 (.03)

In Figuur 2 is het effect van de aanschafprijs op de gemiddelde jaarlijkse faalkans grafisch weergegeven voor een gezin van vier personen. In Figuur 3 is het effect van de gezinsgrootte op de gemiddelde jaarlijkse faalkans grafisch weergegeven voor een aanschafprijs van *f* 1000,–. In beide figuren neemt de faalkans iets meer dan evenredig toe met een stijging van de leeftijd. Ook is te zien dat het effect van de gezinsgrootte beduidend sterker is dan het prijseffect.



Figuur 2. Grootte HH = 4



Figuur 3. Prijs = f.1000

Ter vergelijking met ander onderzoek is ook een constante faalkans berekend. Deze bedraagd gemiddeld .43 per jaar. Huppes en Kanis (1985) rapporteren een faalkans van .39 per jaar voor wasmachines met een leeftijd van maximaal 10 jaar. *Which* (1986) noemt een faalkans van .37 gedurende de eerste vier jaar van de levensduur.

De kans op afdanken, gegeven een defect, bedraagt 31% na de leeftijd van vijf jaar. Dit heeft het effect dat een cohort wasmachines, na het bereiken van de leeftijd van vijf jaar, na gemiddeld twee defecten ongeveer is gehalveerd.

Een pseudo-R^2, gebaseerd op een vergelijking met een model met een constante afdankintensiteit, bedraagt in dit geval 39%. De verwachte levensduur van de wasmachines in het survey, uitgaande van de samengestelde afdankintensiteit, bedraagt gemiddeld 13.4 jaar. Deze verwachting stemt vrij goed overeen met die, gebaseerd op Tabel 2.

Het tweede deel van de resultaten betreft de afdankkans, gegeven een defect, m.b.v. het economisch-psychologische model. De determinanten van de afdankbeslissing zijn onderzocht ná melding van het defect, maar vóór een beslissing was genomen. In feite had ongeveer de helft van de consumenten reeds een beslissing genomen bij de melding van het defect, waardoor bij deze consumenten informatie ontbreekt. Verder is er bij een aantal enquêtes partiële nonrespons opgetreden, waardoor het aantal complete waarnemingen van determinanten gering is. M.b.v. imputatie-procedures (Sande, 1982) is het in de meeste gevallen echter gelukt een schatting van de respons te maken waar deze ontbrak. Tenslotte zijn de parameters van vgl. (8) geschat op basis van 94 observaties van de determinanten. De schattingen worden gepresenteerd in Tabel 4. In 23 gevallen is een afdankbeslissing genomen, in 71 gevallen een reparatiebeslissing.

Tabel 4. Schatting van het economisch-psychologisch afdankmodel.

ξ_0	(constante)	.11 (.32)
ξ_1	(nut van andere goederen, V)	.79 (.41)*
ξ_2	(nut van de wasmachine, U)	1.65 (.84)*
ρ	(discontovoet)	.21 (.16)+

* significant volgens een éénzijdige t-toets op 5% nivo
\+ significant volgens een éénzijdige t-toets op 10% nivo

De parameter ξ_1 betreft het nutsverschil van de andere goederen en diensten, V(1)-V(1′), tussen de alternatieven 1 (reparatie) en 1′ (afdanken). Het effect van deze determinant op een beslissing tot repareren is positief, zoals verwacht.
De parameter ξ_2 betreft het nutsverschil van de wasmachine, U(1)-U(1′), tussen de beide alternatieven. Het effect van deze determinant is eveneens positief, volgens verwachting.
Uit het effect van de beide determinanten blijkt dat de consument i.h.a. een rationele beslissing neemt in de zin dat gekozen wordt voor het alternatief met het hoogste nut. Het model waarin het nut van een enkel produkt, opgeteld bij het nut van alle andere goederen en diensten, de consumenten-beslissing verklaart (hypothese van additieve separabiliteit) wordt niet verworpen.
Enigszins verrassend is de hoge waarde van de discontovoet, 21%. Deze waarde ligt ruim boven de marktrente en stemt overeen met hoge schattingen van de discontovoet bij de aanschaf van air-conditioners in Hausman (1979)

Een uitbreiding van het model betreft de samenhang tussen de discontovoet en de financiële positie van het huishouden. In Hausman (1979) varieert de discontovoet van 89% bij huishoudens in de laagste inkomensklasse tot 5% in de hoogste klasse. Cunningham en Joseph (1978) vinden dat de maximaal aanvaardbare tijd om investeringen in huisisolatie en in zonne-energie installaties korter is voor lagere inkomensgroepen dan voor hogere inkomensgroepen. Dit resultaat duidt er eveneens op dat bij lagere inkomensgroepen toekomstige uitgaven in sterkere mate worden gedisconteerd dan bij hogere inkomensgroepen. Fisher (1930) stelde reeds dat het huidige inkomen, maar ook de verwachte toekomstige verdeling van het inkomen, invloed heeft op het ongeduld van de consument bij de besteding van het inkomen.

In het survey is een vraag gesteld naar de financiële situatie van het huishouden tijdens het defect en in de toekomst. Huishoudens waar de schulden stegen, waar spaargeld werd aangesproken en waar het inkomen juist voldoende was, zijn geklassificeerd als niet-spaarders. Huishoudens die een beetje of veel spaarden zijn geklassificeerd als spaarders. Verder zijn huishoudens geklassificeerd als optimistisch als de verwachte inkomenspositie in de volgende twaalf maanden veel beter, een weinig beter, of hetzelfde als heden was. Als de verwachte inkomenspositie slechter of veel slechter was zijn de huishoudens geklassificeerd als pessimistisch. Het economisch-psychologische afdankmodel is opnieuw geschat met een verschillende discontovoet voor spaarders en niet-spaarders, en voor optimistische en pessimistische huishoudens. De geschatte discontovoet voor de verschillende huishoudens is vermeld in Tabel 5.

Hoewel de geschatte discontovoeten niet significant van elkaar verschillen, is de hoogte van de coëfficiënten steeds volgens verwachting. Optimistische huishoudens disconteren toekomstige uitgaven in sterkere mate (vertonen meer ongeduld bij de besteding van het inkomen) dan pessimistische huishoudens. Wellicht geldt voor pessimistische huishoudens het spreekwoordelijke bewaren van het appeltje voor de dorst.

Tabel 5. Schatting van de discontovoet bij verschillende inkomensposities.

huidige positie	verwachte positie		
	optimistisch	pessimistisch	totaal
sparen	.32 (.28)	.10 (.18)	.17 (.14)
niet sparen	.42 (.35)	.28 (.25)	.31 (.20)
totaal	.36 (.31)	.19 (.20)	.20 (.16)

Huishoudens die niet sparen disconteren eveneens in sterkere mate dan huishoudens die wel sparen. Hier geldt dat huishoudens die niet sparen het gehele inkomen hard nodig hebben en toekomstige uitgaven minder zwaar laten wegen. De hoogste discontovoet is gevonden bij optimistische huishoudens die niet sparen. Deze huishoudens consumeren hun gehele inkomen in de verwachting dat het inkomen in de toekomst even groot of groter zal zijn dan in het heden. De laagste discontovoet benadert de marktrente het meest en is gevonden bij pessimistische huishoudens die sparen.

Een andere uitbreiding van het model betreft het effect van gedragskosten (Verhallen en Pieters, 1984) op de afdankbeslissing. Deze gedragskosten betreffen de moeite die de consument moet doen om er achter te komen wat er met de defecte wasmachine aan de hand is, de moeite die een consument moet doen om er achter te komen wat een repara-

tie zal kosten, en de moeite die een consument moet doen om een nieuwe wasmachine te kopen. Deze gedragskosten zijn toegevoegd aan de functie f(Z) in het economisch-psychologisch model (8). De schattingen van dit model worden gegeven in Tabel 6.

Tabel 6. Schatting van het economisch-psychologisch afdankmodel inclusief gedragskosten.

ξ_0	(constante)	-.453 (.873)
ξ_1	(nut van andere goederen, V)	.815 (.421)*
ξ_2	(nut van de wasmachine, U)	1.884 (.905)*
ρ	(discontovoet)	.216 (.156)+
ξ_3	(moeite om achter de defect-oorzaak te komen)	-.324 (.160)*
ξ_4	(moeite om achter de reparatiekosten te komen)	.129 (.177)
ξ_5	(moete om een nieuwe wasmachine te kopen)	.061 (.175)

* significant volgens een éénzijdige t-toets op 5% nivo
\+ significant volgens een éénzijdige t-toets op 10% nivo

De coëfficiënten in Tabel 6 verschillen niet veel van die in Tabel 4. Verder blijkt de moeite om er achter te komen wat er aan de defecte wasmachine mankeert, de kans op reparatie negatief te beïnvloeden. De moeite om achter de reparatiekosten te komen en de moeite om een nieuwe wasmachine te kopen blijken de beslissing niet significant te beïnvloeden.

De voorwaardelijke afdankkans, een constante in het eerste deel van deze paragraaf, is een resultante van een aantal economisch-psychologische determinanten waaronder het nut van het produkt en het nut van alle andere goederen en diensten, de discontovoet en de moeite om achter de oorzaak van het defect te komen. Het nut van de wasmachine hangt af van de evaluatie van de produkt-attributen en de uitgaven aan het produkt. De waardering van de kosten van het produkt hangt samen met de waardering van het inkomen en de verwachte (resterende) levensduur van het produkt.

6. CONCLUSIES

Met behulp van informatie uit een consumenten-survey is het mogelijk de faalkans van een duurzaam consumptiegoed en de (voorwaardelijke) afdankkans te schatten, zonder het produkt gedurende de gehele levensduur waar te nemen. De faalkans van de wasmachine blijkt gevoelig voor de kwaliteit en het gebruik van het produkt. Tevens wordt hierdoor de overlevingskans van het produkt beïnvloed.
Gegeven de trend van een verminderende gezinsgrootte in de Westerse wereld, zal de faalkans van de wasmachine verkleinen. Dit resulteert in verminderde reparatie-activiteiten per machine, waardoor een illusie van een vergrote betrouwbaarheid kan ontstaan, vooral wanneer deze uitsluitend van de vraag naar reparatiediensten wordt afgeleid.
Voor de vraag naar de service-diensten van wasmachine-dealers of -fabrikanten lijkt kennis over de gezinsgrootte of de wasfrequentie van de consument van groot belang. Van een gegeven cohort wasmachines kan immers de jaarlijkste faalkans binnen zekere marges worden voorspeld m.b.v. deze informatie. Tevens kan hieruit de overlevingsfunctie worden bepaald, waardoor de omvang van de vervangingsmarkt kan worden geschat.

Voor consumentenorganisaties is de vergelijking van de faalkans van verschillende merken en typen van een produkt van belang. Een aanduiding van de gemiddelde faalkans bij gegeven gezinsgrootte en type produkt zal zeker bijdragen tot de kwaliteit van het beslissingsproces bij de aanschaf en ook bij het afdanken van het produkt.
Voor de consument van tweedehands wasmachines is informatie over het gebruik van het produkt in het verleden eveneens van belang. M.b.v. de gezinsgrootte van de laatste eigenaar en de leeftijd en aanschafprijs van de machine kan binnen zekere marges de faalkans en de resterende levensduur worden geschat. Een veel nauwkeuriger schatting zou echter kunnen worden verkegen door een telling van het aantal wasbeurten van de machine, bijv. door het aanbrengen van een (verzegeld) telmechanisme door de fabrikant.

Enkele beperkingen van het onderzoek verdienen hier de aandacht. Een belangrijke vooronderstelling in het onderzoek is die van een exponentiële faalkans. Hoewel deze tamelijk vaak wordt toegepast in de betrouwbaarheidsanalyse van produkten wordt in de praktijk uitgegaan van de 'badkuipkromme' als beschrijving van de faalkans.
De 'badkuipkromme' geeft eerst een daling van de faalkans te zien als gevolg van 'kinderziekten' van het produkt, vervolgens een min of meer constante faalkans gedurende de 'bruikbare periode' en tenslotte een stijging in de laatste jaren van de levensduur. De waarnemingen in het survey betreffen echter machines van vijf jaar of ouder, zodat 'kinderziekten' tijdens het survey niet voorkomen. Na de leeftijd van vijf jaar zal de exponentiële faalintensiteit de 'badkuipkromme' kunnen benaderen. M.b.v. meer realistische specificaties kan de badkuipkromme worden onderzocht. Deze specificaties zullen echter waarschijnlijk niet leiden tot een vereenvoudiging van de schattingsprocedures. Het bovenstaande geldt eveneens voor de veronderstelling van de onderlinge onafhankelijkheid van defecten tijdens de levensduur van een produkt. Onderlinge afhankelijkheid van defecten zal het model aanzienlijk compliceren.
Een andere vooronderstelling is die van de constante voorwaardelijke afdankkans. In werkelijkheid varieert deze kans over huishoudens en over maatschappelijke omstandigheden in de loop der tijd. Vergeleken met bestaande modellen van betrouwbaarheid van produkten, waar in het geheel geen rekening is gehouden met een voorwaardelijke afdankkans, lijkt een constante kans echter een verbetering.

De determinanten van de voorwaardelijke afdankbeslissing van de consument zijn in het consumenten-survey onderzocht. De resultaten wijzen er op dat het nut van een nieuw produkt wordt afgewogen tegen het nut van het oude. Verbeteringen van produkten op de markt leiden via dit mechanisme tot een grotere afdankkans van het bestaande produkt. Een beperkende factor hierbij vormen de uitgaven aan het duurzame goed. Een belangrijke variabele hierbij is de discontovoet die de consument hanteert. De financiële positie van de consument in het heden en in de toekomst beïnvloedt de mate waarin toekomstige uitgaven worden gedisconteerd. Een manier om consumenten te helpen noodzakelijke vervangingsuitgaven te doen is m.b.v. sterk gespreide betalingsregelingen. Dit systeem zal het meeste succes hebben bij consumenten die niet kunnen of willen sparen en optimistisch zijn over hun financiële positie in de toekomst.
De wasmachine is geselecteerd als object van een voorbeeldstudie naar afdanken van een duurzaam consumptiegoed. De ontwikkelde modellen voor de faalkans en de voorwaardelijke afdankkans kunnen echter zonder grote aanpassingen worden toegepast op gebeurtenissen tijdens of ter beëindiging van de levensduur van andere consumptiegoe-

deren. Zonder verlies van algemeenheid kunnen andere verklarende variabelen voor de faalintensiteit worden opgenomen, bijvoorbeeld (bij de stofzuiger) de grootte van de woning of (bij de koelkast) de omgevingstemperatuur. Zoals blijkt uit Tabel 1 zijn er voor andere huishoudelijke toestellen, vaker dan bij de wasmachine, vaak andere redenen om een produkt af te danken. Ook voor deze andere redenen van afdanken (verhuizing, beschadiging, etc.) kan de waarschijnlijkheid van gebeurtenissen worden gemodelleerd. De voorwaardelijke afdankkans van andere produkten dan de wasmachine kan eveneens worden opgevat als een functie van het nut, ontleend aan het produkt in kwestie en aan andere goederen en diensten. De kosten (van reparatie, onderhoud en gebruik) en de baten (restwaarde bij verkoop, en nut ontleend aan diensten geleverd door het produkt) kunnen van produkt tot produkt sterk verschillen naar aard en grootte, de wijze waarop deze in het model zijn opgenomen kan echter algemeen worden toegepast.

De onderzoeksresultaten zijn relevant voor het beleid van consumenten-organisaties (voorlichting en onderzoek), het beleid van producenten en dealers van duurzame consumptiegoederen (duurzaamheid van het ontwerp, reparatiediensten en produktieplanning en verkoop), en het overheidsbeleid m.b.t. milieu en consument.

LITERATUUR

Antonides, G., 1988. *Scrapping a durable consumption good.* Dissertatie, Erasmus Universiteit Rotterdam.

Barlow, R.E. en Proschan, F., 1967. *Mathematical theory of reliability.* New York: Wiley.

Box, J.M.F., 1981. Verlenging van de levensduur van duurzame consumptiegoederen. *Economisch Statistische Berichten* 66, 896-902.

Butlin, J., 1976. *Product durability and product life extension: their contribution to solid waste management.* Parijs: OECD.

C.B.S., 1982. Enkele resultaten van het Consumenten Conjunctuur Onderzoek van October 1981. *Statistisch Bulletin,* 12-2-1982.

Cox, D.R. en Lewis, P.A.W., 1966. *The statistical analysis of series of events.* London: Methuen.

Cunningham, W.H. en Joseph, B., 1978. Energy conservation, price increases and payback periods. *Advances in Consumer Research* 5, 201-205.

Fischer, I., 1930. *The theory of interest.* New York: The Macmillan Company.

Hanson, J.W., 1980. A proposed paradigm for consumer product disposition processes. *Journal of Consumer Affairs* 14, 49-67.

Hausman, J.A., 1979. Individual discount rates and the purchase and utilization of energy-using durables. *The Bell Journal of Economics* 10, 33-54.

Huppes, G. en Kanis, H., 1985. *Bezitsduurkosten.* Den Haag: SWOKA.

Jacoby, J., Berning, C.K. en Dietvorst, T.F., 1977. What about disposition? *Journal of Marketing,* 11, 22-28.

Jorgenson, D.W., 1974. The economic theory of replacement and depreciation. In: Sellekaerts, W. (ed.), *Econometrics and economic theory.* London: MacMillan.

Kalbfleisch, J. en Prentice, R., 1980. *The statistical analysis of failure time data.* New York: Wiley.

Kapteyn, A., Herwaarden, F. van, en Praag, B.M.S. van, 1978. *Direct measurement of welfare functions: methods and results.* Report 77.02. Leiden: Economisch Instituut van de Rijksuniversiteit Leiden.

Kapteyn, A., Wansbeek, T. en Buyze, J., 1979. Maximizing or satisficing? *The Review of Economics and Statistics* 61, 549-563.

Lancaster, T., 1979. Econometric methods for the duration of unemployment. *Econometrica* 47, 939-956.
Laurentius, G., 1987. (red.) *Onderhoudsbewust ontwerpen nu en in de toekomst.* Alphen aan den Rijn: Samson.
Maddala, G.S., 1983. *Limited-dependent and qualitative variables in econometrics.* Cambridge, Cambridge University Press.
Praag, B.M.S. van, 1968. *Individual welfare functions and consumer behaviour.* Amsterdam: North-Holland Publ. Comp.
Raaij, W.F. van, en Wandwossen, K., 1978. Motivation-need theories and consumer behaviour. *Advances in Consumer Research* 5, 590-595.
Sande, I.G., 1982. Imputation in surveys: coping with reality. *The American Statistician* 36 (3), 145-152.
Schelbergen, F. en Serail, S., 1987. *Het besluitvormingsproces bij het afdanken van wasmachines. Deel II.* Den Haag: SWOKA.
Stone, R. en Rowe, D.A., 1957. The market demand for durable goods. *Econometrica* 25, 423-443.
Stone, R. en Rowe, D.A., 1958. Dynamic demand functions: some econometric results. *Economic Journal* 68, 250-270.
Stone, R. en Rowe, D.A., 1960. The durability of consumer's durable goods. *Econometrica* 28, 407-416.
Su, T.T., 1975. Durability of consumption goods reconsidered. *American Economic Review* 65, 148-157.
Verhallen, Th.M.M. en Pieters, R.G.M., 1984. Attitude theory and behavioral costs. *Journal of Economic Psychology* 5, 223-249.
Westin, R.B., 1975. Empirical implications of infrequent purchase behaviour in a stock-adjustment model. *American Economic Review* 65, 384-396.
Which?, 1986. Reliability and servicing of domestic appliances, 67-70.

10. Het specificeren van modellen voor markten

Enkele specificaties en schattingsmethoden*

F.W. PLAT

SAMENVATTING

Dit artikel behandelt de specificatie van marketing modellen waarbij rekening wordt gehouden met de aanwezigheid van fouten in de gebruikte gegevens. Het lineaire fouten-in-variabelen model wordt besproken, alsmede de jackknife en de bootstrap, twee verdelingsvrije methoden waarmee op relatief eenvoudige wijze de standaarddeviaties van schatters kunnen worden bepaald. Genoemde methoden geven bovendien inzicht in de nauwkeurigheid van parameterschattingen.
Het lineaire fouten-in-variabelen model, de jackknife en de bootstrap zijn voor diverse merken toegepast. De resultaten geven inzicht in de kwaliteit van de parameterschattingen alsook in de kwaliteit van de gebruikte gegevens.

1. INLEIDING

Dit artikel is een weergave van een gedeelte uit een uitgebreidere studie betreffende de (numerieke) specificatie van marketing modellen[1]. In de genoemde studie zijn modellen ontwikkeld waarmee expliciet rekening gehouden wordt met factoren als:
- de aanwezigheid van fouten in de gebruikte data,
- de effecten van acties en reacties van concurrenten, en
- mogelijke structurele veranderingen in het consumentengedrag tijdens een periode van stagflatie.

Dit artikel behandelt de specificatie van modellen waarmee expliciet rekening gehouden wordt met de fouten in de gebruikte gegevens (de eerstgenoemde factor). Panelgegevens, zoals consumentenpanel- en detaillistenpanel gegevens, vormen een belangrijke verzameling marktonderzoekgegevens. In de literatuur[2] is reeds uitvoerig aangegeven dat panelgegevens aanzienlijk kunnen afwijken van de werkelijke waarden die variabelen aannemen. Deze verschillen worden ook wel 'data bias' genoemd. De consequenties van de data bias in panelgegevens en methoden om de data bias te reduceren zijn in de literatuur beschreven . Eén van de consequenties van de data bias is dat het gebruik van panelgegevens ertoe kan leiden dat de schatters van de parameters van een marketing model onzuiver en inconsistent zijn[3]. In plaats van de data bias in panelgegevens te reduceren wordt in dit artikel

getracht om modellen te ontwikkelen die rekening houden met de aanwezigheid van fouten in de gebruikte gegevens.
De ontwikkelde modellen zijn zogenaamde *lineaire fouten-in-variabelen modellen*. Deze modellen worden verkregen door in de 'traditionele' marketing modellen de waarden van de foutief gemeten verklarende variabelen uit te drukken in termen van de werkelijke waarden van de verklarende variabelen plus meetfouten. Op deze wijze resulteren kwalitatief betere schatters: de schatters zijn consistent (zij kunnen echter nog wel onzuiver zijn).
De parameters van een fouten-in-variabelen model kunnen echter niet zonder meer worden geschat. De parameters van een dergelijk model zijn namelijk niet allemaal geïdentificeerd[4]. Dit betekent dat additionele informatie benodigd is. In dit artikel is deze informatie voorhanden in de vorm van zogenaamde *instrumentele variabelen* waardoor instrumentele variabelen (IV-) schatters kunnen worden verkregen.
Naast het schatten van de parameters van een fouten-in-variabelen model vereist ook het schatten van de standaarddeviaties van deze schatters bijzondere aandacht.
Na deze inleiding zal in paragraaf 2 het lineaire fouten-in-variabelen model worden beschreven, en wel het enkelvoudige model (met ééń verklarende variabele, paragraaf 2.1) en het meervoudige model (met meer dan één verklarende variabele, paragraaf 2.2). Paragraaf 3 bevat een korte beschrijving van enkele methoden voor het schatten van de standaarddeviaties van de IV-schatters. In het bijzonder worden in paragraaf 3 de *jackknife* en de *bootstrap* beschreven[5].
In paragraaf 4 worden enkele lineaire fouten-in-variabelen modellen numeriek gespecificeerd met behulp van (panel) gegevens betreffende de marktaandelen, prijzen, verkrijgbaarheid en reclame-uitgaven van een aantal merken. De jackknife en de bootstrap zijn daarbij gebruikt om schattingen van de standaarddeviaties te bepalen. Zoals nog uiteen wordt gezet (paragraaf 3), geven beide methoden tevens informatie omtrent de mate van onzuiverheid van de schatters van de parameters van een model.

2. EEN LINEAIR FOUTEN-IN-VARIABELEN MODEL

Deze paragraaf behandelt de specificatie van het lineaire fouten-in-variabelen model, alsmede het schatten van de parameters van het model. Paragraaf 2.1 bevat de beschrijving van het enkelvoudige model. In paragraaf 2.2 komt het meervoudige model aan de orde.
Uit een oogpunt van lees- (en schrijf-)baarheid veronderstellen we dat alle variabelen gemeten zijn in afwijking van hun gemiddelden, en werken we slechts met tweede orde momenten.

2.1 Een enkelvoudig fouten-in-variabelen model

Beschouw de volgende vergelijking:

$$y = \gamma\xi + \zeta \tag{1}$$

waarbij:

y = de waargenomen waarde van de afhankelijke of te verklaren variabele (bijv. marktaandeel, afzet),

ξ = de werkelijke waarde van de verklarende variabele (bijv. prijs, distributie),
ζ = de storingsterm in de relatie,
γ = de response parameter.

Relatie (1) weerspiegelt een veelgemaakte veronderstelling, namelijk dat de waargenomen waarden van de verklarende variabelen overeenkomen met de werkelijke waarden van deze variabelen. Op basis van deze veronderstelling worden de waargenomen waarden voor ξ gesubstitueerd en wordt de relatie vervolgens geschat met behulp van de methode der kleinste kwadraten of één van haar varianten.
Indien de genoemde veronderstelling juist is, zijn de kleinste kwadraten schatters, onder de gebruikelijke assumpties, zuiver en consistent. Een cruciale assumptie is in dit verband het ongecorreleerd zijn van ξ met de gesubsitueerde waarden van de verklarende variabele.
Veronderstel echter dat de waargenomen waarden van de verklarende variabele (x) afwijken van de werkelijke waarden van deze variabele (ξ):

$$x = \xi + \delta_1, \tag{2}$$

waarbij:
δ_1 = de meetfout in x.
Substitutie van x voor ξ in (1) leidt tot:

$$y = \gamma(x-\delta_1) + \zeta = \gamma x + (\zeta-\gamma\delta_1) \tag{3}$$

In (3) is de 'storingsterm' $\zeta-\gamma\delta_1$ niet ongecorreleerd met x, zodat aan de genoemde cruciale assumptie niet is voldaan: de kleinste kwadraten schatters zijn onzuiver en inconsistent.

Relaties (1) en (2) vormen samen het enkelvoudige lineaire fouten-in-variabelen model. Het model bevat twee *waarneembare* variabelen x en y. De variabele ξ wordt wel *latente* of *niet-waarneembare* variabele genoemd (evenals ζ en δ_1). In het navolgende zullen we laten zien op welke wijze consistente schatters van de parameters van dit model kunnen worden verkregen.
Verondersteld wordt dat ξ, δ_1 en ζ elk een gemiddelde gelijk aan nul hebben[6] en onderling ongecorreleerd zijn. Eventuele meetfouten in y zijn bevat in ζ.
Het model bevat vier onbekende parameters. Dat zijn:

γ, σ^2_ξ, $\sigma^2_{\delta_1}$ en σ^2_ζ,

waarbij de laatste drie parameters respectievelijk de variantie van de werkelijke verklarende variabele, de meetfout-variantie en de variantie van de storingsterm representeren.
Relaties (1) en (2), en de gemaakte veronderstellingen, impliceren dat de populatie covariantie-matrix van $(x,y)'$ als volgt als een functie van de onbekende parameters geschreven kan worden[7]:

$$\begin{bmatrix} \sigma^2_x & \sigma_{xy} \\ & \sigma^2_y \end{bmatrix} = \begin{bmatrix} \sigma^2_\xi + \sigma^2_{\delta_1} & \gamma\,\sigma^2_\xi \\ & \gamma^2\,\sigma^2_\xi + \sigma^2_\zeta \end{bmatrix} \tag{4}$$

Het zal duidelijk zijn dat uit de drie vergelijkingen in (4), in het algemeen, de vier onbekende parameters niet eenduidig bepaald kunnen worden. Anders gesteld: de parameters kunnen niet worden *geïdentificeerd* met behulp van de tweede orde momenten (welke laatste consistent zijn te schatten, gegeven een aselecte steekproef).

Desalniettemin bevat (4) toch enige informatie. Uit (4) is namelijk af te leiden dat voor γ geldt:

$$\begin{aligned} &(\sigma_{xy}/\sigma_x^2) < \gamma < (\sigma_y^2/\sigma_{xy}) \text{ als } \sigma_{xy} > 0, \\ &(\sigma_y^2/\sigma_{xy}) < \gamma < (\sigma_{xy}/\sigma_x^2) \text{ als } \sigma_{xy} < 0. \end{aligned} \tag{5}$$

Met andere woorden, γ is begrensd door de hellingen van de twee regressielijnen welke resulteren uit de regressies van y op x en x op y. Echter, γ kan alleen worden bepaald uit de tweede orde momenten als x en y volledig correleren en de grenzen in (5) dus samenvallen.
Voor het oplossen van dit identificatie probleem is additionele informatie vereist[8]. Veronderstel dat nog een waarneembare variabele (z) beschikbaar is, die net als x lineair gerelateerd is aan ξ:

$$z = \lambda\xi + \delta_2, \tag{6}$$

waarbij het residu δ_2 gemiddelde nul heeft en ongecorreleerd is met ξ, δ_1 en ζ. De variabele z wordt wel een *instrumentele variabele* genoemd.
De populatie covariantie-matrix van $(x,y,z)'$ is[7]:

$$\begin{bmatrix} \sigma_x^2 & \sigma_{xy} & \sigma_{xz} \\ & \sigma_y^2 & \sigma_{yz} \\ & & \sigma_z^2 \end{bmatrix} = \begin{bmatrix} \sigma_\xi^2 + \sigma_{\delta_1}^2 & \gamma\sigma_\xi^2 & \lambda\sigma_\xi^2 \\ & \gamma^2\sigma_\xi^2 + \sigma_\zeta^2 & \gamma\lambda\sigma_\xi^2 \\ & & \lambda^2\sigma_\xi^2 + \sigma_{\delta_2}^2 \end{bmatrix} \tag{7}$$

Vergeleken met (4) is in (7) het aantal vergelijkingen toegenomen met drie, ten koste van twee extra parameters (λ en $\sigma_{\delta_2}^2$). De zes parameters in (7) kunnen als volgt in de populatie varianties en covarianties worden uitgedrukt:

$$\gamma = \frac{\sigma_{yz}}{\sigma_{xz}}, \tag{8a}$$

$$\lambda = \frac{\sigma_{yz}}{\sigma_{xy}}, \tag{8b}$$

$$\sigma_\xi^2 = \frac{\sigma_{xy}\,\sigma_{xz}}{\sigma_{yz}}, \tag{8c}$$

$$\sigma_{\delta_1}^2 = \sigma_x^2 - \frac{\sigma_{xy}\,\sigma_{xz}}{\sigma_{yz}}, \tag{8d}$$

$$\sigma_\zeta^2 = \sigma_y^2 - \frac{\sigma_{xy}\,\sigma_{yz}}{\sigma_{xz}}, \tag{8e}$$

$$\sigma_{\delta_2}^2 = \sigma_z^2 - \frac{\sigma_{xz}\,\sigma_{yz}}{\sigma_{xy}}, \tag{8f}$$

Uiteraard moeten we hierbij veronderstellen dat $\sigma_{xy} \neq 0$, $\sigma_{xz} \neq 0$ en $\sigma_{yz} \neq 0$, ofwel dat $\gamma \neq 0$, $\lambda \neq 0$ en $\sigma_\xi^2 \neq 0$. De uitdrukkingen (8a) - (8f) vormen de unieke oplossingen van (7), er zijn geen graden van vrijheid.
Uitgaande van een steekproef van omvang n, met waarnemingen $(x_1, y_1, z_1)'$, ..., $(x_n, y_n, z_n)'$, kunnen consistente schatters van de zes parameters (7) worden verkregen. De steekproef momenten $s_x^2, s_{xy}, \ldots, s_z^2$ zijn consistente schatters van de populatie momenten $\sigma_x^2, \sigma_{xy}, \ldots, \sigma_z^2$. Door nu in (8) steekproef momenten voor populatie momenten te substitueren, resulteren consistente schatters van $\gamma, \lambda, \ldots, \sigma_{\delta_2}^2$. Tevens kunnen eenvoudig consistente schatters van functies van de parameters worden verkregen. Zo wordt bijvoorbeeld de 'ruis-signaal verhouding' voor x, gedefinieerd als $\sigma_{\delta_1}^2/\sigma_x^2$, consistent geschat door $\hat{\sigma}_{\delta_1}^2/s_x^2$. De *ruis-signaal verhouding* meet het gedeelte van de variantie van x dat toe te schrijven is aan meetfouten. Daar de schatters continue differentieerbare functies zijn van steekproef momenten, zijn de schatters asymptotisch normaal verdeeld.
De hier beschreven methode voor het bepalen van schatters staat bekend als de momenten methode. De schatters $\hat{\gamma} = s_{yz}/s_{xz}$ is de gebruikelijke instrumentele variabele (IV-) schatter van γ[9].
Schattingen van de standaarddeviaties van de schatters kunnen op verschillende manieren worden verkregen. De asymptotische variantie van $\hat{\gamma} = s_{yz}/s_{xz}$ is relatief gemakkelijk af te leiden, en gelijk aan[10]:

$$\frac{1}{n}\sigma_\zeta^2 \frac{\sigma^2 z}{\sigma^2 xz} \tag{9}$$

Door nu in (9) een consistente schatter voor σ_ζ^2 te substitueren en de populatie momenten door steekproef momenten te vervangen, resulteert, na het trekken van de wortel, een consistente schatting van de asymptotische standaarddeviatie van $\hat{\gamma}$. Voor de andere schatters kunnen niet zo gemakkelijk uitdrukkingen voor de asymptotische varianties worden afgeleid. Voor deze schatters moeten de schattingen van de standaarddeviaties op andere wijzen worden bepaald. We komen hierop terug in paragraaf 3.

2.2. Een meervoudig fouten-in-variabelen model

Veronderstel nu dat het model k ($k > 1$) verklarende variabelen bevat. Voor het gemak gebruiken we dezelfde symbolen als in paragraaf 2.1, met dien verstande dat x, z, γ, ξ, δ_1 en δ_2 nu vectoren zijn van orde k; y en ζ blijven scalair. Relaties (1), (2) en (6) worden nu:

$$y = \gamma'\xi + \zeta, \tag{10}$$

$$x = \xi + \delta_1, \tag{11}$$

$$z = \Lambda\xi + \delta_2, \tag{12}$$

waarbij Λ een ($k \times k$)-matrix is van regressie coëfficienten van z op ξ. Evenals in paragraaf 2.1 veronderstellen we dat de gemiddelden van de variabelen gelijk zijn aan nul en dat:

$$E(\xi\delta_1') = E(\xi\delta_2') = E(\delta_1\delta_2') = 0, E(\xi\zeta) = E(\delta_1\zeta) = E(\delta_2\zeta) = 0 \tag{13}$$

Zij Σ de populatie covariantie-matrix van $(x', y, z')'$. Σ kan met een voor zich sprekende notatie worden geschreven als[7]:

$$\begin{bmatrix} \Sigma_{xx} & \Sigma_{xy} & \Sigma_{xz} \\ & \sigma_y^2 & \Sigma_{yz} \\ & & \Sigma_{zz} \end{bmatrix} = \begin{bmatrix} \Sigma_{\xi\xi} + \Sigma_{\delta_1\delta_1} & \Sigma_{\xi\xi}\gamma & \Sigma_{\xi\xi}\Lambda' \\ & \gamma'\Sigma_{\xi\xi}\gamma + \sigma_\zeta^2 & \gamma'\Sigma_{\xi\xi}\Lambda' \\ & & \Lambda\Sigma_{\xi\xi}\Lambda' + \Sigma_{\delta_2\delta_2} \end{bmatrix} \tag{14}$$

We veronderstellen dat Σ_{xz} regulier is, ofwel dat $\Sigma_{\xi\xi}$ en Λ regulier zijn, zodat:

$$\gamma = \Sigma_{zx}^{-1}\Sigma_{zy}. \tag{15}$$

Voor de identificatie van de andere parameters in (14) zijn additionele veronderstellingen nodig[11]. We gaan ervan uit dat de covariantie-matrix van de meetfouten in x een diagonaal matrix is:

$$\Sigma_{\delta_1\delta_1} = \text{diag}\,(\Sigma_{\delta_1\delta_1}). \tag{16}$$

Dat wil zeggen dat we veronderstellen dat de meetfouten in de waargenomen waarden van de verklarende variabelen ongecorreleerd zijn. Een veronderstelling die niet onredelijk is, gezien de data welke in paragraaf 4 worden geanalyseerd.
Ten slotte veronderstellen we dat de variabelen in ξ allen relevant zijn, dat wil zeggen: geen enkel element van γ is gelijk aan nul.
Als gevolg van (16) kunnen de niet-diagonaal elementen van $\Sigma_{\xi\xi}$ worden bepaald volgens:

$$\Sigma_{\xi\xi} - \text{diag}\,(\Sigma_{\xi\xi}) = \Sigma_{xx} - \text{diag}\,(\Sigma_{xx}). \tag{17}$$

De diagonaal elementen van $\Sigma_{\xi\xi}$ worden bepaald als volgt:

$$\Sigma_{\xi\xi}\gamma = [\text{diag}\,(\Sigma_{\xi\xi}) + \Sigma_{xx} - \text{diag}\,(\Sigma_{xx})]\gamma = \Sigma_{xy}, \text{ zodat}$$

$$[\text{diag}\,(\Sigma_{\xi\xi})]\gamma = \Sigma_{xy} - [\Sigma_{xx} - \text{diag}\,(\Sigma_{xx})]\gamma. \tag{18}$$

Het i-de diagonaal element van $\Sigma_{\xi\xi}$ is nu gelijk aan het i-de element van het rechterlid van (18), gedeeld door het corresponderende element van γ.
Nadat $\Sigma_{\xi\xi}$ aldus is bepaald, volgt de matrix Λ uit:

$$\Lambda = \Sigma_{zx}\,\Sigma_{\xi\xi}^{-1}. \tag{19}$$

De resterende parameters zijn:

$$\Sigma_{\delta_1\delta_1} = \Sigma_{xx} - \Sigma_{\xi\xi}, \tag{20}$$

$$\sigma_\zeta^2 = \sigma_y^2 - \gamma'\Sigma_{\xi\xi}\gamma, \tag{21}$$

$$\Sigma_{\delta_2\delta_2} = \Sigma_{zz} - \Lambda\Sigma_{\xi\xi}\Lambda' . \tag{22}$$

Door nu in (15), (17), (18)-(22) steekproef momenten te substitueren voor populatie momenten, resulteren consistente schatters van $\gamma, \Lambda, \ldots, \Sigma_{\delta_2\delta_2}$.
De asymptotische covariantie-matrix van $\hat{\gamma}$ is [10]:

$$\frac{1}{n}\sigma_{\zeta}^2 \Sigma_{zx}^{-1} \Sigma_{zz} \Sigma_{zx}^{-1}, \tag{23}$$

en consistente schattingen van de asymptotische standaarddeviaties kunnen eenvoudig worden bepaald. Voor het schatten van de standaarddeviaties van de overige schatters zij verwezen naar paragraaf 3.

3. HET SCHATTEN VAN STANDAARDDEVIATIES VAN SCHATTERS VAN PARAMETERS

Zoals in paragraaf 2 is aangegeven, zijn uitdrukkingen voor de asymptotische varianties van schatters slechts beschikbaar voor de schatters van de response parameters. Voor de overige schatters zijn moeilijk soortgelijke uitdrukkingen af te leiden. Schattingen van de standaarddeviaties zullen dan ook langs andere wegen moeten worden bepaald. In hetgeen volgt worden vier methoden voor het schatten van standaarddeviaties besproken: de delta-methode, het LISREL computerprogramma, de jackknife en de bootstrap.
De klassieke *delta-methode*[12] levert consistente schattingen van de asymptotische standaarddeviaties van de schatters, alsook van de standaarddeviaties van functies van de schatters (zoals b.v. de in paragraaf 2.1 genoemde ruis-signaal verhoudingen). Deze methode is gebaseerd op het ontwikkelen van een schatter (als functie van de steekproef momenten) in een Taylorreeks rond het punt dat gevormd wordt door de populatie momenten. Daarbij worden termen, welke corresponderen met afgeleiden van de orde twee en hoger, verwaarloosbaar geacht.

Relaties (1), (2), (6) en (10)-(12) vormen twee zogenaamde *lineaire structurele vergelijkingen* modellen (LISREL modellen)[13]. Schattingen van de standaarddeviaties kunnen dan ook worden verkregen met behulp van het LISREL computerprogramma. De opties ML (Maximum Likelihood) en GLS (Generalized Least Squares) berekenen consistente schattingen van de asymptotische standaarddeviaties, mits x, y en z normaal verdeeld zijn.
Een derde manier om standaarddeviaties te berekenen bestaat uit het gebruiken van de *jackknife*[14], een verdelingsvrije methode welke vrij gemakkelijk is toe te passen. Teneinde aan te geven hoe de jackknife werkt: beschouw de schatter $\hat{\gamma} = s_{yz}/s_{xz}$. Een schatting van de variantie van $\hat{\gamma}$ wordt in drie stappen verkregen. Eerst worden n zogenaamde *pseudo-waarden* $\tilde{\gamma}_1, \ldots, \tilde{\gamma}_n$ berekend volgens:

$$\tilde{\gamma}_t = n\hat{\gamma} - (n-1)\,\hat{\gamma}\,(t),\ t = 1, \ldots, n, \tag{24}$$

waarbij $\hat{\gamma}$ (t) op dezelfde wijze als $\hat{\gamma}$ uit de steekproef is berekend, echter met weglating van de t-de waarneming[15]. Vervolgens wordt de jackknife-schatting van γ (aangegeven als J ($\hat{\gamma}$)) bepaald en wel als de gemiddelde pseudo-waarde. De variantie van $\hat{\gamma}$ wordt ten slotte geschat door:

$$\frac{1}{n}\left[\frac{1}{n-1}\sum_{t=1}^{n}(\tilde{\gamma}_t - J(\hat{\gamma}))^2\right]. \tag{25}$$

De jackknife kan worden toegepast op elke willekeurige schatter. De jackknife schattingen hebben dezelfde asymptotische verdelingen als de oorspronkelijke schattingen, maar zijn veelal minder onzuiver in kleine steekproeven (Miller 1974)).
In de praktijk worden de pseudo-waarden veelal beschouwd als onafhankelijke en identiek verdeelde waarden. Bij gebruikmaking van het Centrale Limiet Theorema betekent dit, in ons voorbeeld, dat de statistische grootheid

$$\frac{J(\hat{\gamma}) - \gamma}{\left(\frac{1}{n}\left[\frac{1}{n-1}\sum_{t=1}^{n}(\tilde{\gamma}_t - J(\hat{\gamma}))^2\right]\right)^{1/2}} \tag{26}$$

bij benadering standaard normaal verdeeld is. De grootheid (26) kan worden gebruikt voor het toetsen van hypothesen omtrent de waarde van γ
Een vierde methode voor het berekenen van standaarddeviaties is de *bootstrap.* Vergeleken met de jackknife is de bootstrap arbeidsintensiever. De methode levert dan ook meer informatie op.
Evenals de jackknife is ook de bootstrap een verdelingsvrije methode. De bootstrap werkt als volgt: trek, met teruglegging, een groot aantal (B) onafhankelijke aselecte steekproeven van omvang n uit de n beschikbare waarnemingen. Deze steekproeven worden wel *replicaties* genoemd. Bereken vervolgens voor de B aselecte steekproeven van omvang n de waarde van de schatter waarin men geïnteresseerd is, bijvoorbeeld $\hat{\gamma} = s_{yz}/s_{xz}$. Het resultaat voor de i-de steekproef noteren we als γ_i^*. De frequentie verdeling van de B γ_i^*'s is een schatting van de steekproef verdeling van de schatter γ. Het gemiddelde $\frac{1}{B}\sum_{i=1}^{B}\gamma_i^*$ is derhalve een schatting van het gemiddelde van de steekproef verdeling van de schatter $\hat{\gamma}$. Daar $\hat{\gamma}$ een schatting is van γ, en we nu tevens beschikken over een schatting van $E(\hat{\gamma})$, kunnen we de *onzuiverheid* of *bias* in γ schatten met[16]

$$\hat{\gamma} - \frac{1}{B}\sum_{i=1}^{B}\gamma_i^*, \tag{27}$$

waarbij

$$\frac{1}{B-1}\sum_{i=1}^{B}\left(\gamma_i^* - \frac{1}{B}\sum_{i=1}^{B}\gamma_i^*\right)^2 \tag{28}$$

een schatting is van de variantie van de steekproef verdeling van $\hat{\gamma}$.
In feite wordt een Monte Carlo experiment uitgevoerd, waarbij de frequentie verdeling van $(x_1, y_1, z_1)', \ldots, (x_n, y_n, z_n)'$ beschouwd wordt als de werkelijke verdeling van $(x, y, z)'$. Het zal duidelijk zijn dat de schatting $\frac{1}{B}\sum_{i=1}^{B}\gamma_i^*$ niet op dezelfde wijze geïnterpreteerd kan worden als de jackknife schatting $J(\hat{\gamma})$ en de andere hierboven besproken schatters. Laatstgenoemde schattingen hebben betrekking op γ terwijl de bootstrap een schatting oplevert van het gemiddelde van $\hat{\gamma}$ (wat ongelijk aan γ kan zijn).
Als de frequentie verdeling van de B γ_i^*'s voldoende overeenkomt met een normale verdeling, kunnen gemakkelijk toetsgrootheden worden gedefinieerd voor het

uitvoeren van significantie-toetsen. Daar $\frac{1}{B}\sum_{i=1}^{B}\gamma_i^*$ een schatting is van $E(\hat{\gamma})$ kan men grootheden vergelijkbaar met (26) gebruiken voor het toetsen van hypothesen betreffende $E(\hat{\gamma})$. In het algemeen, echter, worden dergelijke toetsen niet uitgevoerd daar men niet zozeer geïnteresseerd is in de waarde van $E(\hat{\gamma})$ als wel in de waarde van de bias in $\hat{\gamma}$: $\gamma - E(\hat{\gamma})$. De hypothese dat de bias in $\hat{\gamma}$ nul is, toetst men met de grootheid

$$\frac{\hat{\gamma} - \frac{\sum_{i=1}^{B}\gamma_i^*}{B}}{\left(\frac{1}{B}\left(\frac{1}{B-1}\sum_{i=1}^{B}(\gamma_i^* - \frac{1}{B}\sum_{i=1}^{B}\gamma_i^*)^2\right)\right)^{1/2}} \tag{29}$$

welke, onder de nul-hypothese van geen bias, bij benadering standaard-normaal verdeeld is.
In de volgende paragraaf worden enkele lineaire fouten-in-variabelen modellen numeriek gespecificeerd. De jackknife en de bootstrap zijn daarbij gebruikt om schattingen van de standaarddeviaties te bepalen. We hebben de voorkeur gegeven aan het toepassen van deze methoden omdat zij in vergelijking met LISREL en de delta-methode tevens informatie opleveren omtrent de bias in de parameterschattingen.

4. HET GEBRUIK VAN PANELGEGEVENS IN MARKETING MODELLEN: EEN TAXATIE VAN DE KWALITEIT VAN SCHATTINGEN EN DATA

Deze paragraaf illustreert het schatten van een lineaire fouten-in-variabelen marktaandeelmodel. Voor het schatten zijn gegevens beschikbaar afkomstig van onder meer een consumenten- en een detaillisten panel. De resultaten illustreren hoe men door middel van het hanteren van instrumentele variabelen gegevens van *verschillende* bronnen kan combineren tot betere (in de zin van consistente) schattingen. De schattingen van de in paragraaf 2 gedefinieerde ruis-signaal verhoudingen geven inzicht in de betrouwbaarheid van de data, terwijl met behulp van de jackknife en de bootstrap de bias in de parameterschattingen wordt bepaald.
In dit artikel laten we de resultaten zien zoals die zijn verkregen voor twee merken A en B van een niet-duurzaam en frequent verkocht consumptiegoed. Waar nodig wordt verwezen naar de meer uitgebreide studie van Plat (1988) waarin de resultaten voor in totaal negen merken worden beschreven.

Uitgangspunt van de analyse is het volgende lineair additieve marktaandeel[17]:

$$mc_{jt} = \gamma_{1j}\, rpc_{jt} + \gamma_{2j} d_{jt} + \gamma_{3j}\, rar_{jt} + \gamma_{4j}\, rat_{jt} + \zeta_{jt}, \tag{30}$$

waarbij:
mc_{jt} = marktaandeel van merk j in periode t, zoals gemeten door een consumentenpanel,

rpc_{jt} = relatieve prijs van merk j in periode t, zoals gemeten door een consumentenpanel, d.w.z. de prijs van merk j gedeeld door de gewogen gemiddelde prijs (weging met hoeveelheden) van alle merken op de markt (e.v.t. segment),

d_{jt} = gewogen distrubutie (fractie) van merk j in periode t, zoals gemeten door een detaillistenpanel,

rar_{jt} (rat_{jt}) = reclame aandeel voor merk j in periode t, berekend voor uitgaven aan de radio (televisie), zoals berekend door het Bureau voor Budgetten Controle (B.B.C.), d.w.z. de reclame-uitgaven voor merk j gedeeld door de uitgaven voor alle merken op de markt (e.v.t. segment)

De beschikbare gegevens bestaan uit tijdreeksen van 30 twee-maandelijkse waarnemingen over de periode maart/april 1978 tot maart/april 1983.
Gezien de opmerkingen in paragraaf 1 betreffende de data bias in panelgegevens, zal het duidelijk zijn dat we veronderstellen dat mc_{jt}, rpc_{jt} en d_{jt} meetfouten bevatten. De variabelen rar_{jt} en rat_{jt} daarentegen worden verondersteld correct te zijn gemeten. De reden voor deze veronderstelling is dat de STER de complete data omtrent radio- en televisie-reclame aanlevert aan het B.B.C. Dit in tegenstelling tot de uitgaven aan reclame via de pers, welke metingen gebaseerd zijn op de registratie van advertenties in de gedrukte media.
Rekening houdend met de meetfouten in rpc_{jt} en d_{jt} specificeren we:

$$mc_{jt} = \gamma_{1j}\xi_{1jt} + \gamma_{2j}\xi_{2jt} + \gamma_{3j}rar_{jt} + \gamma_{4j}rat_{jt} + \zeta_{jt}, \tag{31}$$

$$rpc_{jt} = \xi_{1jt} + \delta_{11jt}, \tag{32}$$

$$d_{jt} = \xi_{2jt} + \delta_{12jt}, \tag{33}$$

waarbij ξ_{1jt}, ξ_{2jt} de ware waarden representeren van rpc_{jt} en d_{jt}, en δ_{11jt}, δ_{12jt} de corresponderende meetfouten zijn.
Als instrumentele variabelen corresponderend met de foutief gemeten verklarende variabelen zijn beschikbaar: de relatieve prijs, zoals gemeten door het detaillistenpanel (rpr_{jt}) en de ongewogen distributie ('market coverage', dm_{jt}), eveneens afkomstig van het detaillistenpanel.
Met behulp van deze additionele informatie kan de momenten methode worden toegepast en kunnen consistente IV-schatters van de parameters in (31)-(33) worden verkregen[18].
Door nu de meetfouten in rpc_{jt} en d_{jt} te negeren en de methode van de kleinste kwadraten ('Ordinary Least Squares', OLS) op (30) toe te passen, resulteren de schattingen vermeld in Tabel 1. Hierbij zij opgemerkt dat voor merk A de variabele rar_{jt} geen significant effect bleek te hebben en derhalve uit de specificatie is verwijderd. Hetzelfde geldt voor merk B voor de variabele rat_{jt}.
Tabel 1 bevat de OLS-schattingen van de γ's, de geschatte standaarddeviaties (tussen haakjes), de schattingen van de varianties van de storingstermen in de relaties, de waarden van de Durbin-Watson toetsgrootheid (D.W.) en de determinatiecoëfficient (R^2).

Tabel 1. OLS-schattingen: relatie (30).

Merk	γ_{1j}	γ_{2j}	γ_{3j}	γ_{4j}	$\sigma^2_{\zeta_j}$	R^2	D.W.
A	-0.656[a] (0.096)	0.775[a] (0.183)	----	0.107[a] (0.039)	0.519×10^{-4}	0.92	1.80
B	-0.256[a] (0.021)	0.163[a] (0.030)	0.022[a] (0.009)	----	0.299×10^{-4}	0.88	1.75

a. Parameterschatting significant op het 5%-niveau (éénzijdige toets)

De schattingen van de standaarddeviaties en de varianties van de storingstermen in Tabel 1 zijn verkregen met behulp van de traditionele formules[19]. De schattingen van de response parameters in Tabel 1 zijn allen significant.
Met behulp van de momenten methode worden de resultaten in Tabel 2 verkregen, waarbij we ons in Tabel 2 in eerste instantie beperken tot de IV-schattingen welke vergelijkbaar zijn met de OLS-schattingen in Tabel 1.

Tabel 2. IV-schattingen: relatie (31).

Merk	γ_{1j}	γ_{2j}	γ_{3j}	γ_{4j}	$\sigma^2_{\zeta_j}$	R^2	D.W.
A	-0.639[a] (0.137)	0.849[a] (0.248)	----	0.116[a] (0.044)	0.399×10^{-4}	0.94	1.78
B	-0.267[a] (0.023)	0.184[a] (0.035)	0.024[a] (0.009)	----	0.176×10^{-4}	0.91	1.89

a. Parameterschatting significant op het 5%-niveau (éénzijdige toets)

De schattingen van de standaarddeviaties in Tabel 2 zijn verkregen door de vierkantswortels te nemen van de diagonaal-elementen van[9]:

$$\frac{n}{n-k}\,\hat{\sigma}^2_{\zeta}\,(Z'X)^{-1}\,(Z'Z)\,(X'Z)^{-1}, \tag{34}$$

waarbij Z de (nxk)-matrix is van waarden van de instrumentele variabelen[20] ($Z = (z_1, \ldots, z_n)'$) en $\hat{\sigma}^2_{\zeta}$ bepaald is volgens de traditionele formules[19]. De in Tabel 2 vermelde schattingen van $\sigma^2_{\zeta_j}$ zijn niet volgens de traditionele formules berekend, maar resulteren uit het substitueren van steekproef momenten voor populatie momenten in (21).
Vergelijken we de schattingen van de γ's in de Tabellen 1 en 2, dan blijkt, zoals we op basis van (5) mogen verwachten, dat de IV-schattingen in het algemeen iets groter zijn in absolute waarde dan de corresponderende OLS-schattingen[21].
Met behulp van de jackknife en de bootstrap (100 replicaties) is de mogelijke bias in de (consistente) IV-schattingen bepaald. De resultaten zijn vermeld in Tabel 3. Voor het gemak zijn in Tabel 3 tevens de IV-schattingen uit Tabel 2, aangegeven met 'geobserveerde schattingen', vermeld.
Zoals aangegeven in paragraaf 3 zijn de jackknife schattingen veelal minder onzuiver dan de originele schattingen. Vergelijking van beide schattingen geeft dan ook

inzicht in de bias in de geobserveerde schattingen. Door gebruik te maken van toetsgrootheid (26) is nagegaan of de jackknife-schattingen significant van nul verschillen. De hypothesen dat de biases in de geobserveerde schattingen gelijk zijn aan nul, zijn getoetst met de grootheid (29). In het geval dat een hypothese is verworpen, is in tabel 3 een voor bias gecorrigeerde schatting (kolom (10)) vermeld. Deze wordt verkregen door de geschatte bias op te tellen bij de geobserveerde schatting. De voor bias gecorrigeerde schatting van bijvoorbeeld γ_{1A} is gelijk aan -0.639 - 0.032 = -0.671. Dat de laatste schatting inderdaad voor bias is gecorrigeerd kan als volgt worden verduidelijkt: de hypothese $\gamma_{1A} - E(\hat{\gamma}_{1A}) = 0$ is verworpen en het verschil $\gamma_{1A} - E(\hat{\gamma}_{1A})$ wordt geschat als -0.032. Bij benadering geldt dus dat $E(\hat{\gamma}_{1A}) = \hat{\gamma}_{1A} + 0.032$, zodat het 'optellen' van -0.032 bij $\hat{\gamma}_{1A}$ een voor bias gecorrigeerde schatting van γ_{1A} oplevert.

Tabel 3. Evaluatie van response parameters en storingsterm varianties: relatie (31).

(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)
		Geobserveerd		*Jackknife*		*Bootstrap (B = 100)*			
Merk	Parameter	IV-schatting	Standaard dev.	Schatting	Standaard dev.	Gemiddelde	Standaard dev.	Bias	Bias gecorrigeerd (3) + (9)
A	γ_{1A}	-0.639[a]	0.137	-0.668[a]	0.137	-0.607	0.149	-0.032[a]	-0.671
A	γ_{2a}	0.849[a]	0.248	0.780[a]	0.301	0.914	0.300	-0.065[a]	0.784
A	γ_{4A}	0.116[a]	0.044	0.132[a]	0.070	0.113	0.059	0.003	----
A[b]	$\sigma^2_{\zeta_A}$	0.399[a]	----	0.466[a]	0.115	0.337	0.100	0.062[a]	0.461
B	γ_{1B}	-0.267[a]	0.023	-0.267[a]	0.023	-0.262	0.024	-0.005[a]	-0.272
B	γ_{2B}	0.184[a]	0.035	0.179[a]	0.038	0.182	0.041	0.002	----
B	γ_{3B}	0.024[a]	0.009	0.022[a]	0.009	0.026	0.009	-0.002	0.022
B[b]	$\sigma^2_{\zeta_B}$	0.176[a]	----	0.207[a]	0.051	0.145	0.038	0.031[a]	0.207

a. Parameterschatting significant op het 5%-niveau (éénzijdige toets).
b. De schattingen van de varianties van de storingstermen zijn vermenigvuldigd met 10^4.

De resultaten van Tabel 3 leiden tot de volgende conclusies:

1. Vergelijkingen van de resultaten in de kolommen (3) en (5) leert dat behalve voor γ_{1B}, γ_{2B} en γ_{3B} de biases in de geobserveerde schattingen aanzienlijk zijn. Behalve voor γ_{4A} wordt dit bevestigd door de resultaten in kolom (9).
   Voor zover de schattingen van de biases in kolom (9) relatief groot en significant zijn, dienen de corresponderende geobserveerde schattingen voor bias te worden gecorrigeerd. Daartoe kan men de jackknife-schattingen in kolom (5) gebruiken, of overgaan tot de berekening van de gecorrigeerde schattingen in kolom (10). Hoewel corresponderende jackknife schattingen en voor bias gecorrigeerde schattingen in het algemeen van elkaar verschillen, zijn in Tabel 3 de richtingen van de modificaties die door deze schattingen worden gesuggereerd dezelfde.
2. Met betrekking tot de response parameters blijkt dat, behalve voor $\hat{\gamma}_{2A}$ en $\hat{\gamma}_{4A}$, geen grote verschillen bestaan tussen de standaarddeviaties in de kolommen (4),

(6) en (8). Zowel de jackknife als de bootstrap geven aan dat voor $\hat{\gamma}_{2A}$ en $\hat{\gamma}_{4A}$ de standaarddeviaties in kolom (4) te klein zijn. De conclusie met betrekking tot de statistische significantie van de geobserveerde schattingen van de response parameters blijven onveranderd als deze significantie toetsen worden uitgevoerd met de geschatte standaarddeviaties in kolom (6) of (8), in plaats van met de standaarddeviaties in kolom (4).
Aangezien de biases in de geobserveerde schattingen van γ_{1A} en γ_{2A} relatief groot zijn, zijn voor deze parameters significantie toetsen op basis van de geobserveerde schattingen niet geschikt. In plaats daarvan lijkt het gebruik van de minder onzuivere jackknife schattingen meer op zijn plaats. Een ander alternatief is het gebruik van de voor bias gecorrigeerde schattingen in kolom (10) en de standaarddeviaties in kolom (8).
Het blijkt dat voor γ_{1A} en γ_{2A}, alsook voor de andere response parameters in Tabel 3 de resultaten van de verschillende significantie toetsen dezelfde zijn. Met andere woorden: significantie toetsen op basis van geobserveerde schattingen geven geen misleidende resultaten.

3. Met betrekking tot de varianties van de storingstermen blijkt dat de standaarddeviaties in kolommen (6) en (8) beide aangeven dat de geobserveerde schattingen significant zijn.
Daar zowel de jackknife als de bootstrap aangeven dat de geobserveerde schattingen te klein zijn, verdienen significantie toetsen op basis van de jackknife schattingen of de voor bias gecorrigeerde schattingen de voorkeur. Evenals hierboven (punt 2) zijn de uitkomsten van de verschillende toetsen dezelfde.

De conclusies 1-3 zijn representatief voor de resultaten in de meer uitgebreide studie van Plat (1988): ondanks dat een aantal IV-schattingen aanzienlijk onzuiver blijkt te zijn, leidt het gebruik van deze schattingen in significantie toetsen niet tot misleidende resultaten.

De schattingen van de meetfout varianties en van de varianties van de werkelijke waarden van de verklarende variabelen, zoals bepaald met de momenten methode, zijn vermeld in kolom (3) van Tabel 4. In deze Tabel zijn tevens schattingen van de ruis-signaal verhoudingen weergegeven. De jackknife en de bootstrap zijn toegepast voor het schatten van standaarddeviaties en het bepalen van de bias in de 'geobserveerde' schattingen.
De significantie van de geobserveerde schattingen in kolom (3) is bepaald met behulp van de standaarddeviaties in kolommen (5) en (7) en de percentielen van de normale verdeling. In alle gevallen leidde het gebruik van de twee standaarddeviaties tot dezelfde conclusies.
Op basis van Tabel 4 kunnen we het volgende concluderen:

1. Vergelijking van de resultaten in de kolommen (3) en (4) leert dat, behalve voor merk B, de geobserveerde schattingen van de varianties en de ruis-signaal verhoudingen aanzienlijk onzuiver kunnen zijn. Dit beeld wordt tot op zekere hoogte weerspiegeld in de resultaten vermeld in kolom (8).
Voor zover de schattingen van de biases in kolom (8) significant zijn verdienen de jackknife schattingen of de voor bias gecorrigeerde schattingen de voorkeur boven de geobserveerde schattingen. Hoewel de eerstgenoemde schattingen in het algemeen verschillend zijn, blijken wederom de richtingen van de gesuggereerde modificaties dezelfde te zijn.

Tabel 4. Evaluatie van meetfout varianties en ware variantie: relaties (31)-(33)

(1)	(2)	(3) Geobs.	(4) Jackknife	(5)	(6) Bootstrap (B = 100)	(7)	(8)	(9)
Merk	Parameter	Schatting	Schatting	Standaard dev.	Gemiddelde	Standaard dev.	Bias	Bias gecorrigeerd (3) + (8)
A[b]	$\sigma^2_{\delta 11A}$	0.137	0.231	0.281	0.028	0.333	0.109[a]	0.246
A[b]	$\sigma^2_{\xi 1A}$	5.685[a]	5.792[a]	1.032	5.506	0.972	0.179[a]	5.864[a]
A[c]	$\sigma^2_{\delta 11A}/\sigma_{x1a}$	2.347	3.777	4.877	0.450	6.278	1.897[a]	4.244
A[b]	$\sigma^2_{\delta 12A}$	0.094	0.133	0.091	0.081	0.074	0.013[a]	0.107
A[b]	$\sigma^2_{\xi 2A}$	1.768[a]	1.793[a]	0.346	1.668	0.326	0.100[a]	1.868[a]
A[c]	$\sigma^2_{\delta 12A}/\sigma^2_{x2A}$	5.053	7.206	4.680	4.591	4.321	0.462	----
B[b]	$\sigma^2_{\delta 11B}$	0.387	0.399	0.544	0.404	0.420	-0.017	----
B[b]	$\sigma^2_{\xi 1B}$	34.605[a]	35.800[a]	4.691	34.020	4.976	0.585	----
B[c]	$\sigma^2_{\delta 11B}/\sigma^2_{x1B}$	1.105	1.105	1.539	1.184	1.249	-0.079	----
B[b]	$\sigma^2_{\delta 12B}$	0.877	0.946	0.874	0.859	0.739	0.018	----
B[b]	$\sigma^2_{\xi 2B}$	20.163[a]	20.819[a]	4.423	19.689	4.008	0.474	----
B[c]	$\sigma^2_{\delta 12B}/\sigma^2_{x2B}$	4.169	4.199	4.245	4.196	3.655	-0.027	----

a. Parameterschatting significant op het 5%-niveau (éénzijdige toets).
b. De schattingen van de varianties (σ^2_δ, σ^2_ξ) zijn vermenigvuldigd met 10^4.
c. De schattingen van de ruis-signaal verhoudingen zijn als percentages vermeld (d.w.z. vermenigvuldigd met 100%).

2. De uitkomsten van de significantie toetsen op basis van de geobserveerde schattingen zijn gelijk aan de uitkomsten verkregen op basis van de jackknife schattingen. Voor zover de schattingen van de biases significant zijn worden deze uitkomsten bovendien bevestigd door de voor bias gecorrigeerde schattingen en de daarbij behorende standaarddeviaties. De geobserveerde schattingen leiden dus opnieuw niet tot misleidende resultaten.
3. De ruis-signaal verhoudingen voor rpc_A, d_A, rpc_B en d_B kunnen ruwweg worden geschat als respectievelijk 2%, 5%, 1% en 4%. Deze percentages zijn verkregen door afronding van de geobserveerde schattingen, of, in het geval van een significante bias, door het gemiddelde te nemen van de jackknife schatting en de voor bias gecorrigeerde schatting.

De conclusies 1-3 komen in grote lijnen overeen met de overige resultaten in de studie van Plat (1988), zij het dat voor één van de merken een negatieve schatting van een meetfout variantie resulteerde. Lawley en Maxwell (1971) refereren naar een dergelijke uitkomst met 'Heywood Case' en geven aan dat het optreden van Heywood Cases duidt op fouten in de modelspecificatie, dan wel het gebruiken van te kleine steekproeven.
De studie laat zien dat er een verband bestaat tussen de correlaties tussen instru-

mentele variabelen en ruis-signaal verhoudingen anderzijds. Des te groter deze correlaties, des te betrouwbaarder de variabelen gemeten zijn en des te kleiner de meetfout varianties en ruis-signaal verhoudingen. Zo zijn voor de merken A en B de correlaties tussen de prijs- respectievelijk de distrubutie variabelen groter dan 0.9 en de meetfout varianties navenant klein. Voor een ander, hier niet beschreven merk, bedroeg de correlatie tussen de prijsmetingen van het consumenten- en detaillisten panel slechts 0.73 en werd een ruis-signaal verhouding gevonden van 27%.

5. ENKELE AFSLUITENDE OPMERKINGEN EN CONCLUSIES

Uit ruimte-overwegingen zijn in paragraaf 4 geen resultaten vermeld voor de relaties tussen de instrumentele variabelen en verklarende variabelen (de relaties (12) in paragraaf 2.2). In (12) wordt elke instrumentele variabele aan alle verklarende variabelen gerelateerd. De storingstermen δ_2 kunnen dan ook niet, zoals de elementen van δ_1, als meetfouten worden geïnterpreteerd. Dit laatste wordt wel mogelijk als de relaties (12) worden gemodificeerd, zodanig dat elke instrumentele variabele slechts wordt gerelateerd aan de verklarende variabele die met de instrumentele variabele wordt gemeten. In concreto betekent dit dat a priori een aantal elementen van de matrix A gelijk aan nul wordt gesteld. Ook deze modellen zijn in Plat (1988) numeriek gespecificeerd. Met behulp van de resultaten kunnen ruis-signaal verhoudingen worden berekend voor zowel de waargenomen waarden van de verklarende variabelen (de x-variabelen) als voor de instrumentele variabelen (de z-variabelen). Voor de onderzochte modellen betekent dit dat de betrouwbaarheid van de relatieve prijsmetingen van het consumentenpanel vergeleken kan worden met die van de prijsmetingen van het detaillistenpanel. Evenzo kunnen gewogen en ongewogen distributie metingen (detaillistenpanel) qua betrouwbaarheid met elkaar worden vergeleken. De rangordes van de metingen op basis van hun betrouwbaarheid blijken over de merken te variëren, dat wil zeggen dat voor bepaalde merken bijvoorbeeld de relatieve prijs gemeten door het consumentenpanel de meest betrouwbare prijsmeting is, terwijl voor andere merken juist de relatieve prijs gemeten door het detaillistenpanel betrouwbaarder is.
Alvorens over te gaan tot de conclusies zij nog gewezen op de mogelijkheid om met behulp van LISREL modellen eveneens de meetfouten in de metingen van het marktaandeel van een merk expliciet te schatten. Op deze wijze kunnen tevens verschillende marktaandeel metingen op hun betrouwbaarheid worden beoordeeld. Voorbeelden van deze modellen zijn te vinden in Leeflang, Plat, Biemans (1984) en Plat (1988).

Ten slotte kunnen we concluderen dat met het lineaire fouten-in-variabelen model consistente schatters van de parameters verkregen kunnen worden met als 'bijprodukt' informatie omtrent de betrouwbaarheid van de gebruikte data. De resultaten van de jackknife en bootstrap laten zien dat de IV-schatters, hoewel consistent, in kleine steekproeven aanzienlijk onzuiver kunnen zijn. Ondanks deze bias blijken significantie toetsen op basis van de IV-schatters nauwelijks tot misleidende resultaten te leiden.
De resultaten van de significantie toetsen op basis van de OLS-schatters lijken bo-

vendien overeen te komen met die op basis van de IV-schatters. Hieruit kunnen we concluderen dat het gebruik van de traditionele panelgegevens in marketing modellen in het algemeen niet tot onjuiste uitspraken omtrent de significantie van de effectiviteit van de gehanteerde marketing instrumenten leidt. De betrouwbaarheid van de gebruikte gegevens speelt hierbij een rol. Naarmate deze betrouwbaarheid afneemt, neemt de kans op onjuiste uitspraken toe. Tevens neemt, bij afnemende betrouwbaarheid, de kans op het optreden van niet-toelaatbare modeluitkomsten (zoals negatieve varianties) bij de specificatie van lineaire fouten-in-variabelen modellen toe.
De beschreven analyse demonstreert de aantrekkelijkheid van de vrij gemakkelijk toe te passen verdelingsvrije methoden jackknife en bootstrap: schattingen van standaarddeviaties worden verkregen, terwijl tevens de mate van onzuiverheid bepaald kan worden. Indien nodig leveren beide methoden bovendien voor onzuiverheid gecorrigeerde schattingen.
De aantrekkelijkheid van de jackknife en de bootstrap beperkt zich niet tot het terrein van de marketing modellen. Ook op andere terreinen van marketing onderzoek, waar het gaat om het vaststellen van de kwaliteit van onderzoeksresultaten, kan zinvol van deze methoden gebruik gemaakt worden. In dit verband dient nog opgemerkt te worden dat, hoewel de auteur van een main frame gebruikt heeft gemaakt, de hier beschreven analyses inmiddels ook met behulp van een PC uitgevoerd kunnen worden.

NOTEN

* *Dit artikel is gebaseerd op een gedeelte van het proefschrift: 'Modelling for Markets: Applications of advanced Models and Methods for Data Analysis', waarop de auteur onlangs is gepromoveerd tot doctor in de Economische Wetenschappen.*

1. Zie Plat (1988).
2. Zie Leeflang, Olivier (1980), Leeflang, Olivier, Plat (1981), Leeflang, Olivier (1985), Olivier (1987), Plat (1988); een Leeflang, Plat (1988), waarin op scanning gebaseerde panelgegevens worden geëvalueerd.
3. Zij $\hat{\gamma}$ een schatter van de parameter van γ; $\hat{\gamma}$ is een zuivere schatter van γ als $E(\hat{\gamma}) = \gamma$; waarbij met E de verwachte waarde van $\hat{\gamma}$ wordt aangegeven; $\hat{\gamma}$ is een consistente schatter van γ als plim $\hat{\gamma} = \gamma$, d.w.z. indien bij toenemende steekproefomvang de kans dat $\hat{\gamma}$ gelijk is aan γ tot 1 nadert (plim is de notatie voor de waarschijnlijkheidslimiet).
4. Dat wil zeggen dat de onbekende parameters van het model niet allemaal te schrijven zijn als functies van de elementen van de populatie covariantie matrix van de waarneembare variabelen in het model. In paragraaf 2 komen wij hierop terug.
5. Zie bijv. Mosteller, Tukey (1971), Thornburn (1976) (jackknife) en Diaconis, Efron (1983), Efron (1979) (bootstrap). Het is opmerkelijk dat beide methoden tot nu toe nog betrekkelijk weinig in het marketing onderzoek zijn toegepast.
6. De veronderstelling dat $E(\xi) = 0$ wordt gemakshalve gemaakt en is niet essentieel.
7. Daar de matrices symmetrisch zijn, worden alleen de diagonaal elementen en de elementen daarboven weergegeven.
8. Deze informatie kan verschillende vormen aannemen. Wij kunnen bijv. a priori bepaalde waarden voor varianties of ratios van varianties veronderstellen op basis van intuïtie, ervaring of resultaten van vergelijkbaar onderzoek. We kunnen ook gebruik maken van instrumentele variabelen, zoals in dit artikel is gedaan.
9. Zie, bijv. Johnston (1972), pp. 278-281.
10. Zie, bijv. Johnston (1972), p. 280 en McCallum (1973).
11. Zie, bijv. Aigner, Hsiao, Kapteyn, Wansbeek (1984).
12. Zie, bijv. Cramér (1946)

13. De afkorting LISREL staat voor LInear Structural RELation. De LISREL-modellen en ook het LISREL computerprogramma, zijn reeds in een eerdere editie van het Jaarboek besproken: zie Kuylen (1984) en Leeflang, Plat, Biemons (1984).
14. Zie noot 5 en Fenwick (1979).
15. Dit is de gebruikelijke formulering. Het is ook mogelijk de data in minder dan n groepen te verdelen en dus minder berekeningen uit te voeren.
16. De onzuiverheid of bias in $\hat{\gamma}$ is gedifinieerd als $\gamma - E(\hat{\gamma})$.
17. De oorspronkelijke specificatie bevatte eveneens de variabele rap (relatieve reclame-uitgaven aan pers). Voor geen van de onderzochte merken werd voor rap een significante schatting van de response parameter gevonden. Bovendien vereist de specificatie van een lineair fouten-in-variabelen model dat het effect van de verschillende variabelen ongelijk aan nul is. Om deze redenen is de variabele rap niet in (30) opgenomen.
18. Omwille van de beschikbare ruimte zullen we ons in het navolgende, wat de IV-schattingen betreft, beperken tot de resultaten voor relaties (31)-(33). De resultaten voor de relaties tussen de instrumentele variabelen en de verklarende variabelen blijven dus buiten beschouwing.
19. Dat wil zeggen: $\hat{\sigma}_\zeta^2 = \frac{1}{n} \sum_{i=1}^{n} (y_i - \hat{\gamma}' x_i)^2$, en de standaarddeviaties in Tabel 1 zijn de vierkantswortels van de diagonaal elementen van $\frac{n}{n-k} \hat{\sigma}_\zeta^2 (X'X)^{-1}$, waarbij X de (nxk)-matrix is van waargenomen waarden van de verklarende variabelen $(X = (x_1, \ldots x_n)')$.
20. Het bepalen van IV-schattingen vereist dat het aantal instrumentele variabelen gelijk is aan het aantal verklarende variabelen. Dit betekent dat naast de rpr_{jt} en dm_{jt} nog een derde instrumentele variabele nodig is. Daar rar_{jt} en rat_{jt} foutloos zijn gemeten kunnen deze voor merk B resp. A als derde instrumentele variabele worden gebruikt.
21. De jackknife en de bootstrap lieten zien dat geen van de verschilen significant was. Deze toepassingen van jackknife en bootstrap, waarbij schattingen en standaarddeviaties van het verschil van twee schatters worden bepaald, zijn ietwat ongebruikelijk. Zie, voor meer details, Plat (1988), hoofdstuk 3.

LITERATUUR

Aigner, D.J., Hsiao, C., Kapteyn, A. en Wansbeek, T.J. (1984), 'Latent Variables in Econometrics', in: *Handbook of Econometrics,* Vol. 2, Griliches, Z. en Intriligator, M.D. (red.), North-Holland Publ. Co., Amsterdam.

Cramér, H. (1946), *Mathematical Methods of Statistics,* Princeton University Press, N.J.

Diaconis, P. en Efron, B. (1983), 'Computer-intensive Methods in Statistics', *Scientific American* (May), pp. 96-108.

Efron, B. (1979), 'Bootstrap Methods: Another Look at the Jackknife', *The Annals of Statistics,* Vol. 7, pp. 1-26.

Fenwick, I. (1979), 'Techniques in Market Measurement: The Jackknife', *Journal of Marketing Research,* Vol. 16, pp. 410-414.

Johnston, J. (1972), *Econometric Methods,* McGraw-Hill Book Company Inc., New York.

Kuylen, A.A.A. (1984), 'LISREL in het marktonderzoek', *Jaarboek Ned. Ver. v. Marktonderzoekers,* De Vrieseborch, Haarlem, pp. 141-168.

Lawley, D.N. en Maxwell, A.E. (1971), *Factor Analysis as a Statistical Method,* American Elsevier, New York.

Leeflang, P.S.H. en Olivier, A.J. (1980), 'What's wrong with the Audit Data We Use for Decision-Making in Marketing', *Proceedings van het 33ste ESOMAR Congress,* Monte Carlo, pp.219-239.

Leeflang, P.S.H. en Olivier, A.J. (1985), 'Bias in Consumer Panel and Store Audit Data', *International Journal of Research in Marketing,* Vol. 2, pp. 27-41.

Leeflang, P.S.H. en Plat, F.W. (1988), 'Scanning Scanning Opportunities', Paper te presenteren tijdens het 41ste ESOMAR Marketing Congress, Lissabon, 4-8 september.

Leeflang, P.S.H., Olivier, A.J. en Plat, F.W. (1981), 'Een Aanzet tot de Verklaring van Dekkingsverschillen in Panelgegevens', *Jaarboek Ned. Ver. v. Marktonderzoekers,* De Vrieseborch, Haarlem, pp.33-42.

Leeflang, P.S.H., Plat, F.W. en Biemans, W.G. (1984), 'Het Simultaan Gebruiken van Verschillende Monitoren bij de Bepaling van de Effectiviteit van Marketing Instrumenten', *Jaarboek Ned. ver. v. Marktonderzoekers,* De Vrieseborch, Haarlem, pp. 41-63.
McCallum, B.T. (1973), 'A Note Concerning Asymptotic Covariance Expressions', *Econometrics,* Vol. 41, pp.581-583.
Miller, R.G. (1974), 'The Jackknife – A Review', *Biometrika,* Vol. 61, pp. 1-15.
Mosteller, F. en Tukey, J.W. (1977), *Data Aanalysis and Regression,* Addison-Wesley, Reading, Massachusetts.
Olivier, A.J. (1987), *Het Samenstellen en Beheer van Gegevens als Onderdeel van een Beslissing-Ondersteunend Systeem ten behoeve van het Marketing Management: een Case Study,* Proefschrift, Rijksuniversiteit Groningen.
Plat, F.W. (1988), *Modelling for Markets: Applications of Advanced Models and Methods for Data Analysis,* Proefschrift, Rijksuniversiteit Groningen.
Thorburn, D. (1976), 'Some Asymptotic Properties of Jackknife Statistics', *Biometrika,* Vol. 63, pp. 305-313.

11. Involvement en reclameverwerking

J.P.H. NILLESEN

SAMENVATTING

In dit 'state of the art' artikel wordt een overzicht gegeven van de ontwikkeling van de voornamelijk psychologische reclameresponsmodellen en de theorieën waarop deze gebaseerd zijn[1]. Veronderstelden de traditionele effektenhiërarchiemodellen een min of meer strikte volgorde van een bepaald aantal stadia, tegenwoordig vindt men zo'n proces een te eenvoudige voorstelling van zaken. Het is beter om te stellen dat in ieder stadium alle drie aspekten van het consumentenbeslissingsproces (cognitie, affektie, conatie) aanwezig zijn, maar dat afhankelijk van het stadium een bepaald aspekt dominant is.
Omdat involvement een van de condities is die de verwerking van reclame – en daarmee het reclameeffekt – nuanceert c.q. bepaalt, wordt daarna dit concept behandeld. De verschillende benaderingen van het involvementconcept en de factoren die de aard en de intensiteit van het concept bepalen worden daarbij uitvoerig besproken. Vervolgens komen aan de orde de gevolgen die dit involvementonderzoek heeft voor de reclametheorie en de recente pogingen om één involvementconcept te ontwikkelen uit de verschillende bestaande visies op het concept.
Tot besluit worden enkele conclusies getrokken ten aanzien van de reclametheorie en de rol die het involvement concept daarin kan spelen.

1. INLEIDING

Er bestaat weinig twijfel over het feit dat het uiteindelijke doel van reclame het beïnvloeden van het menselijk gedrag is. Bij het plannen van een nieuwe reclamecampagne gaan marketers er impliciet vanuit dat een advertentie of commercial de potentiële kopers op een bepaalde manier zal beïnvloeden.
Over het algemeen gaat men hierbij uit van een communicatiemodel met 'hiërarchische' effekten, met andere woorden een bepaalde volgorde van taken of subdoelen (bijv. awareness → attraction → action). Volgens dit model zijn de drie taken van reclame het opmerkzaam maken van het publiek op het produkt en de bijbehorende produkteigenschappen, het produkt aantrekkelijk maken voor de consument (de 'unique selling proposition' kan hierbij een rol spelen) en deze stimuleren in actie te komen.

In de loop der tijd zijn er verschillende van deze reclameresponsmodellen beschreven, zoals we verderop zullen zien, maar het merendeel is op dit principe gebaseerd. Hoewel deze reclameresponsmodellen verschillen, is een gemeenschappelijk en zeer centraal punt de manier waarop de inhoud van een reclameuiting verwerkt wordt, hoe mensen de inhoud ervan begrijpen en hoe zij zich deze herinneren. Niettemin hebben veel auteurs bezwaar tegen deze rationele volgorde die ook zeker niet in alle gevallen door reclameonderzoek wordt ondersteund. In het bijzonder wordt de mate van de betrokkenheid van de consument ('involvement')[2] geacht van invloed te zijn op de wijze waarop

deze reageert op marketing-, in het bijzonder op reclamestimuli. Tevens zou deze mate van involvement de verklaring kunnen zijn van het feit dat bovenstaande reeks soms maar gedeeltelijk of in een andere volgorde of richting wordt gevolgd.

Het doel van dit artikel is op de eerste plaats een overzicht te geven van de ontwikkeling van de reclameresponsmodellen sinds het einde van de vorige eeuw tot heden, waarbij ook de meest recente gezichtspunten worden belicht.
Op de tweede plaats, gezien het belang van de involvement voor de manier waarop consumenten reageren op reclame en derhalve van de implicaties van dit concept voor de reclameresponsmodellen, is een groot deel van dit artikel gewijd aan de discussie van de rol van de involvement met betrekking tot reclame.
Ik wil pogen aan te geven wat de stand van zaken is ten aanzien van de theorie over dit involvement concept en wat de relevantie ervan is voor de ontwikkeling van reclameresponsmodellen. Het is derhalve noodzakelijk eerst een overzicht te geven van de literatuur over beide onderwerpen, waarna tot integratie van deze onderwerpen kan worden overgegaan. Het lijkt verstandig echter te beginnen met een kort overzicht van de in dit artikel voorkomende categorieën modellen.

2. VERSCHILLENDE CATEGORIEËN VAN MODELLEN

Het gedrag van consumenten kan vanuit vele gezichtshoeken benaderd worden, hetgeen dan ook heeft geleid tot een groot aantal verschillende modellen die in een aantal categorieën kunnen worden ondergebracht.
In dit artikel wordt een drietal soorten modellen genoemd, te weten consumentengedragsmodellen, communicatiemodellen en reclameresponsmodellen.

Het doel van de consumentengedragsmodellen is het verloop van het besluitvormings- en koopproces van individuele consumenten en de uitkomsten daarvan te beschrijven (Howard en Sheth 1969; Engel, Blackwell en Miniard 1986). Deze modellen bevatten een aantal relaties die vaak op onderzoek gebaseerd zijn, maar soms ook het theorievormend stadium nog niet zijn gepasseerd. Meestal worden in deze consumentengedragsmodellen de marktinstrumenten niet expliciet opgenomen; de effekten van deze marktinstrumenten kunnen daarentegen wèl deel uit maken van zo'n model.

Communicatiemodellen en reclameresponsmodellen houden zich in tegenstelling tot de consumentengedragsmodellen bezig met de werking van een van die marktinstrumenten. De communicatiemodellen geven op de een of andere manier het communicatieproces weer. In het meest eenvoudige geval worden in dit proces één zender en één ontvanger onderscheiden. We nemen aan dat de zender een idee, een gedachte heeft, welke hij wil overbrengen in de richting van de (potentiële) ontvanger. Deze gedachte moet daarvoor gecodeerd worden in de vorm van een boodschap, die via een bepaalde weg, het kanaal genoemd, bij de ontvanger aankomt. De ontvanger moet dan de boodschap weer decoderen, dat wil zeggen interpreteren op grond van zowel verbale als nonverbale tekens (cues) afkomstig van de zender[3]. De mogelijkheid bestaat zowel bij het coderen als bij het decoderen dat er vervormingen optreden. Feedback (terugkoppeling van informatie van ontvanger naar zender, in de vorm van reclameonderzoek bijvoorbeeld) stelt de zender in staat te controleren of zijn gedachte is overgekomen zoals hij had be-

doeld. In het ontkennende geval kan de zender dan zijn boodschap veranderen, bijstellen.
Over het algemeen leggen communicatiemodellen het accent op een deel van dit proces. De communicatiemodellen die de nadruk leggen op de verwerking van de boodschap door de ontvanger in dit communicatieproces, noemen we ontvangermodellen of – de meer ingeburgerde term – responsmodellen. Responsmodellen hebben tot doel weer te geven hoe consumenten reageren op de hantering van de marktinstrumenten, in ons geval van de communicatie, in het bijzonder de reclame. Hoewel er een trend in de huidige ontvangermodellen is, rekening te houden met de doelgroep (segmentatie dus) en zender- en boodschapvariabelen weer van belang te achten (b.v. Petty en Cacioppo 1986) – deze trend wordt wel aangeduid als 'the return of the power of the media' – ligt in dit artikel de nadruk op de ontvangende partij en daarmee op de reclameresponsmodellen.

3. EEN OVERZICHT VAN DE ONTWIKKELING VAN RECLAMERESPONSMODELLEN

Al heel lang bestaat de uitdaging de verklaring te vinden van de werking van reclame. Dàt reclame werkt is het punt niet, maar van het hoe en waarom weten we nog steeds niet het fijne [4]. Er zijn reclamemensen die vinden dat dat ook te ver gaat (Cowling 1984). Zij zijn er tevreden mee 'to see what advertising actually does, in terms of changing awareness, attitudes and behavior, and to use this information for campaign planning purposes'. Het heeft geen zin voor reclame- of marktonderzoekers de vraag te trachten te beantwoorden hoe onze hersenen informatie verwerken, aldus Cowling. Er bestaan echter ook reclamemensen die een wetenschappelijke aanpak van het reclameonderzoek, een empirisch ondersteunde theorie van de effektiviteit van reclame zouden toejuichen. Een gelukkige omstandigheid is dat deze vraag de laatste decennia ook steeds meer wetenschappers is gaan bezighouden. Een enorme hoeveelheid empirisch materiaal ligt ondertussen verspreid over ontelbare bureauladen en archiefkasten.

In de vijftiger jaren ontstonden de traditionele reclametheorieën uit de economie, de leertheorie, de persoonlijkheidsleer en de sociologie. Tegenwoordig vindt men deze theorieën slechts geschikt om een deel van het consumentengedrag te verklaren.
De economische theorie veronderstelt dat de consument een rationele, berekende produktkeuze maakt op basis van funktionele informatie. Olshavsky en Granbois (1979) maken bezwaar tegen deze afschildering van de consument als een betrokken, informatiezoekende persoon die zijn best doet de optimale keuze te maken. Zij laten zien dat een groot deel van het koopgedrag niets met besluitvorming te maken heeft, zelfs niet wanneer het betreffende produkt voor het eerst wordt aangeschaft.
Volgens de responsieve of leertheorie kopen consumenten het liefst met minimale inspanning. Ze ontwikkelen koopgewoonten door middel van stimulus-respons leren. Dit gedrag is niet-rationeel, maar automatisch, waarbij de herhaling het kopen van routineprodukten versterkt. Informatie heeft in dit geval alleen de funktie van geheugensteuntje.
De persoonlijkheidstheorie verklaart koopgedrag vanuit het begrip ego-involvement. De persoonlijkheid van de consument moet bevestigd en verdedigd worden. Mensen worden vaker en sterker bewogen door impliciete produktattitudes dan door functionele

produktvoordelen ('benefits'). Packard's 'The hidden persuaders' is een bekende illustratie van deze zienswijze.
De vierde traditionele theorie, de sociale theorie, zoekt de verklaring van het consumentengedrag in het imiteren van voor de consument belangrijke personen: mensen kopen wat andere mensen kopen. Status, de rol in de groep en opinieleiders zijn belangrijke begrippen binnen deze sociale theorie.
Alle vier genoemde theorieën hebben helaas slechts een beperkt gezichtspunt en beperkte toepassingsmogelijkheden.

Met name in de zestiger jaren is een aanzienlijk aantal modellen ontwikkeld die qua vorm nogal van elkaar verschilden hoewel zij van dezelfde assumptie uitgingen, namelijk van een geïnvolveerde consument.

Reeds in 1898 had Lewis het zogenaamde AIDA communicatiemodel ontwikkeld dat bestond uit de achtereenvolgende fasen attention, interest, desire en action (Preston 1982). Dit model is vermoedelijk wel het bekendste. Destijds ging men uit van het 'active audience' standpunt omdat er weinig commerciële media waren (men kende eigenlijk alleen drukwerk) en ook de beschikbaarheid van de produkten beperkt was. De consument moest hierdoor wel actief informatie vergaren en een uitgebreid beslissingsproces doormaken.

Colley (1961) schiep een model dat bestond uit de fasen awareness, comprehension, conviction en action en dat bekend is geworden onder de naam DAGMAR. In hetzelfde jaar ontwikkelden Lavidge en Steiner een model waarin zes stappen worden onderscheiden: awareness, knowledge, liking, preference, conviction en purchase. Hoewel zij toegaven dat onder bepaalde omstandigheden (met name in geval van lage involvement) consumenten een of meerdere stappen konden overslaan, wordt de consument in essentie via een aantal achtereenvolgende stappen (effekten) verleid tot aankoop. Lavidge en Steiner waren de eersten die deze stappen groepeerden in drie categorieën: cognitief (mentaal), affektief (emotioneel) en conatief (actie, gedrag). Zo legden zij de link tussen de drie attitudedimensies uit de sociale psychologie en het consumentengedrag. Vanwege de indeling in deze drie categorieën wordt dit model ook vaak de leer-voel-doe volgorde genoemd[5]. Veel modellen, die in de zestiger jaren zijn ontwikkeld zijn afgeleid van het model van Lavidge en Steiner en daardoor van een hiërarchische, chronologische aard. Hoewel zij elkaar deels tegenspreken, soms overlappen of aanvul-

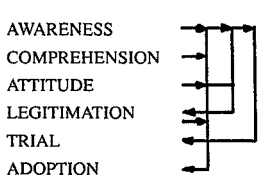
THREE A	AIDA	ADOPTION PROCESS	DAGMAR	LAVIDGE + STEINER	HOWARD + SHETH
Awareness	Attention	Awareness	Awareness	Awareness	Attention
Attraction	Interest	Interest	Comprehension	Knowledge	Brand
Action	Desire	Evaluation	Conviction	Liking	comprehension
	Action	Trial	Action	Preference	Attitude
		Adoption		Conviction	Intention
				Purchase	Purchase

Figuur 1. Enkele traditionele hiërarchische modellen en een vereenvoudigde versie van het Howard en Sheth (1969) model.

len, veronderstellen zij allen dat mensen een aantal stappen doorlopen tussen aandacht (attention) en koop (purchase).
Het was Palda (1966) die dit model voor het eerst een hiërarchie van effecten noemde vanwege het feit dat elke stap een noodzakelijke voorwaarde vormt voor de volgende.
Assael en Day (1968) vonden in hun onderzoek weinig empirische ondersteuning voor het oorzakelijk verband cognitie → affect → gedrag dat inherent is aan het traditionele hiërarchische model.
Palda (1966) twijfelde overigens aan de veronderstelling van een hiërarchische causaliteit in één richting en merkte op dat attitudes het gevolg kunnen zijn van koopgedrag, maar er ook aan vooraf kunnen gaan. Wèl ging hij uit van de hypothese dat ieder volgend bereikt stadium een hogere aankoopwaarschijnlijkheid betekende, waar hij overigens in zijn onderzoek geen bevestiging voor vond.
Tegelijkertijd werd dit causale eenrichtingsverkeer onderzocht in de sociale psychologie (Bem 1972). Meer dan eens moest men concluderen dat gedragsveranderingen kunnen leiden tot attitude- of beliefwijzigingen (Insko en Schopler 1967).

Ondertussen waren er nogal wat varianten ontwikkeld, waaronder het meest volledige en complexe (consumentengedrags)model van Howard en Sheth (1969). In dit model worden zo'n veertig variabelen ingedeeld in input-, perceptie-, leer- en outputvariabelen. Hoewel Van Veldhoven (1979) stelt dat de tijd van de integrale consumentengedragsmodellen voorbij is omdat de modellen van Howard en Sheth, Nicosia en Engel, Kollat en Blackwell niet toetsbaar zijn gebleken, hebben deze modellen toch ook een aantal andere funkties: het integreren van informatie, als checklist van relevante factoren en ten behoeve van de generatie van ideeën.

In figuur 1 zijn een aantal traditionele hiërarchische modellen weergegeven, waaronder een gesimplificeerde versie van dat van Howard en Sheth, waarin alleen de reeks interveniërende (output) variabelen is aangegeven. De stippellijnen stellen feedbacks voor, waardoor de volgorde niet meer volledig vast staat en wederzijdse beïnvloeding mogelijk wordt.
Robertson (1971) ontwikkelde een eenvoudiger variant. Dit gewijzigde hiërarchische model veronderstelde dat sommige consumenten afhankelijk van de omstandigheden de verschillende achtereenvolgende stappen zullen doorlopen. Onder meer feedbacks kunnen deze volgorde veranderen. Robertson handhaafde dus de leer-voel-doe volgorde van de traditionele hiërarchische modellen, maar maakte deze flexibeler. Zijn model hielp mee koopgedrag te verklaren zonder meetbare produktkennis of attitudeformatie.
Vanwege de veranderbare volgorde en het variërende belang van de verschillende stadia werden de effektenhiërarchie-modellen door Van Westendorp (1970) preweten-schappelijk genoemd.



Figuur 2. Robertson's (1971) model van reclame-effekten.

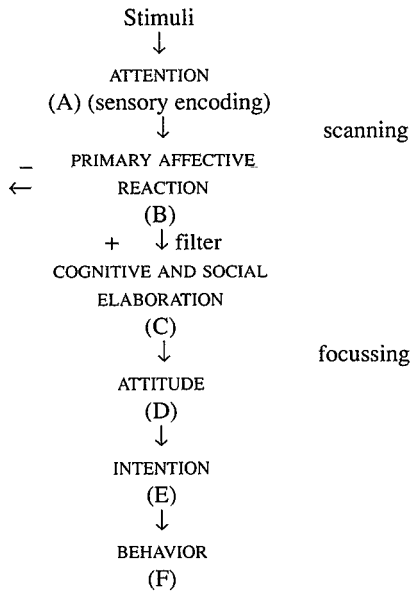
		WAT KOMT EERST?	
		DENKEN	VOELEN
HOGE	INVOLVEMENT	1. INFORMATIEF (DENKEN) Auto-huis-meubels-nieuwe produkten Model: Leer-voel-doe (Economisch)	2. AFFEKTIEF (VOELEN) Juwelen-cosmetica-modekleding Model: Voel-leer-doe (Psychologisch)
LAGE		3. GEWOONTEVORMING (DOEN) Eten-huishoudelijke artikelen Model: Doe-leer-voel (Leertheoretisch)	4. SATISFACTIE (REAGEREN) Sigaretten-sterke drank snoepgoed Model: Doe-voel-leer (Sociaal)

Figuur 3. Een vereenvoudigde versie van het Foote-Cone-Belding model (1980, 1986).

Een ander gezichtspunt op de werking van reclame is afkomstig van Foote, Cone & Belding, een vooraanstaand Amerikaans reclamebureau (Vaughn 1980, 1986; Berger 1986). Het FCB model integreert de vier oorspronkelijke reclametheorieën, het leer-voel-doe model van Lavidge en Steiner en de varianten erop, en twee theorieën die zijn ontstaan in de zeventiger jaren, namelijk de involvementtheorie en de hersenspecialisatietheorie.
De involvementtheorie houdt in dat de mate waarin de consument betrokken is bij het produkt of het beslissingsproces zijn gevoeligheid voor reclame bepaalt. Verderop in dit artikel komen we uitgebreid op dit onderwerp terug.
Hersenspecialisatie (of hersenlateralisatie) wil zeggen dat het reageren van de consument op reclame afhankelijk is van de cognitieve (denk) en affektieve (gevoel) taken die de reclame met zich meebrengt.
De matrix verdeelt de potentiële reacties van de consument in denken en voelen (horizontaal) en in hoge en lage involvement (verticaal). In essentie stelt dit FCB model dat allereerst de relevantie van de informatie (leer), attitude (voel) en gedrags (doen)-issues voor de doelgroep moeten worden vastgesteld. In de matrix vindt men implicaties voor relevante doelstellingen, communicatie- en mediaconsequenties en methodologieën voor pre- en posttesten. De produkten die in de matrix genoemd worden, kan men hooguit zien als voorbeelden. In feite versimpelt het FCB model de communicatiemodellen tot vier stuks, voor elk waarvan mogelijke produkttoepassingen genoemd worden. In figuur 3 staat een vereenvoudigde versie van de FCB matrix afgebeeld.

Vrij recent publiceerde Van Raaij (1984, 1985, 1986) een nieuw model van de werking van reclame, waarin hij beschikbare hypothesen en delen uit vroegere theorieën met betrekking tot massacommunicatie, tot reclame, integreert.
In tegenstelling tot de meer traditionele hiërarchische modellen, waarin een cognitieve fase voorafgaat aan de affektieve/emotionele fase, veronderstelt Van Raaij in zijn model een primaire affektieve reactie vóór de cognitieve reactie op de reclame. Vanwege de steeds toenemende hoeveelheid reclamestimuli kan de consument onmogelijk al deze stimuli waarnemen of verwerken. Daarom moet reclame opvallen en een positieve emotionele respons, de zogenaamde primaire affektieve reactie, oproepen: de onver-

schilligheidsbarrière van de consument moet doorbroken worden. Bovendien moet de ontvanger van de reclameboodschap er de juiste conclusies uit trekken en de door de zender bedoelde boodschap onthouden. Awareness must be created. Daarom is de primaire affektieve reactie (verder PAR te noemen) de tweede fase in Van Raaij's reclameverwerkingsmodel, na de aandacht, maar voor de cognitieve elaboratie. In figuur 4 staat Van Raaij's model van het massacommunicatie proces afgebeeld.



Figuur 4. Een model van het massacommunicatieproces (Van Raaij 1986).

Volgens Van Raaij (1985) geeft veel psychologisch onderzoek aan dat een affektieve reaktie voorafgaat aan een cognitieve respons. Pas daarna volgt een meer uitgewerkte secundaire affektieve reactie, die in Van Raaij (1986) de attitudefase is geworden. In Van Raaij's model heeft de PAR drie funkties. Ten eerste de funktie van aandacht voor stimuli met opvallende, dat wil zeggen emotionele of collatieve (afwijkende, ongebruikelijke) eigenschappen. Ten tweede wordt in reclame emotie als zodanig vaak opgeroepen door associatie van het produkt met een positief emotionele situatie of persoon (een vorm van klassiek conditioneren dus). Ten derde vergemakkelijkt een positieve PAR de rest van de informatieverwerking doordat deze PAR een activatieniveau (arousal) creëert dat verdere cognitieve elaboratie bevordert.

De PAR, die grotendeels onbewust is, bepaalt in aanzienlijke mate de soort en diepte van de verwerking in de volgende fase. Vanwege de onbewuste aard van de PAR wordt deze cruciale factor in veel modellen vergeten of over het hoofd gezien. Van Raaij (1985) merkt op dat '... de reden dat reclame waarschijnlijk effektiever is dan we dachten, is dat de PAR niet of in zeer geringe mate erkend is als een reclameeffekt. Dit model vult deze leemte'. Hij besluit het artikel uit 1985 met een humanistische kijk op reclame, die hij gemeen heeft met Lannon en Cooper (1983). Hoe reclame inwerkt op

mensen is minder belangrijk dan wat mensen met reclame doen. Met andere woorden: op welke manier roept reclame een positieve emotie of interesse op? En hoe verwerkt de consument de informatie ten behoeve van de wereld waarin hij leeft en zijn produktgebruikssituatie.

Op Van Raaij's PAR is voornamelijk kritiek gekomen van de kant van enkele Nederlandse psychologen, die in tegenstelling tot Van Raaij van mening zijn dat er weinig empirische ondersteuning is ten aanzien van het affektieve element van het concept. Psychofysiologisch gezien zou het juister zijn deze PAR een oriëntatierespons te noemen (Zeldenrust 1986; Spijker 1986)[6]. Een oriëntatierespons 'consists of orienting or investigatory movements to stimuli that are novel or that signal a biologically important event' (Butter 1968). Een essentieel aspekt van een oriëntatierespons, een reflex, is het feit dat deze onbewust is. Deze respons 'aids in focussing the organism's attention to the stimulus, it also prepares the organism for appropriate reactions to a potentially significant stimulus'. Butter noemt verder nog een aantal psychofysiologische aspekten van de oriëntatierespons, maar die zijn hier niet relevant.

Vanwege het onbewust zijn van de PAR en het feit dat de PAR helpt in het richting geven aan de aandacht van de consument, wat niet hetzelfde is als emotionele effektiviteit, is de auteur van dit artikel geneigd Zeldenrust's mening te ondersteunen. Laat het duidelijk zijn dat affectie (emotie, gevoel) bij het verwerken van reclame heel belangrijk is; er blijft echter twijfel bestaan aan de affectieve aard van de PAR.
Spijker (1986) kon uit zijn experimenten om de PAR met behulp van de electrodermale respons te meten geen relatie afleiden tussen deze EDR[7] en affectieve reacties. Hij concludeert de EDR te moeten zien als een oriëntatierespons of een maat voor de activatie van de persoon, eerder dan als een maat voor een primaire affectieve reactie.
Verder onderzoek naar de aard van de PAR zou welkom zijn, waarbij het dan vooral zou moeten gaan om de volgende vragen:
- is een oriëntatierespons equivalent aan een primaire affektieve reactie?
- bestaat er een relatie tussen de primaire affektieve reactie en arousal? Arousal is namelijk niet noodzakelijk stimulusgericht, terwijl een oriëntatierespons dat wèl is.

Een recent verschenen reclametekstboek (Rossiter en Percy 1987) komt overigens wel weer dicht bij Van Raaij's PAR-model. We mogen aannemen dat het laatste woord over dit onderwerp nog niet geschreven zal zijn.

4. KRITIEK OP DE TRADITIONELE HIËRARCHISCHE RECLAMEMODELLEN

Zo'n twee decennia geleden ontstond er ernstige kritiek op de noodzaak voor consumenten de fasen in een vaste volgorde te doorlopen. Ieder hiërarchisch model vormde een waardevolle bijdrage aan het systematischer denken over reclame; aan de andere kant gaven empirische resultaten reden tot twijfel aan de validiteit ervan. Een van de eersten die ernstige kritiek leverde op de 'hierarchy of effects' zoals hij de modellen samenvatte, was Palda (1966). Naast zijn bezwaren tegen de onvoldoende mogelijkheden om de verschillende stappen te meten, wees hij op de conceptuele problemen die aan iedere fase verbonden waren, in het bijzonder aan de operationele definities ervan.

	attitudecomponenten		
leer-hiërarchie	cognitief	affektief	conatief
dissonantie-attributie hiërarchie	conatief	affektief	cognitief
lage involvement hiërarchie	cognitief	conatief	affektief

Figuur 5. De drie hiërarchieeën van de attitudecomponenten (Ray 1973).

Nadat Robertson (1971) de leer-voel-doe volgorde flexibeler had gemaakt, zoals we gezien hebben, opperde Ray (1973) dat een reclamerespons drie vormen kon aannemen, zoals uit figuur 5 moge blijken. De leerhiërarchie in deze figuur is het oorspronkelijke Lavidge en Steiner model.

Verschillende auteurs hebben opgemerkt dat niet iedereen noodzakelijk op dezelfde manier het proces doorloopt. Volgens Krugman (1965, 1967, 1977) is de leerhiërarchie alleen van toepassing als de consument in hoge mate geïnvolveerd is om een rationele keuze te maken. In de meeste gevallen echter, waarin slechts sprake is van een lage involvement, leidt reclame en in het bijzonder t.v.-reclame tot vertrouwdheid, die op haar beurt kan leiden tot een proefaankoop (product trial). Niet voor maar ná deze proefaankoop vindt dan waarschijnlijk attitudeformatie of attitudeverandering plaats.
De resterende sequentie in figuur 5, de tweede, is in feite gebaseerd op Festinger's cognitieve dissonantie theorie. Festinger toonde aan dat onder bepaalde condities (b.v. in het geval van onvoldoende rechtvaardiging van een bepaald koopgedrag) attitudeformatie of -verandering pas plaatsvindt nadat de betreffende consumenten het gedrag hebben vertoond. In dergelijke gevallen worden hun cognitieve structuren pas naderhand met het vertoonde gedrag in overeenstemming gebracht.

Er is eigenlijk niet echt sprake van een verwerping van het traditionele hiërarchie-van-effekten idee door Krugman of Festinger maar meer van nuanceringen: zij maakten duidelijk onder welke condities de traditionele theorie geldig is.

Vergelijkingen we Vaughn's (1980) FCB model met dat van Ray, dan valt op dat de omstandigheden waaronder de doe-voel-leer volgorde zich voordoet, verschillen. In dat geval veronderstelt Vaughn lage involvement, terwijl het in dat geval volgens Ray (en Festinger) gaat om een geïnvolveerde, maar onzekere consument. Vandaar de suggestie om eerst eens nader kennis te maken met het concept betrokkenheid ('involvement').

Omdat een aantal van de bedoelde reclame-effekten de nadruk leggen op het gedrag en andere modellen afkomstig zijn uit leer- of attitudetheorie en de meeste van de bovengenoemde effekten een pluriforme afkomst hebben, is het op zich niet onmogelijk dat een aantal van deze effekten simultaan in plaats van achtereenvolgens plaatsvinden en elkaar zo beïnvloeden. Moriarty (1983, 1987) stelt dan ook dat de verschillende fasen alle even belangrijk zijn en zich tenminste voor een deel gelijktijdig afspelen. Bij nader inzien ligt het waarschijnlijk meer voor de hand het parallelle effekt dat door terugkoppeling (feedback) ontstaat, toch als sequentieel op te blijven vatten.

5. HET INVOLVEMENT CONCEPT

Aan het involvement concept zijn nogal wat facetten te onderscheiden, dat wil zeggen: het concept heeft een aantal belangrijke dimensies. Een van de grootste problemen is geweest het beschrijven van deze relevante dimensies en hun belang voor de reclame.

De eersten die de term ego-involvement gebruikten waren waarschijnlijk Sherif en Cantril (1947). Volgens hen betekent involvement een relatie van een onderwerp of object met het unieke cluster attitudes en waarden die samen de ego van een persoon vormen. Sindsdien zijn er veel verschillende definities (Asseal 1981; Kassarjian 1981) verschenen in de psychologische en consumer research literatuur. Day (1970) definieerde involvement als 'the general level of interest in the object or the centrality of the object to the person's structure'.
Tyebjee (1979) veronderstelde dat involvement afhankelijk is van 'the number of values engaged by the product, the centrality of these values and the product's relatedness to these values'. Trayler (1981) stelt dat 'basically, the concept of product involvement is a recognition that certain product classes may be more or less central to an individual's life, his attitudes about himself, his sense of identity and his relationship to the rest of the world'. Krugman (1965) noemde 'the number of concious bridging experiences, connections or personal references per minute that the viewer makes between his own life and the stimulus'. Vanwege deze 'bridging experiences' zou de mate van involvement van de consument zijn reclamerespons beïnvloeden.

Dit veelvoud van definities komt voort uit het feit dat men het niet eens is over de inhoud van het involvement concept: een verheldering van wat het theoretische concept inhoudt is volgens een aantal auteurs (o.m. Kassarjian 1981) allereerst noodzakelijk. Vervolgens komt dan de vraag in ons op of het theoretisch mogelijk is een involvement concept te ontwikkelen door de bestaande zienswijzen met betrekking tot het concept te combineren. Enkele pogingen daartoe zijn recent ondernomen (b.v. Laurent en Kapferer 1985; Kapferer en Laurent 1986).

Bovengenoemde definities kunnen we ruwweg in drie groepen indelen: respons-, individuele toestand- en cognitieve benaderingen (Rajaniemi en Laaksonen 1986).
De eerste categorie definieert involvement in termen van het statische of veranderlijke individuele responspatroon dat een stimulusobject op kan roepen (b.v. Ray 1973; Batra en Ray 1983). Gewoonlijk wordt hierbij slechts een dichotoom onderscheid gemaakt in hoge en lage involvement: de respons is dan hoog- of laag-involvement gedrag. In geval van hoge involvement met een produkt (of dienst) wordt de betreffende consument verondersteld over te gaan tot uitvoering probleemoplossend gedrag, bij lage involvement is het beslissingsproces van de consument voornamelijk gebaseerd op gewoonten. Niettemin wijst Solomon (1986) erop dat ook geïnvolveerde consumenten als zij geconfronteerd worden met belangrijke keuzen, hun zoek- en beslissingsactiviteiten kunnen minimaliseren. Naast een toename van 'gewone' kopers vond Solomon ook een toenemend aantal surrogaat-consumenten, 'agents retained by a consumer to guide, direct and/or transact marketplace activities'.
Daar men het er in brede kring over eens is dat involvement geen responsvariabele is maar een mediërende variabele, en zeker niet de enige gedragsdeterminant (Mitchell 1979; Ray 1979; Park en Young (1984), kunnen we deze responses beter beschouwen

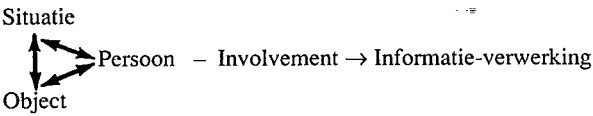
als effekten van involvement en niet als involvement zelf (Rajaniemi en Laaksonen 1986).

De tweede categorie, de individuele toestand benaderingen, concentreren zich op de individuele toestand van de consument die een stimulus op kan wekken (Mitchell 1979; Park en Young 1984). Vanuit dit gezichtspunt wordt involvement beschreven in termen van doelgerichtheid en intensiteit. Vormen van de motivationele toestand waarin het individu verkeert zijn bijvoorbeeld arousal, interesse en drijfveer ('drive') die dus de intensiteit van de involvement kunnen bepalen. Deze intensiteit varieert constant van heel laag tot heel hoog. In deze individuele toestand benaderingen wordt de doelgerichtheid van de involvement bepaald door het stimulusobject, de reclameuiting, of door het te bereiken doel. Als dit doel bereikt is, neemt het involvementniveau weer af. Dit is de reden dat involvement volgens deze benaderingen van tijdelijke aard is.

Enkele jaren geleden heeft Mitchell (1981, 1983) gesuggereerd dat involvement niet twee maar drie niveaus kent – wat in feite een uitwerking en toepassing is van Broadbent (1958) – terwijl Greenwald en Leavitt (1984) vier involvementniveaus onderkennen die relatief verschillend zijn, eerder dan variërend op een intensiteitscontinuum. Zij leggen de nadruk op de interne processen die zich afspelen in de consument als deze wordt blootgesteld aan een bepaalde reclameuiting.
Een hoger involvementniveau correspondeert met een hogere graad van symbolische activiteit die gebruikt wordt om de binnenkomende informatie te analyseren en daardoor met verschillende stappen van de sequentiële informatieverwerking. Mitchell (1981, 1983) en Greenwald en Leavitt (1984) zijn beide cognitieve benaderingen vanwege de nadruk die zij leggen op de manier waarop de produktinformatie wordt verzameld en verwerkt. Involvement slaat hier op de waargenomen persoonlijke belangrijkheid van een object of activiteit, met andere woorden op de relatieve belangrijkheid van de objectgerelateerde attitude. Deze relatieve belangrijkheid wordt geacht afhankelijk te zijn van de sterkte van de relaties van het betreffende object met het zelfconcept van de consument, met belangrijke waarden of motieven. Het involvementniveau is veranderlijk en hangt af van de complexiteit van de structuur van de attitude ten aanzien van het object in kwestie, die weer wordt opgeroepen door de sterkte van bovengenoemde relaties (Rajaniemi en Laaksonen 1986).

Een tweede dimensie van het involvement concept heeft te maken met de omstandigheden die de aard of intensiteit ervan beïnvloeden. Zaichowsky onderscheidt in haar artikel uit 1986 over het involvement concept drie soorten 'antecedents of involvement': persoonlijke factoren, object- of stimulusfactoren en situationele factoren. Rothschild (1978) deed ongeveer hetzelfde, maar noemde de daaruit resulterende soorten involvement respectievelijk responsinvolvement, voortdurende ('enduring') involvement en situationele involvement. Beide gezichtspunten komen voort uit het situationele model van het consumentengedrag dat Belk (1975) introduceerde, volgens welk er drie groepen factoren zijn die de omstandigheden waaronder een reclameuiting wordt waargenomen beïnvloeden en op hun beurt weer verantwoordelijk zijn voor het soort involvement (zie figuur 6).

Dit is de reden dat een aantal auteurs voorstelt het involvement concept te vervangen door de begrippen persoonlijke, situationele en objectrelevantie. Anderen (Poiesz en



Figuur 6. Involvement bij het waarnemen van reclame (Belk 1975; Mühlbacher 1986).

Pieters 1987) stellen voor een beter onderscheid te maken tussen de aan elkaar grenzende begrippen produktbelangrijkheid, situationele relevantie en involvement.

Een nieuwe, zij het verwante interpretatie van involvement is recentelijk verschenen (Petty, Cacioppo en Schuman 1983; Petty en Cacioppo 1986). In hun Elaboration Likelihood Model (ELM) maken zij onderscheid tussen twee 'routes' die reclameinformatie kan volgen, een centrale en een perifere route. De mate van involvement bepaalt welke route gevolgd zal worden. In geval van hoge involvement, als de waarschijnlijkheid dat de reclameinformatie cognitief geëlaboreerd zal worden groter is, is de centrale route effektiever; bij lage involvement, als de elaboratiewaarschijnlijkheid lager is, is de perifere route geschikter. In het laatste geval zal de informatie slechts oppervlakkig beoordeeld worden op basis van extrinsieke en contextuele boodschapcues[8]. Volgens Petty en Cacioppo's studies zullen attitudes als de centrale route wordt gevolgd, beter bestand zijn tegen de overredingskracht van de reclame: de inhoud van de informatie die de reclame verschaft is dan het meest relevant voor de vorming en verandering van deze attitudes.
Indien echter de perifere route gevolgd wordt, spelen andere aspekten van de reclame die niet tot de inhoud behoren zoals de expertise of aantrekkelijkheid van de informatiebron een bepalende rol. Afhankelijk van de omstandigheden kunnen beide soorten van attitudebeïnvloeding voorkomen, het belang van perifere cues moet dus niet onderschat worden. Persoonlijke involvement en produktrelevantie kunnen bepalen welke informatieverwerkingsroute genomen zal worden omdat deze factoren iemands motivatie beïnvloeden om de relevante informatie cognitief te bewerken.

Situaties kunnen vanwege hun motivationele aard evenals individuele verschillen (in persoonlijkheid bijvoorbeeld) eveneens belangrijke routebepalers zijn. 'The critical feature of the central route to persuasion is that an attitude change is based on a diligent consideration of information that the person feels is central to the true merits of an issue or product'. Met Zaichowsky (1986) kunnen we concluderen dat de antecedenten van involvement die gevonden zijn in de Petty en Cacioppo studies voornamelijk veroorzaakt worden door verschillen in omstandigheden en in stimulusobject. Helaas moeten we ook tot de conclusie komen dat Petty en Cacioppo onvoldoende aangeven welke situaties in het bijzonder zullen leiden tot hoge en welke tot lage involvement. Het ELM model ziet centrale en perifere verwerking als alternatieven, maar er zijn situaties denkbaar dat in dezelfde context beide routes werkzaam zijn. Bijvoorbeeld als centrale verwerking leidt tot gelijke preferenties voor twee produktalternatieven, zou perifere verwerking de doorslag kunnen geven.
Een zwakke kant van dit model van Petty en Cacioppo is tenslotte dat het in bepaalde opzichten weinig rekening houdt met andere theorieën en onderzoeksresultaten. Het ELM model onderscheidt slechts twee involvementniveaus, terwijl involvement veeleer

een continue variabele is. Een indeling in vier niveaus (Greenwald en Leavitt 1984) zou een acceptabele tussenoplossing kunnen zijn. Het ELM model moet dan wel aangepast worden. In geval van hoge involvement zou de nadruk dan moeten liggen op de cognitieve verwerking, bij lage involvement zou deze moeten liggen op een meer affectieve verwerking. De markt zou dan vervolgens afhankelijk van het involvementniveau van het betreffende segment, dienovereenkomstig op verschillende wijzen communicatief benaderd moeten worden. Hoewel er een tendens bestaat in de reclame steeds meer de nadruk op het affect te leggen door middel van het oproepen van emoties, zou men rekening houdend met deze verschillen in involvement bij de hogere involvementniveaus volgens mij daarentegen eerder de nadruk op de cognitieve aspekten moeten leggen.

Ten aanzien van involvement in relatie tot reclame kunnen we een aantal conclusies trekken (Zaichowky 1986):

- Het involvement niveau heeft motivationele gevolgen voor de consument. Als de consument geïnvolveerd is, is zijn manier en niveau van aandacht schenken (Krugman 1962, 1965; Weinstein et al. 1980; Wright 1973; Hansen 1981), zijn waarneming van de belangrijkheid van het produkt (Howard en Sheth 1969; Hupfer en Gardner 1971; Lastovicka en Gardner 1978; Zaichowsky 1985) en de informatie erover, en zijn beslissingsstrategie en keuzegedrag (Howard en Sheth 1969; Engel, Blackwell en Miniard 1986; Clark en Belk 1978) verschillend van dat als hij niet geïnvolveerd zou zijn.
- Verschillen tussen personen bepalen voor een belangrijk deel de verschillen in respons tussen consumenten op dezelfde reclameuiting. Krugman (1962, 1965) vond dat t.v.-reclame effektiever was bij laag-geïnvolveerde consumenten, terwijl gedrukte reclameadvertenties meer effect sorteerden bij hooggeïnvolveerde personen. Weinstein et al (1980) toonden aan dat de reclameeffektiviteit niet per definitie afhankelijk is van het gebruikte medium, maar van het involvementniveau van de kijker c.q. lezer. We komen daar verderop op terug.
- Volgens Wright (1973) leidt hoge involvement tot meer tegenargumenten ten aanzien van de inhoud van reclameboodschappen. Hij vond dat gedrukte media de ontvanger in staat stelden de informatie in het eigen gewenste tempo beter cognitief te verwerken. Van Raaij (1979) onderscheidt in een vergelijkbare context self-paced en external-paced media, waarbij in het laatste geval grote groepen consumenten 'onvoldoende toekomen aan de cognitieve elaboratie die nodig is voor een diep niveau van informatieverwerking'. Vreemd genoeg vond Wright geen hoger aantal proargumenten, hetgeen ook een mogelijk gevolg van een meer diepgaande cognitieve elaboratie zou kunnen zijn.
- Uit de ELM studies (Petty en Cacioppo 1979, 1981, 1983, 1986) is gebleken dat afhankelijk van hun mate van involvement met het onderwerp consumenten verschillende soorten boodschapcues gebruiken om tot evaluaties te komen. In geval van hoge involvement vinden attitudeformatie en -verandering plaats op basis van de inhoud van de boodschap, terwijl bij lage involvement de vorming en wijziging van attitudes voornamelijk gebaseerd zijn op de aantrekkingskracht en expertise van de informatiebron.

6. DE IMPLICATIES VAN HET INVOLVEMENTONDERZOEK VOOR DE RECLAMEVERWERKINGSTHEORIE

De vraag komt ongetwijfeld in u op: welke theoretische implicaties heeft het bovenstaande? Hoe beïnvloedt involvement nu precies de manier waarop een reclameboodschap wordt verwerkt?

We hebben natuurlijk de experimenten van Krugman (1965, 1967). Hij hield zich bezig met het effekt van media op involement en vond daarbij dat t.v.-reclame niet erg involverend was. Krugman was in feite een van de eerste onderzoekers die op het idee kwam van een 'low involvement' model. In dit model werd gesignaleerd dat gedrukte media een actief publiek vereisen om de boodschap over te laten komen – maar indien de consument dit wenst hem ook beter in staat stellen informatie juist en grondig te verwerken (Van Raaij 1979). Televisie daarentegen is in staat om een passief publiek te beïnvloeden. De selectieve verdedigingsmechanismen van de passieve t.v. kijker zijn zwak of werken niet vanwege zijn weinige interesse voor triviale produkten. Na een frequente herhaling zal de boodschap toch doorsijpelen, de cognitieve structuur van de kijker veranderen en zo onbewust het produkt in zijn geheugen positioneren. Tot zo ver heeft er geen attitudeformatie of -verandering plaatsgevonden. In een koopsituatie kan de consument het betreffende produkt selecteren op grond van zijn veranderde cognitieve structuur (geheugen) zonder dat hij een of andere vorm van affect ten aanzien van het produkt heeft. Krugman veronderstelde dat affectie pas wordt gevormd nadat de (proef)aankoop heeft plaatsgevonden en het produkt is gebruikt. Dit houdt in dat consumenten deze lage-involvementprodukten niet kopen omdat ze deze fijn vinden, maar dat ze deze produkten fijn vinden omdat ze ze kopen. Deze cognitie → conatie → affect volgorde – ook wel de leer-doe-voel volgorde genoemd – wordt empirisch ondersteund door de onderzoeksresultaten van Ray (1973), Rothschild (1974) en vele anderen. Er werd bovendien aangetoond dat deze lage-involvementvolgorde kan optreden bij lagere reclamefrequenties. Krugman's model is inmiddels het meest bekende lage-involvementmodel geworden en is verwerkt in een aantal recentere theorieën (zie onder meer figuur 3 en figuur 5) zoals die van Vaughn (1980) en Ray (1973).

Smith en Swinyard (1982) hebben een poging gedaan de traditionele leerhiërarchie en de lage-involvementhiërarchie te integreren in een cognitie → trial → commitment model. Consumenten die hoewel ze weinig geïnvolveerd zijn hun onzekerheid willen reduceren, kunnen een proefaankoop doen – in het bijzonder wanneer het gaat om goedkope produkten – om zo tegen lage kosten op eenvoudige wijze een maximum aan informatie te verkrijgen. Dit is een alternatieve manier om de leer-doe-voel volgorde in het geval van lage-involvementprodukten, die over het algemeen goedkoop zijn, te verklaren, waarbij de (proef)aankoop wordt beschouwd als een manier om informatie te verzamelen. Smith en Swinyard veronderstellen dat proefaankoop voorafgaat aan affect, terwijl een 'committed' (betrokken, geïnvolveerde) aankoop het gevolg is van affect.

Als in een lage-involvementsituatie mensen overgaan tot proefaankoop, is die proefaankoop dan, samen met het proefgebruik, inderdaad de belangrijkste informatiebron zoals Robertson (1976) stelt? Tesamen met de resultaten van Rothschild en Ray (1974) dat reclame sterkere effekten heeft in lage-involvementsituaties, kan dit interessante

complicaties hebben. Zelfs als reclame de aanzet is geweest tot de proefaankoop, kan de ervaring met het produkt na deze aankoop vervolgens alle reclame-effekten overschaduwen (Zaichowsky 1986).

Recent zijn er een aantal uitgewerkte involvementtheorieën verschenen zoals die van Petty en Cacioppo (1983, 1986), Greenwald en Leavitt (1984), die beide in het bovenstaande aan de orde zijn geweest, en Laurent en Kapferer (1985). Deze kunnen tot beter begrip leiden over de omstandigheden waaronder sommige reclameboodschappen effektiever zijn dan andere.

Laurent en Kapferer zien involvement als een pot-pourri concept, een combinatie van een aantal bestaande visies op involvement. De aard van involvement moet volgens hun bestudeerd worden in termen van involvementprofielen, die zowel het niveau als de aard van het concept voortdurend laten veranderen. Zij definiëren involvement als een hypothetisch construct met constante en situationele antecedenten. Deze antecedenten, ook wel facetten genoemd, zijn

- de waargenomen produktbelangrijkheid;
- het waargenomen (functionele) risico dat verbonden is aan de produktaankoop, waaraan twee facetten worden onderscheiden:
  1. de waargenomen belangrijkheid van de negatieve consequenties van een eventuele miskoop, en
  2. de waargenomen kans op het nemen van een dergelijke verkeerde koopbeslissing;
- de symbolische of teken-waarde (het psychologische risico) dat geassocieerd wordt met produkt, de aanschaf of de consumptie ervan;
- de hedonische waarde, het 'emotional appeal' van het produkt.

Hoewel deze facetten zelfstandige, zij het verwante dimensies zijn, kan involvement alleen worden beschouwd als een combinatie van deze vijf facetten. Daarom heeft involvement behalve emotionele en funktionele ook constante en situationele eigenschappen. Volgens Kapferer en Laurent (1986) was de involvementtheorie te veel gesimplificeerd en is kennis van de antecedenten van involvement noodzakelijk om de consequenties ervan te kunnen overzien en te voorspellen. In hun experiment uit 1986 maken hoge- en lage-involvementsituaties maar 25% uit van het totaal aantal koopsituaties, de resterende 75% wordt getypeerd door andere, genuanceerde vormen van involvement. Op deze manier zal het mogelijk zijn meer exact te omschrijven welk involvementaspekt in feite wordt gemanipuleerd, waarbij we het kunnen hebben over het involvementniveau, maar ook over 'cognitieve' en 'affektieve' involvement (Park en Young 1984). Involvement wordt dan een attitude.

Is het nu mogelijk een vergelijking te maken tussen zo'n involvementprofiel en het Foote-Cone-Belding matrixmodel van Vaughn (1980)? Volgens Laurent en Kapferer (1985) is het moeilijk vast te stellen welke facetten of combinatie ervan de involvement uit het FCB model omvat, wat een bepaling van de analogie tussen beide ernstig bemoeilijkt.

We kunnen concluderen dat het nog niet duidelijk is welke consequenties de recente ontwikkelingen in de involvementtheorie precies zullen hebben voor de cognitieve, informatieverwerkingsbenadering van de reclameresponstheorie. De belangrijkste vraag

blijkt echter niet te zijn welke mate van involvement, maar welke soort. Laurent en Kapferer (1985) en de andere genoemde recente theorieën zijn veelbelovend vanwege hun onderzoek naar de componenten van involvement.

Nauw verwant aan het bovenstaande is het onderzoek op het gebied van de hersenspecialisatie, ook wel hersenlateralisatie genoemd. Hypothesen die zijn gebaseerd op onderzoek op dit gebied zijn het laatste decennium vrij populair geweest. Men gaat er daarbij van uit dat de rechter hemisfeer (hersenhelft) onbewust en nonverbaal is en in sterkere mate onder invloed staat van onze emoties dan de linker hemisfeer die gespecialiseerd is in verbale en analytische taken (Lindsay en Norman 1977; Wittrock 1977). Wat betreft de discussie over hoge en lage involvement zou dit mogelijk inhouden dat lage-involvement verwerking in de rechter hersenhelft plaatsvindt, terwijl hoge-involvement verwerking (de centrale route) loopt middels processen in de linker hemisfeer (Krugman 1977; Hansen 1980). Hoewel split-brain onderzoek het voorkomen van cognitieve (linker hersenhelft) en affektieve/emotionele (rechter hemisfeer) stijlen bevestigt, weten we ook dat onze hersenen in feite één systeem vormen (Levy 1981) dat in staat is complexe stimuli te integreren en zowel met informatie als emotie om te gaan (Vaughn 1986). Er zijn wel een aantal verbanden te leggen tussen de hersenspecialisatietheorie en het Elaboration Likelihood Model (Petty en Cacioppo 1986). Aangezien het grootste deel van de discussie over hersenspecialisatie geen empirische grondslag heeft en we evenmin iets weten van de fysiologische basis van de twee routes van het ELM, is het moeilijk aan te geven in welke richting onderzoek zou moeten gaan. Maar omdat onderzoek wel heeft aangetoond dat de linker en rechter hersenhelften volledig verschillen qua manieren van en mogelijkheden tot het verwerken van informatie, kunnen deze verschillen mogelijk verklaren waarom een deel van de verwerkte informatie beter bewaard blijft dan informatie die verwerkt is via de andere route.
De suggesties van Krugman (1965, 1977) dat t.v.-reclame minder involverend is dan gedrukte reclame vanwege hersenspecialisatie effekten moeten we verwerpen. Weinstein, Appel en Weinstein's (1980) onderzoek heeft aangetoond dat televisie niet noodzakelijk een minder involverend medium is: het involvementniveau van het publiek hangt in belangrijke mate af van de inhoud van de reclameboodschap.

7. CONCLUSIES

In dit artikel is getracht een overzicht te geven van de ontwikkeling van reclameresponsmodellen vanuit een meer cognitief, informatieverwerkingsstandpunt en van de zeer grote hoeveelheid literatuur die daarop betrekking heeft. Dit houdt vanzelf in dat andere onderzoeksbenaderingen minder aandacht hebben gekregen.
In dit overzicht ligt de nadruk op de sequentiële modellen. Aangezien consumentenvariabelen, waaronder involvement, en omgevingsvariabelen bepalen welke fasen achtereenvolgens worden doorlopen, is tevens geprobeerd de 'state of the art' weer te geven van de involvementtheorie en aan te geven wat de relevantie van deze theorie is voor de reclameresponsmodellen.
Aangezien een aantal onderzoeksvelden sterk in beweging zijn en bovendien zeer veel omvattend, is het noodzakelijk gebleken op sommige onderwerpen meer de nadruk te leggen dan op andere. Hierbij is de relevantie voor het belangrijkste onderwerp, de reclameresponsmodellen, de voornaamste gids geweest.

Op dit moment is het nog niet precies te zeggen welk effekt de recente trends in de involvementtheorie en het involvementonderzoek voor de toekomstige ontwikkelingen van reclameresponsmodellen (en consumentengedragsmodellen) zullen hebben. Daarom zijn enkele suggesties voor te verrichten onderzoek gedaan en een aantal mogelijke verbanden met andere, verwante theorieën aangegeven.

Wat opvalt aan de beschreven reclameresponsmodellen is dat deze ophouden als het 'doel' (de actie, adoptie of aankoop, afhankelijk van het model) bereikt is. Naar mijn mening moet er echter nog een fase komen, waarin reclame een zeer belangrijke rol kan spelen, namelijk het tevredenheids- of satisfactiestadium.
In dit stadium kan reclame een aantal uiteenlopende funkties vervullen. In de eerste plaats het onder de aandacht van de consument brengen van de produkteigenschappen; deze funktie wordt vaak over het hoofd gezien. Tweedens: als de consumenten die het produkt gekocht hebben niet meer voor dit gedrag beloond worden, vindt er inhibitie plaats. Dat houdt in dat zij het produkt niet meer zullen kopen.
Op de derde plaats heeft reclame in dit laatste stadium de zeer belangrijke funktie van het verminderen van de eventuele cognitieve dissonantie.
Vanuit een marketing standpunt is het erg vreemd dat deze mogelijkheid om de tevredenheid van de klant te doen toenemen, in alle reclameresponsmodellen met uitzondering van het PAR model van Van Raaij, wordt genegeerd.

Het involvementconcept is erg belangrijk gebleken om nuanceringen aan te brengen ten aanzien van de omstandigheden waaronder bepaalde reclameresponsmodellen van toepassing zijn. Aan het involvement concept zijn nogal wat facetten te onderscheiden. Een van de grootste problemen is geweest het beschrijven van deze relevante dimensies en hun belang voor de reclame.
Om duidelijkheid te scheppen in de veelheid van meningen over wat involvement nu precies is, lijkt de formulering van Antil (1984) goed aan te geven waar het om gaat: '...the level of perceived personal importance and/or interest evoked by a stimulus (or stimuli) within a specific situation'. Ook volgens Antil zijn er dus drie groepen factoren die de omstandigheden waaronder een reclameuiting wordt waargenomen beïnvloeden en op hun beurt weer verantwoordelijk zijn voor het soort involvement.
Dit is de reden dat een aantal auteurs voorstelt het involvementconcept te vervangen door de begrippen persoonlijke, situationele en objectrelevantie, terwijl anderen (Poiesz en Pieters 1987) voorstellen een beter onderscheid te maken tussen de aan elkaar grenzende begrippen produktbelangrijkheid, situationele relevantie en involvement.

Recent zijn er een aantal uitgewerkte involvementtheorieën verschenen zoals die van Petty en Cacioppo (1983, 1986), Greenwald en Leavitt (1984)) en Laurent en Kapferer (1985) die tot beter begrip kunnen leiden over de omstandigheden waaronder sommige reclameboodschappen effektiever zijn dan andere.
Helaas moeten we ook tot de conclusie komen dat Petty en Cacioppo onvoldoende aangeven welke situaties in het bijzonder zullen leiden tot hoge en welke tot lage involvement.
Laurent en Kapferer (1985; Kapferer en Laurent 1986) hebben een zeer verdienstelijke poging gedaan een involvementconcept te ontwikkelen door de bestaande zienswijzen met betrekking tot het concept te combineren.

Ten aanzien van het involvement concept zijn er ook een aantal andere dingen die we nog niet precies weten, bijvoorbeeld omdat verschillende onderzoeken verschillende resultaten laten zien[9]. Voordat deze bij kunnen dragen aan de verdere ontwikkeling van reclameresponsmodellen, zullen deze onderzoeken eerst herhaald moeten worden en waar mogelijk uitgebreid naar andere omstandigheden en produktcategorieën.

Een groot probleem blijft de moeilijke operationalisatie van het involvement concept; een aantal oplossingen hiervoor zijn wel aangedragen (b.v. Antil 1984), maar blijven toch onbevredigend.
Dan nog een laatste opmerking. Hoewel er een tendens bestaat in de reclame steeds meer de nadruk op het affect te leggen door middel van het oproepen van emoties, zou men rekening houdend met de verschillen in involvement tussen verschillende groepen consumenten bij de hogere involvementniveaus volgens de schrijver van dit artikel daarentegen eerder de nadruk op de cognitieve aspekten moeten leggen.

NOTEN

1. De auteur is K.E. Kristian Möller van de Helsinki School of Economics, Jan C. Reuijl van de Rijksuniversiteit Groningen en W. Fred Van Raaij van de Erasmus Universiteit Rotterdam zeer erkentelijk voor hun waardevolle opmerkingen met betrekking tot eerdere versies van dit artikel.
2. Daar dit concept bekender is onder de Engelse naam dan onder de Nederlandse, wordt in dit artikel de voorkeur gegeven aan het gebruik van de Engelse term.
3. Cues zijn letterlijk vingerwijzingen, aanwijzingen. Bedoeld worden in dit verband de eigenschappen van de reclameboodschap in de meest ruime zin – ook wel 'tekens' genoemd – die de ontvanger gebruikt bij het verwerken van de boodschap. Deze tekens kunnen zowel verbaal als nietverbaal zijn en impliciet deel uitmaken van de inhoud van de boodschap of (expliciet) te maken hebben met de omstandigheden waaronder de boodschap wordt gecommuniceerd (b.v. de aantrekkelijkheid, betrouwbaarheid, onafhankelijkheid of expertise van degene die de boodschap overbrengt; denk hierbij aan de commotie rond het optreden van Koos Postema voor de Postbank).
4. Op geaggregeerd niveau is wel degelijk uitgebreid onderzoek gedaan naar de effektiviteit van reclame, naar de invloed van reclame (in het bijzonder van reclamebestedingen) op de verkoop. Zie hiervoor onder meer Reuijl (1982) en Leeflang en Reuijl (1985).
5. Het gebruik van de term 'leren' als men cognitie bedoelt, kan hier misverstanden oproepen. Naast cognitief leren (informatieverwerking) onderscheiden we namelijk associatief leren (klassiek en operant conditioneren). Dit kan gemakkelijk tot verwarring leiden. Omdat echter in een groot aantal artikelen de beide omschrijvingen cognitief-affectief-conatief en leer-voel-doe naast elkaar worden gebruikt zullen wij dat hier ook doen.
6. Zeldenrust schreef in 1984 ook zelf een artikel over de werking van reclame. Dit artikel, noch het daarin geponeerd reclamemodel, hebben sindsdien veel aandacht gekregen. Zie ook Roomer (1987) en Nillesen (1988).
7. De EDR (electrodermale respons) is een verandering van de electrische huidweerstand als gevolg van opwinding (arousal). Deze verandering wordt gemeten door middel van electroden aan de binnenzijde van de handen van de proefpersoon. De EDR registreert de activiteit van het perifere zenuwstelsel.
8. Zie noot 3.
9. De lezer die ondanks de niet-eenduidigheid van de onderzoeksresultaten op dit gebied, hier meer over wil lezen, zij verwezen naar Zaichowsky (1985, 1986). Met name het laatste artikel is zeer de moeite waard.

LITERATUUR

Antil, J.H. (1984), 'Conceptualization and operationalization of involvement'. *Advances in consumer research* X, 203-209.
Assael, Henry (1981), *Consumer behavior and marketing action*, New York: Kent.

Assael, Henry and George S. Day (1968), 'Attitudes and Awareness as Predictors of Marketing Share', *Journal of Advertising Research,* 8 (October), 3-10.
Batra, Rajeev and Michael L. Ray (1983), 'Operationalization involvement as depth and quality of cognitive response', In: *Advances in consumer research,* (Ed.) Bagozzi, Richard P. and Alice M. Tybout, Ann Arbor: Association for Consumer Research, 309-313.
Bauer, R. and S. Greyser (1968), *Advertising in America: the consumer view.* Boston: Harvard University.
Belk, Russel W. (1975), 'Situational variables and consumer behavior', *Journal of Consumer Research,* 2 (December), 157-164.
Belk, Russel W. (1982) 'Effects of Gift-Giving Involvement on Gift Selection Strategies, In: *Advances in Consumer Research,* (Ed.) Mitchell, Andrew A., Ann Arbor: Association for Consumer Research, 9, 408-411.
Bem, D.J. (1972), 'Self-Perception Theory', In: *Advances in Experimental Social Psychology,* 6, (Ed.) Berkowitz, L., New York: Academic Press, 1-62.
Berger, David (1986), 'Theory into practice: the FCB grid', *European research,* 35-46.
Bloch, Peter H. and Marsha L. Richins (1983), 'A Theoretical Model for the Study of Product Importance Perceptions', *Journal of Marketing,* 47 (Summer), 69-81.
Broadbent, Donald E. (1958), *Perception and communication,* New York: Pergamon Press.
Broadbent, Donald E. (1977). The hidden preattentive processes, *American Psychologist* (February), 109-118.
Burnkrant, Robert E. and Alan G. Sawyer (1983), 'Effects of Involvement and Message Content on Information-Processing Intensity', In: *Information Processing Research in Advertising,* (Ed.) Harris, Richard J., Hillsdale, NJ: Lawrence Erlbaum.
Butter, C.M. (1968), *Neuropsychology: the study of brain and behavior.* Belmont CA: Brooks/Cole.
Cacioppo, John T. and Richard E. Petty (1985), 'Central and peripheral routes to persuasion: the role of message repetition', In: (Eds.) Alwitt, Linda F. and Andrew A. Mitchell, *Psychological processes and advertising effects. Theory, research and applications,* Hillsdale NJ: Lawrence Erlbaum.
Clarke, Keith and Russel Belk (1978), 'The Effects of Product Involvement and Task Definition on Anticipated Consumer Effort', In: *Advances in Consumer Research,* (Ed.) Hunt, H. Keith, Ann Arbor: Association for Consumer Research, 5, 313-318.
Colley, R. (1961), *Defining Advertising Goals for Measured Advertising Results,* New York: Association of National Advertisers.
Cowling, A.B. (1984), 'How advertising works – what a damn fool question', *Admap* (June), 288-293.
Craik, Fergus I.M. and Robert S. Lockhart (1972), 'Levels of processing: a framework for memory research', *Journal of verbal learning and verban behavior* 11, 671-684.
Day, George S. (1970), *Buyer attitudes and brand choice,* New York: Free Press.
Engel, J.F., R.D. Blackwell and P.W. Miniard (1986), *Consumer behavior,* Chicago: the Dryden Press.
Festiger, L. (1957), *A theory of cognitive dissonance,* Stanford: Stanford University Press.
Greenwald, A.G. and C. Leavitt (1984), 'Audience involvement in advertising: four levels'. *Journal of Consumer Research,* 11 (June) 581-592.
Hansen, Fleming (1981), 'Hemispheral Lateralization: Implications for Understanding Consumer Behavior'. *Journal of Consumer Research,* 8 (June), 23-26.
Hansen, Fleming (1984), 'Towards an alternative theory of the advertising communication process?', *International Journal of Consumer Research in Marketing,* 1, 69-80.
Holbrook, Morris B., and Johan O'Shaughnessy (1984), 'The Role of Emotion in Advertising', *Psychology & Marketing* 1, 2 (Summer), 45-64.
Houston, Michael J. Michael L. Rothschild (1978), 'Conceptual and Methodological Perspectives in Involvement', In: *1978 Educator's Proceedings,* (Ed.) Jain, S.C., Chicago: American Marketing Association, 184-187.

Howard, John A. and Jagdisch N. Sheth (1969), *The Theory of Buyer Behavior,* New York: John Wiley.
Hupfer, Nancy T. and David M. Gardner (1971), 'Differential Involvement with Products and Issues: An Exploratory Study', In: *Proceedings Association for Consumer Research,* ed. David M. Gardner, College Park, MD: Association for Consumer Research.
Insko, Chester A. and John Schopler (1967), 'Triadic Consistency: A Statement of Affective-Cognative-Conative Consistency', *Psychological Review,* 74 (5), 361-376.
Kapferer, J.N. and G. Laurent (1985), 'Consumer involvement profiles: a new practical approach to consumer involvement', *Journal of Advertising Research,* 25, 6, 48-56.
Kassarjian, Harold H. (1981), 'Low Involvement: A Second Look', In: *Advances in Consumer Research,* V. 7, (Ed.) Monroe, Kent B., Ann Arbor: Association for Consumer Research, 31-34.
Katz, William A. (1983), 'Point of view: A critique of split-brain theory', *Journal of Advertising Research,* 23, 2, 63-66.
Krugman, Herbert E. (1965), 'The Impact of Television Advertising: Learning Without Involvement', *Public Opinion Quarterly,* 29 (3), 349-356.
Krugman, Herbert E. (1967), 'The Measurement of Advertising Involvement', *Public Opinion Quarterly,* 30, 583-596.
Krugman, Herbert E. (1977), 'Memory without Recall, Exposure without Perception', *Journal of Advertising Research* 17, 4 (August), 7-12.
Krugman, Herbert E., (1980), 'Point of View: Sustained Viewing of Television', *Journal of Advertising Research* 20, 3, 65-68.
Lannon, Judie and Peter Cooper (1983), 'Humanistic Advertising: A Holistic Cultural Perspective', *International Journal of Advertising* 2, 195-213.
Lastovicka, John L. and David M. Gardner (1978), 'Low Involvement versus High Involvement Cognitive Structures', In: *Advances in Consumer Research,* V. 5, (Ed.) Hunt, Keith H., Ann Arbor: Association for Consumer Research, 87-92.
Laurent, Gilles and Kapferer, Jean-Noël (1985), 'Measuring Consumer Involvement Profiles', *Journal of Marketing Research* 22 (February), 41-53.
Lavidge, R., and G.A. Steiner (1961), 'A Model For Predictive Measurements of Advertising Effectiveness', *Journal of Marketing* 25, 59-62.
Lazarus, Richard S. (1984), 'On the primacy of cognition', *American Psychologist,* (February), 124-129.
Leeflang, Peter S.H. and Jan C. Reuijl (1985), 'Advertising and industry sales: an empirical study of the West German cigarette market', *Journal of marketing* 49 (fall), 92-98.
Levy, Jerre (1981), '*Children Think With Whole Brains: The myth and Reality of Hemisphere Difference and Inter-hemispheric Integration*'. Paper presented at the conference of The National Association of Secondary School Principles.
Lindsay, Peter H. and Donald A. Norman (1977), *Human Information Processing,* New York: Academic Press.
Mitchell, Andrew A. (1979), 'Involvement: a potentially important mediator of consumer behavior', In: *Advances in consumer research,* (Ed.) Wilkie, W.L., Ann Arbor: Association for Consumer Research, 191-196.
Mitchell, Andrew A. (1981), 'The Dimensions of Advertising Involvement', In: *Advances in Consumer Research,* V. 8, (Ed.) Monroe. Kent B., Ann Arbor: Association for Consumer Research, 25-30.
Mitchell, Andrew A. (1983), 'Cognitive processes initiated by exposure to advertising', In: *Information processing research in advertising,* (Ed.) Harris, Richard J., Hillsdale, NJ: Lawrence Erlbaum.
Moriarty, Sandra E. (1983), 'Beyond the Hierarchy of Effects: A Conceptual Model', In: *Current Issues and Research in Advertising,* (Eds.) Leigh, James H. and Claude R. Martin Jr., Ann Arbor, MI.: Graduate School of Business Administration, 45-56.
Moriarty, Sandra E. (1987), *Creative advertising: Theory and Practice,* Englewood Cliffs, NJ.: Prentice-Hall.

Mühlbacher, Hans (1986), 'An involvement model of advertising information acquisition and processing motivation', In: *Proceedings of the XVth Annual EMAC Conference,* Helsinki, Finland, 461-480.
Neisser, Ulric (1967), *Cognitive Psychology,* New York: Appleton-Century-Crofts.
Nicosia, F.M. (1966), *Consumer decision processes,* Englewood Cliffs, NJ.: Prentice-Hall.
Nillesen, J.P.H. (1985), *Het merkbegrip in de reclame,* scriptie vakgroep Psychologische funktieleer, Katholieke Universiteit Nijmegen.
Nillesen, J.P.H. (1988), Boekbespreking van J. Roomer 'De praktijk van de marketing-communicatie. Achtergronden van concepten, teksten en ontwerpen', *Maandblad voor accountancy en bedrijfshuishoudkunde* 62, 3 (maart), 104-105.
Olshavsky, R.W. and D.H. Granbois (1979), 'Consumer decision making – fact or fiction?' *Journal of Consumer Research* 6, 93-100.
Packard, Vance (1958), *The hidden persuaders,* New York.
Palda, Kristian S. (1966), 'The Hypothesis of a Hierarchy of Effects: A Partial Evaluation', *Journal of Marketing Research,* 3 (February), 13-24.
Park, C. Whan and S. Mark Young (1983), 'Types and Levels of Involvement and Brand Attitude Formation'. In: *Advances in Consumer Research,* V. 10, (Eds.) Bagozzi, Richard P. and Alice M. Tybout, Ann Arbor: Association for Consumer Research, 320-324.
Petty, Richard E. and John T. Cacioppo (1979), 'Issue Involvement Can Increase or Decrease Persuasion by Enhancing Message-Relevant Cognitive Responses', *Journal of Personality and Social Psychology,* 37 (October), 1915-1926.
Petty, Richard E. and John T. Cacioppo (1981), *Attitudes and Persuasion: Classic and contemporary Approaches,* Dubuque, IA: Wm.C. Brown.
Petty, Richard E. and John T. Cacioppo (1983), 'Central and peripheral routes to persuasion: application to advertising' In: (Eds.) Percy, L. and A.G. Woodside, *Advertising and Consumer psychology,* Lexington (Mass.): Lexington Books.
Petty, Richard E. and John T. Cacioppo (1986), 'The elaboration likelihood model of persuasion', In: (Ed.) Berkowitz, L., *Advances in experimental social psychology,* Vol. 19., New York: Academic Press.
Petty, Richard E., John T. Cacioppo and David Schumann (1983), 'Central and peripheral routes to advertising effectiveness: The moderating role of involvement', *Journal of Consumer Research* 10 (September), 135-146.
Pieters, Rik G.M. and W. Fred Van Raaij (1988, in druk), 'The role of affect in economic behavior', In: *Handbook of Economic Psychology,* (Eds.) Raaij, W. Fred van, Gery M. van Veldhoven, Theo M.M. Verhallen and Karl-Erik Wärneryd, Deventer: Kluwer.
Poiesz, Theo and Rik Pieters (1987), 'Consumer involvement: Theoretical illusion of practical fact?', In: *Contemporary research in marketing,* Proceedings of the XVIth Annual EMAC Conference, Toronto (1-4 June).
Preston, Ivan L. (1982), 'The Association Model of the Advertising Communication Process', *Journal of Advertising* 11, 2, 3-15.
Preston, Ivan L., and Esther Thorson (1984), 'The Expanded Association Model: Keeping the Hierarchy Concept Alive'. *Journal of Advertising Research* 24, 1, 59-65.
Raaij, W. Fred van (1979), 'Consumenteninformatie: hoeveel en in welke vorm?', *Tijdschrift voor marketing* 13(5), 10-16.
Raaij, W. Fred van (1984), *Affective and cognitive relations to advertising.* Marketing Science Institute, Cambridge, MA, Report 84-111.
Raaij, W. Fred van (1985), 'Een nieuw model van de werking van reclame', *Markeur* 1, 6-11.
Raaij, W. Fred van (1986), *'The affective value of brand name and campaign style in advertising',* Proceedings of the International Conference on Economics and Psychology, Kibbutz Shefayim, Israel (9-11 July).
Rajaniemi, Pirjo and Martti Laaksonen (1986), 'The role of profile analyses in the conceptual development of involvement', In: *Proceedings of the XVth Annual EMAC Conference,* Helsinki (3-6 June).

Ray, Michael L. (1973), 'Marketing Communications and the Hierarchy-of-Effects', In: *New Models for Mass Communication Research*, (Ed.) Clarke, P., Beverly Hills, CA: Sage Publishing, 147-176.
Ray, Michael K. (1979), 'Involvement and other variables mediating communication effects as opposed to explaining all consumer behavior', In: *Advances in consumer research*, (Ed.) Wilkie, W.L., Ann Arbor: Association of consumer research, 197-199.
Ray, William J. and Jerry C. Olson (1982), *'Perspectives on psychophysiological assessment of psychological responses to advertising'*, ESOMAR congres, Vienna.
Reuijl, Jan C. (1982), *On the determination of advertising effectiveness. An empirical study of the German cigarette market.* Dissertatie Rijksuniversiteit Groningen. Leiden/Antwerpen: Stenfert Kroese.
Robertson, Thomas S. (1971), *Innovative behavior and communication*, New York: Holt, Rinehart and Winston.
Rossiter, John R. and Larry Percy (1987), *Advertising and promotion management*, New York: McGraw Hill Inc.
Rothschild, M.L. (1974), *'The Effects of Political Advertising on the Voting Behavior of a Low-Involvement Electorate'*, Ph. D. dissertation, Stanford University, Graduate School of Business.
Rothschild, M.L. (1978), 'Advertising Strategies for High Involvement and Low Involvement Situations', In: *Attitude Research Plays for High Stakes*, (Ed.) Maloney, John C., Chicago: American Marketing Association.
Roomer, Joop (1987), *De praktijk van de marketing-communicatie. Achtergronden van concepten, teksten en ontwerpen*, Deventer/Antwerpen: Kluwer.
Rothschild, M.L. and M.L. Ray (1974), 'Involvement and Political Advertising Effect: An Exploratory Experiment', *Communication Research*, 1 (July), 264-286.
Sherif, Muzafer and Hadlai Cantril (1947), *The Psychology of Ego-Involvement*, New York: Wiley.
Slama, Mark E. and Armen Tashchian (1985), 'Selected socioeconomic and demographic characteristics associated with purchasing involvement', *Journal of Marketing* 49 (Winter), 72-82.
Smith, Robert E. and William R. Swinyard (1982), 'Information Response Models: An Integrated Approach', *Journal of Marketing* 46, (Winter), 81-93.
Solomon, Michael R. (1986), 'The missing link: surrogate consumers in the marketing chain', *Journal of Marketing* 50, 4, 208-218.
Spijker, J. (1986), *De elektrodermale respons als indicator voor primaire affectieve reacties*, doctoraalscriptie vakgroep Economische Psychologie, Katholieke Universiteit Brabant.
Traylor, M.B. (1981), 'Product involvement and brand commitment', *Journal of Advertising Research*, 21, 6 (December), 51-56.
Tyebjee, Tyzoon T. (1979). 'Refinement of the Involvement Concept: An Advertising Planning Point of View', In: *Attitude Research Plays for High Stakes*, (Eds.) Maloney, John C. and Bernard Silverman, Chicago: American Marketing Association, 94-111.
Vaughn, Richard (1980), 'How Advertising Works: A Planning Model', *Journal of Advertising Research* 20, 5, 27-33.
Vaughn, Richard (1986), 'How Advertising Works: A Planning Model Revisited', *Journal of Advertising Research* 20, (Febr./March), 57-66.
Veldhoven, G.M. van (1979), 'Wetenschappelijk onderzoek m.b.t. consumentengedrag', *Tijdschrift voor Marketing* (Mei), 3-8.
Wärneryd, Karl-Erik and Kjell Nowak (1967), *Mass communication and advertising*, Stockholm: The Economic Research Institute at the Stockholm School of Economics.
Weinstein, Sidney, Valentine Appel and Curt Weinstein (1980), 'Brain Activity Responses to Magazine and Television Advertising', *Journal of Advertising Research*, 20 (June), 57-63.
Westendorp, P.H. van (1970), *Reclame, consument en psychologie*, *Pre*-advies VEA, Amsterdam.
Wright, Peter (1973), 'The Cognitive Processes Mediating Acceptance of Advertising', *Journal of Marketing Research*, 10, (February), 53-62.
Wittrock, Merlin C. (1977), *The human brain*, Englewood Cliffs, NJ: Prentice-Hall.

Zaichkowsky, Judith L. (1985), 'Measuring the involvement construct', *Journal of Consumer Research* 12, (December), 341-352.
Zaichkowsky, Judith L. (1986), 'Conceptualizing involvement', *Journal of Advertising* 15, 2, 4-15.
Zajonc, R.B. (1984), 'On the primacy of affect', *American Psychologist* 39, 2, 117-123.
Zajonc, R.B., and H. Markus (1982), 'Affective and Cognitive Factors in Preferences', *Journal of Consumer Research* 9, 123-131.
Zeldenrust, D. (1984), 'De werking van reclame', *Massacommunicatie* 12, 5/6, 197-203.
Zeldenrust, D. (1986), 'Noem Van Raaij's affectieve reactie liever oriëntatieresponse', *Adformatie* 28/29 (17-7), p. 30.

12. Modellen ter verklaring van het industriële koopproces

M.J. BRAND

SAMENVATTING

De aandacht voor industriële marketing en tevens het gebruik van modellen hierbij blijft groeien. In dit artikel wordt een overzicht gegeven van een bijzondere groep modellen binnen de industriële marketing, de industriële inkoopmodellen. Na enkele inleidende paragrafen over modellen, industriële marketing en inkoopmodellen in het bijzonder worden enkele groepen inkoopmodellen besproken. In paragraaf 4 zijn dat de beslissingsprocesmodellen, in paragraaf 5 de algemene modellen en in paragraaf 6 de interactiemodellen. Vervolgens wordt in paragraaf 7 een totaalmodel gepresenteerd waarna een aantal submodellen nader besproken wordt (paragraaf 8).

1. INLEIDING

1.1. Inleiding[1]

In dit artikel wordt een aantal modellen besproken die allen een verklaring trachten te geven van (een deel van) het industriële koopproces. De laatste jaren is gebleken dat onderzoek naar industriële markten o.a. inzicht kan geven in:
- inkoopgedrag;
- vertegenwoordigersmotivatie;
- produktontwikkeling;
- invloed van de technologie;
- marketing mix beslissing, en
- relatie beheer ('management').

De toegenomen aandacht voor industrieel marketing onderzoek werpt zichtbaar vrucht af. Over de noodzaak tot industrieel onderzoek in tegenstelling tot consumentenonderzoek wordt in de literatuur echter verschillend gedacht. In paragraaf 2 zal dit uitgebreid besproken worden.
Het industriële *koopproces* neemt binnen de bestaande industriële marketing literatuur een belangrijke plaats in. Gedurende de afgelopen 20 jaar is een veelheid aan inkoopmodellen ontwikkeld en dat is niet verwonderlijk daar koopprocessen zowel voor aanbiedende als vragende ondernemingen van wezenlijk belang zijn. Voor de *producent* van industriële produkten vormen de karakteristieken van de potentiële kopers en van de markt (aantal vragers en aanbieders, aantal internationale transacties) het kader waarbinnen de afzet gerealiseerd moet worden. Het afstemmen van korte en lange termijn marketingstrategiën op dit kader is een voorwaarde voor bevredigende resultaten voor de aanbiedende onderneming. Ook voor de *afnemer* is het echter zeer belangrijk

inzicht te krijgen in de eigen inkoopactiviteiten. De uitgaven van ondernemingen aan inkoop van industriële goederen en diensten beslaan *gemiddeld* zo'n 50%[2] van de omzet. Een ander punt dat steeds belangrijker wordt is het feit dat toelevering en produktie direct op elkaar aan moeten sluiten. Dit hangt samen met de invoering van zgn. 'integrated logistic systems' als M.R.P. en J.I.T.. Tenslotte is ook het kwaliteitsaspect belangrijk voor de afnemer. Hoe kan men een kwaliteif goed eindprodukt maken als de ingekochte grondstoffen en/of halffabrikaten niet deugen?

1.2. Waarom modellen?

Onderzoek binnen het industriële koopproces heeft zich tot nu toe voornamelijk geconcentreerd op de volgende onderwerpen:
1. stappen in het aankoopproces;
2. het belang van leverancierskenmerken en van de marketing mix elementen;
3. de samenstelling en invloed van koopcentra[3];
4. beschrijving van een bepaald inkoopproces.

Om een compleet beeld te krijgen van het inkoopproces is het zaak zoveel mogelijk van deze invalshoeken te integreren. Het beoogde resultaat is dan een algemeen model, een 'framework', waar tevens vervolgonderzoek op gebaseerd zal kunnen worden. Bij de pogingen een dergelijk framework te ontwikkelen wil dit artikel ondersteuning geven door een opsomming en beschrijving te geven van de belangrijkste modellen die de afgelopen 20 jaar in de literatuur besproken zijn.
Ter afsluiting van deze eerste paragraaf willen we de nadruk leggen op het nut dat inkoopmodellen kunnen hebben in de praktijk. Naast het verbeteren van (generaliseerbare) kennis kunnen modellen namelijk ook dienen ter ondersteuning van het management-beslissingsproces. Deze 'decision support' kan in 3 punten worden samengevat:
1. hulp bij het ontwikkelen van een marketingstrategie;
2. hulp bij het analyseren van de beschikbare informatie;
3. hulp bij het onderzoeken van de behoefte aan meer informatie.

2. INDUSTRIËLE VERSUS CONSUMENTENMARKETING

2.1. De controverse

Zoals in de inleiding al werd aangegeven verschillen de menigen in de literatuur over het bestaansrecht van de dichotomie industriële versus consumentenmarketing[4]. Met name Fern en Brown (1984) hebben een felle aanval op deze veel gehanteerde indeling gedaan. Alvorens deze verschillende meningen te bespreken, is het nuttig aan te geven wat nu precies onder industriële marketing wordt verstaan en waar dat dan verschilt van consumentenmarketing.

def. 'Industrial (—) marketing is the process of (1) determining the needs (—) of commercial enterprises, governments and institutions, and (2) developing the appropriate products (—) to statisfy those requirements'.
(Hutt en Speh, Industrial marketing management, 1985, p. 4).

De elementen die in deze zienswijze het onderscheid met consumentenmarketing aangeven zijn:
1. organisaties als afnemers;
2. het gebruik van het produkt voor produktie of doorverkoop in plaats van gebruik voor consumptie.

We zullen nu allereerst de argumenten van de voorstanders van de classificatie industriële versus consumentenmarketing bespreken. Daarna schenken we enige aandacht aan de tegenstanders van dit onderscheid en maken we enkele afsluitende opmerkingen.

2.2. *Verschillen en overeenkomsten*

Een veel gehanteerde indeling binnen het probleemgebied marketing is die in consumentenmarketing *versus* industriële marketing. Het criterium hierbij is welke subjecten in een specifiek probleem centraal staan; consumenten of bedrijven. De meeste onderzoeken in de marketing-literatuur beperken zich tot óf de consumentenmarkt óf de bedrijvenmarkt. Ook de meeste leerboeken beginnen met dit onderscheid. Vervolgens be-

Tabel 1. Enkele verschillen tussen consumenten- en industriële marketing.

	industriële marketing	consumenten marketing
1. marketing planning	• geen of globaal plan • lange termijn horizon	• meerdere plannen • korte termijn horizon
2. markt		
a. structuur	• aantal afnemers klein • afgeleide vraag	• aantal afnemers zeer groot • autonome vraag
b. koper	• motieven rationeel • meest mannen in een team die overleggen • vaak onderhandeling	• motieven emotioneel • meest individuele vrouwen die (impulsief) aankopen • weinig onderhandeling
3. marktonderzoek	• veel desk research • meest kwalitatief autoriteitenonderzoek met kleine steekproef • beslisser moeilijk te identificeren en bereiken • weinig externe bureau's	• veel field research • grootschalige enquêtes en ingewikkelde kwantitatieve technieken • beslisser makkelijk te identificeren en te bereiken (huisvrouw) • veel externe bureau's
4. marketing mix		
a. produkt	• invloed klant op ontwikkeling • technisch complex	• geen invloed • eenvoudig
b. distributie	• enkele korte kanalen	• veel lange kanalen
c. prijs	• prijs ondoorzichtig en van verschillend belang	• prijs duidelijk en belangrijk
d. communicatie	• weinig reclamebestedingen, -onderzoek en -bureaus • l.t. doelstellingen • rationele technische info • vakbladen, d.m., beurzen	• veel reclamebestedingen, -onderzoek en -bureaus • k.t. doelstellingen • sfeer merk benadrukt • massamedia

handelen ze beide apart of kiezen een van de twee benaderingen voor de rest van het boek. De argumenten voor deze strikte scheiding worden door de Rijcke en Matthijssens (1982) uitgebreid opgesomd. De belangrijkste elementen uit deze opsomming staan in tabel 1 vermeld. Hierbij wordt een verdeling gemaakt in: marketing planning, marktkenmerken, marktonderzoek en marketing mix.

Het moge duidelijk zijn dat de tegenstellingen in deze tabel nogal ongenuanceerd zijn en door recente ontwikkelingen achterhaald.

- Het beeld van de impulsieve, emotionele, slecht geïnformeerde huisvrouw die verantwoordelijk is voor het merendeel van de consumentenaankopen wordt steeds minder actueel.
- Op televisie en in andere massa media wordt steeds vaker reclame gemaakt voor industriële ondernemingen (image) en/of produkten (denk aan Canon en Overtoom).
  En zo zijn nog veel meer tegenvoorbeelden te geven. Verschillende ontwikkelingen zorgen er blijkbaar voor dat de grezen tussen industriële en consumentenmarketing vervagen. In de visie die we nu gaan bespreken zijn deze grenzen altijd al vaag geweest en zijn er andere meer zinvolle onderscheiden te maken.

De uitspraak dat consumenten- versus industriële marketing geen zinvol onderscheid is wordt uitvoerig onderbouwd door Fern en Brown (1984). Vooral met betrekking tot koopgedrag hebben al verschillende auteurs in het verleden duidelijke *overeenkomsten* beschreven tussen consumenten- en industriële markten. Enkele voorbeelden staan in tabel 2 weergegeven.

Tabel 2. Overeenkomsten tussen consumenten- en industrieël koopgedrag.

onderzoeksgebied	overeenkomsten
– beslissingseenheden	• zowel het koopcentrum als de familie beslist voor belangrijke aankopen gezamenlijk
– participanten rollen	• te onderscheiden rollen: initiator, influencer, decider, buyer, user, plus bij IM de gatekeeper
– procesfasen	• te onderscheiden o.a.: awareness, interest, evaluation, trial, adoption
– modellen	• gedragsmodellen met psychologische details en motivatie • soort aankoop als determinant van oplossingsmethodiek
– invloeden	• sociale normen en verwachtingen spelen een rol • economische omstandigheden hebben een grote invloed

Fern en Brown verwerpen de genoemde dichotomie op grond van verschillende criteria waaronder:

1. Het is niet moeilijk tegenvoorbeelden te geven van genoemde verschillen in tabel 1. (Bijvoorbeeld: de vraag naar typepapier in huishoudens is afgeleid van het gebruik van typemachines).
2. De huidige indelingen van goederen naar 'verbruikt worden bij consumptie', soort aankoop (zie tabel 2) en diensten versus tastbare produkten leiden tot grotere verschillen *binnen* IM resp. CM als *tussen* deze twee.

Na de huidige goederenindelingen verworpen te hebben komen Fern en Brown met een alternatief. Zij stellen voor markten te classificeren op basis van produkteigenschappen, gebruiksgedrag en behoeftebevrediging. Met behulp van deze classificatie en een groot aantal bekende marketing principes formuleren zij een aantal stellingen. Eén van deze stellingen is de volgende:

> 'Als de kennis bij afnemers van een bepaald produkt klein is (dus in het bijzonder bij technisch complexe produkten) moet de communicatie zo direkt mogelijk zijn.'

Die geldigheid van deze stellingen in zowel IM als de CM is een illustratie en tevens een bevestiging van de betoogde generaliseerbaarheid van marketingkennis.
Andere alternatieve indelingen die men in de literatuur kan vinden berusten vaak op technologische verschillen tussen produkten. Een voorbeeld van een onderzoek waar een uitspraak werd gedaan die geldig is voor zowel de IM als de CM is het *meta-onderzoek* van Gemünden (1985). In dit onderzoek werden onderzoeksresultaten met betrekking tot zowel consumenten als industriële beslissers op één noemer gebracht (zie ook 8.4.).
Na deze uiteenzetting is het duidelijk dat universele uitspraken over marketing zeer goed mogelijk zijn. De waargenomen verschillen tussen CM en IM zou men op moeten vatten als slechts beschrijvend en niet normatief van aard. Eén onderzoeksinspanning i.p.v. een opdeling tussen IM en CM lijkt vruchtbaar, efficiënter, en meer mogelijkheden te bieden voor creatieve oplossingen.
Kotler schrijft in 1972 (p. 8) al:

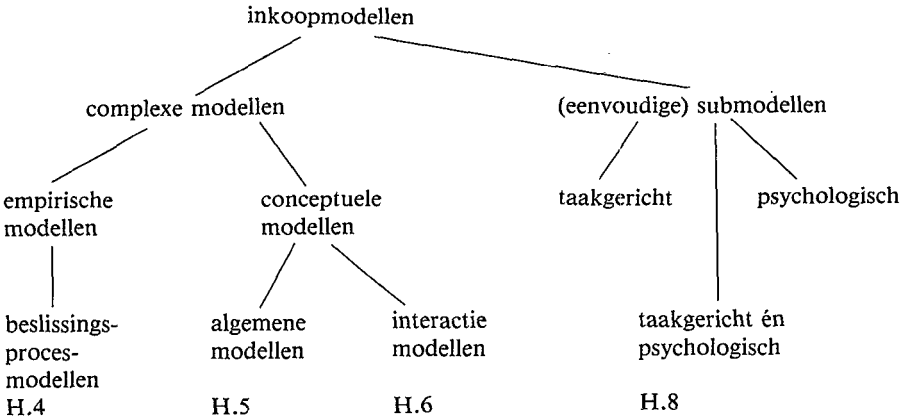
> 'An important means of achieving progress in a science is to try to increase the generality of its concepts.'

Waarom dan een heel artikel gewijd aan *industriële* koopprocessen?
De reden hiervoor is de volgende. Voordat een onderzoeksprogramma geformuleerd kan worden waarbinnen consument én ondernemer een gelijkwaardige plaats innemen moeten de beide deelgebieden zelf een bepaalde ontwikkeling doorgemaakt hebben. De huidige kennis op industrieel marketinggebied is nog te beperkt en fragmentarisch vergeleken met de kennis van consumentenmarkten. Meer kennis is nodig om een onderzoeksterrein af te kunnen bakenen waar grote overeenkomsten liggen, zodat een zinvolle integratie mogelijk is en gezamenlijk onderzoek haalbaar.

3. OPBOUW VAN HET ARTIKEL

3.1. De opbouw

De modellen die in de marketingliteratuur beschreven zijn met betrekking tot het inkoopproces van ondernemingen zijn onderling zeer verschillend. Een indeling die we kunnen maken is in figuur 1 weergegeven.
De eerste scheiding tussen complexe modellen en (eenvoudige) submodellen vindt hoofdzakelijk plaats o.b.v. het aantal relaties dat door het model verklaard wordt. Bij de complexe modellen spelen vele variabelen een rol die onderling allerlei verbanden vormen. De eenvoudige submodellen daarentegen bevatten slechts enkele variabelen. We noemen deze laatste groep modellen *sub*modellen omdat ze een rol spelen binnen de veelomvattende complexe modellen. De onderverdeling binnen deze twee hoofdgroepen inkoopmodelllen zullen we hieronder afzonderlijk behandelen.



Figuur 1. Een indeling van de inkoopmodellen.

De complexe modellen

De eerste groep die binnen deze verzameling modellen is het oog springt bestaat uit beschrijvingen van het inkoopproces binnen één of een aantal gelijksoortige ondernemingen. Het zijn empirische modellen; op basis van observaties wordt het proces stap voor stap schematisch weergegeven. Wij noemen deze groep modellen de *beslissingsprocesmodellen*. In paragraaf 4 zullen deze uitgebreid besproken worden. De tweede groep complexe modellen is wat heterogener van samenstelling dan de eerste. Het bindend element in deze verzameling modellen is hun conceptuele aard. In tegenstelling tot de zojuist genoemde beslissingsprocesmodellen berusten deze modellen op theoretische constructen en moet hun geldigheid nog getoetst worden.

De eerste subgroep conceptuele modellen is al redelijk oud en zeer bekend. We noemen deze de *algemene modellen*. Deze modellen geven een opsomming van factoren die het koopgedrag beïnvloeden en bovendien geven ze ruwweg aan hoe deze invloeden plaatsvinden. Hierbij wordt de omgeving van de onderneming óf het individu binnen de onderneming benadrukt. Zie voor een beschrijving van deze modellen paragraaf 5. Bij de laatste subgroep die we onderscheiden binnen de complexe modellen wordt expliciet rekening gehouden met het interactie aspect binnen en/of tussen ondernemingen. We noemen deze modellen dan ook de *interactiemodellen*. Het belang van relaties tussen leverancier, producent en afnemer voor het succes van een onderneming wordt de laatste jaren steeds meer erkend. De interactiemodellen die we zullen behandelen dateren dan ook allen uit de jaren tachtig. In paragraaf 6 zal dit alles nader aan bod komen.

De eenvoudige submodellen

Zoals hiervoor al werd aangegeven beperken de eenvoudige submodellen zich tot het verklaren van één verschijnsel met behulp van een zeer beperkt aantal variabelen. Het onderscheid dat straks in paragraaf 7 verder uitgewerkt wordt berust in deze groep modellen op de aard van de onafhankelijke variabelen. Zo kunnen voornamelijk taakgerichte variabelen een rol spelen of psychologische variabelen de meeste invloed hebben. Evenzo is het mogelijk dat de afhankelijke variabele zowel door psychologische als taakgerichte variabelen beïnvloed wordt.

3.2. Optiek van de modellen

Bij het formuleren van een model kan men kiezen voor een bepaalde invalshoek. Bij inkoopmodellen zijn er twee mogelijkheden: de optiek van de vrager en de optiek van de aanbieder. De stappen in het koopproces van de vrager en de aanbieder lopen deels parallel maar verschillen sterk op andere punten. De stappen die globaal kunnen worden onderscheiden staan in tabel 3.

Tabel 3. De stappen in het ruilproces vanuit verschillende gezichtspunten bekeken.

stappen gezien vanuit de aanbieder	stappen gezien vanuit de vrager
1. informatie verzamelen te aanzien van potentiële afnemer;	1. probleemherkenning;
2. contact leggen;	2. probleembeschrijving, specificatie, oplossing;
3. probleem formuleren en opdracht krijgen	3. potentiële leveranciers en verkrijgen offertes;
4. probleemonderzoek, indien mogelijk oplossingen formuleren;	4. evaluatie en selectie; eventueel onderhandelen;
5. presentatie en goedkeuring voorstel	5. uitvoering;
6. uitvoering en begeleiding.	6. evaluatie van de aankoop.

De modellen die in dit artikel besproken zullen worden zijn modellen vanuit de vrager. Het is de taak van de aanbieder om te analyseren hoe de verschillende stappen met elkaar corresponderen. Hierna kan de aanbieder dan gevolgtrekkingen afleiden voor het eigen beleid in de verschillende fasen van het (ver)koopproces.

4. BESLISSINGSPROCESMODELLEN

4.1. Inleiding

In paragraaf 2 werd reeds aangegeven dat de beslissingsprocesmodellen empirische beschrijvingen van het koopproces geven. Door het stapsgewijze karakter van deze modellen wordt steeds expliciet rekening gehouden met veranderingen die in de tijd plaatsvinden. In dit hoofdstuk zullen we vier voorbeelden geven van deze soort modellen waarbij gebruik gemaakt wordt van drie methodes: 1. protocol analyse; 2. script analyse en 3. beslissingsreeks analyse.
De twee modellen die hieronder allereerst besproken worden maken gebruik van protocol analyse, zij het niet op exact dezelfde wijze. Hierbij wordt o.a. inzicht verkregen in de gehanteerde beslissingsregels. Daarna behandelen we een model van Leigh en Rethans dat gebaseerd is op script analyse. Tot slot geven we een voorbeeld van een benadering waarbij het inkoopproces als een reeks beslissingen beschouwd wordt.

4.2. Crow, Olshavsky en Summers (1980)

Crow e.a. willen een gedetailleerde analyse geven van het gebruik van bepaalde criteria door de inkoper wanner hij/zij offertes aanvraagt of een koopbeslissing neemt. Het

gaat dus om keuzestrategieën in de laatste fasen van het inkoopproces. De methode die Crow, Olshavsky en Summers hiervoor gebruikt hebben is protocol analyse; de ondervraagde personen verwoorden hun gedachten terwijl ze denkbeeldige koopsituaties doorlopen. In dit onderzoek werden 14 inkopers van middelgrote tot grote ondernemingen ondervraagd. De uitkomsten van het onderzoek tonen aan dat het heel goed mogelijk is met behulp van protocol analyse gedetialleerde stroomschema's te maken voor de selectie van potentiële en definitieve leveranciers.
In dit model worden twee fasen onderscheiden bij het aanvragen van offertes:
1. schifting aanvaardbare leveranciers;
2. keuze 3 à 4 beste leveranciers

Daarna vindt dan de uiteindelijke leverancierskeuze plaats.
Uit het onderzoek blijkt dat in het algemeen wordt gestreefd naar 3 uitgebrachte offertes. Men bereikt dit aantal door acceptatienivo's te verhogen of juist te verlagen. Wanneer de methode correct uitgevoerd wordt zullen de variabelen in het model zowel taakgericht als psychologisch van aard kunnen zijn. Of beslissers hun irrationele beweegredenen hardop zullen weergeven valt echter ernstig te betwijfelen. Het model is zo gedetailleerd en simpel dat het op basis van cijfermatige data tot een eenduidige leverancierskeuze leidt.

4.3. Vyas en Woodside (1984)

Het onderzoek van Vyas en Woodside sluit duidelijk aan bij dat van Crow e.a. uit de vorige paragraaf. Ook zij gebruiken protocol analyse maar dit keer in echte aankoopsituaties. In aanvulling hierop werden ontmoetingen van de twee partijen geobserveerd en documenten (notulen, offertes) geanalyseerd. Met behulp van 'snowballing'[5] werden 62 personen ondervraagd die samen 18 koopcentra vormden (dit is dus gemiddeld 3,5 persoon per centrum). Uit het onderzoek van Vyas en Woodside blijkt dat de leverancierssselectie als volgt in fasen is op te delen:

1. identificatie potentiële leveranciers; 2. kwalificatie potentiële leveranciers;	stap I
3. bid invitation; 4. bid evaluation; 5. bid selection.	stap II

Voor elke fase construeerden zij vervolgens uitvoerige stroomschema's. Opvallend zijn de overeenkomsten met de resultaten van Crow e.a. (IV.2), met name het variëren van de criteria om het gewenste aantal offertes (=3) te verkrijgen en het gebruik van verschillende beslissingsregels gedurende het beslissingsproces. Het bijzondere van dit onderzoek is dat niet alleen de inkoper maar de gehele DMU in het onderzoek is betrokken. Het model bevat taakgerichte (bv. prijs) en psychologische (bv. source loyalty) variabelen wat de werkelijkheid goed benaderd. Bovendien zijn de gegevens die gebruikt worden betrouwbaar door de combinatie van dataverzamelingsmethoden. De stroomdiagrammen waarin het onderzoek van Vyas en Woodside uitmondt zijn redelijk gedetailleerd en nogal complex. Ze leiden echter niet rechtstreeks tot een beslissing doordat men rekening houdt met de invloed van subjectieve oordelen en ervaring.

4.4. Leigh en Rethans (1984)

Een iets andere benadering om te komen tot gedetailleerde beschrijvingen van industrieel koopgedrag is door gebruik te maken van de script analyse. Deze analyse berust op een concept uit de cognitieve psychologie te weten het cognitieve script. Het cognitieve script is een vaste reeks van gedragingen die een persoon in het geheugen opslaat wanneer hij of zij herhaaldelijk eenzelfde situatie meemaakt. In het onderzoek van Leigh en Rethans hebben 36 inkopers hun cognitieve scripts op papier gezet voor vier verschillende hypothetische situaties:
– 'new task'[6] aankoop van computerterminals;
– eerste vertegenwoordigersbezoek;
– onderhandelingsproces na offerte-ontvangst;
– 'modified rebuy' van computerterminals.

De ondervraagden kwamen uit zowel grote als kleine ondernemingen en bleken zeer goed in staat hun cognitieve scripts te verwoorden. Ook uit dit onderzoek blijkt wederom dat men streeft naar drie uitgebrachte offertes. Bovendien is ook duidelijk dat bij heraankopen het informatieproces vaak anders verloopt dan bij nieuwe aankopen. Bij heraankopen is in veel gevallen de oude leverancier de enige die geraadpleegd wordt over mogelijke oplossingen van het probleem. Helaas vormt het script van alle respondenten samen geen logisch geheel. Men kan zich voorstellen dat elke beslisser uit het totale script maar een deel van de stappen doorloopt. Het is dus niet mogelijk uit het script een duidelijke stroomdiagram af te leiden. De bruikbaarheid van het model ligt vooral in het persoonlijke vlak. Men probeert dan bijvoorbeeld vast te stellen wat het verwachte script van een inkoper is t.a.v. een eerste afspraak met een vertegenwoordiger. Deze vertegenwoordiger kan van deze kennis gebruik maken door er positief van af te wijken of juist aan bepaalde verwachtingen te voldoen. Ook kunnen scripts gebruikt worden bij de opleiding van nieuwe medewerks. De gevallen die in dit onderzoek bekeken worden, zijn geen praktijksituaties maar hypothetische aankopen. Wat dit voor gevolg heeft voor de validiteit van de uitkomsten is onduidelijk. Ook vindt er op geen enkele wijze verificatie plaats wat bij de eerder besproken protocol analyse wel het geval is. Evenals bij het model van Crow e.a. is het waarschijnlijk dat zowel taakgerichte als psychologische variabelen een rol spelen, maar dat deze allen boven water komen is zeer onwaarschijnlijk. De mate van detail van een script kan zeer groot zijn en afhankelijk van de vraagstelling kan het script een zeer klein deel tot het hele traject van het koopproces beslaan.

4.5. Cardozo (1983)

Als laatste voorbeeld van een beslissingsprocesmodel zullen we in deze subparagraaf een hele andere benadering laten zien. Cardozo onderscheidt net als de meeste andere auteurs verschillende stappen in het aankoopproces (zeven om precies te zijn). Nu stelt hij dat elke stap afzonderlijk als een beslissing op te vatten is die positief of negatief voor een bepaalde leverancier kan uitvallen. De reeks beslissingen die moet worden genomen begint met het instellen van een DMU voor een bepaalde aankoop en eindigt bij de acceptatie van een offerte of een heraankoop. Door de kansen op een gunstige uitkomste voor iedere stap afzonderlijk te schatten (op basis van ervaring, historische data of onderzoek) kan men de uiteindelijke aankoopkans bepalen:

P(actie) x P(specificaties) x P(offerte) x P(keus) x P(goedgekeurd) x P(acceptatie) = P(aankoop)

P(aankoop) x (P(heraankoop) = P(serie aankopen)

De marketing manager kan dit model op twee manieren gebruiken:
1. Om de verwachte resultaten van een bepaald marketing plan te voorspellen.
2. Om de huidige marketingproblemen te diagnostiseren.

Bij deze tweede toepassing zoekt men naar die schakel in het beslissingsproces die met marketinginspanningen het meest efficiënt bewerkt kan worden teneinde P(aankoop) te verhogen. In het model worden geen afzonderlijke variabelen onderscheiden. Alle taakgerichte en psychologische variabelen worden samengenomen in één waarschijnlijkheid. De gevoeligheid van de uitkomst van het model kan geanalyseerd worden door van elke stap een pessimistische, een beste en een optimistische schatting van de kansen te maken. Met behulp van dit model heeft Cardozo geen beschrijving gegeven van het beslissingsproces van de koper, wat wél zijn bedoeling was. Het model is beter te beschouwen als een instrument voor de verkoper die met de stappen in het beslissingsproces van zijn afnemers rekening wenst te houden.

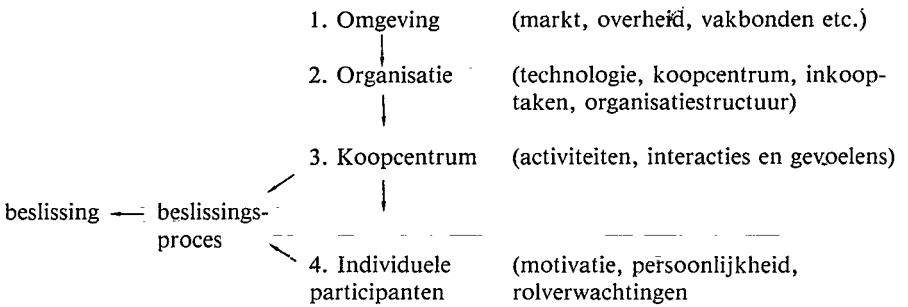
5. ALGEMENE MODELLEN

5.1. Inleiding

Bij de bespreking van de opbouw van dit artikel in paragraaf 3 hebben we de plaats van de algemene inkoopmodellen reeds aangegeven. Evenals de interactiemodellen zijn de algemene modellen complex en conceptueel van aard. De complexiteit geeft aan dat er vele variabelen en verbanden in de modellen een rol spelen. De conceptuele aard houdt in dat de verbanden theoretisch verondersteld worden en pas daarna (eventueel) aan de werkelijkheid getoetst.
In deze paragraaf zullen twee bekende algemene modellen worden besproken. Als eerste het model van Webster en Wind (1972) en vervolgens dat van Sheth (1973). Beide modellen zijn in principe geldig voor alle produkten, markten en ondernemingen. Bovendien hebben ze betrekking op het hele inkoopproces. Deze algemeenheid heeft als keerzijde dat daadwerkelijk meten van de variabelen niet goed mogelijk is. De bruikbaarheid van deze modellen ligt dan ook bij het verkrijgen van inzicht in welke invloeden een rol kunnen spelen bij het industriële koopproces. Voor een specifiek probleem kan een deel van een algemeen model nader uitgewerkt worden.

5.2. Webster en Wind (1972)

Het model dat Webster en Wind hebben ontwikkeld noemen zij: 'A General Model for Understanding Buying Behavior'. In deze naam geven zij al aan dat het een *algemeen* model is ten einde meer *inzicht* in het organisatorische koopgedrag te krijgen. Het totale model ziet er enigszins vereenvoudigd als in figuur 2.
In dit model is wordt de invloed van de omgeving nogal sterk benadrukt; het individuele handelen is voornamelijk een resultante van invloeden van buitenaf. Bovendien

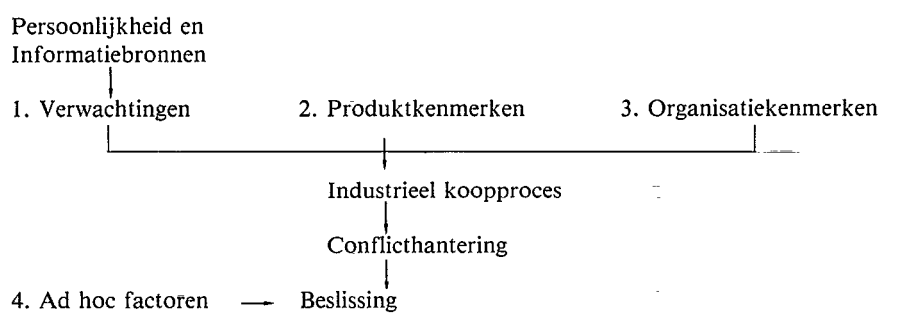


Figuur 2. Vereenvoudigd Webster en Wind-Model.

wordt niet expliciet rekening gehouden met de invloed van de tijd. Het model neemt wel allerlei soorten variabelen, taakgericht en psychologisch, mee. Van enig detail is echter geen sprake. Een heel andere invalshoek heeft het model van Sheth dat we hierna behandelen, dat juist het individu als bepalende factor ziet.

5.3. Sheth (1973)

Zoals gezegd ziet Sheth het individu, met zijn/haar eigen opleiding, persoonlijkheid en verwachtingen als centraal punt in het industrieel koopproces. De omgeving heeft vooral invloed via de verwachtingen van het individu. Mensen zijn echter geen objectieve waarnemers en dus wordt slechts een deel van de omgeving waargenomen, op vaak onzuivere wijze. Het totale model van Sheth is redelijk gedetailleerd al wordt niet ingegaan op het informatieverwerkingsproces zelf. Sterk vereenvoudigd ziet het 'Model of Industrial Buyer Behavior' eruit als in figuur 3:



Figuur 3. Het Sheth-model.

De vier groepen van variabelen in dit model kunnen zowel taakgerichte als psychologische variabelen bevatten. Het Sheth-model heeft geen specifieke aandacht voor het tijdselement. Door de mate van detail is het in principe mogelijk toetsbare hypothesen af te leiden waarbij dan wel psychometrische methoden nodig zijn.

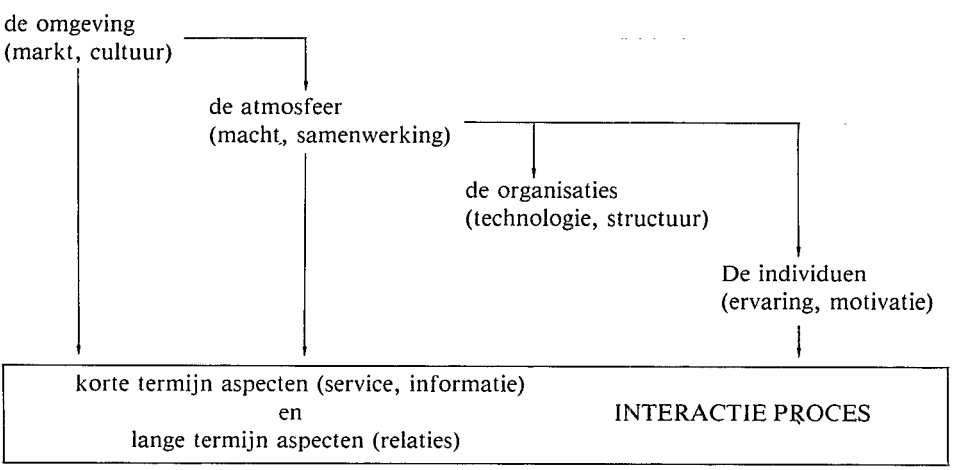
6. INTERACTIEMODELLEN

6.1. Inleiding

De algemene modellen die we zojuist besproken hebben, hebben beide het nadeel dat ze een belangrijke invloed op het koopproces verwaarlozen. Deze invloed is de interactie die plaatsvindt tussen een onderneming, haar handelspartners en haar concurrenten, maar ook binnen de onderneming zelf. Deze interactie kan leiden tot langdurige relaties op bedrijfsnivo maar ook tussen medewerkers persoonlijk.
De in dit hoofdstuk te bespreken inkoopmodellen proberen allen op een of andere manier dit interactieaspect mee te nemen. De modellen zijn conceptueel van aard terwijl enkelen empirisch getoetst zijn. Doordat in deze modellen met meerdere ondernemingen en vele variabelen rekening wordt gehouden zijn ze al snel zeer complex.
De volgorde waarin de interactiemodellen in dit hoofdstuk zullen worden besproken hangt samen met de verschillende abstractienivo's van deze modellen. We beginnen met een model ontwikkeld door de I.M.P.-Group waarin interactie centraal staat. De waarde van dit model berust meer op de ideeën die erachter zitten dan op de praktische toepasbaarheid. Ook het netwerkmodel dat in paragraaf 3 besproken wordt berust op een 'nieuwe' zienswijze en is nogal abstract. Er zijn echter wel toepassingen van dit model afgeleid die empirisch getoetst kunnen worden. In het daarop volgende model staat de interactie niet meer centraal maar is deze expliciet opgenomen in een model. Dit model kan wel empirisch getoetst worden.

6.2. IMP-Group (1982)

De IMP-Group bestaat uit een aantal academische onderzoekers uit Frankrijk, Duitsland, Italië, Zweden en Engeland. Gezamenlijk doen zij onderzoek op het gebied van 'Industrial Marketing and Purchasing'. In een groots opgezet onderzoek in 1982 werd o.a. onderzocht hoe managers het inkoop- en marketinggedrag van hun relaties waarnemen en hoe de interactie plaatsvindt. Het onderzoek was opgehangen aan een



Figuur 4. Het interactiemodel van de IMP-group.

redelijk eenvoudig model waarin koopgedrag en marketing worden geïntegreerd: een interactiemodel. In dit model is de interactie de afhankelijke variabele die bepaald wordt door vier invloeden (zie figuur 4).
Het onderzoek betrof de aan- en verkopen van grondstoffen, halffabrikaten en duurzame produktiemiddelen. In totaal werden 876 interviews gehouden met marketing- en inkoopfunctionarissen uit verschillende industrieën. Het model uit figuur 4 is te globaal om relaties te specificeren. Zoals uit het onderzoek bleek kan het echter nuttig zijn voor de manager bij het bepalen van de te volgen strategie met betrekking tot het al of niet bevorderen van relaties. Het model is zeer compleet; het bevat taakgerichte en psychologische variabelen en heeft betrekking op het gehele koopproces. Mede hierdoor is het model weer erg globaal zodat geen duidelijk onderscheid tussen de verschillende fasen van het inkoopproces kan worden gemaakt. Ten slotte kunnen we opmerken dat het model, hoewel het moeilijk kwantificeerbaar is, met behulp van vergelijkingen tussen organisaties wel degelijk tot normatieve uitspraken kan leiden.

6.3. Netwerkmodel

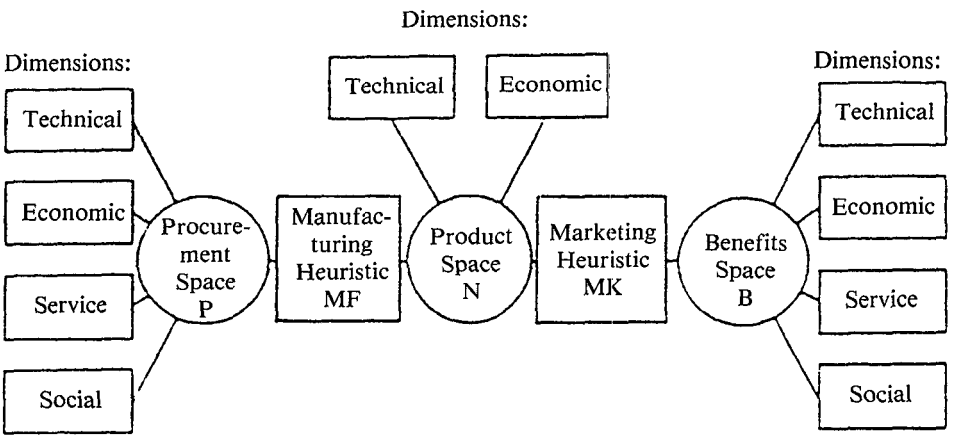
De idee dat het geheel van relaties tussen industriële ondernemingen op te vatten is als één netwerk is in de jaren tachtig in Zweden ontwikkeld. Deze benadering wil de gevolgen van marketingactiviteiten op *lange* termijn expliciet beschrijven. De kern van de idee is dat iedere industriële producent een element is van verschillende netwerken waarbinnen hij relaties heeft die:
1. achterwaarts (met leveranciers);
2. voorwaarts (met afnemers) en
3. zijdelings (met concurrentie/partners)

gericht zijn. Deze direkte contacten kunnen onderling ook weer relaties vormen zodat een heel netwerk ontstaat. De positie van een onderneming in een netwerk op een bepaald moment wordt bepaald door:
a) de identiteit van de overige ondernemingen binnen het netwerk;
b) de rol van de organisatie in het netwerk;
c) het belang van de organisatie binnen het netwerk, en
d) de sterkte van de relaties met de andere ondernemingen.

Het netwerk concept is toegepast op verschillende gebieden, zoals: het verdedigen en aanpassen van marktposities; stabiliteit van relaties en de invloed van technologische ontwikkeling. In deze toepassingen wordt het verloop van de netwerkpositie van één onderneming bestudeerd, of worden vergelijkingen getrokken tussen de posities van meerdere ondernemingen.
Het netwerkmodel is redelijk compleet; de interactie en afhankelijkheden tussen de verschillende ondernemingen staan centraal. De gedetailleerdheid van het model, de aard van de gebruikte variabelen en de toetsbaarheid van het model zijn afhankelijk van de specifieke toepassing. In principe wordt het gehele aankoopproces meegenomen, waarbij echter nog geen onderscheid wordt gemaakt tussen de verschillende fasen binnen het koopproces. Het enige tijdsaspect wat dan ook meegenomen wordt is dat tussen verschillende aankopen in plaats van binnen een aankoop.

6.4. Bender (1983)

Het model dat door Bender ontwikkeld is noemt hij 'Industrial Conversion Framework' (ICF). Het model geeft een systematische beschrijving van een 'produktketen' waarbij koopgedrag van organisaties, marketing en innovatie geïntergreerd worden. In het ICF model worden drie fasen in het industriële produktieproces onderscheiden. Elk van deze fasen in op te vatten als een kenmerkenruimte ('space'). Voor één produkt van een onderneming is het ICF model in figuur 5 weergegeven.



Figuur 5. Enkelvoudig ICF. Bron: Bender (1983, p. 32a).

In werkelijkheid bestaat er een keten van dergelijke produktieprocessen waarbij de 'benifit space' van de aanbieder zo goed mogelijk moet aansluiten bij de 'procurement space' van de afnemer. Op deze wijze ontstaan er materiële stromen en informatiestromen. De produktstroom van grondstofleverancier tot eindgebruiker wordt gecompenseerd door een stroom geld (of goederen in het geval van ruilhandel). Bovendien is er een continue uitwisseling van informatie. Ook hier wordt dus duidelijk rekening gehouden met het interactie-aspect. De wijze waarop de interactie tussen inkoop, produktie en marketing plaatsvindt is in dit model van groot belang. Het ICF model is getoetst met gebruik van LISREL modellen. In een dergelijk model kan men hypothesen formuleren tussen variabelen die niet direkt meetbaar hoeven te zijn. Zo zijn de drie 'spaces' uit figuur 5 niet zelf meetbaar al kunnen er wel indicatoren voor gevonden worden. Wanneer we het ICF model evalueren valt allereerst op dat het zeer duidelijk en compleet is. De hele produktieketen wordt in beschouwing genomen en met velerlei verschillende variabelen wordt rekening gehouden. Jammer genoeg wordt de rol van het individu onderbelicht en ook de sociale interactie komt nauwelijks aan bod. Ook wordt niet expliciet rekening gehouden met het tijdselement binnen het koopproces.

7. EEN TOTAALMODEL, DE BUYGRID

7.1. Inleiding

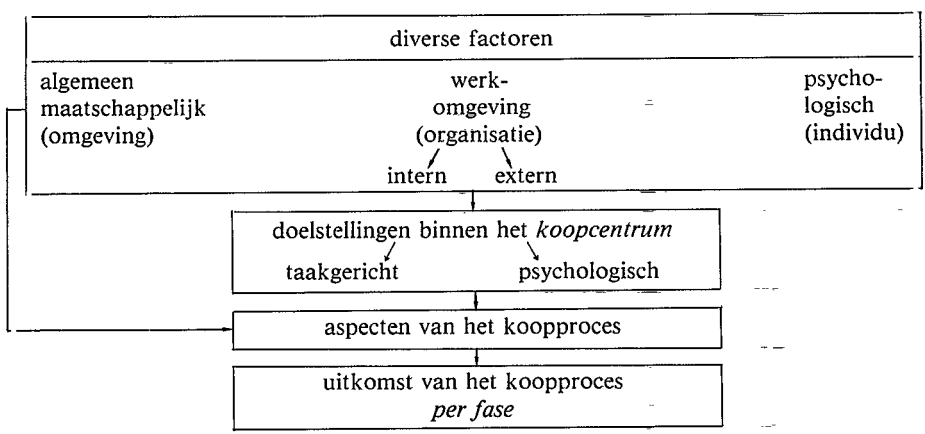
Wanneer we de kenmerkende eigenschappen van de conceptuele modellen uit hoofdstuk 5 en 6 bekijken krijgen we het volgende beeld:

– Algemene modellen:	Webster en Wind (1972)	– omgeving centraal
	Sheth (1973)	– individu centraal
– Interactiemodellen:	diversen	– interactie centraal

Willen we een zo compleet mogelijk model formuleren dan zullen we echter al deze drie elementen mee moeten nemen. In dit hoofdstuk willen we een raamwerk aangeven aan de hand waarvan een redelijk compleet totaalbeeld verkregen kan worden.
In dit raamwerk komt de variabele 'nieuwheid van het produkt' voor. Deze variabele is afkomstig uit het zeer bekende Buygrid model van Robinson e.a. (1967). Dit model zal in 7.3. nader besproken worden.

7.2. Een totaalmodel

In deze paragraaf wordt een totaalmodel besproken waarbij zowel de omgeving, het individu als de interactie aan bod komen. Het model is in fig. 6 weergegeven.



Figuur 6. Een totaalmodel van invloeden op het koopgedrag van ondernemingen.

Dit model moet men niet zien als een nieuwe zienswijze maar meer als een integratie van bestaande ideeën. We zullen de verschillende nivo's in de figuur eerst apart bespreken en vervolgens een vergelijking maken met het bestaande model van Webster en Wind (1972).
In het totaalmodel is het ten eerste van belang de bestaande situatie in kaart te brengen. Hier spelen *diverse factoren* een rol. Deze factoren zijn te groeperen als: omgevings-, organisatie- en individuele factoren. Voorbeelden van deze invloeden worden in tabel 4 gegeven. Ten tweede is het van belang op welke wijze deze diverse factoren de *doel-*

stellingen binnen het koopcentrum bepalen. Hierbij spelen de submodellen die in het volgende hoofdstuk behandeld zullen worden een grote rol. Vervolgens komt de belangrijke derde stap in het model. Hier wordt gesteld dat het uiteindelijke koopproces een resultante is van de diverse factoren en de afgeleide doelstellingen. Het koopproces beschrijven we dan aan de hand van de volgende 5 *aspecten:*

1. Het koopcentrum: – samenstelling;
   – rolverdeling.
2. Het criteria die gebruikt worden.
3. Het informatiegedrag.
4. Het aantal alternatieven dat overwogen wordt.
5. De beslissingsregels die gehanteerd worden.

Deze aspecten kunnen veranderen naar gelang de fase van het koopproces waarin men zich bevindt. Vandaar dat het vierde blok in het model de uitkomst *per fase* van het koopproces voorstelt. Voor een aanbieder zijn vooral het moment waarop de bid-list vastgesteld wordt en de toekenning van het contract van belang.

Het totaalmodel uit figuur 6 kunnen we het beste vergelijken met het algemente model van Webster en Wind (1972). Ook in dit model wordt een systematische opsomming gegeven van de invloeden op het koopproces.Webster en Wind onderscheiden hierbij vier nivo's namelijk de omgeving, de organisatie, het koopcentrum en het individu (zie ook 5.2).

Tabel 4. Diverse factoren die in het totaalmodel een rol spelen.

Diverse factoren:	Bijvoorbeeld:
alg. maatschappelijk:	– alg. prijsnivo
	– onzekerheid afzet
	– onzekere inkoopmarkten
werkomgeving intern:	– additionele kosten (onderhoud)
	– belang van onafhankelijkheid
	– benodigde technische kennis
	– capaciteitstekort cq. -overschot
	– kennis van of ervaring met leverancier
	– nieuwheid van het produkt (buyclass)
	– organisatiestructuur (formeel en informeel)
extern:	– aangeboden prijs
	– aangeboden service
	– aangeboden produkteigenschappen
psychologisch/ persoonlijk:	– (im)materiële geschenken
	– reputatie
	– persoonlijkheid (verwachtingen, risicoafkeer)
deels werkomgeving/ deels psychologisch:	– switching costs
	– relaties

In ons totaalmodel zijn, mede ter wille van de eenvoud, de omgeving en de organisatie in één nivo ondergebracht. Ook het individu wordt niet als apart nivo gezien maar is verweven in het geheel als een van de 'diverse factoren' en als lid van het koopcentrum (zie fig. 6). Zo wordt op meerdere plaatsen expliciet aandacht aan het individu besteed. Ook de interactie is als steeds terugkerend element meegenomen. Zoals uit tabel 4 blijkt spelen relaties op verschillende plaatsen een rol: in de werkomgeving, bij het individu maar ook bij het bepalen van de taakgerichte en psychologische doelstellingen binnen het koopcentrum. Het in deze paragraaf besproken totaalmodel kunnen we omschrijven als een vereenvoudigde versie van het algemene model van Webster en Wind dat uitgebreid is met het individu en met het interactieaspect.
Een bijzonder geval is de factor *nieuwheid van het produkt* (buyclass) die we in tabel 4 onder de interne werkomgeving geplaatst hebben. Buyclass is een term afkomstig uit het 'Buygrid Model' van Robinson, Faris en Wind (1967). Dit Buygrid model is echter geen model in de betekenis zoals die in dit artikel gebruikt wordt. Het is eerder een classificatie van industriële aankopen o.b.v. hun nieuwheid voor de organisatie. Dit model heeft echter aanleiding gegeven voor veel vervolgonderzoek en de Buygrid is zeer bekend. Vandaar dat dit model in de volgende paragraaf uitgebreid toegelicht en bekritiseerd zal worden.

7.3. De Buygrid

Het eerste model dat op basis van observaties een beschrijving gaf van het industrieel koopgedrag was de 'Buygrid'. Het aantal observaties was zeer beperkt maar toch werd het model een groot succes. De voornaamste redenen hiervoor waren de eenvoud van het model, de toepasbaarheid op alle produktklassen en de mogelijkheid toetsbare hy-

dimensie / buyclass	nieuwheid van het probleem	informatie-behoefte	aantal overwogen alternatieven
new task[a]	nieuw	groot	groot
modified[b]	matig	beperkt	enkele (2 à 3)
straight rebuy[c]	routine	nihil	klein, meestal 1

Figuur 7. De dimensies van de drie buyclasses.

a. Deze aankoop is voor de onderneming totaal nieuw. Er is onzekerheid op het gebied van bijvoorbeeld specificaties, leverancierskeuze en prijsnivo, zodat het 'perceived risk' (zie V.1) aanzienlijk is. Als gevolg hiervan zal het koopcentrum uit veel personen bestaan waarbij vooral voor de technici een belangrijke rol is weggelegd.
b. Een modified rebuy is zoals de naam als zegt een heraankoop waarbij een verandering is aangebracht in de specificaties. De eerdere aankoop kan zowel een new task als een routine aankoop zijn geweest. Voor een potentiële leverancier is de categorie van de gewijzigde heraankoop van het grootste belang. Hier heeft hij tenslotte de kans om vaste leverancier te worden. Dit kan zijn van een produkt dat tamelijk nieuw is maar in de toekomst vaker gekocht kan gaan worden, of van een frequent gekocht produkt dat iets veranderd is maar tot nu toe door een concurrent werd aangeleverd.
c. De heraankoop tenslotte is voor de potentiële leverancier het minst interessant. Deze aankopen zijn zo routinematig dat vaak slechts één alternatief, namelijk de huidige leverancier, in overweging wordt genomen. Vanuit de kopende organisatie gezien is het echter wel nuttig om wat meer aandacht aan deze categorie te besteden. Bij grote hoeveelheden routinematige aankopen is heel wat geld te besparen.

pothesen te vormen. Robinson e.a. onderscheiden drie 'buyclasses' (new task, modified rebuy, straight rebuy) die invloed kunnen hebben op de activiteiten binnen de verschillende fasen van het koopproces (zij onderscheiden 8 fasen). De buyclasses verschillen onderling op drie dimensies te weten: (1) nieuwheid van het probleem, (2) informatiebehoefte van de inkoper en (3) het aantal overwogen alternatieven. Op deze wijze ontstaat een matrix waarin de buyclass variabele de ene ingang vormt en de drie dimensies van het koopproces de andere (figuur 7).

De kern van het Buygrid model is dat bij onderzoek naar industriële aankopen:
a. onderscheid gemaakt moet worden tussen de verschillende buyclasses, en tevens
b. rekening moet worden gehouden met de verschillende fases in het koopproces.

Er zijn dus 24 verschillende situaties (3 buyclasses x 8 fases) waarin verschillende factoren belangrijk zijn.
Het eerder genoemde 'succes' van dit model heeft geleid tot veel onderzoeken naar de geldigheid ervan. Een deel van deze onderzoeken steunt het Buygrid model en een deel heeft duidelijke kritiek. Zo heeft het model geen aandacht voor persoonlijke en organisatorische (psychologische) variabelen. Min of meer indirekte kritiek zal gegeven worden in hoofdstuk 8. Zoals in dat hoofdstuk zal blijken is de relatie tussen onzekerheid en informatiezoekgedrag niet bepaald eenduidig (zie 8.4.) en kan de samenstelling van het koopcentrum mede bepaald worden door variabelen die onafhankelijk zijn van de soort aankoop (zie 8.5.).
De direkte kritiek op het model van Robinson e.a. concentreert zich op twee punten:
1. Het is niet mogelijk te generaliseren over produktklassen.
2. Er zijn andere factoren die grote invloed hebben op het koopproces zoals financiëel of technologisch belang van het produkt of technische complexiteit.

Verschillende beschrijvende onderzoeken blijken de indeling van Robinson, Farris en Wind soms wel en soms niet als zinvol te bestempelen. *Bevestiging* is te vinden op het gebied van variaties in het aantal overwogen alternatieven, functionele invloed en variaties in informatiebehoefte. *Niet bevestigend* echter zijn onderzoeken met betrekking tot de invloed van de verschillende leden van het koopcentrum, maar ook wat betreft het aantal overwogen alternatieven. Hierbij willen we erop wijzen dat onderzoek van het koopcentrum speciale problemen oplevert. In 8.5. zullen we de methoden bespreken waarmee men het probleem van verschillende meningen in één koopcentrum tracht te omzeilen.
Een deel van de kritiek op het Robinson, Farris en Wind model is te ondervangen door gebruik te maken van de zgn. *Fisher matrix*. Fischer (1974) benadrukt dat groepsbeslissingen (waaronder dus ook beslissingen van koopcentra vallen)sterk beïnvloed worden door onzekerheid en de complexiteit van het probleem. Deze variabelen kunnen we vertalen in commercieel risico van de aankoop, respectievelijk complexiteit van het produkt zodat de matrix ontstaat, zoals weergegeven in figuur 8.

Op deze wijze kan duidelijk gemaakt worden dat een heraankoop van relatief lage waarde (zoals mappen voor computeruitvoer) en een heraankoop van bijv. bedrijfsauto's niet te vergelijken zijn m.b.t. informatiebehoefte en koopgedrag, doordat ze in de Fisher-matrix in een verschillende cel thuishoren.
Een laatste opmerking die we hier willen plaatsen is de volgende. In bijna alle onder-

		produktcomplexiteit	
		hoog	laag
commercieel risico	hoog	– new task – modified rebuy – straight rebuy	– new task – modified rebuy – straight rebuy
	laag	– new task – modified rebuy – straight rebuy	– new task – modified rebuy – straight rebuy

Figuur 8. Fisher-matrix

zoeken naar de geldigheid van de Buygrid wordt de new task tegenover de straight rebuy geplaatst of wordt slechts een van de 3 buyclasses bekeken. De grenzen tussen de 3 categoriën zijn dan ook niet echt duidelijk en daarom pleiten wij ervoor de buyclass variabele als een continuum te zien. Een continuum met aan het ene uiteinde de onbekende nieuwe aankoop, aan het andere uiteinde de routine aankoop zonder uitvoerige besluitvorming en daartussen een range met als middelpunt de moeilijk definieerbare modified rebuy.

8. ENKELE SUBMODELLEN

8.1. Inleiding

Zoals in hoofdstuk 3 werd aangegeven en ook in tabel 4 in hoofdstuk 7 naar voren kwam kunnen we de submodellen in drie groepen indelen:
1. taakgerichte submodellen;
2. psychologische submodellen en
3. submodellen met taakgerichte én psychologische elementen.

In tabel 5 wordt aangegeven welke 'doelstellingen' in de submodellen nagestreeft worden, wat de afhankelijke en wat de onafhankelijke variabelen zijn.

8.2. Wederkerige aankopen

Onder de noemer wederkerigheid (reciprocal buying) kunnen we een aantal transactiemethoden samenbrengen die nader op te delen zijn in 2 groepen:
1. tegenaankopen ('countertrade') en
2. ruilhandel ('barter').

Het grote verschil tussen deze 2 groepen is dat bij tegenaankopen formeel 2 transacties plaatsvinden (nu of in de toekomst) waarbij geld tegen goederen of diensten geruild wordt. Bij ruilhandel komt er helemaal geen geld op tafel en is er dus sprake van een eenmalige uitwisseling van produkten. De voor- en nadelen van dergelijke transacties worden bijvoorbeeld bepaald door het land van herkomst van de handelspartners. In het

Tabel 5. De verschillende categorieën submodellen.

doelstelling	onafh. variabelen/ determinanten	afh. var./fenomeen
1. Taakgericht		
A. minimale kosten	• prijs resp. kosten als enige aankoopcriterium	• kosten
B. wederkerige aankopen	• zekerheid m.b.t. eigen afzet • lagere prijs d.m.v. afspraken	• wederkerige aankopen
C. source loyality	• kosten info-zoeken, doorbreken routine • risico nieuwe leverancier	• source loyality
D. multiple sourcing	• wens onafhankelijkheid • lagere prijs door concurrentie	• multiple sourcing
E. co-makership	• behoefte technologische kennis • kosten en risico delen	• co-makership
F. make or buy	• beperkte of juist over-capaciteit • kennis, kosten en risico	• make/buy beslissing
2. Psychologisch		
G. ego-enhancement	• schenken van materiële (bv. geld) of immateriële (vlijerij)	• gedragaanpassing
H. sociale interactie	• interactie tussen 2 of meer subjecten (personen, funct. eenheden), rolverwachtingen	• gedragsaanpassing
3. Gemengd		
I. relaties: taakgericht	• onzekere inkoopmarkten • specialisatiemogelijkheden • bijhouden technologische ontwikkelingen	• relatievorming
psychologisch	• voorkeur voor vrienden en/of bekenden	• relatievorming
J. risicoreductie:		
taakgericht	• onvoldoende kennis t.a.v. leverancier	• extra informatie • source loyalty • meer info • minder info
psychologisch	• gewenste verlaging 'perceived risk'	• reputatie v. belang

ad B. De wederkerige aankopen zullen in subparagraaf 8.2 nader besproken worden.

ad C. Het derde model (source loyalty) door Wind bestudeerd in 1970. Hij concludeerde dat naast bepaalde organisatiekarakteristieken de volgende factoren source loyalty versterken:
a. De prijs van het produkt is de afgelopen periode gedaald.
b. De totale uitgaven aan het produkt zijn klein.
c. De totale kosten zijn de vorige periode al gedaald zodat de druk laag is.
d. De gebruiker binnen de organisatie beveelt een bepaalde bron aan.

ad.H. Aan dit submodel zal in 8.3. uitgebreid aandacht worden besteed.

ad I en J. Daar deze gemengde submodellen nogal gecompliceerd zijn is veel onderzoek op deze gebieden verricht. Aan beide modellen zal dan ook een aparte paragraaf gewijd worden (8.4. en 8.5.). Door de aard van de belangrijkste variabelen is in de psychologische submodellen voor promotie als marketinginstrument een belangrijke rol weggelegd. Dit in tegenstelling tot de taakgerichte modellen waar de enige functie van promotie het verstrekken van objectieve informatie is.

algemeen leveren het kwaliteits- en levertijdsrisico en de noodzaak tot doorverkopen de meeste problemen op. De voordelen op het gebied van capaciteitsbenutting en het omzeilen van valutaproblemen kunnen echter aanzienlijk zijn.
Dacht men in 1978 nog dat wederkerige transacties bijna niet meer voorkwamen (Finney, 1978), recent onderzoek in de Verenigde Staten en in Groot Brittannië wijst in een heel andere richting. Zo is het aandeel van tegenaankopen en ruilhandel in de wereldhandel zeer sterk gestegen. Van ± 2% in 1976 tot wel 30% in 1982. Bij de Oost-West handel ligt dit percentage nog veel hoger namelijk op ± 50%! (Huszagh en Barksdale, 1986).
Het Amerikaanse onderzoek van Huszagh en Barksdale (1986) betreft alleen internationale transacties. De reden hiervoor is waarschijnlijk het feit dat wederkerige transacties binnen de V.S. niet veel voorkomen omdat deze al snel vallen onder de anti-trust wetgeving. Deze wetgeving vereist o.a. dat er tegen 'eerlijke' prijzen gehandeld wordt en dat er geen concurrentievervalsing optreedt. Huszagh en Barksdale hebben uiteindelijk 18 ondernemingen onderzocht die gezamenlijk gedurende 12 jaar 147 (internationale) wederkerige transacties hadden uitgevoerd. Het grootste deel van deze ondernemingen bleek marktpenetratie of marktontwikkeling als ondernemingsstrategie te voeren, te handelen met centraal geplande landen en produceerde zelf kapitaalgoederen.
Het onderzoek in Engeland werd in 1987 gerapporteerd door Shipley en Neale. Zij stuurden een vragenlijst aan 1000 marketing directeuren, ontvingen 217 bruikbare antwoorden terug, waarvan 57 respondenten (26,3%) bevestigden enige vorm van wederkerigheid te praktiseren. De respondenten waren te verdelen in een groep die nationaal 'ruilde' en een groep die buitenlandse partners had. Opvallend hierbij is dat geen enkele respondent in beide categoriën viel. Bij een vergelijking tussen de 2 groepen bleek dat vooral de hoofdreden voor de toename van wederkerigheid verschilde. De internationale handelaars wezen het tekort aan harde valuta bij hun klanten als belangrijkste reden aan; Binnenlandse transacties namen vooral toe doordat betere ruilprodukten werden aangeboden.
Dat ook in Nederland wederkerige transacties plaatsvinden is wel zeker, hoeveel is echter onbekend. Een feit is dat er in Nederland verschillende 'Barter-clubs' bestaan die als intermediair fungeren voor hun leden (zoals Bartering B.V. te Heusden en Dutch Bartering House te Veendam).

8.3. Relaties

In de literatuur zijn er twee verschillende benaderingen van het fenomeen relaties. De eerste benadering legt de nadruk op de 'seller/buyer dyad' als studie object. De tweede groep auteurs ziet een organisatie als onderdeel van verschillende netwerken en analyseert dus verschillende relaties tegelijk. We zullen hier alleen de simpele relatie bekijken, het netwerk is uitgebreid in 6.3 aan de orde geweest bij de interactiemodellen.
De opvatting dat een transactie in de marketing als een *'exchange relationship'* gezien kan worden, begon in het begin van de jaren '70 post te vatten. De veel gebruikte term 'dyad' is afkomstig van het Griekse woord 'dyas' dat betekent twee gezien als één. Deze zienswijze impliceert dat een marketeer niet alleen inzicht moet hebben in de onderneming als koper maar ook als verkoper. In de loop der tijd is bij de inkoop de nadruk steeds meer komen te liggen op niet-financiële opbrengsten en is het belang van relaties alleen maar toegenomen. Deze opbrengsten kunnen bestaan uit: produktontwikkeling, technische informatie, statusverhoging en het vergroten van zekerheid.

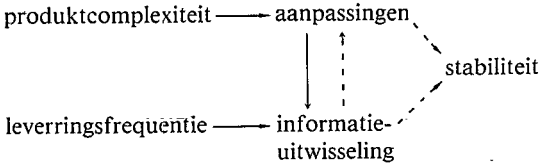
Een auteur die de relatie tussen kopende en verkopende onderneming centraal stelt is Ford (1980). Ford onderscheidt en beschrijft 5 *ontwikkelingsfasen* in de buyer/seller relatie, variërend van pre-relationship tot een stabiele eindfase. Hierna kan eventueel nog een 'dissolutionfase' optreden. Tevens benadrukt Ford het belang van 'relatie-management' met betrekking tot internationale afnemers die immers vaak veel moeite hebben een buitenlandse onderneming goed naar waarde te schatten.
Arndt (1974) wijst op het feit dat ook in de consumentenmarketing relatie-management noodzakelijk kan zijn. Denk hierbij bijvoorbeeld aan boekenclubs, creditcards, 'brand loyalty' en seizoenkaarten voor het voetbal of het theater. Dit is weer een element van het probleem dat in hoofdstuk 2 uitvoerig besproken is, het al dan niet terechte onderscheid tussen consumenten- en industriële marketing.

Men kan zich afvragen wat nu de *stabiliteit* van een industriële relatie bepaalt. Jackson (1985) legt de nadruk op de omschakelingskosten en onderscheidt op deze wijze uitersten:
1. zeer lage omschakelingskosten: transaktiemarketing;
2. zeer hoge omschakelingskosten: relatiemarketing.

De omschakelingskosten zijn echter moeilijk te kwantificeren. Zij kunnen bestaan uit: geld, mensen, activa en procedureveranderingen. In deze zienswijze wordt de 'macht' in de relatie sterk bij de afnemer gelegd wat naar onze mening niet terecht is. Er zijn echter ook benaderingen waarbij stabiliteit op een meer evenwichte wijze bekeken wordt.
Gebaseerd op een grootscheeps Europees onderzoek uit 1981 (IMP group) maken Paliwoda en Druce (1987) een onderscheid tussen de verschillende soorten bindingen die een relatie tussen 2 ondernemingen kunnen verstevigen. Deze soorten bindingen ('bonds') zijn:
1. technologie: aanpassen produktietechnologie;
2. tijd: levertijden en voorraden;
3. kennis: sterkte/zwakte compensatie, werkwijze;
4. sociale binding: vertrouwen, persoonlijke relatie en
5. financiële en wettelijke bindingen: kredietfaciliteiten, gezamenlijke deelnemingen.

De kosten die gepaard gaan met het afbreken van deze bonds (de verloren gegane voordelen inbegrepen) zijn ook op te vatten als omschakelingskosten. Paliwoda en Druce hebben in verschillende bedrijfstakken in Engeland en Frankrijk onderzocht, wat nu als belangrijkste en/of sterkste binding werd ervaren. Zoals verwacht waren hier duidelijke verschillen waar te nemen die zij echter niet eenduidig konden verklaren.
Een andere benadering wordt gevolgd door Hallèn, Johanson en Mohamed (1987). Hallèn e.a. onderzochten de verbanden tussen informatie-uitwisseling, wederzijdse aanpassing en stabiliteit van een relatie. Hun bevindingen staan in figuur 9 samengevat.

Opvallend is dat bij internationale relaties geen wisselwerking bestaat tussen informatieuitwisseling en aanpassingen. Dit impliceert dat het moeilijker is voor een dergelijk contact om te 'groeien'. Bovendien blijkt dat een sterke relatie (hoge mate van aanpassingen) niet noodzakelijk leidt tot een stabiele relatie. Sterkte en stabiliteit zijn blijkbaar verschillende fenomenen!



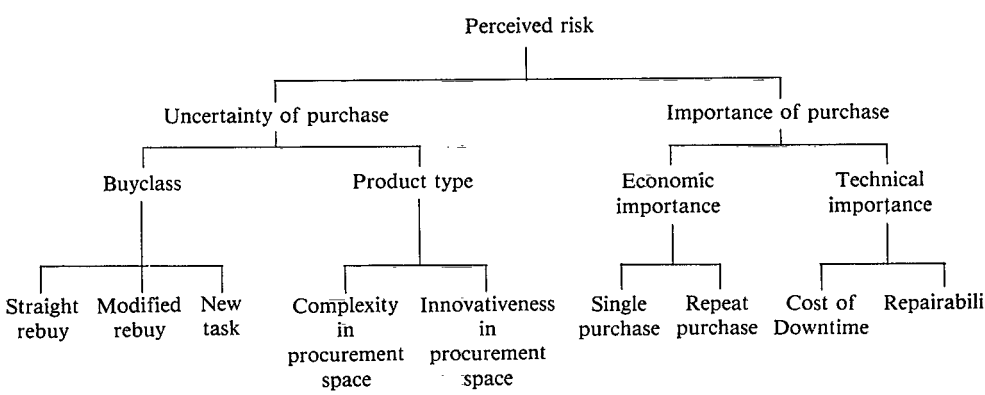
—— : in nationale en internationale steekproef
- - - : alleen in nationale steekproef

Figuur 9. Structuurmodel van nationale en internationale relaties, geschat m.b.v. LISREL. Bron: Hallèn e.a. (1987)

Het ontwikkelen van een *basis* voor relaties kan gezien worden als een strategische keuze. Een andere strategisch keuze die een onderneming moet maken betreft de *leveranciersportfolio*. Een onderneming kan namelijk besluiten zich uitsluitend te richten op het initiëren van nieuwe relaties of op het ontwikkelen van bestaande relaties. Een andere mogelijkheid is te proberen een evenwichtige portfolio te verkrijgen of te behouden. De IMP-group (1982) adviseert dan ook om een ontwikkelingsplan voor de relaties op langere termijn te maken en in het strategisch organisatieplan op te nemen.

8.4. Risicoreductie

Risicoreductie heeft, zoals in 8.1. reeds is aangegeven, zowel taakgerichte als psychologische aspecten. Het taakgerichte element houdt in dat extra informatie gezocht wordt met betrekking tot bepaalde produkten of leveranciers. Het psychologische element berust op beperking van de persoonlijke onzekerheid t.a.v. een bepaalde aankoop. Deze onzekerheid noemt men 'percieved risk' ofwel risicoperceptie. De risicoperceptie is sterk afhankelijk van het zelfvertrouwen van de koper en de (subjectieve) geloofwaardigheid van informatiebronnen. Welke aspecten samen met deze persoonsgebonden factoren de onzekerheid van een aankoop beïnvloeden is weergegeven in figuur 10.



Figuur 10. Elementen van risicoperceptie in het koopcentrum. Bron: Bender (1983).

Perceived risk is opgebouwd uit onzekerheid t.a.v. het resultaat en het belang dat op het spel staat. In figuur 10 worden deze respectievelijk 'uncertainty of purchase' en 'importance of purchase' genoemd. In deze paragraaf zullen een aantal aspecten van risicoreductie besproken worden. Ten eerste is dat de risicohantering van het individu, vervolgens die op organisatienivo en tot slot zullen we aandacht besteden aan de informatiebronnen die gebruikt worden in het koopproces.

Het individu
Hoe *personen* voor zichzelf omgaan met hun risicoperceptie is onderzocht door Hawes en Barnhouse (1987). Zij stellen dat risicoperceptie voor een individu het hoogst is in een modified rebuy situatie (zie 7.3.). Dit is begrijpelijk daar bij een nieuwe aankoop de verantwoordelijkheid vaak op een hoger nivo in de organisatie gedragen wordt, terwijl een regelrechte heraankoop routinematig afgehandeld kan worden. De onderzoeksresultaten in tabel 6 zijn dan ook gebaseerd op een hypothetische modified rebuy.

Tabel 6. Rangordening van i) persoonlijke risico situaties en ii) risico-hanteringsstrategieën. Gebaseerd op: Hawes en Barnhouse (1987)

i) risico situaties rangnummer = f (kans x ernst v.d. situatie)	rang-nummer	ii) risico-hanteringsstrategieën rangnummer = f (belang)
• zelf ontevreden	– 1 –	• persoonlijk fabrieksbezoek
• interne relatie met gebruikers verstoord	– 2 –	• huidige afnemers vragen
• status inkoopafdeling daalt	– 3 –	• multiple sourcing
• persoonlijke beoordeling slechter	– 4 –	• strafclausule in contract
• promotiekansen dalen	– 5 –	• bij collega's vragen

Een aspect van persoonlijke risicoreductie wordt door Jackson (1985, p. 12) als volgt getypeerd:
'For the short run, a manager — may feel considerably safer — with not fixing, something that is not demonstrably broken.'

De organisatie
Over de elementen uit figuur 10 zal *een onderneming* zoveel mogelijk informatie verzamelen (hierbij de kosten niet vergetend!), om het risico van de aankoop beperkt te houden. De informatiebronnen die de onderneming hierbij raadpleegt kunnen variëren in aantal en in soort. De *mate van inspanning* bij het zoek naar informatie over verschillende alternatieven is sterk afhankelijk van 2 dingen:
1. de risicoperceptie;
2. de gevolgde risicoreductiestrategie.

Hierboven zijn we al uitgebreid ingegaan op de elementen die de risicoperceptie kunnen beïnvloeden. Hoe groter deze onzekerheid hoe meer kans dat men bewust of onbewust een van de volgende reductiestrategieën gaat volgen:
a. het ontkennen van het probleem door de beslissing voor zich uit te schuiven;
b. het zoeken van meer informatie over vaak meer alternatieven om zo een beter overwogen beslissing te kunnen nemen;

c. het raadplegen van minder informatiebronnen over minder alternatieven (misschien zelfs maar één) (Zie ook 6.1.1.).

Het verband tussen perceived risk en informatiegedrag is al vaak onderzocht. Gemünden (1985) heeft in een meta-analyse van ca. 100 van deze (deel-) onderzoeken getracht te achterhalen welke achterliggende factoren dit onduidelijke verband beïnvloeden.Gemünden noemt verschillende oorzaken die kunnen zorgen dat meer risico niet altijd tot meer informatieverzameling leidt. Dat wil dus zeggen dat in een dergelijk geval de niet rationele reductiestrategie c gevolgd wordt. Een bekend voorbeeld is dat het informatiezoeken soms slechts dient ter vermindering van de cognitieve dissonantie zodat alleen gezocht wordt naar informatie die de genomen beslissingen rechtvaardigt.

De *soorten informatiebronnen* de geraadpleegd worden zijn door verschillende auteurs onderzocht. Een bekend onderzoek dateert uit 1984 en is uitgevoerd door Moriarty en Spekman in de Verenigde Staten. Het bijzondere van dit project is dat alle leden van het koopcentrum (via 'extensive snowballing')[7] ondervraagd werden over hun informatiegedrag. Vervolgens werden de verschillende informatiebronnen gekoppeld aan karakteristieken van beslissers, van de organisatie, van de koopsituatie en van de fasen in het inkoopproces. Hun voornaamste bevindingen zijn de volgende:

1. *Commerciële persoonlijke* informatie is in alle fasen van het inkoopproces de belangrijkste informatiebron.
2. *Commerciële onpersoonlijke* informatie zoals direct mail wordt vooral gebruikt bij kleinere ondernemingen en met name in de zoekfase.
3. *Niet commerciële persoonlijke* informatie is in elke fase van het inkoopproces van belang.
4. *Niet commerciële onpersoonlijke* informatie zoals onafhankelijke 'rating services' en publicaties zijn vooral voor het hogere personeel van belang.

Uit een recent onderzoek naar het gebruik van informatiebronnen in Nederland blijkt dat catalogi en vakbladen het meest geraadpleegd worden. Het meeste belang wordt echter gehecht aan vertegenwoordigers en collega's uit de eigen onderneming. Bij hoge complexiteit van de aankoop wordt bovendien meer gebruik gemaakt van niet-commerciële bronnen en meer belang gehecht aan persoonlijke informatie-verschaffers (Meijn, 1987).
Tot slot van deze paragraaf bespreken we nog een recent onderzoek naar informatiezoekgedrag bij het kopen van micro-computers (McTavish en Guillery, 1987). De resultaten van dit onderzoek waren verrassend:

1. Er bleek *geen* verband te zijn tussen de fase in het inkoopproces en de aard van de gebruikte bronnen.
2. Het *aantal* gebruikte informatiebronnen *daalde* in de tijd. Een mogelijke verklaring voor deze daling is de overweldigende hoeveelheid informatie die een leek op computergebied over zich heen krijgt, zodat deze wel gedwongen is slechts enkele bronnen serieus te raadplegen.
3. Wél verwacht was de conclusie dat *het belang* van een bepaalde informatiebron verandert gedurende het inkoopproces. Zo worden vertegenwoordigers in alle fasen als informatieverschaffers genoemd, maar slechts in de oriëntatie- en evaluatiefase zijn ze de belangrijkste bron.

Het model van de sociale interactie speelt een rol op verschillende plaatsen binnen het inkoopproces. In de eerste plaats bij de ontmoetingen tussen de *vertegenwoordiger* van de verkoper en de *inkoopfunctionaris*. Uit onderzoek blijkt dat persoonlijke verkoop effectiever is als de verkoper en inkoper veel op elkaar lijken wat betreft persoonlijkheid en achtergrond (dit verschijnsel noemt men ook wel 'similarity'). Na verloop van tijd kan tussen de koper en verkoper een relatie ontstaan. Men kent elkaar en weet wat men aan de ander heeft. Merk op dat deze *persoonlijke* band vaak min of meer emotioneel is en verschilt van de relatie die in paragraaf 8.3. is besproken. Deze laatste is namelijk een relatie tussen *ondernemingen* en heeft rationele achtergronden die soms sterk kunnen lijken op hun psychologische tegenhangers (zoals risicoreductie).
Een tweede bron van interactie die invloed heeft op de uitkomsten van het inkoopproces is die tussen *functionele eenheden* binnen de onderneming. De motieven zijn hier te herleiden tot een streven naar status en macht. In een van oorsprong technisch georiënteerd bedrijf zal bijvoorbeeld de R&D- of produktieafdeling zijn machtspositie ten opzichte van de andere functies willen handhaven en zich steeds met o.a. de inkoop blijven bemoeien.
De laatste bron van interactie is het *koopcentrum*. Binnen dit koopcentrum moeten functionarissen uit verschillende functionele eenheden samen tot een beslissing komen die uiteindelijk van veel variabelen afhankelijk kan zijn. Naar de verschillende kanten van het koopcentrum is al veel onderzoek gedaan.

- In de eerste plaats blijkt de *omvang* van het koopcentrum bij een vergelijkbare aankoop in een grote onderneming groter te zijn dan in een kleine onderneming. Hiernaast is het aantal personen in het koopcentrum afhankelijk van de onzekerheid van de aankoop (zie ook 8.4.).
- Ten tweede varieert ook de invloed van de diverse functionele gebieden. De *samenstelling* van het koopcentrum is bijv. afhankelijk van de grootte van de onderneming, het type produkt, de fase in het aankoopproces, de duur van de betreffende relatie en de gebruikte technologiën.
- Tot slot kunnen binnen het koopcentrum verschillende *dimensies* onderscheiden worden. Deze dimensies zijn bijvoorbeeld:
  1. vertikale betrokkenheid : aantal betrokken hierarchische nivo's;
  2. laterale betrokkenheid : aantal betrokken functionele gebieden;
  3. uitgebreidheid : aantal leden koopcentrum;
  4. verbondenheid : aantal gebruikte communicatielijnen;
  5. centralisatie : aantal gebruikte communicatielijnen via de purchasing manager als % van het totale aantal mogelijke lijnen.

  Daar dergelijke dimensies goed meetbaar zijn leveren zij een bruikbaar instrument om verschillen en/of overeenkomsten tussen bijvoorbeeld goederencategoriën of ondernemingen te bestuderen. Johnston en Bonoma (1981) vergeleken de dimensies van het koopcentrum bij aankopen van kapitaalgoederen resp. van diensten. De verschillen die zij vonden staan in tabel 7 weergegeven.

In een onderzoek uitgevoerd door Meijn (1987) bleek de gemiddelde omvang van het koopcentrum te variëren tussen 2 en 4,28 personen afhankelijk van de produktgroep, maar onafhankelijk van de complexiteit van het produkt.
Bij pogingen om de in deze paragraaf genoemde invloeden te *meten* rijzen verschillen-

Tabel 7. Resultaten onderzoek naar koopcentrumdimensies. De cijfers zijn medianen. Bron: Johnston, Bonoma (1981)

	kapitaalgoederen	diensten	verschillen significant
1. Vertikale betrokkenheid	3 nivo's	2 nivo's	nee
2. Laterale betrokkenheid	4 funkties	3 funkties	ja
3. Uitgebreidheid	7 personen	5 personen	ja
4. Verbondenheid	45%	50%	ja
5. Centralisatie	– sterk fluctuerend	–	–

de problemen. Ten eerste kan de geïnterviewde persoon slechts een subjectieve mening geven. Om toch een zo *objectief* mogelijk *totaalbeeld* te krijgen worden in de literatuur de volgende oplossingen aangedragen:

1. Het gebruik van (extensieve) snowballing (zie noot 5).
2. Het ondervragen van sales managers van leveranciers i.p.v. de BC-leden zelf. Deze sales managers zullen een beter overzicht hebben van het functioneren van een koopcentrum.

Een tweede probleem is dat veel zaken als invloed, procedures en verantwoordelijkheden gedurende het inkoopproces veranderen. Het is dus raadzaam *per fase* informatie te verzamelen en conclusies te trekken.

9. CONCLUSIE

9.1. Een overzicht

In dit artikel is een overzicht gegeven van een aantal soorten modellen met betrekking tot de inkoop van ondernemingen. Binnen deze modellen is een onderscheid gemaakt tussen complexe modellen die een relatief groot aantal variabelen en verbanden verklaren en eenvoudige submodellen. Deze submodellen geven een verklaring voor bepaalde fenomenen, die binnen meer omvattende complexe modellen een rol kunnen spelen.
In paragrafen 4, 5 en 6 worden de verschillende groepen complexe modellen besproken. Dit zijn achtereenvolgends de beslissingsprocesmodellen, de algemene modellen en de interactiemodellen.
In paragraaf 7 wordt getracht de kernelementen van de verschillende conceptuele modellen in een totaalmodel te integreren. Deze elementen zijn: omgeving; individu en interactie. Tevens wordt in deze paragraaf het Buygrid model van Robinson, Faris en Wind (1967) besproken. Alhoewel we stellen dat de Buygrid geen echt 'model' is, heeft het in de industriële marketing een vaste plaats veroverd en is er veel secundair onderzoek op gebaseerd.
Ten slotte worden in paragraaf 8 de submodellen behandeld. hier wordt onderscheid gemaakt tussen taakgerichte, psychologische en gemengde submodellen. Vier belangrijkere submodellen worden in afzonderlijke paragrafen uitgebreid toegelicht.

9.2. *Conclusie*

Industriële inkoopprocessen zijn zeer complex. Naast invloeden uit de organisatie zelf spelen ook de omgeving, het individu en de interactie tussen ondernemingen en personen een rol. Om tot een compleet beeld te komen van het inkoopproces zullen al deze aspecten bekeken moeten worden. Alhoewel een groot aantal mechanismen het industriële koopproces beïnvloedt (de submodellen), zullen deze niet allemaal in dezelfde situatie van belang zijn.
Het probleem bij de constructie van een algemeen inkoopmodel is de onverenigbaarheid van de verschillende modeleigenschappen. Zo is een zeer gedetailleerd model nauwelijks te generaliseren. Evenzo wordt een gedetailleerd én compleet model te complex om hanteerbaar te blijven.

Indien men meer inzicht in het inkoopproces van een bepaalde onderneming of groep ondernemingen wenst te verkrijgen kan men in principe twee wegen bewandelen:
1. Men ontwikkelt een empirisch model dat een nauwkeurige weergave van het beslissingsproces geeft. De methoden die hiervoor gebruikt kunnen worden staan in paragraaf 4 beschreven onder de noemer 'beslissingsprocesmodellen';
2. Men ontwikkelt een conceptueel model dat in de gekozen onderneming(en) getoetst wordt. Dit model zal afgeleid kunnen worden van het in paragraaf 7 beschreven totaalmodel. Men kiest dan o.b.v. de in de literatuur reeds aangetroffen verbanden die variabelen die in de specifieke situatie van belang zijn. Na toetsing kan het model eventueel worden bijgesteld.

Indien men niet zozeer geïnteresseerd is in het inkoopproces zelf maar meer in de invloed van relaties tussen ondernemingen op het inkoopgedrag zijn ook twee benaderingen mogelijk:
1. Men neemt de interactie tussen twee ondernemingen mee als variabele in een groter model.
2. Men gaat uit van het netwerk model en probeert het ondernemingsgedrag te analyseren aan de hand van de positie die de onderneming inneemt in haar omgeving.

Zowel voor de aanbieders van produkten voor industriële afnemers, als voor deze afnemers zelf is het van groot belang meer begrip te krijgen van de factoren die bij het inkoopproces een rol spelen. Op langere termijn zal in vele gevallen de kwaliteit van de relaties met leverancier en afnemer van strategisch belang blijken te zijn. Op korte termijn is het voor vele ondernemingen bittere noodzaak de kosten van de inkoop te beheersen en/of van een betrouwbare toelevering verzekerd te zijn. Het belang van relaties wordt in de recente literatuur steeds vaker benadrukt. Wij willen hier twee voorbeelden geven:
1. Mayer (1983, p. 24): 'The buyer will... demand that performance on price competitiveness, product quality and service remain consistently high over the period of repeat purchases.'
2. Storm (1987) onderscheidt naast de bekende vier P's nog de drie R-en. Deze R-en staan voor reputatie, relatie en ruil. Zoals Storm stelt kunnen deze drie aspecten zorgen voor een bevredigende transactie ('satisfizers'). De vier P's worden echter meer gezien als nevenvoorwaarden die slechts als 'dissatisfizers' kunnen werken wanneer ze niet voldoen.

In dit artikel is herhaaldelijk naar voren gekomen dat verbanden tussen variabelen erg complex kunnen liggen en empirische resultaten niet altijd duidelijk zijn. 'Het Industriële Inkoopmodel' bestaat duidelijk (nog) niet. Om onderlinge vergelijkingen tussen situaties te kunnen maken zullen nog vele empirische onderzoeken moeten volgen. Misschien dat dit zal leiden tot de identificatie van onderliggende dimensies die bepalen welke delen van ons totaalmodel bij welke aankoop van belang zijn.

NOTEN

1. Een uitgebreid verslag van deze literatuurstudie verschijnt najaar 1988 als I.E.O.-Memorandum van de Economische Faculteit te Groningen. Ook voor een complete literatuurlijst wordt naar deze publicatie verwezen.
2. Bij gebruik van hoogwaardige halffabrikaten kan dit percentage nog aanzienlijk hoger liggen.
3. Synoniemen die wij zullen gebruiken zijn buying center (BC) en decision making unit (DMU).
4. In het vervolg zullen we ook wel spreken van IM (industriële marketing) en CM (consumenten marketing).
5. Bij deze techniek wordt bij elke gespreksronde gevraagd welke personen nog meer invloed op het koopproces gehad hebben. Deze personen worden vervolgens ook in het onderzoek betrokken. Stopt men na bijv. vier respondenten dan spreken we van 'snowballing'; wordt doorgegaan tot geen nieuwe namen meer genoemd worden, dan is dit 'extensive snowballing'.
6. De precieze inhoud van de begrippen new task, modified rebuy en straight rebuy zal in 7.3. besproken worden.
7. Zie noot 5.

LITERATUUR

Arndt, J. (1979), Toward a Concept of Domesticated Markets, *Journal of Marketing*, 43, Fall, p. 69-75.

Bender, H.O. (1983), *Industrial Conversion Framework: A Theory of Organizational Marketing*, doctoral thesis, Columbia University.

Brand, M.J. (1988), *Industriële inkoopmodellen; Een literatuurstudie*, I.E.O.-Memorandum Economische Faculteit, R.U. Groningen.

Cardozo, R.N. (1983), Modelling Organizational Buying as a Sequence of Decisions, *Industrial Marketing Management*, 12, p. 75-81.

Crow, L.E., R.W. Olshavsky en J.O. Summers (1980), Industrial Buyers' Choice Strategies: A Protocol Analysis, *Journal of Marketing Research*, 17, no. 2, p. 34-44.

Cunningham, M.T. (1982), International Marketing and Purchasing of Industrial Goods-Features of a European Research Project, *European Journal of Marketing*, 14, 5/6, p. 322-338.

Fern, E.F. en J.R. Brown (1984), The Industrial/Consumer Marketing Dichotomy: A Case of Insufficient Justification, *Journal of Marketing*, 48, Spring, p. 68-77.

Finney, F.R. (1978), Reciprocity: Gone but not Forgotten, *Journal of Marketing*, 42, January, p. 54-59.

Fisher, B.A. (1974), *Small group decision making*, New York: McGrawHill.

Ford, D. (1980), The Development of Buyer-Seller Relationships in Industrial Markets, *European Journal of Marketing*, 14, 5/6, p. 339-353.

Gemünden, H.G. (1985), Perceived Risk and Informations Search. A Systematic Meta-analyses of the Empirical Evidence, *International Journal of Research in Marketing*, 2, no. 2, p. 79-100.

Hallén, L., J. Johanson en N.S. Mohamed (1987), Relationship Strength and Stability in International and Domestic Industrial Marketing, *Industrial Marketing and Purchasing*, 2, no. 3, p. 22-37.