1. Wat bepaalt de effectiviteit van markt-onderzoek informatie: Een gebruikersonderzoek[1]

M. VAN BIRGELEN, K. DE RUYTER en M. WETZELS

SAMENVATTING

Zonder enige twijfel is goede informatie, verkregen via marktonderzoek, van cruciaal belang voor het succes of falen van hedendaagse ondernemingen. Ondanks dit belang van marktonderzoek en de toenemende hoeveelheid informatie die tegenwoordig beschikbaar is, lijkt het echter nog maar de vraag of en op welke manier hiervan door het management van organisaties daadwerkelijk gebruik wordt gemaakt bij het nemen van beslissingen. In dit artikel wordt het effectief gebruik van marktonderzoek informatie door organisaties nader onder de loep genomen en wordt een poging ondernomen dit gebruik door managers en factoren die het beïnvloeden in kaart te brengen. Hierbij wordt het uitvoeren van marktonderzoek door marktonderzoeksbureaus als een dienst beschouwd die verleend wordt aan een opdrachtgever. Door middel van een empirisch onderzoek onder NVMI leden die regelmatig marktonderzoek laten uitvoeren wordt een aantal hypothesen betreffende het effectief gebruik van marktonderzoek informatie getoetst.

1. INLEIDING

Tegenwoordig is het gebruikelijk dat zowel profit- als non-profit organisaties uitgebreid informatie verzamelen over de markt waarop ze actief zijn. Organisaties huren veelal een marktonderzoeksbureau in om marktonderzoek voor hen uit te voeren en hen te voorzien van informatie en aanbevelingen waarmee ze aan de slag kunnen gaan. Het is duidelijk dat informatie, verzameld door het doen van marktonderzoek, belangrijk is voor bedrijven. Tegenwoordig hangt het succes van organisaties voor een groot deel af van de mogelijkheid om te beschikken over informatie op basis waarvan toekomstige activiteiten gepland kunnen worden. Effectief gebruik van marktinformatie stelt organisaties in staat om meer markt- en klantgericht te opereren en hun kans van slagen te vergroten in markten waar de concurrentie steeds heviger wordt. Reeds in de zestiger jaren stelde Newman (1962) dat indien een bedrijf voordeel wil halen uit marktinformatie, zij deze informatie zal moeten omzetten in beslissingen en vervolgens haar activiteiten moet afstemmen op deze informatie. Dit lijkt een open deur. De praktijk laat echter zien dat ondanks de erkenning van het belang van marktonderzoek informatie en de enorme stroom informatie die verzameld wordt, het niet altijd duidelijk is in welke

[1] Met dank aan Suzanne Keularts voor haar hulp bij het verzamelen van de onderzoeksgegevens.

mate managers daadwerkelijk gebruik maken van deze informatie, bijvoorbeeld tijdens het nemen van beslissingen. Vaak lijkt het alsof managers een snelle blik op marktonderzoek rapporten werpen en ze daarna wegstoppen in een bureaulade zonder er serieus gebruik van te maken. Om deze reden is het interessant om aandacht te besteden aan factoren die de effectiviteit van marktonderzoek informatie bepalen vanuit een gebruikersperspectief. Het feit dat het gerenommeerde Marketing Science Institute (MSI) 'Management use of information and market research' op het prioriteitenlijstje heeft gezet is waarschijnlijk niet zonder reden (MSI, 1996). Dit betekent dat niet alleen in de marktonderzoek praktijk, maar ook in de academische wereld het belang van onderzoek naar het gebruik van marktkennis en -informatie en het effectief gebruik hiervan als belangrijk gezien wordt. De vraag is nu wat het gebruik van marktonderzoek informatie precies inhoudt en welke factoren dit gebruik doen toe- dan wel afnemen.
Marktonderzoek dat in opdracht van een bedrijf wordt verricht door een marktonderzoeksbureau kan gezien worden als een dienst die aan deze opdrachtgever verleent wordt. Hierbij is het goed mogelijk dat indien deze dienst correct geleverd wordt er wellicht meer opdrachten zullen volgen en er daardoor ook op de langere termijn een relatie tussen organisatie en onderzoeksbureau kan ontstaan. In dit artikel zullen we het uitvoeren van marktonderzoek dan ook benaderen vanuit een dienstenmarketing en een relatiemarketing perspectief, waarbij factoren als tevredenheid en vertrouwen van doorslaggevend belang zijn. Hiervoor is het allereerst belangrijk om te kijken naar wat het gebruik van marktonderzoek informatie door opdrachtgevers nu eigenlijk inhoudt en vervolgens naar welke factoren het effectief gebruik van deze informatie kunnen beïnvloeden.
Dit artikel is als volgt opgebouwd. Ten eerste zal nader worden ingegaan op het gebruik van marktonderzoek informatie. Dit zal plaatsvinden aan de hand van een bespreking van de bestaande literatuur over dit onderwerp. Vervolgens zullen verschillende factoren die van invloed zijn op het gebruik van marktonderzoek informatie worden besproken. Dit stelt ons in staat om een aantal hypothesen met betrekking tot effectief informatiegebruik te formuleren en een conceptueel raamwerk te ontwikkelen. Na het empirisch onderzoek besproken te hebben volgt tenslotte aan de hand van de resultaten een conclusie met een aantal theoretische en praktische implicaties.

2. WAT IS INFORMATIE GEBRUIK NU EIGENLIJK?

In de laatste twee decennia is in de academische literatuur veel aandacht besteed aan het gebruik van marktonderzoek informatie door organisaties (Deshpandé en Zaltman, 1982, 1984, 1987; Lee et al., 1987; Krishnan, 1989; Raphael en Parket, 1991; Moorman et al., 1992, 1993; Menon en Wilcox, 1994; Maltz en Kohli, 1995). Dit geeft duidelijk aan dat het gebruik van marktonderzoek informatie van enorm belang lijkt te zijn voor succesvolle bedrijfsactiviteiten. Zoals reeds gezegd, zijn bedrijven die proberen hun concurrentie positie te versterken en meer marktgericht te opereren, sterk afhankelijk van effectief gebruik van markt informatie. Het management ziet haar omgeving met hoge snelheid veranderen. Dit resulteert in een hoge mate van onzekerheid, onduidelijkheid en dynamiek (Deshpandé en Zaltman, 1982). Marktonderzoek informatie is van groot belang om een dergelijke situatie het hoofd te kunnen bieden.
Als we de academische literatuur erop naslaan wordt al snel duidelijk dat het begrip van wat gebruik van onderzoeksinformatie nu eigenlijk inhoudt niet eenduidig is. John

en Martin (1984) zien informatie gebruik als de mate waarin een rapport direct gebruikt wordt voor het sturen van gedrag en het nemen van beslissingen en hebben het derhalve over een gedragsmatige manier van informatie gebruik. Maltz en Kohli (1995) beschouwen informatie gebruik als de mate waarin de ontvanger (lees de opdrachtgever) de informatie, verspreidt door de zender (lees het marktonderzoeksbureau), gebruikt om zijn/haar werkomgeving te begrijpen en om beslissingen te nemen en te implementeren. Naast een gedragsmatige vorm houden deze auteurs dus ook rekening met een meer cognitieve dimensie van informatie gebruik, gericht op een vergroting van kennis. Tenslotte zien Menon en Wilcox (1994) informatie gebruik vanuit drie verschillende dimensies: (a) actie georiënteerd gebruik, gericht op het veranderen van concrete activiteiten, gebruiken of beleid, (b) kennis vergrotend gebruik dat kennis en begrip van de onderzochte onderwerpen verandert, en (c) affectief gebruik door managers om 'zich goed te voelen', bijvoorbeeld ter bevestiging van eerder genomen beslissingen. Hier wordt dus naast een gedragsmatige en cognitieve dimensie nog rekening gehouden met een affectieve vorm van informatie gebruik. Deze ontwikkeling van gedrag naar cognitie naar affectie is ondertussen reeds op velerlei onderzoeksgebieden waar te nemen geweest.

De verschillende vormen van informatie gebruik kunnen worden samengevat in de volgende, meest gebruikte, typologie: instrumenteel gebruik, conceptueel gebruik, en symbolisch gebruik. Instrumenteel gebruik van informatie heeft betrekking op de directe toepassing van onderzoeksbevindingen en conclusies met als doel een bepaald probleem op te lossen (Deshpandé en Zaltman, 1982). Hierbij valt bijvoorbeeld te denken aan de informatie die nodig is bij de introductie van een nieuw produkt. In dit geval bestaat er dus een concreet probleem en de oplossing daarvan is afhankelijk van de informatie die verzameld is door het uitvoeren van een marktonderzoek. Volgens Menon en Wilcox (1994) kan instrumenteel gebruik verder onderverdeeld worden in zogenaamd consistent (congruous) en inconsistent (incongruous) instrumenteel gebruik. Consistent instrumenteel gebruik houdt in dat de informatie gebruikt wordt in lijn met het doel en de implicaties van de onderzoeksbevindingen. Een voorbeeld hiervan kunnen de aanpassingen aan het nieuwe produkt zijn aan de hand van de meningen van een aantal consumenten die hebben deelgenomen aan een onderzoek. Inconsistent instrumenteel gebruik heeft betrekking op een opzettelijk verstoord gebruik. Dit laatste type van informatie gebruik kan bijvoorbeeld optreden als informatie uit de context gehaald of vervormd wordt voor politieke doeleinden en vervolgens instrumenteel gebruikt wordt bij het vaststellen van beleid of strategie.

Indien informatie niet voor directe actie in een gegeven situatie wordt gebruikt maar ter vergroting van het inzicht, bijvoorbeeld in de markt, wordt dit conceptueel gebruik genoemd (Deshpandé en Zaltman, 1982). Menon en Wilcox (1994) noemen deze vorm van gebruik kennis vergrotend gebruik, zowel van het eindprodukt van het onderzoek in de vorm van de verkregen informatie, als van het onderzoeksproces zelf, bijvoorbeeld omdat managers dit proces leerzaam en interessant vinden. Door kennis vergrotend gebruik verandert het begrip over de onderwerpen en thema's van een bepaald onderzoek.

Tenslotte wordt informatie soms gebruikt op een manier die afwijkt van het doel van het onderzoek. Dit heet dan symbolisch gebruik, hetgeen in tegenstelling staat tot instrumenteel en conceptueel gebruik waarbij de bevindingen gebruikt worden op een manier die overeenkomt met het gestelde onderzoeksdoel. Symbolisch gebruik doet zich voor als informatie gebruikt wordt door conclusies uit de context te halen, bijvoor-

beeld om bepaalde acties die om andere redenen ondernomen zijn te verdedigen of te onderbouwen, of indien conclusies slechts selectief gebruikt worden (Beyer en Trice, 1982). Om nu het effectief gebruik van marktonderzoek informatie door organisaties beter te kunnen begrijpen en verklaren is het belangrijk om aandacht te schenken aan factoren die dit gebruik beïnvloeden. De volgende paragraaf gaat aan de hand van de bestaande academische literatuur op dit gebied dieper op enkele van deze factoren in.

3. FACTOREN DIE HET GEBRUIK VAN MARKTONDERZOEK INFORMATIE BEINVLOEDEN

Het uitvoeren van een marktonderzoek voor een opdrachtgever kan als een dienst worden beschouwd. Net zoals bij andere typen van dienstverlening kan er een onderscheid worden gemaakt in technische en functionele kwaliteitsaspecten van dienstverlening (Grönroos, 1990). Technische kwaliteit heeft dan betrekking op de uitkomst van de dienstverlening, dus het resultaat voor de opdrachtgever. Functionele kwaliteit is gerelateerd aan het proces waarmee de dienst wordt verleend, dus de wijze waarop de opdrachtgever dit resultaat verkrijgt. Deze vormen van kwaliteit worden bepaald door een aantal zaken waaraan reeds in het verleden in verschillende studies gedeeltelijk aandacht is besteed in de context van marktonderzoek informatie gebruik. Deshpandé en Zaltman (1982, 1987) concentreerden zich in hun studies bijvoorbeeld voornamelijk op het instrumenteel gebruik van informatie. Zij onderscheidden verschillende factoren die gerelateerd konden worden aan het instrumenteel gebruik van marktonderzoek informatie door managers. Deze factoren waren o.a. het doel van het onderzoeksproject, de organisatie structuur van het bedrijf waarin de manager werkzaam is, karakteristieken van het onderzoeksrapport, de mate van verrassing over de resultaten van het onderzoek, de fase in de levenscyclus van het produkt dat het object vormt van de studie en de mate van interactie tussen managers en onderzoekers. Enkele van deze factoren kunnen mede bepalend zijn voor de mate van technische kwaliteit, bijvoorbeeld de karakteristieken van het onderzoeksrapport, en functionele kwaliteit, bijvoorbeeld de wijze waarop het onderzoeksbureau en de opdrachtgever met elkaar omgaan. Dit laatste is vaak sterk afhankelijk van bepaalde interpersoonlijke kenmerken van het onderzoeksbureau, zoals integriteit en flexibiliteit.

Het leveren van kwaliteit zal uiteindelijk moeten leiden tot tevredenheid van de opdrachtgever. Tevredenheid kan worden omschreven als de cognitieve en affectieve evaluatie door de consument van een produkt of dienst dat aan hem/haar geleverd wordt door een specifieke aanbieder (Oliver, 1996). Met andere woorden, tevredenheid komt dus voort uit een bepaalde mening en gevoel over een produkt of een dienst. Hanan en Karp (1989) stellen dat bevrediging van de behoeften die klanten ertoe bewegen om zaken te doen het ultieme doel is van iedere organisatie. Zonder tevredenheid over het onderzoek zal het waarschijnlijk voor het laatst zijn dat een bureau een bepaalde opdracht heeft mogen uitvoeren voor een klant. Tevredenheid kan dan ook worden gezien als een belangrijk criterium waaraan voldaan moet worden wil een zakelijke relatie in de toekomst voortgezet worden. Een relatie tussen twee partners kan alleen blijven voortbestaan als zich een bepaalde mate van vertrouwen tussen beide partijen heeft ontwikkeld, zoals uit eerder onderzoek is gebleken (Wetzels et al., 1998). Tevredenheid over het onderzoek zal derhalve bepalend zijn voor de mate van vertrouwen dat de opdrachtgever heeft in het marktonderzoeksbureau. Tevredenheid over een af-

zonderlijke transactie tussen marktonderzoeksbureau en opdrachtgever vormt als het ware een bouwsteen voor vertrouwen van de klant in het bureau op de langere termijn. Vertrouwen is bepalend voor de relatie tussen een opdrachtgever en het bureau en kan dan ook gezien worden als een voorwaarde voor een eventuele volgende opdracht. Naast vertrouwen in het onderzoeksbureau op de langere termijn, zal tevredenheid op korte termijn natuurlijk dienen te resulteren in acceptatie en daadwerkelijk gebruik van de geleverde informatie. Dit betekent dat tevredenheid over het onderzoek dus van cruciaal belang is, zowel voor de korte als de lange termijn relatie tussen opdrachtgever en onderzoeksbureau.
Eerder onderzoek heeft verder aangetoond dat de mate van vertrouwen die organisaties hebben in de onderzoekers ook een significante invloed heeft op het informatie gebruik. Verschillende studies concentreerden zich op deze relatie tussen vertrouwen en het gebruik van informatie (Moorman et al., 1992; Moorman en Austin, 1995; Morgan en Hunt, 1994). Eén van de conclusies van Moorman et al. (1993) is dat de mate van vertrouwen die gebruikers hebben in de leveranciers van informatie bepaalt of bedrijven de informatie alleen maar bezitten of deze ook daadwerkelijk gebruiken. Verder werd er ook geconcludeerd dat de mate van vertrouwen van invloed is op de waargenomen kwaliteit van de gebruiker-leverancier interacties, de mate van betrokkenheid van de onderzoeker en de betrokkenheid van de gebruiker bij de relatie (Moorman et al., 1992). Vertrouwen werd beïnvloed door verschillende interpersoonlijke kenmerken van de onderzoeker, in plaats van dat het afhankelijk is van individuele factoren (Moorman et al., 1993). Volgens dit onderzoek wordt de mate van vertrouwen tussen gebruikers en onderzoekers o.a. significant bepaald door de expertise, de integriteit, de vertrouwelijkheid, de stiptheid, de mate waarin de gebruiker bij een onderzoek wordt betrokken en de oprechtheid van de onderzoekers. Het is dus duidelijk dat een aantal factoren een belangrijke rol spelen bij het effectief gebruik van marktonderzoek informatie. De grote vraag is nu hoe deze factoren zich tot elkaar verhouden. Om dat uit te werken gaan we in de volgende paragraaf over tot het opstellen van een aantal hypothesen.

4. DE ONTWIKKELING VAN HYPOTHESEN

In deze paragraaf zullen op basis van het voorgaande overzicht van wat gebruik van marktonderzoek informatie nu eigenlijk inhoudt en welke factoren hieraan gerelateerd kunnen worden een aantal hypothesen geformuleerd worden. Deze hypothesen kunnen dan vervolgens worden samengevat in een conceptueel raamwerk. In dit onderzoek wordt gebruik van marktonderzoek informatie door organisaties beschouwd als de mate waarin de opdrachtgever de door het marktonderzoeksbureau geleverde informatie gebruikt om zijn werkomgeving te begrijpen en voor het nemen en implementeren van beslissingen. Hierbij richten we ons dus op zowel conceptueel als instrumenteel gebruik van marktonderzoek informatie. Deze uitwerking van gebruik is gebaseerd op de overeenstemming tussen verschillende studies naar gebruik van onderzoek, dat er iets goeds is aan onderzoeksinformatie die gebruikt wordt, en dat er iets mis is met onderzoeksinformatie die niet gebruikt wordt (Menon en Varadarajan, 1992; Dunn, 1986).
Omdat klanttevredenheid gezien kan worden als één van de belangrijkste variabelen in relatie management neemt het een centrale plek in ons onderzoek in. Zoals we gezien hebben zal tevredenheid over afzonderlijke transacties dienen te leiden tot gebruik van

de informatie en tot een hogere mate van vertrouwen in het onderzoeksbureau op de langere termijn. Verder wordt verwacht dat naast tevredenheid over het onderzoek ook vertrouwen een groter gebruik van de informatie tot gevolg zal hebben. Om deze redenen formuleren we de volgende hypothesen:

H_1: Er bestaat een positieve relatie tussen tevredenheid over het onderzoek en het gebruik van informatie.
H_2: Er bestaat een positieve relatie tussen tevredenheid over het onderzoek en het vertrouwen in het marktonderzoeksbureau.
H_3: Er bestaat een positieve relatie tussen vertrouwen in het marktonderzoeksbureau en het gebruik van informatie.

Technische kwaliteit van de dienstverlening kan worden bepaald door een aantal factoren, die via tevredenheid het gebruik van informatie beïnvloeden. Deze factoren zijn o.a. inhoudelijke kwaliteit, kwaliteit van de presentatie, de mate waarin acties naar aanleiding van de informatie genomen kunnen worden (toepasbaarheid) en de mate waarin de resultaten aanvaardbaar zijn voor de opdrachtgever. Inhoudelijke kwaliteit heeft betrekking op zaken die door de ogen van opdrachtgever de totale kwaliteit van het marktonderzoek vormen, zoals de gehanteerde onderzoekstechniek, het gebruik van wetenschappelijke meetinstrumenten, de structuur van de data verzameling en de prijswaarde afweging (Rich, 1979). Menon en Varadarajan (1992) suggereren dat de perceptie van de kwaliteit van de informatie de mate beïnvloedt waarin de ontvanger activiteiten onderneemt op basis van die informatie. Wij verwachten dat een hogere kwaliteit van de inhoud resulteert in een grotere klanttevredenheid en dus indirect in een hoger gebruik van de informatie:

H_4: Er bestaat een positieve relatie tussen inhoudelijke kwaliteit en tevredenheid over het onderzoek.

Niet alleen de kwaliteit van de inhoud, maar ook de kwaliteit van de presentatie, schriftelijk dan wel mondeling, zal van invloed zijn op het gebruik van de door het onderzoeksbureau geleverde informatie. De fysieke representatie van het marktonderzoek is vaak het eindrapport, dat schriftelijk en/of mondeling door de onderzoekers aan de managers wordt gepresenteerd (Deshpandé en Zaltman, 1982). Deshpandé en Zaltman (1984) vonden empirisch bewijs voor het belang van kwaliteit van de presentatie voor het effectief gebruik van marktonderzoek informatie. Om deze reden stellen we de volgende hypothese op:

H_5: Er bestaat een positieve relatie tussen kwaliteit van de presentatie en tevredenheid over het onderzoek.

Informatie en aanbevelingen naar aanleiding van een onderzoek moeten toepasbaar zijn. De waarde van onderzoeksinformatie wordt hoger als managers ervan overtuigd zijn dat er naar aanleiding van de informatie concrete acties ondernomen kunnen worden en aanbevelingen in de praktijk geïmplementeerd kunnen worden (Deshpandé en Zaltman, 1982). Toepasbaarheid kan worden gezien als de mate waarin de informatie door organisaties als praktisch implementeerbaar beschouwd wordt.

H6: Er bestaat een positieve relatie tussen de toepasbaarheid van de informatie en tevredenheid over het onderzoek.

Aanvaardbaarheid kan worden gedefinieerd als de compatibiliteit van de marktonderzoek informatie met de doelstellingen van de organisatie. Volgens Kohli en Jaworski (1990) kan de bereidheid om informatie te gebruiken een functie zijn van factoren zoals de politieke aanvaardbaarheid van de informatie en de mate waarin het de status-quo aantast. Wij verwachten dat een hogere mate van aanvaardbaarheid van de informatie resulteert in hogere tevredenheid van de opdrachtgever en dus indirect in hoger informatie gebruik:

H7: Er bestaat een positieve relatie tussen de aanvaardbaarheid van de informatie en tevredenheid over het onderzoek.

Functionele kwaliteit van de dienstverlening wordt met name bepaald door de kenmerken van het marktonderzoeksbureau. Deze kenmerken kunnen gedefinieerd worden als een breed scala van kwaliteiten die gedemonstreerd worden in de interacties tussen opdrachtgever en onderzoeksbureau (Moorman et al., 1993). Belangrijke kwaliteiten zijn expertise, integriteit, flexibiliteit, en vertrouwelijkheid. Moorman et al. (1993) en Griffin (1986) vonden een empirisch verband tussen deze vier kwaliteiten en vertrouwen. Wij veronderstellen dat dit verband niet rechtstreeks maar via tevredenheid zal verlopen:

H8: Er bestaat een positieve relatie tussen de kenmerken van het marktonderzoeksbureau en tevredenheid over het onderzoek.

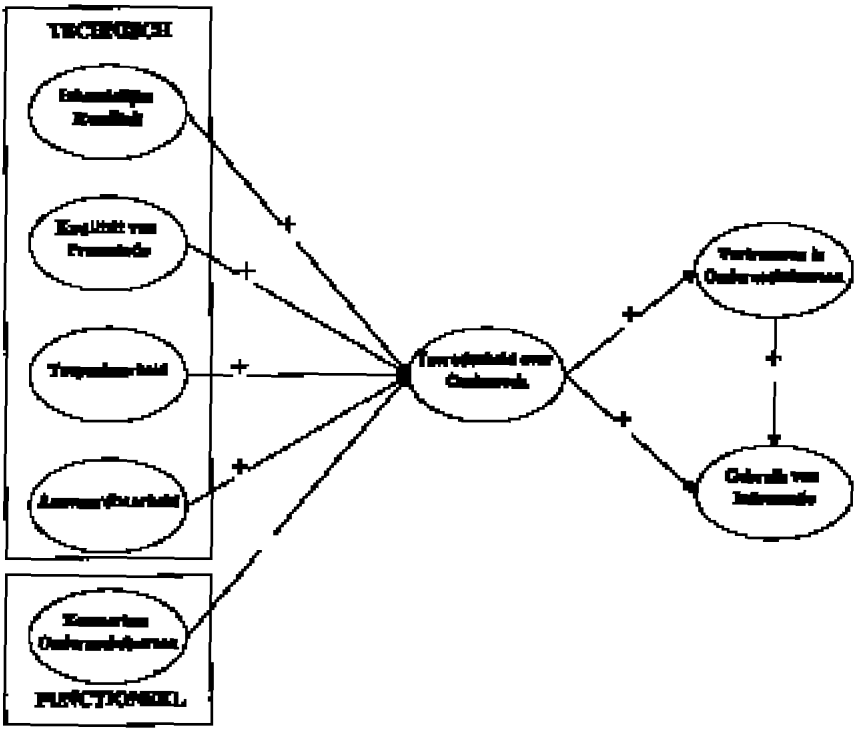
Bovenstaande hypothesen kunnen worden samengevat door middel van het raamwerk dat wordt weergegeven in figuur 1.

In de volgende paragraaf zal het onderzoek besproken worden dat is verricht om de bovengenoemde hypothesen empirisch te toetsen.

5. EEN EMPIRISCH ONDERZOEK

5.1 Onderzoeksontwerp en gegevensverzameling

De populatie voor dit onderzoek bestond uit managers van organisaties in zowel de consumenten als de industriële markt. Deze managers moesten gebruikers van marktonderzoek zijn dat door externe, onafhankelijke marktonderzoeksbureaus werd verricht. Het ledenbestand van de Nederlandse Vereniging voor Marktonderzoek en Informatiemanagement (NVMI) vormde de lijst waaruit de uiteindelijke steekproef van respondenten is getrokken. Het is duidelijk dat dit niet zal resulteren in een volledig representatieve steekproef. Echter voor dit onderzoek waarin we op zoek zijn naar theoretische generaliseerbaarheid en we willen kijken of verbanden op basis van de theorie in de praktijk voorkomen lijkt representativiteit niet noodzakelijk. Alle organisaties op de lijst werden telefonisch benaderd met het verzoek om deel te nemen aan het onderzoek. Respondenten die geschikt en bereid waren om deel te nemen, werd



Figuur 1. Conceptueel raamwerk van informatie gebruik.

vervolgens een schriftelijke vragenlijst opgestuurd. Deze vragenlijst moest dan ingevuld worden naar aanleiding van één specifiek marktonderzoek dat de organisatie recentelijk had laten uitvoeren door een extern marktonderzoeksbureau. In totaal retourneerden 140 respondenten de vragenlijst. Na een kritische beschouwing van de teruggestuurde vragenlijsten, waarbij o.a. gelet werd op incomplete vragenlijsten, bleven er uiteindelijk 120 respondenten over. Op de antwoorden van deze 120 respondenten hebben we een test uitgevoerd waarmee je kunt controleren of er verschillen bestaan tussen respondenten die hun ingevulde vragenlijst snel terugsturen en respondenten die hier langer over doen. De aanname bij deze test is dat respondenten die minder snel reageren meer lijken op niet-respondenten (Armstrong en Overton, 1977). Het bleek dat er op geen van de variabelen in de vragenlijst significante verschillen bestonden tussen vroege en late respondenten en ons databestand dus niet onderhevig is aan respons-gerelateerde problemen.

De vragenlijst bestond uit een aantal gedeelten met daarin vragen die de verschillende constructen in het conceptuele raamwerk meten. De items werden gemeten met een zeven-punts Likert-schaal die varieerde van 1 = helemaal mee oneens tot 7 = helemaal mee eens. De technische kwaliteit constructen van ons model, inhoudelijke kwaliteit, kwaliteit van de presentatie, toepasbaarheid en aanvaardbaarheid werden geoperationaliseerd op basis van de schalen, zoals deze door Deshpandé en Zaltman (1982) gebruikt werden. Om de kenmerken van het marktonderzoeksbureau te meten werd gebruik gemaakt van verscheidene items van Moorman et al. (1993). Sommige items van de gebruikte schalen moesten enigszins aangepast worden voor het specifieke doel van dit

onderzoek. De schaal voor tevredenheid over het onderzoek werd speciaal ontwikkeld voor dit onderzoek en vertrouwen in het onderzoeksbureau werd gemeten middels de schaal van Moorman et al. (1992). Als laatste werd informatie gebruik gemeten door middel van de schaal van Maltz en Kohli (1995), die ook aangepast werd voor het doel van dit onderzoek. Ter illustratie worden in onderstaande tabel 1 een aantal voorbeelditems voor de verschillende constructen weergegeven.

5.2 De resultaten

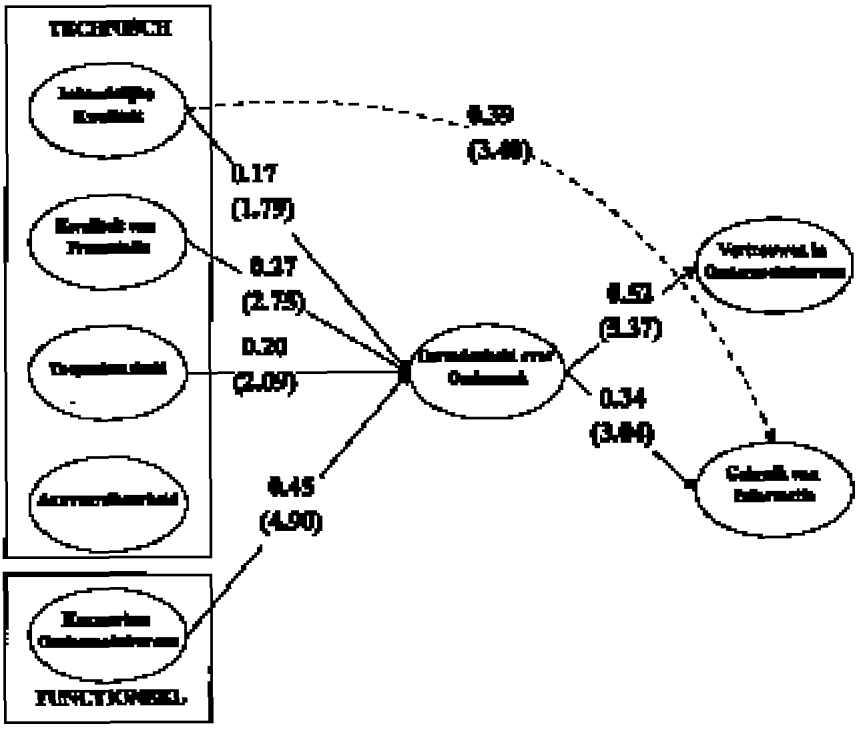
Het conceptuele model van gebruik van marktonderzoek informatie en de corresponderende hypothesen zijn getoetst door gebruik te maken van Structural Equation Modeling (SEM) middels LISREL 8. Met behulp van deze techniek is het mogelijk een zogenaamde padanalyse uit te voeren die laat zien welke constructen significant aan elkaar gerelateerd zijn en in welke mate, vergelijkbaar met de regressiecoëfficiënten die verkregen worden via een multivariate regressieanalyse. Alvorens echter de resultaten van deze padanalyse te bespreken dienen eerst de betrouwbaarheden van de verschillende constructen geëvalueerd te worden. De resultaten van de betrouwbaardheidsanalyse worden samen met een aantal voorbeelditems weergegeven in tabel 1.

Tabel 1. Betrouwbaarheden en voorbeelditems van de constructen.

Construct	*Betrouwbaarheid (α)*	*Voorbeelditem*
Inhoudelijke kwaliteit	0.69	In het onderzoek werden de onderzoekstechnieken goed toegepast.
Kwaliteit van presentatie	0.68	Het taalgebruik van/tijdens de presentatie was duidelijk.
Toepasbaarheid	0.84	Alle aanbevelingen konden in de praktijk verwerkt worden.
Aanvaardbaarheid	0.67	De aanbevelingen van het marktonderzoek leken 'politiek' aanvaardbaar.
Kenmerken onderzoeksbureau	0.82	Het marktonderzoeksbureau was zeer deskundig.
Tevredenheid over onderzoek	0.89	Ik ben zeer tevreden met de inhoud van het onderzoek.
Vertrouwen in onderzoeksbureau	0.76	Ik vertrouw het onderzoeksbureau zaken toe die ik zelf niet kan doen.
Gebruik van informatie	0.69	De informatie uit het onderzoek zal tot gevolg hebben dat ons bedrijf concrete acties zal ondernemen.

Voor alle constructen blijkt de Cronbach's Alpha rond de kritieke waarde van 0.70, zoals voorgesteld door Nunnally en Bernstein (1994), te liggen of deze te overschrijden. Dit duidt op voldoende betrouwbaarheid van de gebruikte meetinstrumenten.
Na schatting van het model bleken de verschillende 'fit' maatstaven goed te noemen: ($\chi^2_{(11)}$ = 18.82 (p = 0.06), GFI = 0.97, AGFI = 0.89, RMR = 0.05, TLI = 0.91, CFI = 0.97. De resultaten van de uiteindelijke padanalyse worden gepresenteerd in

figuur 2. De coëfficiënten zijn gestandaardiseerd en tussen haakjes staan tevens de bijbehorende t-waarden vermeld. Alleen de significante relaties zijn in de figuur opgenomen ($\alpha = 0.05$).



Figuur 2. Eindmodel van informatie gebruik.

Zoals uit figuur 2 geconcludeerd kan worden bestaat er een significante positieve relatie tussen de mate van tevredenheid van de gebruiker over de inhoud, presentatie en het onderzoeksbureau en het gebruik van de informatie die door middel van het onderzoek verzameld is. Hypothese 1 wordt dan ook empirisch ondersteund in deze studie. In navolging van hypothese 2 wordt er tevens een positieve relatie gevonden tussen tevredenheid en het vertrouwen in het onderzoeksbureau, iets wat essentieel is om een lange termijn relatie tussen onderzoeksbureau en informatie gebruiker te bewerkstelligen. Hypothese 3 blijkt echter niet ondersteund te worden: er bestaat geen significant verband tussen de mate van vertrouwen in het onderzoeksbureau en informatiegebruik. Wanneer gekeken wordt naar de technische kwaliteit van de dienstverlening in de vorm van een marktonderzoek blijken er ook een drietal significante verbanden te bestaan. Allereerst blijken hypothesen 4 en 5 niet verworpen te kunnen worden: het leveren van zowel inhoudelijke kwaliteit bij een marktonderzoek als zorg dragen voor kwaliteit van de presentatie draagt inderdaad significant bij tot een hogere tevredenheid van de informatiegebruiker. Indien de opdrachtgever verder het idee heeft dat de informatie en aanbevelingen die hem door het onderzoeksbureau aangereikt worden voldoende toepasbaar en implementeerbaar zijn zal dit ook resulteren in een hogere tevredenheid, zoals weergegeven door hypothese 6. Alleen de mate van aanvaardbaarheid en overeenstem-

ming met organisatiedoelstellingen blijkt niet te leiden tot een significante toename in de tevredenheid over het onderzoek. Hypothese 7 dient derhalve verworpen te worden. Verder kan geconcludeerd worden dat ook functionele, proces-gerelateerde, kwaliteit van de dienstverlening belangrijk is bij de uitvoering van een marktonderzoek. Zoals hypothese 8 doet vermoeden zullen interpersoonlijke eigenschappen van de kant van het marktonderzoeksbureau inderdaad de tevredenheid van de gebruiker over het uitgevoerde onderzoek significant doen toenemen. Naast deze verbanden is er bij het schatten van het model op basis van modificatie indices ook nog een andere relatie gevonden die niet terugkomt in de geformuleerde hypothesen. Inhoudelijke kwaliteit blijkt namelijk ook direct het gebruik van marktonderzoek significant positief te beïnvloeden. Klaarblijkelijk is de rol van tevredenheid in dit geval van ondergeschikt belang.

6. CONCLUSIE

6.1 Interpretatie van de resultaten

Opvallend aan de resultaten van dit onderzoek is dat gebruik van marktonderzoek informatie door organisaties het sterkst gerelateerd is aan de inhoudelijke kwaliteit van het onderzoek en de informatie. Deze relatie verloopt rechtstreeks en niet via de tevredenheid van de informatiegebruiker, een resultaat dat de eerdere bevindingen van Menon en Varadarajan (1992) bevestigt. Het lijkt erop dat het verrichten van inhoudelijk kwalitatief hoogwaardig onderzoek van cruciaal belang is voor marktonderzoeksbureaus om ervoor te zorgen dat hun informatie daadwerkelijk door de opdrachtgever gebruikt wordt, hetzij voor het ondernemen van concrete acties (instrumenteel gebruik), hetzij om een bepaalde situatie beter te begrijpen (conceptueel gebruik). Indien de klant het idee krijgt dat de informatie met grote zorgvuldigheid en betrouwbaarheid verkregen is, doet zijn of haar tevredenheid over zaken als de presentatie of de kenmerken van het bureau minder ter zake. Dit wil echter absoluut niet zeggen dat tevredenheid onbelangrijk is. Integendeel, ook klanttevredenheid zal het gebruik van de informatie zeker doen toenemen. Aangezien inhoudelijke kwaliteit tevens een positief effect heeft op tevredenheid, is een focus op deze vorm van technische kwaliteit dubbel zo belangrijk.

Tevredenheid heeft niet alleen een significante invloed op het gebruik van de informatie, het zal tevens leiden tot een andere factor, die belangrijk is voor het opbouwen van een duurzame relatie, namelijk vertrouwen in het onderzoeksbureau. Zorg dragen voor tevreden opdrachtgevers zal derhalve zowel op korte als lange termijn zijn vruchten afwerpen. Op de korte termijn omdat de informatie daadwerkelijk gebruikt zal worden, op de lange termijn omdat het vertrouwen van de klant in het onderzoeksbureau zal toenemen Dit laatste is des te belangrijker omdat uit eerder onderzoek naar relaties tussen zakelijke partners is gebleken dat een hogere mate van vertrouwen in een partner er significant toe bijdraagt dat een klant van plan is om de relatie met deze partner voort te zetten (Wetzels et al., 1998). Het feit dat we tijdens dit onderzoek geen significant verband vinden tussen vertrouwen in het onderzoeksbureau en het gebruik van de informatie geeft weer dat deze twee factoren onafhankelijke uitkomsten van tevredenheid zijn, waarvan de ene betrekking heeft op een lange termijn relatie en de andere op korte termijn gebruik van de informatie.

Als tevredenheid van de opdrachtgever op korte en lange termijn zo belangrijk is, is het

nuttig om te weten welke factoren deze tevredenheid beïnvloeden. Hierbij kan een dienstenmarketing benadering, waarbij onderscheid wordt gemaakt in technische en functionele kwaliteitsaspecten, zeer bruikbaar zijn. Uit dit onderzoek blijkt dat functionele, proces gerelateerde, kwaliteit in de vorm van de kenmerken van het onderzoeksbureau het sterkst gerelateerd is aan de tevredenheid van de opdrachtgever. Enkele belangrijke kenmerken die betrekking hebben op de interpersoonlijke relatie tussen onderzoeksbureau en klant zijn expertise, deskundigheid, integriteit, flexibiliteit en vertrouwelijkheid. Naast het leveren van kwaliteit tijdens het dienstverlenende proces dient tevens de technische kwaliteit niet uit het oog verloren te worden. De uitkomst van de dienstverlening is zeker zo belangrijk, zoals we reeds zagen bij het belang van de inhoudelijke kwaliteit van de informatie. Van de technische kwaliteitsaspecten is de kwaliteit van de presentatie, schriftelijk dan wel mondeling, het sterkst van invloed op tevredenheid. Hierbij valt bijvoorbeeld te denken aan duidelijk taalgebruik en een presentatiewijze met niet onnodig veel complexe analyses en statistieken. Behalve inhoudelijke kwaliteit en kwaliteit van de presentatie is het belangrijk om als onderzoeksbureau te zorgen voor resultaten en aanbevelingen die in de praktijk van de opdrachtgever toepasbaar zijn, aangezien dit de klanttevredenheid zal verhogen. Alleen met bruikbare informatie kan de klant daadwerkelijk aan de slag. Tenslotte blijkt de aanvaardbaarheid van de resultaten en aanbevelingen tijdens dit onderzoek niet significant van belang te zijn voor de tevredenheid en dus het gebruik van de informatie. Een verklaring hiervoor kan zijn dat opdrachtgevers het idee krijgen dat de resultaten niet correct kunnen zijn indien deze perfect aansluiten bij gangbare opvattingen binnen de organisatie. Een te rooskleurig resultaat zou op deze manier kunnen leiden tot ongeloofwaardigheid van het onderzoeksbureau.

6.2 Theoretische implicaties

Met dit onderzoek naar de verschillende factoren die ervoor zorgen dat organisaties wel of geen gebruik zullen maken van informatie verkregen uit marktonderzoek hebben we getracht een interessant probleem in kaart te brengen. Zoals bij ieder onderzoek echter het geval is bestaan er ook in dit geval een aantal beperkingen die weer kansen bieden voor toekomstig onderzoek. Allereerst zijn we van mening dat de operationalisering van de verschillende constructen in ons model voor verbetering vatbaar is. Voor een aantal constructen geldt dat de betrouwbaarheid weliswaar voldoende te noemen is, maar een verdere ontwikkeling van de meetinstrumenten zou zeker kunnen bijdragen tot een verhoging van de betrouwbaarheid. Een tweede punt voor toekomstig onderzoek heeft betrekking op het grote belang van de inhoudelijke kwaliteit van marktonderzoek. In de praktijk blijkt echter dat opdrachtgevers van marktonderzoek deze inhoudelijke kwaliteit vaak slecht kunnen beoordelen. Een objectieve maat voor de inhoudelijke kwaliteit van het onderzoek zou derhalve in vervolgonderzoek uitkomst kunnen bieden. Hierbij valt te denken aan additionele vragen, bijvoorbeeld of er een non-respons verantwoording wordt gegeven, of er multivariate analysetechnieken zijn benut, etc. Ten derde is de mate van verassing van de opdrachtgever over de uitkomsten van het onderzoek een variabele die zeker interessant genoeg is om mee te nemen in toekomstig onderzoek. Hiermee kan achterhaald worden of de uitkomsten overeenkomstig of tegengesteld waren aan de verwachtingen van de opdrachtgever vooraf en hoe dit van invloed is op de tevredenheid over het onderzoek en het gebruik van de informatie. Het is zeer goed denkbaar dat bij verrassende uitkomsten, die de opdracht-

gever niet leuk vindt, de tevredenheid afneemt. Een voorbeeld hiervan is een hele dure, veel geplaatste, tv-spot die in het onderzoek slecht beoordeeld wordt, hetgeen de opdrachtgever natuurlijk slecht uitkomt. Naar aanleiding hiervan zou hij de onderzoeksaanpak dan achteraf ter discussie kunnen gaan stellen. Vervolgonderzoek zou verder rekening kunnen houden met verschillende typen van informatiegebruik. Hebben we ons hier nog vooral gericht op een combinatie van instrumenteel en conceptueel gebruik, verder onderzoek zou echter een duidelijker onderscheid kunnen maken in de soorten gebruik. Op deze manier kan een nog genuanceerder beeld ontwikkeld worden, waarbij het gebruiksdoel van de informatie dan voorop staat en meer specifieke relaties tussen de constructen onderzocht kunnen worden. Tevens zou er een onderscheid gemaakt kunnen worden in verschillende soorten van informatie. De informatie die centraal stond in dit onderzoek was overwegend primaire informatie, dus voor de eerste keer verzameld voor een specifiek probleem. Wellicht dat een focus op secundaire gegevens en het gebruik hiervan een ander beeld zou opleveren, aangezien in dit geval het marktonderzoeksbureau geen directe controle heeft over de betrouwbaarheid van de informatie afkomstig van een derde partij en daardoor over de inhoudelijke kwaliteit van het onderzoek. Dit laatste construct zou dan wellicht een andere rol in deze onderzoeksopzet kunnen gaan vervullen. Een andere beperking van het onderzoek is dat er geen onderscheid is gemaakt tussen het gebruik van informatie in industriële markten en consumenten markten. Wellicht dat afzonderlijk vervolgonderzoek in beide settings tot andere en meer gedifferentieerde conclusies zou kunnen leiden. Tenslotte kan worden genoemd dat de empirische relaties slechts op één moment in de tijd zijn gemeten. Door meerdere metingen gedurende een langere periode te verrichten kan de dynamiek van de gevonden verbanden beter bepaald worden en kunnen exacte oorzaak-gevolg verbanden beter achterhaald worden.

6.3 Praktische implicaties

Tenslotte hebben de onderzoeksresultaten natuurlijk enkele implicaties voor de managementpraktijk van met name marktonderzoeksbureaus die regelmatig onderzoek verrichten voor zakelijke opdrachtgevers. Deze implicaties zijn reeds voor een deel af te leiden uit onze voorgaande interpretatie van de resultaten. Zoals reeds vermeld zou het streven naar een inhoudelijk kwalitatief hoogwaardig produkt een zeer belangrijk doel moeten zijn voor marktonderzoeksbureaus. Inhoudelijke kwaliteit zal namelijk niet alleen leiden tot hogere klanttevredenheid maar tevens direct tot een hoger gebruik van de gegeven informatie door de opdrachtgever. Inhoudelijke kwaliteit kan bewerkstelligd worden door zorg te dragen voor een goede toepassing van de onderzoekstechnieken, een correcte gegevensverzameling, geen tegenstrijdige resultaten te leveren, de klant het gevoel te geven dat de informatie het bestede geld waard is en niet teveel de nadruk leggen op getallen en statistieken. Deze zaken zijn o.a. terug te vinden in de kwaliteitseisen die de Vereniging van Marktonderzoeksbureaus (VMO) heeft geformuleerd voor de uitvoering en rapportage van marktonderzoek. Ook op de lange termijn is een tevreden klant belangrijk voor het voortbestaan van een aanbieder van marktonderzoek informatie. Tevredenheid zorgt ervoor dat opdrachtgevers het marktonderzoeksbureau gaan vertrouwen, iets wat van cruciaal belang is voor de intentie om de relatie op langere termijn voort te zetten. Tevreden klanten kunnen zowel door het leveren functionele als technische kwaliteit verkregen worden. Functionele kwaliteit heeft dan betrekking op het dienstverlenende proces waarover de controle voor een groot gedeel-

te bij het marktonderzoeksbureau ligt. Interpersoonlijke eigenschappen van het bureau in termen van expertise, integriteit, flexibiliteit en vertrouwelijkheid zijn daarbij het belangrijkst. Ook de technische kwaliteit van het resultaat van de dienstverlening draagt bij tot tevredenheid. Hierbij valt te denken aan het op een duidelijke manier schriftelijk en/of mondeling presenteren van de onderzoeksresultaten, zonder dit al te complex te maken. Naast inhoudelijke kwaliteit is het tevens zeer belangrijk om de opdrachtgever van bruikbare informatie te voorzien waarvan hij de conclusies en aanbevelingen zonder al te grote problemen in de praktijk kan toepassen. Het in acht nemen van deze implicaties zal er uiteindelijk voor moeten zorgen dat gebruikers van informatie de leveranciers hiervan als een belangrijke partner gaan beschouwen, een partner die zeker een toegevoegde waarde kan hebben voor de activiteiten van moderne ondernemingen.

LITERATUUR

Armstrong, J.S. en T.S. Overton, 1977. Estimating non-response bias in mail surveys. *Journal of Marketing Research,* Vol. 14, 396-402.

Beyer, J.M. en H.M. Trice, 1982. The utilization process: A conceptual framework and synthesis of empirical findings. *Administrative Science Quarterly* 27, December, 591-622.

Deshpandé, R. en G. Zaltman, 1982. Factors affecting the use of market research information: A path analysis. *Journal of Marketing Research,* Vol. 19, February, 14-31.

Deshpandé, R. en G. Zaltman, 1984. A comparison of factors affecting research use. *Journal of Marketing Research,* Vol. 21, February, 32-38.

Deshpandé, R. en G. Zaltman, 1987. A comparison of factors affecting use of marketing information in consumer and industrial firms. *Journal of Marketing Research,* Vol. 14, February, 114-118.

Dunn, W., 1986. "Studying knowledge use: *A profile of procedures and issues".* In: G. Beal (ed.), Knowledge generation, exchange, and utilization. The Westview Press.

Griffin, K., 1986. The contribution of studies of source credibility to a theory of interpersonal trust in the communication process. *Psychological Bulletin* 68, April.

Grönroos, C., 1990. *Service management and marketing: Managing moments of truth in service competition. Lexington,* Masachusetts & Toronto: Lexington Books.

Hanan, M. en P. Karp, 1989. *Customer Satisfaction.* New York: Amacom American Management Association.

John, G. en J. Martin, 1984. Effects of organisational structure of marketing planning on credibility and utilization of plan output. *Journal of Marketing Research,* Vol. 21, May, 170-183.

Kohli, A. en B.J. Jaworski, 1990. Market orientation: The construct, research propositions, and managerial implications. Report No. 90-113, November, 3-25. Cambridge, MA: Marketing Science Institute.

Krishnan, S., 1989. Making more effective use of market information. Report No. 90-113, November, 1-22. Cambridge, MA: Marketing Science Institute.

Lee, H., F. Acito en R.L. Day, 1987. Evaluation and use of marketing research by decision makers: A behavioral simulation. *Journal of Marketing Research,* Vol. 24, May, 187-196.

Maltz, E. en A.K. Kohli, 1995. Market intelligence dissemination across functional boundaries. Report No. 95-115, November, 3-24. Cambridge, MA: Marketing Science Institute.

Marketing Science Institute, 1996. Research priorities 1996-1998: *A guide to MSI research programs and procedures.* Cambridge, MA.

Menon, A. en R. Varadarajan, 1992. A model of marketing knowledge use within firms. *Journal of Marketing,* Vol. 56, October, 53-71.

Menon, A. en J. Wilcox, 1994. USER: A scale to measure use of market research. Report No. 94-108, May, 1-20. Cambridge, MA: Marketing Science Institute.

Moorman, Ch., R. Deshpandé en G. Zaltman, 1992. Relationships between providers and users of market research: The dynamics of trust within and between organisations. *Journal of Marketing Research,* Vol. 29, August, 314-328.
Moorman, Ch., R. Deshpandé en G. Zaltman, 1993. Factors affecting trust in market research relationships. *Journal of Marketing,* Vol. 57, January, 81-101.
Moorman, Ch. en J. Austin, 1995. The paradox of low quality and high use: How research trust impacts market research outcomes. Report No. 95-116, December, 1-24. Cambridge, MA: Marketing Science Institute.
Morgan, R.M. en S.D. Hunt, 1994. The commitment-trust theory of relationship marketing. *Journal of Marketing,* Vol. 58, July, 20-38.
Newman, J., 1962. Put research into marketing decisions. *Harvard Business Review,* 40 (2), 105-112.
Nunnally, J.C. en I.H. Bernstein, 1994. *Psychometric Theory.* New York, etc.: McGraw-Hill, Inc.
Oliver, R.L., 1996. Satisfaction: *A behavioral perspective on the consumer.* New York: McGraw-Hill, Inc.
Raphael, J. en R. Parket, 1991. The need for market research in executive decision making. *Journal of Business & Industrial Marketing,* Vol. 6, Iss. 1, 2, Winter/Spring, 15-21.
Rich, R., 1979. *The Power of Social Science Information and Public Policy Making.* San Francisco: Jossey-Bass.
Wetzels, M., K. de Ruyter en M. van Birgelen, 1998. Marketing service relationships: The role of commitment. *Journal of Business and Industrial Marketing,* Te verschijnen.

2. *Overstag?!*[1] *Reductie van het weigerpercentage in telefonisch onderzoek*

L.M. PONDMAN

SAMENVATTING

Sociaal-wetenschappelijk onderzoek vormt een belangrijke informatiebron. Helaas neemt het aantal mensen dat om welke reden dan ook weigert om mee te doen nog steeds toe. Een hoog percentage weigeraars brengt de representativiteit in gevaar waardoor de resultaten van menig onderzoek een vertekend beeld van de werkelijkheid geven en niet generaliseerbaar zijn naar de gehele onderzoekspopulatie.
In de onderhavige onderzoek wordt getracht een bijdrage te leveren aan de oplossing van het bovenstaande probleem. Het onderzoek is uitgevoerd om inzicht te verkrijgen in de invloed van een aantal aspecten van het verbale gedrag van een interviewer op de beslissing om al dan niet deel te nemen aan een *telefonisch* interview. Om dit inzicht te verkrijgen is de interviewer-subject interactie die plaats vindt voordat het eigenlijke interview begint onder de loep genomen. Dit heeft geresulteerd in de ontwikkeling van een overhaalstrategie. Er is een experiment uitgevoerd om te toetsten of toepassing van de overhaalstrategie een reductie van het aantal weigeringen tot gevolg heeft.
Het onderzoek toont aan dat toepassing van de overhaalstrategie de kans op deelname verhoogt. Verder blijkt motivatie van interviewers een heel belangrijke rol te spelen. Interviewers die gemotiveerd zijn om hun uiterste best te doen om zoveel mogelijk mensen over te halen, behalen een significant lager weigerpercentage dan interviewers waarbij dit niet is gebeurd.

1. INLEIDING

Het aantal mensen dat om welke reden dan ook het verzoek tot deelname aan onderzoek afwijst neemt nog steeds toe. Deze afnemende bereidheid tot deelname heeft een negatieve invloed op de kwaliteit van de data en vormt daardoor een bron van zorg voor veel survey-onderzoekers.
Uit onderzoek blijkt dat participanten, mensen die meedoen aan onderzoek, verschillen van weigeraars, mensen die hun medewerking niet toezeggen (Brehm, 1993; Visscher, 1995; De Kruijk & Hermans, 1998). De verschillen hebben niet alleen betrekking op

1 Dit artikel is een verkorte weergave van promotie-onderzoek dat heeft plaats gevonden aan de vakgroep Methoden en Technieken van de Vrije Universiteit. Voor een volledig verslag van dit onderzoek wordt verwezen naar: Pondman, L.M. (1998). The influence of the interviewer on the refusal rate in telephone surveys. Deventer: Print Partners Ipskamp

demografische kenmerken als leeftijd, sekse, opleiding en inkomen, maar ook op attitude en gedrag. Het ontbreken van gegevens van weigeraars heeft tot gevolg dat de resultaten van menig onderzoek een vertekend beeld van de werkelijkheid geven. De representativiteit van de data komt in gevaar waardoor generaliseerbaarheid naar de gehele onderzoekspopulatie moeilijk wordt.
Het bovenstaande probleem heeft geleid tot de uitvoering van vele studies waarbij het probleem van de toenemende nonrespons ten gevolge van weigeringen centraal staat. Het merendeel van deze studies is gericht op de ontwikkeling van strategieën die een reductie van het weigerpercentage tot gevolg hebben. Voorbeelden van dit soort strategieën zijn het sturen van een aankondigingsbrief (b.v. Dillman e.a., 1976; Groves & Lyberg, 1988; Porst, 1991; Brehm, 1994), het aanbieden van een beloning (b.v. Gunn & Rhodes, 1981; Goetz e.a., 1984; Yammarino e.a., 1991; Church, 1993) en het manipuleren van de introductietekst (b.v. Singer & Frankel, 1982; Morton-Williams, 1991 en 1993; Houtkoop-Steenstra & Van den Bergh, 1994).
In het hier beschreven onderzoek wordt een aantal aspecten van het verbale gedrag van de interviewer nader onderzocht. Het onderzoek is uitgevoerd om inzicht te verkrijgen in de invloed van deze aspecten op de beslissing van de beoogde respondent, ook wel het subject genoemd, om al dan niet deel te nemen aan een telefonisch interview. Het heeft als doel de identificatie van technieken die interviewers kunnen gebruiken bij het overhalen van subjecten om mee te doen. Het artikel is met name gericht op de vraag of het mogelijk is om interviewers dusdanig te instrueren, te motiveren en te trainen zodat het percentage weigeringen gereduceerd wordt.

Om de onderzoeksvraag te kunnen beantwoorden is een experiment opgezet. Het experiment is gebaseerd op de resultaten van twee niet-experimentele studies. Deze niet-experimentele studies zijn uitgevoerd om de op weigeren gerichte opmerkingen van subjecten en de reacties van interviewers op deze opmerkingen te onderzoeken. De resultaten hebben geleid tot de ontwikkeling van een overhaalstrategie. Deze strategie bestaat uit het vermijden van reacties die een negatief effect op deelname blijken te hebben en het toepassen van reacties die een positief effect blijken te hebben:

- vermijd het stellen van ‘waarom-vragen’ (“waarom wilt u niet meedoen?”);
- herhaal op geen enkele wijze de weigering: niet door deze op vragende wijze (al dan niet letterlijk) opnieuw weer te geven; niet door middel van een op weigeren gericht vragend “ja” of “nee”;
- vraag of je op een ander tijdstip terug kunt bellen als het subject aangeeft geen tijd te hebben;
- geef in reactie op weigeringen die niet refereren aan tijdgebrek informatie die op deelname is gericht.

2. HET EXPERIMENT[2]

Het experiment is opgezet om te onderzoeken of toepassing van de overhaalstrategie een reductie van het aantal weigeringen tot gevolg heeft. Het experiment bestaat uit drie condities: (1) een controlegroep en twee experimentele groepen, te weten, (2) een

2 Met dank aan Edith Smit. Zij is werkzaam bij de vakgroep Communicatiewetenschap van de Universiteit van Amsterdam en doet onder andere onderzoek naar de effecten van reclamecampagnes. Het experiment is in nauwe samenwerking met haar uitgevoerd.

motivatiegroep en (3) een trainingsgroep. De condities verschillen met betrekking tot de manier waarop de interviewers tijdens de voorbereiding van hun eigenlijke taak zijn benaderd. In tabel 1 wordt de opbouw van het experimentele design kort weergegeven.

Tabel 1. Het experimentele design.

conditie 1: controlegroep	**conditie 2**: motivatiegroep	**conditie 3**: trainingsgroep
INSTRUCTIE	INSTRUCTIE MOTIVATIE	INSTRUCTIE MOTIVATIE TRAINING

In totaal hebben er 35 interviewers aan het experiment meegedaan; 12 in de controlegroep, 12 in de motivatiegroep en 11 in de trainingsgroep. De interviewers waren allemaal eerstejaars studentes aan de vakgroep Communicatiewetenschap van de Universiteit van Amsterdam in de leeftijd van 19 tot 22 jaar.
De steekproef bestond uit een groot aantal door de computer gegenereerde telefoonnummers. Deze zijn random toegewezen aan de interviewers.

De interviewers uit de controlegroep zijn kort geïnstrueerd. Ze hebben alleen die informatie gekregen die noodzakelijk is bij het uitvoeren van hun taak; de vragenlijst is kort toegelicht, het CATI (Computer Assisted Telephone Interviewing) - programma is gedemonstreerd en ze zijn korte tijd in de gelegenheid gesteld om het programma uit te proberen.

De interviewers uit de motivatiegroep zijn eveneens geïnstrueerd. Daarnaast is bij deze groep extra aandacht besteed aan het verhogen van de motivatie. Door de interviewers te motiveren is getracht een bepaalde mate van betrokkenheid te creëren. Hierbij is uitgegaan van de veronderstelling dat interviewers die zich betrokken voelen bij het onderzoek en bij de onderzoekers, hun uiterste best zullen doen om zoveel mogelijk mensen over te halen tot deelname aan het telefonisch interview.
Om de motivatie van de interviewers te verhogen zijn de volgende acties ondernomen: (a) de interviewers hebben uitgebreid informatie gekregen over de inhoud van het onderzoek; (b) het belang van het onderzoek is benadrukt; (c) de interviewers is duidelijk gemaakt dat het erg belangrijk is dat de verzamelde gegevens representatief zijn; (d) de interviewers is geleerd dat een hoog weigerpercentage een slechte invloed heeft op de kwaliteit van het onderzoek; (e) de onderzoekers hebben het vertrouwen in de interviewers met betrekking tot hun overtuigingskracht uitgesproken; (f) er zijn rollenspellen gespeeld om de interviewers het vertrouwen te geven dat ze in staat zijn om mensen over te halen. De rollenspellen waren gericht op het overhaalproces *zonder dat werd aangegeven hoe er in een bepaalde situatie gereageerd moet worden*.

De interviewers uit de trainingsgroep zijn, op dezelfde wijze als de interviewers uit de motivatiegroep, geïnstrueerd en gemotiveerd om hun best te doen om zoveel mogelijk mensen over te halen tot deelname. Bovendien *hebben zij aanwijzingen gekregen hoe ze een laag percentage weigeringen kunnen bereiken*; Ze zijn getraind in het toepassen van de overhaalstrategie.

Voor het welslagen van het experiment was het erg belangrijk dat de interviewers zich aan de vier overhaalregels zouden houden. Daarom zijn zij bij de uitoefening van hun taak geassisteerd door de computer. Op het scherm verscheen een lijst van weigeringen; als de interviewer de weigering aanklikte die correspondeerde met de opmerking van het subject kwam er een voorgeschreven reactie te voorschijn. Een voorbeeld van zo'n voorgeschreven reactie wordt hieronder gegeven:

S: "Oh, dat duurt me veel te lang"
I: "Het hoeft niet zo lang te duren hoor! Dat hangt ook een beetje van u af. Als u de vragen vlot beantwoordt zijn we er zo doorheen. Zullen we gewoon beginnen?"

De interviewers waren niet verplicht de voorgeschreven reactie te gebruiken. Het was toegestaan naar eigen inzicht van deze reactie af te wijken, *mits* de overhaalregels in acht werden genomen.

De verwachting was dat de interviewers uit de beide experimentele groepen significant lagere weigerpercentages zullen behalen dan de interviewers uit de controlegroep. Verder werd verondersteld dat de weigerpercentages van de interviewers uit de trainingsgroep significant lager zullen zijn dan de weigerpercentages van de interviewers uit de motivatiegroep.

3. RESULTATEN

Weigerpercentages

In tabel 2 worden de gemiddelde weigerpercentages van de interviewers uit de drie verschillende condities gepresenteerd. De weigerpercentages zijn berekend op basis van het aantal weigeraars en het aantal participanten. Subjecten die het interview afgebroken hebben nadat ze al een of meer vragen hebben beantwoord, worden beschouwd als participanten. Zij hebben tenslotte in eerste instantie het verzoek tot deelname geaccepteerd. Variantie analyse (ANOVA) laat zien dat het resultaat van de interactie tussen de interviewer en het subject inderdaad afhankelijk is van de manier waarop de interviewers tijdens de voorbereiding van hun eigenlijke taak zijn benaderd ($F = 10.0$; $df = 2/32$; $p < .001$). De conditie waartoe de interviewers behoren verklaard 39% van de variantie tussen de drie groepen.

Tabel 2. Gemiddeld weigerpercentage per conditie.

controle groep	motivatie groep	trainings groep	F	df	p
59.8	46.8	48.6	10.0	2/32	< .001
$R^2 = 0.39$	S = 0.08				

De interviewers uit de controlegroep hebben de hoogste weigerpercentages behaald; een gemiddelde van 59.8%. In vergelijking met deze interviewers hebben de interviewers uit de motivatiegroep en de interviewers uit de trainingsgroep het een stuk beter

gedaan. Van de mensen die zij hebben benaderd hebben gemiddeld respectievelijk 46.8% en 48.6% het verzoek tot deelname geweigerd.
T-toetsen tonen aan dat zowel het verschil tussen de eerste en de tweede conditie ($t = 4.0$; $df = 22$; $p < .001$) als het verschil tussen de eerste en de derde conditie ($t = 4.0$; $df = 21$; $p < .001$) significant is. Hiermee is de verwachting dat de interviewers uit de experimentele condities, waarbij extra aandacht is besteed aan de motivatie tot overhalen, een lager weigerpercentage zullen behalen dan de interviewers uit de controlegroep die deze aandacht niet hebben gekregen uitgekomen. Het positieve effect van motivatie wordt niet afgezwakt door een toename van de kans op een afgebroken interview. Hoewel het niet ondenkbaar is dat mensen die overgehaald zijn het interview eerder zullen afbreken dan mensen die 'uit eigen vrije wil' het verzoek tot deelname hebben geaccepteerd, komen er in de experimentele condities zelfs relatief minder afgebroken interviews voor dan in de controlegroep (zie tabel 3).

Tabel 3. Gemiddeld percentage afgebroken interviews per conditie.

controle groep	motivatie groep	trainings groep	F	df	p
5.3	2.2	2.3	2.2	2/32	ns

De weigerpercentages van de interviewers uit de trainingsgroep zijn, tegen de verwachting in, niet lager dan de weigerpercentages van de interviewers uit de motivatiegroep. De interviewers uit de trainingsgroep hebben zelfs een iets hoger gemiddeld weigerpercentage bereikt (48.6%) dan de interviewers uit de motivatiegroep (46.8%). Dit verschil is echter klein en niet significant.
Dat er geen verschil is tussen de experimentele condities wordt mogelijk veroorzaakt door de ineffectiviteit van de ontwikkelde overhaalstrategie die alleen aan de interviewers uit de trainingsgroep is aangeboden. Om deze mogelijkheid te onderzoeken is de inhoud van de interviewer-subject interacties van 9 interviewers, 3 per conditie, nader geanalyseerd. Deze inhoudsanalyse is gericht op interacties waarin één of meer overhaalpogingen worden gedaan. De resultaten worden hieronder beschreven.

Inhoudsanalyse

In de eerste plaats is onderzocht of de interviewers uit de trainingsgroep zich vaker aan de regels hebben gehouden dan de interviewers uit de andere twee groepen. De resultaten worden gepresenteerd in tabel 4. In de tabel wordt onderscheid gemaakt tussen interacties waarin alle overhaalpogingen in overeenstemming zijn met de overhaalregels (+), interacties waarin sommige pogingen in overeenstemming met de regels zijn en andere in strijd met de regels (+/–), en interacties waarin alle pogingen in strijd zijn met de overhaalregels (–).
Uit tabel 4 komt naar voren dat de interviewers uit de trainingsgroep zich doorgaans vrij goed aan de regels houden. In 81% van de interacties waarin een of meer overhaalpogingen worden gedaan handelen zij volledig in overeenstemming met de regels. In 18% van de interacties reageren ze soms in overeenstemming met en soms in strijd met de regels. Het komt slechts een keer voor dat een interviewer uit de trainingsgroep zich gedurende de gehele interacties niet aan de regels houdt.

Tabel 4. Toepassing van de overhaalregels per conditie.

CONDITIE	OVERHAALREGELS							
	+		+/–		–		totaal	
	n	%	n	%	n	%	n	%
controlegroep	10	28	11	30	15	42	36	100
motivatiegroep	23	43	22	42	8	15	53	100
trainingsgroep	63	81	14	18	1	1	78	100

+ = in overeenstemming met de overhaalregels
+/– = zowel in strijd met als in overeenstemming met de overhaalregels
- = in strijd met de overhaalregels

De interviewers uit de controlegroep gedragen zich veel vaker in strijd met de regels; in 42% van hun interacties bestaan de overhaalpogingen uit het stellen van een 'waaromvraag' of uit het herhalen van de weigering. Dit is natuurlijk niet verwonderlijk daar de interviewers uit deze groep niet in kennis zijn gesteld van de negatieve invloed die dit soort reacties kunnen hebben op de beslissing van het subject om al dan niet mee te doen. Desondanks komt het voor dat zij zich uit eigen beweging toch in overeenstemming met de regels gedragen; in 28% van de interacties doen zij dit volledig en in 30% van de interacties soms wel en soms niet.
Net als de interviewers uit de controlegroep zijn de interviewers uit de motivatiegroep niet getraind in het toepassen van de overhaalregels. Gezien dit gegeven is hun prestatie uitzonderlijk. Slecht in 15% van hun interacties heeft de interviewer alle overhaalpogingen gedaan in strijd met de regels. Daartegenover staat dat in niet minder dan 43% van de interacties de interviewer zich, zonder zich hiervan bewust te zijn, volledig aan de regels heeft gehouden. De overige 42% van hun interacties bevat een mix van 'goed' en 'slecht' gedrag.

De interviewers uit alle drie de condities hebben zich dus in meer of mindere mate aan de regels gehouden. Om te achterhalen of toepassing van de overhaalregels een reductie van het weigerpercentage tot gevolg heeft, zijn de interviewer-subject interacties geanalyseerd onafhankelijk van de conditie waartoe ze behoren. De resultaten van deze analyse worden gepresenteerd in tabel 5.

Tabel 5. Effect van het toepassen van de overhaalregels, los van de conditie.

OVERHAALREGELS	RESULTAAT					
	weigering		deelname		totaal	
	n	%	n	%	n	%
+	51	53	45	47	96	100
+/–	38	81	9	19	47	100
–	23	96	1	4	24	100

+ = in overeenstemming met de overhaalregels
+/– = zowel in strijd met als in overeenstemming met de overhaalregels
– = in strijd met de overhaalregels

Tabel 5 laat zien dat 53% van de interacties waarin de interviewer zich volledig volgens de regels gedraagt, heeft geresulteerd in een weigering. Dit percentage ligt veel hoger voor interacties waarin de interviewer zich soms wel en soms niet aan de regels houdt (81%). Als alle overhaalpogingen zijn uitgevoerd in strijd met de regels nadert het weigerpercentage zelfs de 100%. Uit deze cijfers kan worden geconcludeerd dat de ontwikkelde overhaalstrategie wel degelijk effectief is: de kans op een weigering neemt af als de interviewer zich bij haar overhaalpogingen aan de regels houdt.

Het blijkt dus dat de overhaalstrategie effectief is en dat de interviewers uit de trainingsgroep zich, in vergelijking met de interviewers uit de motivatiegroep, vaker aan de overhaalregels hebben gehouden. Toch is het zo dat de interviewers uit de trainingsgroep geen lager weigerpercentage hebben bereikt. Een mogelijke verklaring voor dit gegeven wordt gesuggereerd door de gegevens in tabel 6. In deze tabel wordt het effect van het toepassen van de overhaalregels op het resultaat van de interviewer-subject interactie uitgesplitst naar conditie gepresenteerd.

Tabel 6. Effect van het toepassen van de overhaalregels per conditie.

OVERHAALREGELS	RESULTAAT					
	weigering		deelname		totaal	
	n	%	n	%	n	%
controlegroep						
+	2	20	8	80	10	100
+/–	9	82	2	18	11	100
–	14	93	1	7	15	100
motivatiegroep						
+	10	43	13	57	23	100
+/–	17	77	5	23	22	100
–	8	100	0	0	8	100
trainingsgroep						
+	39	62	24	38	63	100
+/–	12	86	2	14	14	100
–	1	100	0	0	1	100

+ = in overeenstemming met de overhaalregels
+/– = zowel in strijd met als in overeenstemming met de overhaalregels
– = in strijd met de overhaalregels

De interviewers uit de controlegroep en de interviewers uit de motivatiegroep zijn veel succesvoller in het toepassen van de overhaalstrategie dan de interviewers uit de trainingsgroep. Als zij zich volledig gedragen in overeenstemming met de regels dan behalen zij respectievelijk een weigerpercentage van 20% en 43%. Het weigerpercentage van de interviewers uit de trainingsgroep ligt ver boven het gemiddelde: 62% van de interacties waarin een of meer overhaalpogingen gedaan zijn met in achtneming van de regels resulteert in een weigering. Om dit opmerkelijke resultaat te kunnen verklaren is de inhoud van de interviewer-subject interacties die deel uitmaken van de inhoudsanalyse nader bestudeerd. Hieruit is gebleken dat er een belangrijk verschil is tussen enerzijds de interviewers die getraind zijn en anderzijds de interviewers die niet getraind zijn in het toepassen van de overhaalstrategie.

De interviewers uit de controlegroep en de motivatiegroep hebben goed geluisterd naar het subject. Hun overhaalpogingen sluiten nauw aan bij de geuitte weigeringen. Het lijkt erop dat de interviewers uit deze twee groepen op spontane wijze hun tegenargumenten hebben aangepast aan de argumenten van het betreffende subject.
De overhaalpogingen van de interviewers uit de trainingsgroep sluiten veelal niet aan bij de weigeringen van het subject. Deze interviewers passen de regels zelfs toe als het subject heel duidelijk heeft gemaakt dat hij echt niet mee wil doen. De indruk bestaat dat zij hebben geprobeerd zich krampachtig vast te houden aan de overhaalregels door zich te concentreren op de voorgeschreven reacties die ze via de computer kregen aangeboden. Hierdoor hebben ze te weinig aandacht gehad voor het subject en is hun inschattingsvermogen zeer waarschijnlijk aangetast.
Het krampachtig vasthouden aan de overhaalregels is vermoedelijk versterkt door een bepaalde mate van 'learned helplessness'. Gevoed door hun hoge weigerpercentages kunnen de interviewers het gevoel hebben gekregen dat zij geen positieve invloed kunnen uitoefenen op het resultaat van de interactie. Met andere woorden: "het maakt niet uit wat ik doe, het helpt toch niet!" Dit gevoel van machteloosheid kan het motivatie-effect hebben verstoord waardoor het effect van de overhaalpogingen nog meer is afgenomen.

4. DISCUSSIE

Het onderhavige onderzoek heeft aangetoond dat het mogelijk is om door motivatie van interviewers het percentage nonrespons ten gevolge van weigeringen te reduceren. Ten allen tijde is het van belang om interviewers uitgbreide informatie te geven over de inhoud van het onderzoek, om het belang van het onderzoek te benadrukken, om ze te vertellen dat het erg belangrijk is dat het onderzoek representatieve gegevens oplevert, om ze te informeren over de gevolgen van een hoog weigerpercentage en om ze het vertrouwen te geven dat ze in staat zijn om mensen over te halen.
Verder lijkt het verstandig om interviewers de door middel van deze studie ontwikkelde overhaalstrategie aan te leren. Toepassing van deze strategie verhoogt de kans op deelname. Het betekent het vermijden van reacties die een negatieve invloed op deelname hebben en het toepassen van reacties die een positieve invloed hebben. De niet-experimentele studies hebben echter aangetoond dat de negatieve reacties zoals het stellen van een 'waarom-vraag' en het herhalen van de weigering, veelvuldig door interviewers geuit worden. Verondersteld wordt dan ook dat dit soort negatieve reacties natuurlijk gedrag representeren. Deze veronderstelling wordt onderschreven door de studie van Smit en Dijkstra (1991). Ook zij hebben gevonden dat de meeste interviewers op weigeringen reageren door deze simpel te herhalen.
Toepassing van de overhaalstrategie vraagt van de interviewers dus dat ze natuurlijk gedrag (het stellen van waarom-vragen en het herhalen van de weigering) afleren en onnatuurlijk gedrag (het geven van informatie die op deelname is gericht) aanleren. Dit is een moeilijk en tijdrovend proces dat een intensieve training vereist. Deze training is bij voorkeur gericht op het internaliseren van de overhaalregels zodat de interviewers zich volledig kunnen concentreren op het subject. Het aanbieden van voorgeschreven reacties via een computerscherm wordt dan ook afgeraden. Op deze manier hoeven de interviewers hun aandacht niet te verdelen tussen het subject en de computer. Dit stelt ze in de gelegenheid om goed te luisteren naar het subject en om eventuele overhaalpogingen aan te laten sluiten bij de weigeringen die het subject uit.

Bij de ontwikkeling van de beschreven overhaalstrategie is (nog) geen aandacht besteed aan de gevolgen voor de validiteit van de gegevens. Dit is echter wel van groot belang en eventueel vervolg onderzoek zou zich hierop kunnen richten.
Ten eerste is het de vraag of overgehaalde subjecten nog wel betrouwbare informatie geven. De toepassing van de overhaalstrategie kan gevolgen hebben voor het gedrag van het subject tijdens het interview zelf. Het is niet ondenkbaar dat overgehaalde subjecten zich tijdens het interview plotseling misleid voelen of dat ze het gevoel krijgen dat ze zich hebben laten verleiden tot iets wat ze eigelijk liever niet willen. Dit gevoel kan resulteren in een toename van item-nonrespons en van het aantal sociaal-wenselijke antwoorden. Bovendien kan het een afname van de hoeveelheid informatie in reactie op open vragen tot gevolg hebben.
De implementatie van de overhaalstrategie kan eveneens van invloed zijn op het gedrag dat de interviewer tijdens het interview vertoont. Tijdens het overhaalproces wordt de interviewer geacht zich aan te passen aan het subject. Dit betekent dat de interviewer moet kunnen improviseren en flexibel en creatief moet zijn om adequaat te kunnen reageren op de weigeringen van het subject in kwestie. Tijdens het interview zelf is het echter de bedoeling dat de interviewer zich zo veel mogelijk houdt aan een voorgeschreven interview-scenario. De overgang van improvisatie naar een voorgeschreven gedragspatroon is niet eenvoudig. De mogelijkheid bestaat dat de interviewer blijft hangen in de vrije stijl van het overhaalproces en de antwoorden van het subject op de vragen uit het interview beïnvloedt door het stellen van directieve of sturende vragen.

Het bovenstaande leidt tot de gedachte dat de perfecte overhaalstrategie niet alleen een reductie van het percentage weigeringen tot gevolg heeft. De perfecte overhaalstrategie oefent eveneens geen negatieve invloed uit op het gedrag van zowel het subject als de interviewer tijdens de ondervraging. Het is bovendien van belang dat de strategie duurzaam is.
De duurzaamheid wordt ondere andere bepaald door de motivatie van de interviewers om de overhaalregels keer op keer toe passen. Het moment dat ze er genoeg van krijgen waardoor ze weer terugvallen op hun oude, ineffectieve gedragingen, dient zolang mogelijk te worden uitgesteld, zoniet helemaal vermeden te worden.
Een ander element van duurzaamheid is gerelateerd aan het subject. De populariteit van het telefonisch interview als dataverzamelingsmethode neemt nog steeds toe. Hierdoor worden subjecten regelmatig blootgesteld aan interviewers die hen proberen over te halen. De kans is aanwezig dat subjecten gewend raken aan de strategie en hier niet meer vatbaar voor zijn. Subjecten kunnen behoed worden voor een overdosis aan telefonische interviews door de instelling van een onderzoekspanel. Een panel dat bestaat uit mensen die plezier beleven aan het meedoen aan telefonische interviews. De leden van het onderzoekspanel kunnen ingezet worden bij onderzoek dat nog in de ontwikkelingsfase is of bij onderzoek waarbij representativiteit niet van belang is. Mensen die zich niet beschikbaar hebben gesteld voor het onderzoekspanel hoeven dan alleen om medewerking te worden gevraagd als representatieve data noodzakelijk zijn.

LITERATUUR

Brehm, J. (1993). *The phantom respondents. Opinion surveys and political representation*. Ann Arbor: The University of Michigan Press.
Brehm, J. (1994). Stubbing our toes for a foot in the door? Prior contact, incentives and survey response. *International Journal of Public Opinion Research*, 6 (1), 45-63.
Church, A.H. (1993). Estimating the effect of incentives on mail survey response rates: A meta-analysis. *Public Opinion Quarterly*, 57, 62-79.
De Kruijk, M.D., Hermans, E. (1998). Effect van non-response op representativiteit. In: A.E. Bronner, P. Ester, P.S.H. Leeflang, A.J. Olivier, W.F. van Raaij en B. Wierenga (red.), *Jaarboek van de Nederlandse Vereniging van Marktonder-zoekers 1997-1998*. (pp. 55-69). Haarlem: De Vrieseborch.
Dillman, D.A., Gallegos, J.G. & Frey, J.H. (1976). Reducing refusal rates for telephone interviews. *Public Opinion Quarterly*, 40 (1), 66-78.
Goetz, E.G., Tyler, T.R. & Lomax Cook, F. (1984). Promised incentives in media research: A look at data quality, sample representativeness and response rate. *Journal of Marketing Research*, 21 (2), 148-154.
Groves, R.M. & Lyberg, L.E. (1988). An overview of nonresponse issues in tele-phone survey. In R.M. Groves, P.P. Biemer, L.E. Lyberg, J.T. Massey, W.L. Nicholls & J. Waksberg (red.), *Telephone survey methodology* (pp. 191-212). New York: Wiley.
Gunn, W.J. & Rhodes, I.N. (1981). Physician response rates to a telephone survey: Effects of monetary incentive level. *Public Opinion Quarter-y*, 45 (1), 109-115.
Houtkoop-Steenstra, H. & van den Bergh, H. (1994). Response-verhoging in telefonische interviews. In: A.E. Bronner, P. Ester, P.S.H. Leeflang, A.J. Olivier, W.F. van Raaij en B. Wierenga (red.), *Jaarboek van de Nederlandse Vereniging van Marktonderzoekers 1994-1995* (pp. 61-69). Haarlem: De Vrieseborch.
Morton-Williams, J. (1991). *Obtaining co-operation in surveys – the development of a social skills approach to interview training in introducing surveys*; Working Paper Series Number 3. London: Joint Centre for Survey Methods.
Morton-Williams, J. (1993). *Interviewer approaches*. Brookfield: Dartmouth Publishing Company.
Pondman, L.M. (1998). *The influence of the interviewer on the refusal rate in telephone surveys.* Deventer: Print Partners Ipskamp
Singer, E. & Frankel. M.R. (1982). Informed consent procedures in telephone interviews. *American Sociological Review*, 47 (3), 416-427.
Smit, J.H., & Dijkstra, W. (1991). Persuasion strategies for reducing refusal rates in telephone surveys. *Bulletin de Méthodologie Sociologique*, 33, 3-19.
Visscher, G. (1995). *Kiezersonderzoek op een dwaalspoor*. Den Haag: Sdu Uitgeverij Koninginnegracht.
Yammarino, F.J., Skinner, S.J. & Childers, T.L. (1991). Understanding mail survey response behavior; A meta-analysis. *Public Opinion Quarterly*, 55, 613-639.

3. *Onderzoek naar relevante kenmerken van gebruikers van het World Wide Web*

H.C.J. VROLIJK, K.R.E. HUIZINGH en J.C. HOEKSTRA

SAMENVATTING

Het doel van het in dit artikel beschreven empirisch onderzoek is het in kaart brengen van de kenmerken van een bezoeker die invloed hebben op de beleving van de een Web site. Hiervoor zijn theorieën ten aanzien van reclame effectiviteitsonderzoek, onderzoek naar informatiezoekgedrag en onderzoek naar consumenten navigatie gedrag in on-line computer omgevingen samengebracht in een onderzoeksmodel. Uit dit onderzoek blijkt dat attentie met name verklaard wordt door de productkennis, de betrokkenheid, de uitdaging en de flow. Bij het verklaren van cognitie, affectie en conatie spelen betrokkenheid en flow een belangrijke rol.

1. INLEIDING

De enorme groei van het aantal gebruikers van het World Wide Web maakt het Web een steeds interessanter medium voor marketeers. (Cronin, 1995; Quelch and Klein, 1996; Paul, 1996). Dit komt onder andere tot uitdrukking in het aantal commerciële Web sites dat de afgelopen jaren sterk is gegroeid tot meer dan 650,000 Web sites in 1997. Daarnaast laten de verkopen die via Web sites worden gerealiseerd, een stijgende lijn zijn. Ook de opbrengsten van advertenties op Internet laten een groeiende tendens zien (Ducoffe, 1996).

Deze en andere ontwikkelingen en voorspellingen maken organisaties bewust van het toenemende belang van het WWW als een gereedschap voor electronic commerce. In een zeer korte tijd hebben een groot aantal organisaties een aanwezigheid op het Web ontwikkeld. Ondanks de grote hoeveelheid aandacht die aan het World Wide Web wordt besteed, is de hoeveelheid empirisch onderzoek zeer beperkt. Richtlijnen ten aanzien van de optimale inrichting van een Web site zijn met name gebaseerd op ervaring en niet zo zeer op empirisch onderzoek. Richtlijnen hebben vaak de vorm van do's en don't's bij het implementeren van een Web site. Recentelijk komt de nadruk meer te liggen op empirisch onderzoek (zie bijvoorbeeld: Briggs and Hollis, 1997; Maddox and Mehta, 1997; Dreze and Zufryden, 1997; Mehta et al. 1995).

Onderzoek dat tot nu toe is uitgevoerd heeft met name betrekking op het ontwerp en de inrichting van de Web site. In dit deelonderzoek willen wij ons niet zo zeer concentreren op het ontwerp van de Web site maar op de kenmerken van de bezoekers en hun percepties van de Web sites die zij bezoeken. De doelstelling van dit onderzoek is dan ook:

'het in kaart brengen van de kenmerken van een bezoeker die invloed hebben op de beleving van een Web site.'

In paragraaf 2 wordt op basis van een beschrijving van voorgaand onderzoek het onderzoeksmodel voor dit onderzoek geformuleerd. In paragraaf 3 wordt de onderzoeksopzet beschreven waarmee wij het in het paragraaf 2 beschreven model willen toetsen. In paragraaf 4 komen de resultaten van dit onderzoek aan de orde. In paragraaf 5 volgt een samenvatting en de belangrijkste conclusies.

2. ONDERZOEKSMODEL

Het World Wide Web vertoont overeenkomsten met andere media maar vertoont ook belangrijke verschillen. Zo kan het Web gebruikt worden als een advertentie medium in de vorm van bijvoorbeeld banners (Berthon et al., 1996), maar ook als communicatie medium om een boodschap over de organisatie of product over te brengen door middel van een Web site. Daarnaast is het Web een enorme informatiebron waar informatie over tal van onderwerpen kan worden gevonden. Het Web biedt, in tegenstelling tot vele andere media, de mogelijkheid tot interactiviteit. Om met al deze aspecten rekening te houden hebben wij verschillende bestaande theorieën in ons onderzoeksmodel ondergebracht. Dit zijn:

- Reclame effectiviteitsonderzoek (bijvoorbeeld Belch et al. (1987), Gelb et al. (1985))
- Informatiezoekgedrag (bijvoorbeeld Schmidt and Spreng (1996))
- Navigatiegedrag in een on-line computer omgeving (bijvoorbeeld Hoffman en Novak. (1996))

In dit onderzoek zijn een drietal modellen uit voorgaand onderzoek geïntegreerd. Onderzoek naar de effectiviteit van reclame en communicatie in het algemeen is relevant omdat het Web wordt gebruikt voor reclame uitingen (bijvoorbeeld in de vorm van banners) maar ook als een breder communicatie medium (bijvoorbeeld de Web site van een organisatie). Naast een reclame- en communicatie medium is het Web een enorme bron van informatie. Op basis van een bepaalde informatiebehoefte kan een gebruiker van het Web op zoek gaan naar relevante informatie. Onderzoek naar het informatie zoekgedrag van consumenten probeert inzicht te verschaffen in dit zoekgedrag. Tenslotte is het Web een interactief medium. Een bezoeker hoeft niet passief te wachten wat hij of zij voorgeschoteld krijgt, maar kan actief interacteren met de Web site. Dit interactieve karakter is het onderwerp van het onderzoek naar het navigatie gedrag in on-line computer omgevingen. Deze drie stromen van onderzoek zullen in de volgende paragrafen nader worden toegelicht.

2.1. Informatie zoekgedrag

Een belangrijke fase in het aankoopbeslissingsproces is het verzamelen van relevante informatie. Het zoeken naar informatie kan zowel intern als extern plaatsvinden. Intern zoeken omvat het gebruiken van informatie die al in een eerder stadium is verzameld. Deze informatie is dus al in het bezit van de consument. Indien intern zoeken niet de benodigde informatie oplevert zal door middel van extern zoeken extra informatie uit

de omgeving worden gezocht. Bronnen van informatie die bij dit externe zoekproces kunnen worden gebruikt zijn bijvoorbeeld: persoonlijke verkoop, advertenties, brochures, productinformatie op de verpakking, catalogi, onafhankelijke organisaties zoals de consumentenbond en persoonlijke bronnen zoals vrienden en bekenden. De mate waarin een consument op zoek zal gaan naar informatie, de zoekintensiteit, is niet in alle gevallen even groot. Bij een belangrijke aankoop zal een consument eerder geneigd zijn om een grote hoeveelheid informatie te verzamelen. Zo zijn er een groot aantal factoren die de zoekintensiteit beïnvloeden. Spreng en Schmidt (1996) hebben deze factoren in kaart gebracht. Een belangrijke indeling die zij hanteren is factoren die de mogelijkheid tot zoeken beïnvloeden en factoren die de motivatie tot zoeken beïnvloeden.

Kennis beïnvloedt zowel de mogelijkheid tot zoeken als de motivatie tot zoeken. Hierbij is het belangrijk een onderscheid te maken tussen objectieve en subjectieve productkennis. Objectieve kennis is wat een consument daadwerkelijk weet, subjectieve kennis is wat de consument denkt te weten en de mate waarin hij zeker is van deze kennis. Onderzoek heeft aangetoond dat de inschatting van de kennis lang niet altijd overeenkomt met de daadwerkelijke kennis (Schacter, 1983). Een hoge subjectieve productkennis leidt tot een hogere mogelijkheid tot het zoeken van informatie. Echter, een hoge subjectieve productkennis leidt ook tot een lagere motivatie tot zoeken, immers, waarom zou je extra informatie zoeken indien je 'meent' al heel veel te weten. De toegevoegde waarde van nieuwe informatie wordt steeds lager.

Betrokkenheid en de neiging tot nadenken hebben met name invloed op de motivatie tot zoeken. Betrokkenheid is de gepercipieerde persoonlijke relevantie van een object of situatie (Celsi and Olson, 1988). De consument zal meer informatie zoeken indien het onderwerp of product voor hem of haar belangrijk is. De neiging tot nadenken heeft eveneens een positieve invloed op de intensiteit van informatie zoeken. De neiging tot nadenken is de neiging van een individu om na te denken en dit te leuk te vinden. Mensen met een hoge neiging tot nadenken worden verondersteld een grotere behoefte aan informatie te hebben en meer plezier te beleven aan het zoeken en verwerken van informatie.

Naast de hiervoor genoemde factoren noemen Spreng en Schmidt diverse andere factoren die de zoekintensiteit beïnvloeden. Wij nemen slechts een deel van deze factoren mee. In dit onderzoek richten wij ons op de kenmerken van de consument. Tot deze kenmerken behoren de reeds genoemde variabelen zoals opleiding, leeftijd, betrokkenheid, neiging tot nadenken en productkennis. Factoren die bijvoorbeeld specifiek gerelateerd zijn aan de aankoopsituatie, zoals het aantal alternatieven en de mate waarin alternatieven van elkaar verschillen zullen wij hierbij buiten beschouwing laten. Ook situationele factoren zoals de tijdsrestricties en het gepercipieerde risico zullen wij niet meenemen.

2.2. Navigatie gedrag

Flow is essentieel voor een goed begrip van de navigatie van consumenten in on-line computer omgevingen zoals het World Wide Web (Hoffman en Novak , 1996). Flow kan worden omschreven als een toestand die optreedt tijdens het navigeren in een computeromgeving die wordt gekenmerkt door een naadloos aaneensluitende stroom van interacties met de computer, intrinsiek plezierig is, vergezeld gaat met een vermindering van het zelfbewustzijn en zelfversterkend is. Belangrijke beinvloeders van flow zijn de vaardigheden van een consument om met het medium om te gaan en de uitda-

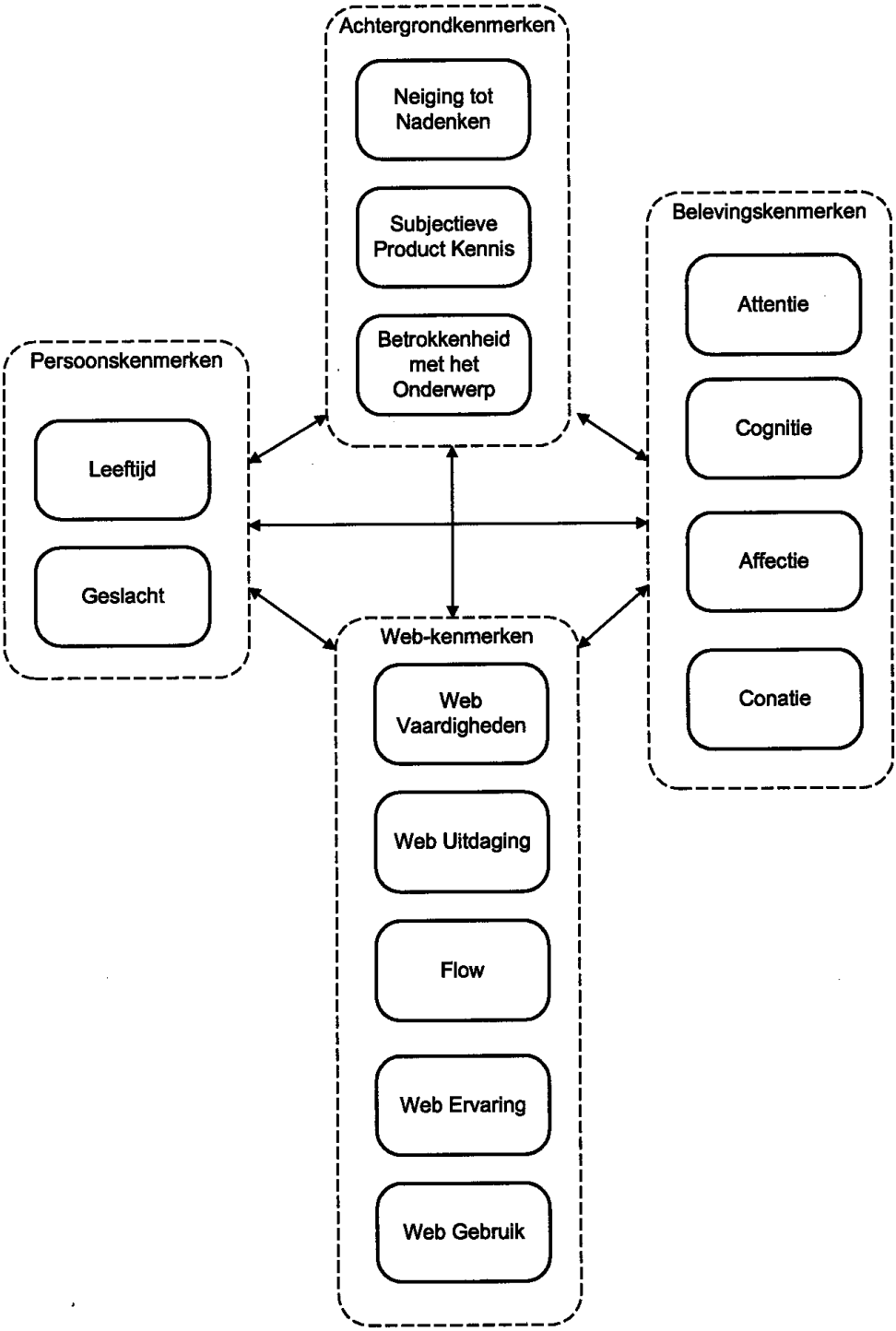
ging die het medium biedt. Flow wordt verondersteld een positief effect te hebben op de hoeveelheid informatie die iemand opneemt en te leiden tot een positieve affectie.

2.3. Reclame effectiviteits onderzoek

In de laatste decennia is veel onderzoek verricht naar de effecten van reclame-uitingen. Diverse studies zijn verricht om de invloed van bepaalde elementen van een communicatie uiting op de reactie van de consument vast te stellen (Gelb et al., 1985, Belch et al., 1987, Percy and Rossiter, 1983). Deze reactie wordt vaak gecategoriseerd aan de hand van een hiërarchie van effecten. Deze hiërarchie stelt dat een commerciële boodschap de aandacht van een respondent moet krijgen (attentie), en vervolgens leidt tot cognitieve, affectieve en conatieve effecten. Cognitie heeft betrekking op hetgeen iemand weet van een reclame- of communicatie-uiting, affectie heeft betrekking op de vraag of de kennis heeft geleid tot positieve of negatieve gevoelens van die persoon ten aanzien van de reclame uiting, het product of de organisatie. Conatie beschrijft of deze affectie ook leidt tot daadwerkelijk gedrag, bijvoorbeeld het kopen van een product. De volgorde van deze effecten verschilt per model. Naast dit model zijn andere effectmodellen beschreven in de literatuur. In de verschillende modellen zijn veel overeenkomstige factoren opgenomen die de effecten beïnvloeden. Een mogelijke indeling is: bron karakteristieken, aard van de boodschap, structuur van de boodschap, psycholinguïstiek en het effect van muziek en plaatjes.

De elementen waarvan de effectiviteit in voorgaand onderzoek is vastgesteld hebben, vertaald naar een onderzoek op het Web, betrekking op het ontwerp van de Web site. Omdat wij in dit onderzoek met name geïnteresseerd zijn in de effecten van kenmerken van de bezoeker op de beleving van de Web site, gebruiken wij de concepten die het effect van een reclame- of media-uiting tot uitdrukking brengen. Wij hanteren dan ook de concepten: attentie, cognitie, affectie en conatie.
De drie deelonderzoeken zijn geïntegreerd tot een onderzoeksmodel om de bestaande theorieën toe te passen en aan te passen op de specifieke kenmerken van het World Wide Web. Als afhankelijke variabelen hanteren wij attentie, cognitie, affectie en conatie uit het reclame onderzoek. De kenmerken zoals die in kaart zijn gebracht, zijn vervolgens in een viertal groepen ingedeeld. Deze groepen zijn persoonskenmerken, achtergrondkenmerken, Web kenmerken en belevingskenmerken. Het totale onderzoeksmodel is weergegeven in figuur 1.
De kenmerken die in nevennstaande figuur zijn genoemd zullen kort worden gedefinieerd:

- de *neiging tot nadenken* is de neiging van een individu om na te denken en dit leuk te vinden,
- *subjectieve productkennis* is wat de consument denkt te weten over een product en de mate waarin hij zeker is van deze kennis,
- *betrokkenheid* is de gepercipieerde persoonlijke relevantie van een object of situatie,
- *Web vaardigheden* betreft de capaciteiten van een persoon om met het medium World Wide Web om te gaan,
- *Web uitdaging* is de uitdaging die het Web biedt voor een consument,
- *Web ervaring* is de periode dat een consument reeds gebruik maakt van het Web,
- *Web gebruik* is de intensiteit van het gebruik uitgedrukt in het aantal uren Web gebruik per week,



Figuur 1. Onderzoeksmodel gebruikersonderzoek.

- *attentie* is de aandacht die een uiting krijgt van een consument,
- *cognitie* heeft betrekking op hetgeen iemand weet c.q. heeft 'geleerd' van een reclame- of communicatie uiting,
- *affectie* heeft betrekking op de vraag of aandacht of de kennis heeft geleid tot positieve of negatieve gevoelens ten aanzien van de reclame uiting, het product of de organisatie,
- *conatie* beschrijft of deze affectie ook leidt tot daadwerkelijk gedrag, bijvoorbeeld het kopen van een product.

In dit onderzoek willen wij nagaan op welke manier de diverse kenmerken en groepen van kenmerken samenhangen en welke persoons-, achtergrond- en Web-kenmerken de grootste invloed hebben op de beleving van Web sites.

3. ONDERZOEKSOPZET

Om het onderzoeksmodel te testen is gebruik gemaakt van een groep van gebruikers van het World Wide Web. Deze groep gebruikers bestond uit studenten van de universiteit. Deze mensen werden uitgenodigd naar het Rekencentrum van de Rijksuniversiteit Groningen te komen, waar zij werden gevraagd te surfen op het World Wide Web. Hiertoe kregen zij een aantal opdrachten. Deze opdrachten worden beschreven in het volgende kader.

Opdracht 1. Op vakantie...
Je hebt net tentamens gedaan en nu heb je zin om plannen te gaan maken voor de zomervakantie. Heerlijk er even tussenuit. Maar waar naartoe? Naar de zon, naar de sneeuw, naar het strand of naar de bergen? Of liever zeilen over de Friese meren, de eindeloosheid van de Spaanse hoogvlakte bewonderen, met een rugzak op een trektocht maken door de Alpen, of gewoon lekker naar Schier? Bedenk, wellicht na raadpleging van Internet, een aantrekkelijke bestemming. De volgende stap is: hoe kom ik daar en hoe verblijf ik daar. Zoek dus uit hoe je bij je bestemming kunt komen (wijze van vervoer, naam vervoerder, kosten, etc.) en waar je kunt overnachten. Bij het samenstellen van je reis kun je onder meer gebruik maken van de onderstaande sites. …..

Opdracht 2. Kadootjes geven
Het is eind november, het weer is guur, de regen slaat tegen de ramen en binnen snort de kachel. Helaas, je moet er nog uit, want je moet voor nog drie personen in je kennissen- en/of familiekring een kadootje voor Sinterklaas kopen. Elk van deze personen wil graag een boek of cd. Gelukkig, er is Internet. Zoek voor drie vrienden c.q. familieleden een boek of cd die zij leuk zullen vinden. Per persoon mag je maximaal 50 gulden besteden. Bij je zoektocht kun je onder meer gebruik maken van de onderstaande sites. ……

Opdracht 3 Voor jezelf
Iedereen heeft wel een hobby. Kies één van je favoriete hobbies en probeer op Internet hier meer informatie over te vinden. Zoek naar sites met algemene informatie, met specifieke producten, met aanvullende diensten, met tips, met ervaringen van mensen met dezelfde hobby als jij, etc. Bij je zoektocht kun je onder meer gebruik maken van de hieronder genoemde zoekmachines (search engines). ……..

Voor elke opdracht was een half uur beschikbaar. Het zoekproces werd na 20 minuten onderbroken. Op dat moment werd aan de deelnemers gevraagd een vragenlijst in te vullen. Deze vragenlijst had betrekking op de Web site specifieke concepten uit het onderzoeksmodel. De concepten in het onderzoeksmodel, zoals productkennis, flow en neiging tot nadenken zijn op zich moeilijk direct te meten. Daarom is gebruik gemaakt van stellingen die aan de respondenten zijn voorgelegd. De stellingen zijn voor een groot deel afkomstig van onderzoek op de desbetreffende deelgebieden. Enkele stellingen zijn aangepast aan het Web en een aantal stellingen zijn zelf ontwikkeld.

De respondenten konden op deze stellingen aangeven of ze het eens of oneens waren met de stelling en de mate waarin men het oneens of eens was[1]. Op basis van de beoordelingen van een aantal stellingen behorende bij een concept is vervolgens een score voor elke respondent op elk concept vastgesteld. Een voorbeeld van de stellingen die behoren bij het concept vaardigheden zijn:

1. Ik ben zeer vaardig in het gebruik van het Web.
2. Ik weet veel van goede zoektechnieken op het Web.
3. Ik weet iets minder van het gebruik van het Web dan de meeste andere gebruikers (R).
4. Ik weet hoe ik iets kan vinden op het Web.

De scores van een respondent op deze 4 stellingen die betrekking hebben op de vaardigheden van het Web gebruik leidden tot een score op het concept vaardigheid. Met behulp van confirmatory factor analysis en reliability analysis is nagegaan of de stellingen daadwerkelijk het concept op een eenduidige manier meten[2].

80 deelnemers voerden drie opdrachten uit. Dit resulteerde in 240 ingevulde vragenlijsten. Naast de Web specifieke vragenlijsten was er één vragenlijst waarin een aantal achtergrondkenmerken van de deelnemers werd gevraagd.

4. RESULTATEN

In dit hoofdstuk komen de resultaten van het onderzoek aan de orde. In de eerste paragraaf zal een beschrijving worden gegeven van de kenmerken van de deelnemers aan het onderzoek. In de daaropvolgende paragrafen zullen achtereenvolgens de Web kenmerken, de achtergrondkenmerken en de beleving worden beschreven. In de laatste paragraaf wordt aandacht besteed aan de samenhang tussen al deze kenmerken.

Bij het lezen van de resultaten dient te worden gerealiseerd dat de resultaten niet representatief zijn voor de gehele Internet populatie. In dit onderzoek staat niet de beschrijving van de Internet gebruiker centraal; het gaat om het verband tussen kenmerken van de gebruiker en de beleving van de site. De verdeling van de persoons- en achtergrond kenmerken hoeft dus niet representatief te zijn voor de Internet populatie.

4.1. Persoonskenmerken

Van het totaal aantal deelnemers was 1/3 deel vrouw en 2/3 deel man. De gemiddelde leeftijd van de deelnemers was 21.7 jaar. De leeftijd is verder niet in de analyses betrokken vanwege de geringe spreiding in leeftijd in de onderzoekspopulatie.

4.2. Web kenmerken

Het gemiddelde aantal maanden dat de deelnemers reeds van Internet gebruik maakten was 19,3. Hierbij is een duidelijk verschil waar te nemen tussen mannen en vrouwen (zie tabel 1). Mannen maakten gemiddeld reeds bijna 22 maanden gebruik van Internet, vrouwen daarentegen maakten nog maar een kleine 15 maanden gebruik van Internet.

Tabel 1. Aantal maanden Internet gebruik.

Geslacht	Aantal	Gemiddelde	Std. Deviation
Man	52	21.6	13.3
Vrouw	26	14.6	14.0

p = .035

Het gemiddelde aantal uren per week dat de deelnemers het Web gebruikten bedroeg 3,95. Ook hier is een klein verschil tussen mannen en vrouwen waar te nemen (zie tabel 2). Mannen maakten gemiddeld 4,1 uur gebruik van het Web, vrouwen maakten gemiddeld 3,6 uur gebruik van Internet. Dit verschil is echter niet significant. Dit betekent dat alhoewel mannen al duidelijk langer gebruik maken van Internet, dit geen invloed heeft op de intensiteit van het gebruik.

Tabel 2. Aantal uren per week gebruik van Internet.

Geslacht	Aantal	Gemiddelde	Std. Deviation
Man	52	4.1	3.3
Vrouw	25	3.6	3.8

Not significant

Tevens werd de vraag gesteld of men ooit iets gekocht heeft via Internet. Een grote meerderheid, 96,2%, heeft nog nooit iets aangeschaft via Internet. 3,8% van de respondenten heeft wel eens iets gekocht via Internet. Het opvragen van informatie via Internet is een stuk populairder dan het doen van aankopen via Internet. 87% van de respondenten gaf aan wel eens extra informatie via Internet aangevraagd te hebben.

Naast de ervaring en de intensiteit van het Web gebruik zijn Web vaardigheden, uitdaging, en flow elk in de vorm van een aantal stellingen gemeten als Web kenmerken.
Vervolgens is met behulp van de correlatie coefficient gekeken hoe de kenmerken onderling samenhangen. Uit tabel 3 blijkt dat er een positieve relatie bestaat tussen het aantal uren per week dat men van het Web gebruik maakt en de periode dat men reeds

het Web gebruikt. Deze beide kenmerken, ervaring en gebruik, hangen op hun beurt samen met een hogere vaardigheid. Het is aannemelijk te veronderstellen dat een langer en intensiever gebruik van het Web leidt tot een hogere vaardigheid in dit gebruik. Deze verhoogde vaardigheid blijkt tenslotte samen te hangen met een hogere flow die men ervaart bij het gebruik van het Web. Het is interessant om te constateren dat de intensiteit en lengte van het gebruik van het Web niet samenhangen met de uitdaging die het Web biedt.

Tabel 3. Samenhang Web kenmerken.

	Vaardigheid	Uitdaging	Ervaring	Gebruik
Uitdaging	-.077			
Ervaring	.526**	-.034		
Gebruik	.615**	.020	.260**	
Flow	.156*	.060	-.038	.191**

** Correlation is significant at the 0.01 level (2-tailed).
* Correlation is significant at the 0.05 level (2-tailed).

De samenhang tussen de vaardigheid in het gebruik van het Web en de periode dat men het Web gebruikt is nader toegelicht in de onderstaande figuur. Het is niet onaannemelijk te veronderstellen dat mensen die het Web intensiever gebruiken ook een grotere vaardigheid in het gebruik van het medium ontwikkelen.

In tabel 4 is de score op de Web kenmerken gesplitst naar mannen en vrouwen. Uit deze tabel valt af te leiden dat mannen significant hoger scoren op vaardigheden. Dit betekent dat mannen zichzelf gemiddeld vaardiger vinden in het gebruik van het Web dan vrouwen. Voor de overige kenmerken zijn kleine maar niet significante verschillen te constateren.

Tabel 4. Gemiddelde score op kenmerken.

Construct	Man	Vrouw
Vaardigheden	3.6	3.1*
Uitdaging	2.8	3.0
Flow	3.0	3.2

* significant difference at the 0.05 level

4.3. Achtergrondkenmerken

Vervolgens zijn de achtergrond kenmerken van de deelnemers geanalyseerd. De achtergrondkenmerken hebben betrekking op de productkennis, de betrokkenheid bij het onderwerp en de neiging tot nadenken. Uit tabel 5 blijkt dat er een significante samenhang bestaat tussen de productkennis en de betrokkenheid. Mensen die veel van een product denken af te weten, voelen zich hier ook meer bij betrokken.

Tabel 5. Samenhang achtergrondkenmerken.

	Neiging tot nadenken	Productkennis
Productkennis	-.020	
Betrokkenheid	.081	.470**

** Correlation is significant at the 0.01 level (2-tailed).

In tabel 6 is de score op de achtergrondkenmerken gesplitst naar mannen en vrouwen. Uit deze tabel valt af te leiden dat mannen significant hoger scoren op subjectieve productkennis. Dit betekent dat mannen gemiddeld beweren meer over het product te weten dan vrouwen. Vrouwen hebben een significant hogere neiging tot nadenken. Voor betrokkenheid is een klein maar niet significant verschil te constateren.

Tabel 6. Gemiddelde score op achtergrondkenmerken.

Construct	Man	Vrouw
Betrokkenheid	2.2	2.2
Productkennis	3.3	2.9*
Neiging tot nadenken	3.8	4.0*

* significant difference at the 0.05 level

4.4. Achtergrondkenmerken versus Web kenmerken

Op basis van tabel 7 is gekeken of er een samenhang bestaat tussen de hiervoor genoemde achtergrond- en Web kenmerken. Uit de tabel blijkt dat de neiging tot nadenken positief samenhangt met de tijdsperiode dat men het Web gebruikt. Mensen met een hogere neiging tot nadenken zullen volgens de theorie meer informatie gebruiken. Vanuit deze theorie is een eerder en daardoor langer gebruik van het Web, als grote informatiebron, te verklaren. Deze langere periode van Web gebruik leidt weer tot een hogere vaardigheid in het Web gebruik. Neiging tot nadenken hangt overigens niet samen met de intensiteit van het Web gebruik. Het is opvallend dat neiging tot nadenken negatief samenhangt met de uitdaging. Dit is wellicht te verklaren uit een rationeler gebruik van het Web. Het Web wordt wellicht door personen met een hogere neiging tot nadenken meer gezien als een van de alternatieve informatiebronnen dan als een medium dat echt een uitdaging biedt. Verder bevestigt tabel 7 het idee dat een hogere productkennis en daarmee een hogere betrokkenheid leiden tot een hogere flow.

Tabel 7. Samenhang achtergrond- en Web kenmerken.

	Vaardigheid	Uitdaging	Ervaring	Gebruik	Flow
Neiging tot nadenken	.140*	-.142*	.329**	-.058	-.054
Productkennis	.180**	-.038	.095	.089	.301**
Betrokkenheid	.023	.010	.054	-.071	.465**

** Correlation is significant at the 0.01 level (2-tailed).
* Correlation is significant at the 0.05 level (2-tailed).

4.5. Beleving

Uit tabel 8 blijkt dat de belevingskenmerken allemaal onderling sterk samenhangen. Met name flow en attentie; attentie en affectie; en affectie en cognitie hangen sterk met elkaar samen. Voor elk van deze combinaties geldt dat er sprake is van een positieve samenhang. Dat wil zeggen dat een hoge score op de ene variabele samenvalt met een hoge score op de andere variabele. Wel wordt de hiërarchie van effecten in zoverre bevestigd dat conatie het sterkst samenhangt met affectie (en attentie), affectie met cognitie (en attentie), cognitie met attentie en flow met attentie.

Tabel 8. Samenhang belevingskenmerken.

	Flow	Attentie	Cognitie	Affectie
Attentie	.607**			
Cognitie	.353**	.529**		
Affectie	.411**	.658**	.622**	
Conatie	.426**	.588**	.478**	.536**

** Correlation is significant at the 0.01 level (2-tailed).

4.6. Web-, achtergrond- en persoonskenmerken versus beleving

Een belangrijke doelstelling van dit onderzoek is vast te stellen welke kenmerken van invloed zijn op de beleving van een Web site. Hiertoe zijn de in de voorgaande paragrafen behandelde Web- en achtergrondkenmerken uitgezet tegen de beleving. De samenhang is weergegeven in tabel 9. Uit deze tabel blijkt dat met name de subjectieve productkennis, flow en de betrokkenheid bij het onderwerp sterk samenhang met alle aspecten van de beleving.

Tabel 9. Samenhang achtergrond en Web kenmerken en beleving.

	Flow	Attentie	Cognitie	Affectie	Conatie
Vaardigheid	.156*	.098	.145*	.140*	.074
Uitdaging	.060	.119	.104	.033	.050
Ervaring	-.038	.044	-.048	.013	-.034
Gebruik	.191**	.081	.138*	.066	.127
Flow	1.000	.607**	.353**	.411**	.426**
Neiging tot nadenken	-.054	.076	-.067	.093	-.032
Productkennis	.301**	.429**	.193**	.351**	.320**
Betrokkenheid	.465**	.636**	.330**	.526**	.550**

** Correlation is significant at the 0.01 level (2-tailed).
* Correlation is significant at the 0.05 level (2-tailed).

In tabel 10 is weergegeven in hoeverre er een samenhang bestaat tussen het geslacht van de deelnemers aan het onderzoek en de diverse gemeten belevingskenmerken.

Tabel 10. Gemiddelde score op belevingskenmerken.

Construct	Man	Vrouw
Flow	3.0	3.2
Attentie	3.6	3.7
Cognitie	3.7	3.6
Affectie	3.5	3.5
Conatie	3.0	3.0

Uit tabel 10 blijkt dat er ten aanzien van de meeste belevingskenmerken kleine verschillen bestaan tussen mannen en vrouwen. Deze verschillen zijn echter niet dusdanig groot dat men met een vrij grote zekerheid kan zeggen dat er daadwerkelijke verschillen zijn. Zoals in tabel 4 en tabel 6 is aangegeven zijn er wel significante verschillen ten aanzien van subjectieve productkennis, Web vaardigheden en neiging tot nadenken. Deze effecten zijn echter niet zo sterk, en heffen elkaar ten dele op, zodat ze geen invloed hebben op de beleving van een Web site en de gedragsintenties. De gemiddelde scores ten aanzien van flow, attentie, cognitie, affectie en conatie zijn bijna identiek voor mannen en vrouwen.

In de voorgaande analyses is per kenmerk gekeken in hoeverre deze samenhangt met een hoge dan wel lage score op de beleving. In de volgende analyses zal gekeken worden welke kenmerken de grootste verklarende waarde hebben in het verklaren van de beleving van de site. Dit is gedaan middels een regressie analyse waarbij de afhankelijke variabele, de beleving, wordt verklaard uit een of meerdere onafhankelijke variabelen, de kenmerken. Flow is hierbij zowel als afhankelijke als onafhankelijke variabele opgenomen door de centrale rol die dit concept speelt.

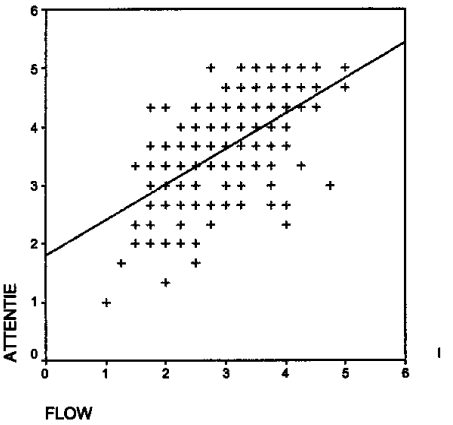
Tabel 11. Regressie-analyse beleving vs. achtergrond en Web kenmerken.

Afhankelijk Onafhankelijk	Flow	Attentie	Cognitie	Affectie	Conatie
Neiging tot nadenken					
Subjectieve productkennis		**			
Betrokkenheid	**	**	**	**	**
Ervaring	**		**		
Gebruik	**				**
Geslacht	**				
Vaardigheden	**				
Uitdaging		**			
Flow		**	**	**	**
R^2	.320	.548	.198	.337	.366

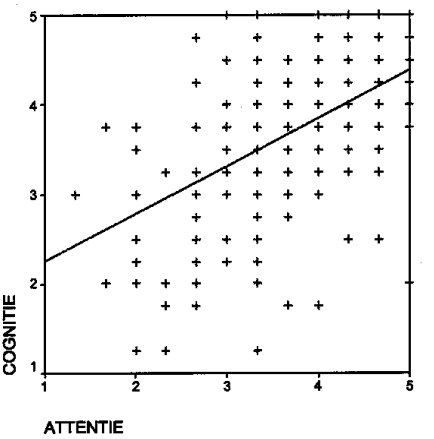
In de vorige tabel is aangegeven welke concepten met elkaar samenhangen. Middels een regressie analyse is gekeken welke onafhankelijke concepten de grootste verklarende kracht hebben. Uit de tabel blijkt dat attentie met name verklaard wordt door de productkennis, de betrokkenheid, de uitdaging en de flow. Bij het verklaren van cognitie,

affectie en conatie spelen betrokkenheid en flow een belangrijke rol. Betrokkenheid heeft hierbij zowel een directe invloed als een indirecte invloed via het concept flow. De R^2 geeft aan welk percentage van de variantie in de afhankelijke variabele wordt verklaard door de onafhankelijke variabelen.

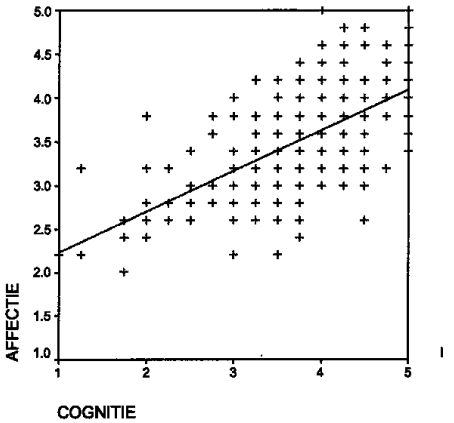
Tenslotte is naar het verband tussen de afhankelijke variabelen attentie, cognitie, affectie en conatie gekeken. In tabel 8 is aangegeven dat er een sterke samenhang bestaat tussen de verschillende kenmerken van de beleving. Volgens bepaalde hiërarchie van effect modellen gaat attentie vooraf aan cognitie. Cognitie leidt tot een bepaalde affectie, welke weer leidt tot een bepaald gedrag oftewel conatie. In de volgende grafieken is de relatie tussen deze concepten weergegeven



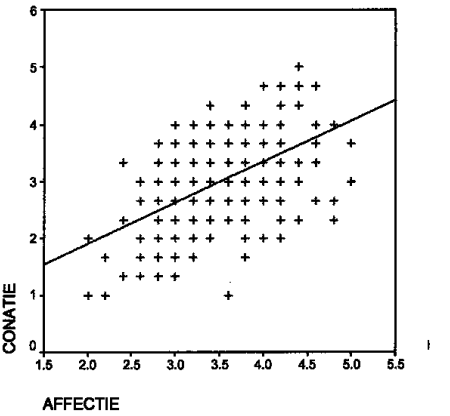
Figuur 2. Flow → attentie



Figuur 3. Attentie → cognitie



Figuur 4. Cognitie → affectie



Figuur 5. Affectie → conatie

Uit de voorgaande vier figuren blijkt dat er daadwerkelijk een positieve relatie bestaat tussen flow en attentie; attentie en cognitie; cognitie en affectie en tenslotte affectie en conatie.

5. SAMENVATTING EN CONCLUSIES

In dit onderzoek zijn verschillende theorieën samengebracht die relevante elementen van het Web gebruik beschrijven. Deze theorieën hebben betrekking op reclame effectiviteitsonderzoek, onderzoek naar informatiezoekgedrag en onderzoek naar consumenten navigatie gedrag in on-line computer omgevingen. Deze theorieën zijn relevant omdat het Web kenmerken vertoond van een advertentie medium en een informatiebron en bovendien interactief is.

Vervolgens is onderzocht welke kenmerken van gebruikers invloed hebben op het informatiezoekgedrag en de beleving van een Web site. De neiging tot nadenken van een persoon blijkt positief samen te hangen met de tijdsperiode dat men het Web gebruikt. Mensen met een hogere neiging tot nadenken zullen volgens de theorie meer informatie gebruiken. Vanuit deze theorie is een eerder en daardoor langer gebruik van het Web, als grote informatiebron, te verklaren. Uit de resultaten van dit onderzoek blijkt dat er een positieve relatie bestaat tussen het aantal uren per week dat men van het Web gebruik maakt en de periode dat men reeds het Web gebruikt. Deze beide kenmerken, ervaring en gebruik, hangen op hun beurt samen met een hogere vaardigheid. Deze verhoogde vaardigheid blijkt tenslotte samen te hangen met een hogere flow die men ervaart bij het gebruik van het Web. Ook een hogere productkennis en daarmee een hogere betrokkenheid leiden tot een hogere flow. Deze flow in combinatie met betrokkenheid spelen een belangrijke rol in het voorspelen van de belevingskenmerken.

Een belangrijke doelstelling van dit onderzoek is vast te stellen welke kenmerken van invloed zijn op de beleving van een Web site. Uit een regressie analyse is gebleken dat attentie met name verklaard wordt door de productkennis, de betrokkenheid, de uitdaging en de flow. Bij het verklaren van cognitie, affectie en conatie spelen betrokkenheid en flow een belangrijke rol. Ook heeft dit experiment laten zien dat er duidelijke relaties bestaan tussen de concepten attentie, cognitie, affectie en conatie. Verder onderzoek is nodig om de oorzakelijke gevolgrelaties tussen deze concepten aan te tonen. Tevens is verder onderzoek nodig om een koppeling te leggen met de kenmerken van de Web site. Naast de kenmerken van de gebruiker is het van belang te weten welke kenmerken van de Web site invloed hebben op de beleving om op basis van deze kennis richtlijnen ten aanzien van het ontwerp van een Web site te kunnen geven.

NOTEN

1. De antwoordcategorieën die zijn gebruikt zijn: helemaal oneens, oneens, neutraal, eens, helemaal eens.
2. Zie Appendix A voor een nadere beschrijving van de meting van de constructs aan de hand van de stellingen.

LITERATUUR

Belch, G.E., M.A. Belch and A. Villarreal (1987), Effects of advertising communications: review of research, *Research in Marketing,* 9, pp. 59-117.
Berthon, P., L. Pitt en R.T. Watson (1996), The World Wide Web as an Advertising Medium:

Toward an Understanding of Conversion Efficiency, *Journal of advertising research,* 36(1), p. 43-54.
Briggs, R. and N. Hollis (1997), Advertising on the Web: is there Response before click through, *Journal of Advertising Research,* march/april, pp. 33-45.
Cronin, M.J. (1995), *Doing more business on the Internet,* Van Nostrand Reinhold, New York
Dreze, X and F. Zufryden (1997), Testing Web site design and promotional content, *Journal of Advertising Research,* march/april, pp. 77-91.
Ducoffe, R.H. (1996), Advertising value and advertising on the Web, *Journal of Advertising Research,* September/october.
Gelb, B.D., J.W. Hong and G.N. Zinkhan (1985), *Communication effects of specific advertising elements: an update,* Current issues and research in advertising.
Maddox, L.M. and D. Mehta (1997), The role and effect of Web adresses in advertising, *Journal of Advertising Research,* march/april, pp. 47-59.
Mehta, R., R. Grewal en E. Sivadas (1995), International Direct Marketing on the Internet: do Internet users form a global segment, *Journal of Direct Marketing,* Vol. 10, no. 1, pp. 45-58.
Hoffman, D.L. en T.P. Novak (1996), Marketing in Hypermedia Computer Mediated Environments: Conceptual Foundations, *Journal of Marketing,* 60, pp. 50-68.
Paul, P. (1996), Marketing on the Internet, *Journal of Consumer Marketing,* Vol 13, no. 4, pp. 27-39.
Percy, L. en J.R. Rossiter (1983), Mediating effect of visual and verbal elements in print advertising upon belief, attitude and intention responses, in Advertising and Consumer behavior, L.Percy en A. Woodside, New York, Lexington Press.
Quelch, J. en L. Klein (1996), Internet en International Marketing, *Holland Management Review,* nr. 48, pp. 44-57.
Schmidt, J.B. en R.A. Spreng (1996), A Proposed Model of External Consumer Information Search, *Journal of the Academy of Marketing Science,* Vol. 24, pp. 246-256.

Bijlage

In de onderstaande tabel staan de resultaten van de confirmatory factor analysis en de reliability analysis weergegeven.

	Items	GFI	AGFI	alpha
Neiging tot nadenken	5	0.95	0.85	0.76
Productkennis[3]	2	0.99	0.92	0.84
Betrokkenheid[3]	2	0.99	0.92	0.84
Vaardigheden	4	0.99	0.97	0.89
Uitdaging	3	0.99	0.96	0.80
Flow	4	0.98	0.90	0.73
Attentie	3	1.00	0.98	0.81
Cognitie	4	0.98	0.92	0.83
Affectie	5	0.98	0.94	0.67
Conatie	3	1.00	0.98	0.60

Items: het aantal items / stellingen waaruit een construct bestaat
GFI: goodness of fit statistiek
AGFI: adjusted goodness of fit statistiek
Alpha: reliability coefficient

3. Productkennis en betrokkenheid zijn in een model middels confirmatory factor analysis getoetst waarbij een schatting van de correlatie tussen beide constructen is gemaakt.

4. Het verzamelen van survey gegevens bij oudere respondenten

J.H. SMIT en W. DIJKSTRA

SAMENVATTING

De afgelopen decennia is de belangstelling voor het oudere deel van de bevolking (55 jaar en ouder) sterk gegroeid. Dit vindt ondermeer zijn weerslag in een toename van het onderzoek onder ouderen. Wat daarbij opvalt is dat het methodologisch onderzoek naar de condities waaronder kennis bij ouderen optimaal verzameld kan worden nagenoeg ontbreekt. In deze studie wordt summier een aantal basis principes besproken met betrekking tot nonrespons en instrumentontwikkeling. Tevens wordt met behulp van empirische gegevens een indicatie gegeven voor de datakwaliteit in survey onderzoek bij ouderen. De gevonden resultaten wijzen erop dat de datakwaliteit wel eens een serieus probleem kan zijn in onderzoek onder ouderen. Er zijn aanwijzingen dat de datakwaliteit afneemt naarmate mensen ouder worden.

1. INLEIDING

Het laatste decennium is er een sterk groeiende belangstelling voor het oudere deel van de bevolking. In dit artikel verstaan we daar gemakshalve personen van 55 jaar en ouder onder. Door de toename van het aantal ouderen groeit ook het besef dat ouderen een steeds belangrijkere rol in de samenleving (gaan) spelen. Vanuit dat besef ontstaat er tevens behoefte aan een verruiming van wetenschappelijke kennis van de factoren die met ouder worden samenhangen en worden antwoorden gezocht op vragen als: "Onder welke omstandigheden leven ouderen?", "In hoeverre zijn deze omstandigheden van invloed op gedrag, meningen en het welbevinden van ouderen?" en "Welke factoren beïnvloeden de autonomie en het welbevinden van ouderen?".
Daarnaast is er vanuit de overheid behoefte om informatie te verzamelen over de ouderen in de samenleving teneinde beleid te ontwikkelen dat optimaal inspeelt op de rol die ouderen (kunnen) vervullen in de samenleving (bijvoorbeeld tot op welke leeftijd men in het arbeidsproces actief wil zijn), en het beroep dat ouderen doen op voorzieningen (gebruik openbare instellingen, vervoer, zorgvoorzieningen, etc.). Een belangrijk signaal voor de behoefte aan meer kennis over ouderen werd in 1988 afgegeven toen bij zeven universiteiten het Nederlands Stimuleringsprogramma Ouderenonderzoek -NESTOR- van start ging. Voor het NESTOR programma werd vanuit overheid en universiteiten 15 miljoen gulden beschikbaar gesteld teneinde het ouderenonderzoek in Nederland goed op gang te brengen (voor een korte samenvatting en evaluatie van het programma zie van Santvoort, 1998). In 1998 werd een nieuw stimuleringsprogramma met betrekking tot onderzoek naar factoren die met ouder worden samenhan-

gen gestart, met als titel "Succesvol Ouder Worden". Hierbij investeert de overheid in samenwerking met NWO ca. 10 miljoen gulden om de autonomie en zelfredzaamheid van ouderen te laten onderzoeken en te bevorderen (voor een korte inleiding van dit stimuleringsprogramma zie Brandsma, 1998). De opvatting dat het belangrijk is om de ouderen goed in kaart te brengen wordt dan ook algemeen ondersteund.
Over het algemeen worden in sociaal-wetenschappelijk onderzoek gegevens verzameld door middel van persoonlijke, telefonische, dan wel schriftelijke vragenlijsten. Er bestaat veel literatuur over de methoden en technieken van deze verschillende vormen van dataverzameling en de kwaliteit van de met verschillende methoden verzamelde gegevens. Wat daarbij opvalt is dat slechts weinig aandacht besteed wordt aan methodologisch onderzoek naar de wijze waarop informatie bij ouderen zo goed mogelijk kan worden verkregen en welke technieken daarbij kunnen worden ingezet om informatie van hoge kwaliteit in termen van betrouwbaarheid en validiteit te verkrijgen. Een voorbeeld: de belangrijkste database met betrekking tot methodologisch onderzoek in sociaal-wetenschappelijk onderzoek wordt gevormd door de SRM database (SRM database of Social Research Methodology). In dat bestand (versie 1997) waarin meer dan 40,000 methodologische referenties zitten waren er slechts 34 traceerbaar als specifiek betrekking hebbend op oudere respondenten.
Deze geringe aandacht voor ouderen als zijnde een speciale groep respondenten roept de vraag op of de methodologie voor onderzoek onder ouderen inderdaad een aparte benadering behoeft. Daar is naar onze mening voldoende aanleiding voor. Ouderen hebben meestal een wat lagere opleiding, hebben meer te kampen met gezondheidsproblemen en hebben meer problemen met het gehoor en de visus die een goede informatie-uitwisseling kunnen verstoren. Ook de cognitieve vaardigheden die nodig zijn om informatie uit het geheugen op te halen nemen bij het ouder worden af, vaardigheden die bijvoorbeeld nodig zijn om retrospectieve vragen te beantwoorden.

Dit artikel heeft uiteraard niet de pretentie om deze tekortkomingen op methodologisch gebied op te lossen. Met behulp van een aantal voorbeelden uit de aanwezige literatuur en aan de hand van empirische gegevens verzameld in het kader van een groot bevolkingsonderzoek onder ouderen, zullen een aantal aspecten die van belang zijn bij het verzamelen en verkrijgen van kwalitatief goede gegevens bij oudere respondenten worden besproken.
Eerst bespreken we hoe het in het ouderenonderzoek gesteld is met de (non)respons bij verschillende dataverzamelingsmethoden en geven we enige tips inzake instrumentgebruik. Tenslotte geven we een indicatie van de datakwaliteit van gegevens verzameld in een empirisch onderzoek onder ouderen.

2. NONRESPONS

Zoals eerder gesteld worden surveygegevens over het algemeen verzameld middels persoonlijke dan wel telefonische interviews of door middel van schriftelijke vragenlijsten.
Bij survey-onderzoek is er vrijwel altijd sprake van het trekken van een steekproef uit de doelpopulatie teneinde gegevens te verzamelen waarmee in een later stadium algemene uitspraken over de doelpopulatie gedaan kunnen worden. Een van de factoren die het doen van algemene uitspraken kan beïnvloeden is de mate waarin nonrespons

optreedt. Het meeste onderzoek naar het voorkomen van nonrespons leidt tot de conclusie dat bij ouderen een hogere nonrespons wordt gevonden dan bij jongeren, ongeacht de manier waarop de gegevens verzameld worden (Kaldenberg, 1994, Herzog en Kulka, 1989). Vergelijken we de nonrespons bij verschillende vormen van dataverzameling, dan is binnen het ouderenonderzoek een vergelijkbare tendens aanwezig als bij ander survey-onderzoek. De respons is het hoogst bij persoonlijke interviews en het laagste bij schriftelijke vragenlijsten (Groves en Kahn, 1979; Dillman, 1978). Bij ouderen worden echter specifieke motivaties over het waarom van deze trend gegeven. Als reden wordt wel genoemd dat ouderen vaker geneigd zouden zijn om te weigeren bij een verzoek om deel te nemen aan een telefonisch interview of een schriftelijke vragenlijst dan aan een persoonlijk interview, omdat ouderen meer zicht- en gehoorproblemen hebben en bijvoorbeeld minder vaak gewend zijn om informatie te verschaffen via de telefoon (Corso, 1977).
Een tegengestelde opvatting is dat juist persoonlijke interviews vaker geweigerd worden door ouderen. Zij zouden uit angst om slachtoffer te worden van oplichting etc. minder snel geneigd zijn om relatief vreemden (interviewers) in hun huis toe te laten, met als meest genoemde argumentatie "reluctance to open the door to a stranger". Empirische ondersteuning van deze beide opvattingen wordt echter slechts zelden gevonden.

3. INSTRUMENTONTWIKKELING

In gerontologisch onderzoek wordt nog al eens gekozen voor een dichotoom antwoordformaat (ja/nee, eens/oneens, waar/onwaar) met als argument dat de taak van de respondent wordt vergemakkelijkt omdat er slechts een keuze gemaakt hoeft te worden uit twee alternatieven. Deze vooral op 'common sense' gebaseerde argumentatie werd echter weerlegd in een studie van Carp en Carp (1984). In deze studie blijkt dat oudere respondenten vragen met een dichotoom antwoordformaat niet prettig vinden. Met name het restrictieve karakter van de keuze die gemaakt moet worden, werd onaangenaam gevonden. Vragen met een antwoordformaat bestaande uit vijf antwoordcategorieën werden het meest gewaardeerd en leverden ook psychometrisch de beste resultaten op (het laagste percentage ontbrekende antwoorden en de beste beantwoordverdelingen in termen van normale verdeling en variantie).
In persoonlijke interviews kan gebruik gemaakt worden van hulpmiddelen. Een hulpmiddel waarvan bekend is dat het de taak van de interviewer en respondent vergemakkelijkt en waarbij aantoonbaar minder fouten worden gemaakt tijdens het stellen van de vraag en het geven van een antwoord, is de toonkaart waarop de antwoord-alternatieven staan aangegeven (Van der Zouwen en Dijkstra, 1996; Smit 1995). In een studie van Hoinville (1983) blijkt echter dat het niet goed kunnen lezen van de toonkaarten een probleem kan zijn in onderzoek bij oudere respondenten. Een mogelijke oplossing voor dit probleem kan worden gevonden in het fabriceren van toonkaarten in verschillend letterformaat. Voor het interview begint kan de respondent vervolgens een selectie maken welk formaat goed leesbaar is (Carp en Carp, 1984; Andrews en Withey, 1976).
Uit het bovenstaande blijkt dat de combinatie van vragen met meerdere antwoordcategorieën waarbij voor het aanbieden van de antwoordalternatieven ook nog eens gebruik gemaakt wordt van toonkaarten, relatief gezien gegevens van de beste kwaliteit oplevert.

4. DATAKWALITEIT

Een volgende belangrijke vraag is hoe het gesteld is met de datakwaliteit van de bij ouderen verzamelde gegevens. Als een indicatie voor de datakwaliteit hanteren we een aantal psychometrische indicatoren. De eerste indicator is de proportie item-nonrespons. Naarmate de item-nonrespons toeneemt neemt de bruikbaarheid van de gegevens af. De tweede indicator is de schaalbaarheid van respondenten op een bepaald construct. Afname van de schaalbaarheid werd geoperationaliseerd als het toenemen van de inter-itemvariantie van respondenten (de variantie tussen de items van een schaal binnen respondenten) en een lagere betrouwbaarheidsscore in termen van Cronbach's alpha.
Een andere kwaliteitsindicator is de invloed van de interviewer op de verkregen antwoordverdelingen. Het is bekend dat degene die de data verzamelt (de interviewer) invloed uit kan oefenen op de verkregen antwoorden (Smit, 1995). De ene interviewer zal de vragen duidelijker voorlezen, sneller een antwoord interpreteren in termen van een van de antwoordalternatieven, meer uitleg geven, instructies aanpassen, etc. Naarmate de taak van interviewer moeilijker is, neemt deze invloed toe maar zullen ook de interviewers systematisch van elkaar gaan verschillen wat betreft de uitvoering van hun taak. Interviewereffecten worden gewoonlijk gekwantificeerd in de zogenaamde interviewervariantie, oftewel de mate waarin verschillende interviewers verschillende gemiddelde scores verkrijgen op een vraag. In dat geval is een gedeelte van de variantie toe te schrijven aan de interviewer en niet aan de respondenten. Naarmate de hoeveelheid interviewervariantie toeneemt, neemt de onzekerheid over de validiteit van de verkregen score toe. Voor het bepalen van de kwaliteit van de gegevens kiezen we instrumenten waarbij de kans op fouten zo klein mogelijk is (lijkt). Gezien de discussie in paragraaf 3 kiezen we voor vragen waarbij meerdere antwoordalternatieven op een toonkaart worden aangeboden. (Uitgangspunt voor deze keuze is dat de kwaliteit van gegevens verkregen met meer complexe instrumenten altijd minder zal zijn.) De gebruikte empirische gegevens en gekozen instrumenten zullen eerst kort worden geïntroduceerd.

4.1. Respondenten

De gegevens voor deze studie werden verzameld in het kader van de Longitudinal Aging Study Amsterdam (LASA). LASA is een logitudinaal, multidisciplinair onderzoeksproject over veranderingen in autonomie en welbevinden bij ouderen (Deeg, Knipscheer en Van Tilburg, 1993). LASA volgt op de NESTOR studie naar leefvormen en sociale netwerken (Knipscheer, et al, 1995) en is gestart in 1990. Het betreft een grote, naar leeftijd en sexe gestratificeerde steekproef van personen van 55-85 jaar uit de bevolkingsregisters van 11 gemeenten in Nederland (n=3107). De steekproef is elders gedetailleerd beschreven (Smit en De Vries, 1994). In deze studie zullen gegevens gebruikt worden die verzameld zijn tijdens de eerste LASA meting die plaatsvond in de periode 1992/1993.

4.2. Meetinstrumenten

De leeftijd van de respondenten op het tijdstip van het interview werd berekend door het verschil te nemen tussen de afnamedatum van het interview en de geboortedatum van de respondenten verkregen via de bevolkingsregisters. Vervolgens werd deze va-

riabele gedichotomiseerd (55 tot en met 74 jaar en 75 tot en met 85 jaar). Twee variabelen gemeten met meerdere indicatoren vormen de basis voor het onderzoeken van de datakwaliteit. De eerste variabele is een indicator met betrekking tot hoe mensen de eigen competentie inschatten (algemene self efficacy of ALCOS; voor een beschrijving zie Bosscher, Smit en Kempen, 1997). Deze zogenaamde ALCOS schaal wordt gevormd door twaalf items met een vijfpunts antwoordformaat aangeboden op een toonkaart (bijvoorbeeld: "Ik twijfel aan mijzelf. 1 'zeer mee oneens', 2 'mee oneens', 3 'niet mee oneens/niet mee eens', 4 'mee eens', 5 'zeer mee eens'). De tweede variabele is een indicator voor depressie, de zogenaamde Center for Epidemiologic Depression Scale (CES-D; voor een beschrijving zie Beekman, Van Limbeek, Deeg, Wouters en Van Tilburg, 1994). De respondent krijgt twintig uitspraken voorgelegd waarna hij moet aangeven in hoeverre de uitspraak op hem van toepassing is (bijvoorbeeld: "De afgelopen week had ik geen zin in eten, was mijn eetlust slecht. 0 'zelden of nooit', 1 'soms', 2 'vaak', 3 'altijd of bijna altijd'). Ook deze antwoorden stonden vermeld op een toonkaart.

De interviews werden afgenomen door 43 verschillende interviewers. De interviewers werden 5 dagen zeer intensief getraind (zie voor de gehanteerde procedures Smit, Van den Eeden, Deeg en Beekman, 1995). Zowel van de ALCOS als van de CES-D zijn achtereenvolgens voor beide groepen respondenten (55-74 jaar versus 75-85 jaar) de item nonrespons, de gemiddelde inter-itemvariantie, de betrouwbaarheid in termen van Chronbach's alpha en de interviewervariantie als kwaliteitsindicatoren vastgesteld en vergeleken. Een overzicht kan worden gevonden in tabel 1.

Tabel 1. Data kwaliteitsindicatoren voor de ALCOS en CES-D voor twee groepen respondenten.

	55-74 jaar	75 t/m 85 jaar
CES-D schaal		
Gemiddelde proportie item-nonrespons*	0,12	0,19
Gemiddelde inter-item variantie*	0,49	0,61
Cronbach's alpha*	0,88	0,85
Percentage interviewer variantie*	5	8
ALCOS schaal		
Gemiddelde proportie item-nonrespons*	0,05	0,13
Gemiddelde inter-item variantie	0,86	0,90
Cronbach's alpha	0,69	0,67
Percentage interviewer variantie*	3	6

* verschil tussen beide groepen significant ($p < 0,05$).

De proportie item-nonrespons bleek hoger te zijn bij de oudste groep respondenten voor beide instrumenten. Ook de gemiddelde inter-item variantie was hoger voor de oudste groep en in het verlengde daarvan de Cronbachs alpha's lager. Tenslotte bleek ook de interviewervariantie in beide groepen aanzienlijk te zijn (in het meeste onderzoek wordt een interviewervariantie van minder dan 1% gevonden). Ook hier was deze aan de interviewer toe te schrijven variantie in de oudste groep aanzienlijk hoger dan in de groep van 55 t/m 74 jaar. Naarmate respondenten ouder zijn lijkt de kwaliteit van de verzamelde gegevens af te nemen.

5. SAMENVATTING EN AANBEVELINGEN

Het nadeel van het vergelijken van responscijfers is dat men niet precies weet waar de verschillende responscijfers voor staan. Naast de inspanningen van de onderzoeker om de responscijfers te verbeteren, is ook het onderwerp van belang. Surveys waarin het onderwerp medisch georiënteerd is, hebben onder ouderen vaak hogere responscijfers dan onder andere groepen. In dit kader zou experimenteel opgezet onderzoek, waarin deelname aan verschillende dataverzamelingsmethoden met een vast onderwerp over diverse leeftijdsgroepen wordt onderzocht, meer gefundeerde conclusies opleveren. Met betrekking tot de datakwaliteit zijn de bevindingen verontrustend. Zelfs bij twee relatief makkelijk af te nemen instrumenten en zeer goed getrainde interviewers is de kwaliteit van de verzamelde gegevens niet optimaal. De resultaten duiden erop dat de schaalbaarheid van instrumenten bij oudere respondenten afneemt. De hoge interviewervariantie bij beide groepen kan duiden op problemen bij het afnemen van vragenlijsten bij oudere respondenten in het algemeen, terwijl de relatief forse toename van de interviewervariantie bij de oudste respondenten erop kan duiden dat deze problematiek toeneemt naarmate respondenten ouder zijn. Vermoeidheid van de respondenten kan een mogelijke oorzaak zijn. De kwaliteit van gegevens, verzameld met complexere vragen dan hier zijn gebruikt en instrumenten waarbij geen visuele hulpmiddelen worden gebruikt (zoals bij telefonische interviews) zullen waarschijnlijk nog grotere verschillen laten zien tussen jonge en oudere respondenten. Om precies na te gaan wat de redenen zijn voor deze verschillen in datakwaliteit zijn meer gedetailleerde studies nodig. In deze studies kunnen dan vragen worden beantwoord als: “Verloopt het afnemen van een standaardvragenlijst bij oudere respondenten anders dan bij jongere?”, “Wijken bij oudere respondenten de interviewers vaker af van het vaste stramien, in die zin dat er bijvoorbeeld meer uitleg gegeven wordt, er meer wordt geïnterpreteerd?”, etc. Nauwkeurig nagaan wat er gebeurt tijdens de afname van de vragenlijsten kan meer inzicht verschaffen in de processen die tot verschillen leiden in datakwaliteit.
Bij het opzetten van een survey-onderzoek onder ouderen dient men zich goed te realiseren dat ouderen niet zonder meer met de doorsnee methoden en technieken kunnen worden benaderd. Er bestaan grote verschillen tussen ouderen onderling waarmee rekening moet worden gehouden. Een instrument ontwikkeld voor een gezonde 60-jarige is niet op dezelfde manier inzetbaar bij een 80-jarige met visus- en gehoorproblemen. Indien men aan dit soort problematiek niet voldoende aandacht besteedt bij instrumentontwikkeling en interviewtrainingen zullen de tekortkomingen zich vertalen in gegevens van mindere kwaliteit met alle gevolgen van dien voor het doen van algemene uitspraken. Immers het ontwikkelen van overheidsbeleid gebaseerd op dit soort gegevens wordt dan uiteraard ook dubieus.

LITERATUUR

Andrews, F.M. en S.B. Whithey, 1976, *Social indicators of well-being: Americans perception of life quality,* New York: Plenum.
Beekman, A.T.F., J. van Limbeek, D.J.H. Deeg, L. Wouters en W. van Tilburg, 1994, Screeningsinstrumenten voor depressie bij ouderen in de algemene bevolking: De bruikbaarheid van de Center for Epidemiologic Studies Depression Scale (CES-D). *Tijdschrift voor Gerontologie en Geriatrie,* 25 (3), pp 95-104.

Brandsma, M., 1998, Stimuleringsprogramma 'Succesvol ouder worden' (SOW), in: *Ouder Worden,* idnr 32.124. Utrecht: Nederlands Instituut voor Gerontologie.
Bosscher, R.B., J.H. Smit en R. Kempen, 1997, Algemene competentie verwachtingen bij ouderen: een onderzoek naar de psychometrische kenmerken van de Algemen Competentieschaal (ALCOS), in: *Nederlands tijdschrift voor de psychologie,* 52, pp 239-248.
Carp, F.M. en A. Carp, 1984, A complementary/congruence model of well-being or mental health for the community elderly, in: *Elderly people and the environment: Advances in theory and research.* vol. 7, I.
Corso, J.F., 1977, Auditory perception and communication, in: *Handbook of the psychology of aging,* J.E. Birren en K.W. Schaie (eds), New York: Van Nostrand Reinhold.
Deeg, D.J.H., C.P.M. Knipscheer en W. van Tilburg, 1993, Autonomie and Well-being in the Aging Population: Concepts and design of the Longitudinal Aging Study Amsterdam. *NIG-overzichtstudie 7.* Bunnik: Nederlands Instituut voor Gerontologie.
Dillman, D.A., 1978, Mail and Telephone Surveys: the total design method. New York: John Wiley.
Groves, R.M. en R.L. Kahn, 1979, *Surveys by telephone: A national comparison with personal interviews.* New York: Academic Press.
Herzog, A.R. en R.A. Kulka, 1989, Telephone and mail surveys with older populations: a methodological overview, in: M.P. Lawton en A.R. Herzog (eds). *Special research methods for gerontology,* pp 63-89. New York: Baywood Publishing Company.
Hoinville, G., 1983, Carrying out surveys among the elderly: Some problems of sampling and interviewing, in: *Journal of Market Research Society,* 25, pp 223-237.
Kaldenberg, D.O., Koenig, H.F. en B.W. Becker, 1994, Mail Survey Response Rate Patterns in a Population of the Elderly: Does Response Detoriate with Age? *Public Opinion Quarterly,* 58, pp 68-76.
Knipscheer, C.P.M., J. de Jong Gierveld, T.G. van Tilburg en P.A. Dijkstra (1995) Living Arrangements and Social Networks of Older Adults. Amsterdam: VU- Uitgeverij.
Rogers, T.F., 1976, Interviews by telephone and in person: Quality of responses and field performance, in: *Public Opinion Quarterly,* 40, pp 51-65.
Santvoort, M.M. van, 1998, NESTOR: resultaten en evaluatie na zeven jaar, in: *Ouder Worden,* idnr 32.121. Utrecht: Nederlands Instituut voor Gerontologie.
Smit, J.H., 1995. Suggestieve vragen in survey-interviews. Vóórkomen, oorzaken en gevolgen. Academisch proefschrift. Amsterdam: Vrije Universiteit. *(Suggestive questioning in Survey-Interviews, occurences, causes and consequences* (in Dutch, summary in English)).
Smit, J.H. en M.Z. de Vries, 1994. Procedures and Results of the Fieldword, in: D.J.H. Deeg en M. Westendorp – de Serière (eds). *Autonomy and well-being in the aging population,* pp 7-14. Amsterdam: VU- Uitgeverij.
Smit, J.H., P. van den Eeden, D.J.H. Deeg en A.T.F. Beekman, 1995, De effecten van interviewer- en respondentenkenmerken op antwoordgedrag in survey-onderzoek: een multilevel-benadering, in: *Sociologische Gids,* 4-5, pp 285-297.
Zouwen, J. van en W. Dijkstra, W., 1996. *Trivial and Non-Trivial Question-Answer Sequences; Types, Determinants and Effects on* DATA *Quality.* Proceedings of the International Conference On Survey Measurement and Process Quality, pp 81-86. Alexandria: American Statistical Association.

5. Uiteindelijk gaat het om de smaak

P.H. PUNTER

SAMENVATTING

Stijgende concurrentie, het wegvallen van handelsbarrières en een groeiende wereldmarkt noodzaakt de voedingsmiddelenindustrie tot steeds snellere vernieuwing, innovatie en kwaliteitsverbetering. Het slaagpercentage van innovaties en vernieuwingen is echter erg laag. Dit wordt voor een deel veroorzaakt door het feit dat veel nieuwe producten niet voldoen aan de verwachting van de consument. De mate van succes wordt met name bepaald door het inzicht van de industrie in de verwachtingen en belevingen van consumenten ten aanzien van voedingsmiddelen en door het vermogen om deze kennis te implementeren. Hoewel het traditionele marktonderzoek hierin deels voorziet blijkt dit een onvoldoende garantie voor succes. Sensorisch onderzoek is van origine een discipline binnen productontwikkeling en kwaliteitscontrole. Met behulp van experts of getrainde beoordelaars worden de sensorische karakteristieken van producten gemeten en kunnen de effecten van bijvoorbeeld veranderingen in grondstof of processing worden vastgesteld. Binnen het sensorisch onderzoek zijn een groot aantal methodieken ontwikkeld om de zintuiglijke reacties van mensen op producten te meten welke in de afgelopen 50 jaar hun nut hebben bewezen. Door de associatie van sensorisch onderzoek met experts of getrainde beoordelaars is haar mogelijke bijdrage aan het marktonderzoek (met name daar waar het concept- of producttests betreft) nooit echt duidelijk geworden. In de afgelopen 10 jaar hebben sensorici zich echter steeds meer gericht op onderzoek met consumenten en pogen zij een brug te slaan tussen marketing en R&D. In dit artikel wordt uitgelegd wat sensorisch onderzoek is, welke methodieken men gebruikt en hoe deze kunnen worden toegepast. Daarnaast wordt een voorbeeld gegeven van 'consumer guided product development'.

1. INLEIDING

Stijgende concurrentie, het wegvallen van handelsbarrières en een groeiende wereldmarkt noodzaakt de voedingsmiddelenindustrie tot steeds snellere vernieuwing, innovatie en kwaliteitsverbetering. Het slaagpercentage van innovaties in de voedings- en genotsmid-delenindustrie is echter bedroevend laag. Wereldwijd haalt slechts 25% van de nieuwe voedingsmiddelen het derde jaar, in Nederland is dit zelfs minder dan 20% (Eilander, 1997; Smit, 1998). Daarnaast blijkt de omzetbijdrage van deze innovaties erg klein.

Hoewel er een groot aantal redenen zijn voor deze slechte performance, is een van de belangrijke oorzaken het feit dat het fysieke product de gedane belofte niet waarmaakt: de smaak voldoet niet aan de verwachting van de consument.

De vraag of een product lekker smaakt lijkt zeer eenvoudig: smaak is uiteindelijk iets waar we allemaal verstand van hebben, 'the proof is in the eating'. Voor een deel ligt deze schijnbare eenvoud ten grondslag aan het hoge faalpercentage. Veel product-

ontwikkelaars en marketeers gaan, waar het hun eigen producten betreft, teveel op hun eigen ervaring af en maken een verkeerde inschatting van de behoeften en de wensen van de consument ten aanzien van de smaak (de 'gutfeeling-benadering'). Daarnaast blijkt regelmatig dat er een grote kloof is tussen de briefing van marketing aan product-ontwikkeling en de 'maakbaarheid' van het product op basis van deze briefing.
Hoewel het smaakaspect in het traditionele marktonderzoek meestal wordt meegenomen geeft dit in de praktijk onvoldoende aanknopingspunten voor de oplossing van dit probleem. Sensorisch onderzoek is vanuit haar achtergrond veel beter in staat om een bijdrage leveren aan de vertaling van de wensen en behoeften van de consument naar concrete product-specificaties (Sidel and Stone, 1993).

1.1. Wat is sensorisch onderzoek?

Sensorisch onderzoek is gedefinieerd als de wetenschappelijk methodiek die zich bezighoudt met het opwekken, meten, analyseren en interpreteren van de zintuiglijke reacties op producten zoals deze worden waargenomen door de tastzin, de reukzin, de smaak, gezichtsvermogen en gehoor (Amerine et. al., 1965; Stone and Sidel, 1985; O' Mahony, 1986; Meilgaard et. al., 1990; Lawless and Heymann, 1998). In dit artikel beperken wij ons tot voedingsmiddelen, er zijn echter tal van andere toepassingen van sensorisch onderzoek mogelijk.
Sensorisch onderzoek heeft betrekking op de vier activiteiten die in de definitie zijn genoemd.
1. Het opwekken van de zintuiglijke reacties.
Het bereiden en aanbieden van de te beoordelen producten gebeurt volgens nauwkeurig omschreven regels waardoor mogelijk storende invloeden (zoals temperatuur, uiterlijk of hoeveelheid) zoveel mogelijk worden geëlimineerd. Producten worden meestal blind aangeboden met random codes, de beoordelaars (of dit nu consumenten zijn of experts) worden in afgescheiden proefcabines geplaatst om te verhinderen dat ze elkaars reacties kunnen zien en elkaars oordeel beïnvloeden, de aanbiedingsvolgorde van de producten is gebalanceerd en de tijd tussen aanbiedingen is constant.
2. Het meten van de respons.
De waargenomen sterkte (of lekkerte) wordt uitgedrukt in een getal. In het sensorisch onderzoek wordt tegenwoordig meestal gebruik gemaakt van lijnschalen (een lijn met een lengte van 100 mm waarop de respondent door middel van een kruisje de waargenomen sterkte aangeeft). Deze lijnschaal heeft in de meeste gevallen ankers op 10% en 90% van de schaal (bijvoorbeeld zwak en sterk). In tegenstelling tot het marktonderzoek, waar meestal slechts één of twee producten worden beoordeeld, gaat het in sensorisch onderzoek over veel grotere aantallen (afhankelijk van de vraagstelling kunnen 30 of meer producten worden beoordeeld). In veel gevallen gaat de dataverzameling via het beeldscherm.
3. Analyse van de data
De data die wordt verkregen van menselijke beoordelaars is niet erg nauwkeurig. Er is een groot aantal bronnen van variantie die niet volledig kunnen worden gecontroleerd (de stemming van de respondent, zijn of haar zintuiglijke gevoeligheid, ervaringen uit het verleden, hun ervaringen met de te beoordelen producten etc.). Om vast te kunnen stellen of een gevonden verschil het gevolg is van verschillen tussen de producten of van random fluctuaties moet gebruik worden gemaakt van statistische methoden en technieken (Dijksterhuis, 1997). Daarnaast is het van essentieel belang dat de data

wordt verzameld met behulp van een goede proefopzet (experimenteel design), om effecten van volgorde en 'carry-over' te voorkomen wordt tegenwoordig bijna altijd gebruik gemaakt van gebalanceerde designs (ieder product komt even vaak op iedere positie en even vaak na ieder ander product: MacFie et. al., 1989 en Wakeling en MacFie, 1995).
4. Interpreteren
Data en statistische informatie hebben alleen waarde wanneer ze goed worden geïnterpreteerd. Wat was de context waarin het onderzoek is opgezet?, wat was de onderzoekshypothese?, welke additionele informatie is voorhanden?, welke generalisaties kunnen worden gemaakt op basis van de steekproef.

1.2. Betrouwbaarheid, accuraatheid, gevoeligheid

Sensorisch onderzoek is een exacte wetenschap, met als belangrijkste steekwoorden betrouwbaarheid, accuraatheid, gevoeligheid (Muñoz et. al., 1992; Meiselman, 1993). Betrouwbaarheid betekent dat we hetzelfde resultaat krijgen als de test wordt herhaald op een ander moment met andere mensen. De belangrijkste voorwaarde om hieraan te voldoen is strikte controle van de testcondities en een voldoende grote steekproef om het effect van willekeurige fluctuaties te minimaliseren. Wanneer wordt gewerkt met getrainde beoordelaars is de minimale steekproefgrootte 10 tot 15 personen, wanneer gewerkt wordt met ongetrainde consumenten is een steekproefgrootte van minimaal 40 tot 50 personen vereist (Moskowitz, 1997).
Accuraatheid verwijst naar de eigenschap van een instrument om de 'werkelijke waarde' zo goed mogelijk te benaderen. In de gedragswetenschappen wordt vaak gesproken van de validiteit van de meting, meet de test wat hij ook werkelijk moet meten. Bij sensorisch onderzoek is een belangrijke vraag of de resultaten, verkregen met een groep respondenten, ook representatief zijn voor andere groepen respondenten of de populatie.
De generaliseerbaarheid is in sterke mate afhankelijk van het type onderzoek, wanneer de waargenomen intensiteit wordt bepaald van producteigenschappen (de zoetheid, zoutheid, bitterheid etc.) is het minder relevant welke steekproef is getrokken dan wanneer de waardering of lekkerte wordt bepaald. Mensen verschillen aanzienlijk minder qua sensorische capaciteiten dan qua voorkeur of preferentie.
Tenslotte moet de testopzet voldoende gevoelig zijn om verschillen te detecteren wanneer deze in werkelijkheid bestaan. Bij de toetsing van hypothesen wordt voornamelijk gekeken naar de kans dat het gevonden resultaat op toeval berust. Wanneer deze kans kleiner is dan 5% (1 op de 20 gevallen) nemen we aan dat er sprake is van een verschil. Aangezien de meeste beslissingen over het al dan niet bestaan van een verschil meestal op meerdere gronden worden genomen, speelt dit risico niet zo'n grote rol. Een veel groter risico is het type 2 fout, ten onrechte concluderen dat er geen verschil is terwijl dit er in werkelijkheid wel is. Als wijziging van de samenstelling van een product een negatief effect heeft op de waardering of beoordeling en dit niet wordt geconstateerd kan dat zeer vervelende consequenties hebben. De vraag of de test voldoende gevoelig was om een bestaand verschil te detecteren is veel moeilijker te beantwoorden dan de kans dat het resultaat op grond van toeval is ontstaan. Belangrijke factoren bij het voorkomen van dit type fout is de nauwkeurigheid van de meting en de steekproefgrootte. Aangezien dit belangrijke kostenbepalende factoren zijn zien we dat hieraan in de praktijk vaak concessies worden gedaan.

2. TESTMETHODEN IN SENSORISCH ONDERZOEK

Binnen het sensorisch onderzoek worden drie typen van methoden onderscheiden.

1. Verschiltests (zijn twee productvarianten op een of ander aspect afwijkend van elkaar).
2. Beschrijvende tests (in welke mate zijn specifieke sensorische eigenschappen in een product aanwezig).
3. Hedonische tests (hoe lekker wordt een product gevonden of welk product heeft de voorkeur).

Dit onderscheid betreft zowel de testprocedure als het type respondenten.

2.1. Verschiltests

Bij de verschiltests gaat het om de vraag of twee varianten (sensorisch) van elkaar verschillen (O' Mahony, 1990; Punter en Kolb, 1994). Hierbij kan het gaan om de verlaging van het percentage vet in een product, vervanging van een grondstof of veranderingen in het productieproces. De verschiltest is de meest populaire en meest gebruikte testvorm in het sensorisch onderzoek. Deze populariteit komt met name door de (schijnbare) eenvoud van de methodiek (O' Mahony, 1995) en de eenvoud van dataanalyse (op basis van het totale aantal oordelen en het aantal correcte oordelen kan men het significantie niveau in een tabel aflezen). De bekendste verschiltest is de driehoekstest, hierbij krijgt de beoordelaar drie producten aangeboden, twee identieke en één afwijkende. De taak is relatief simpel: welke van de drie is de afwijkende (men is gedwongen een keuze te maken). Op basis van het proportie goede antwoorden en na correctie voor gokken kan op basis van de kansberekening worden bepaald of het gevonden verschil op toeval berust of echt bestaat. Bij de verschiltest kunnen twee vraagstellingen worden onderscheiden: 'is er verschil waarneembaar' en 'hoeveel mensen nemen verschil waar'. Voor de eerste vraag is het voldoende om aan te tonen dat er één persoon is die consequent de afwijkende detecteert. Hierbij speelt de gevoeligheid van de beoordelaar een cruciale rol en gaat het om de prestaties van individuen. Voor de tweede vraag is individuele gevoeligheid minder relevant maar gaat het om een voldoende grote en representatieve steekproef uit de relevante populatie. In de praktijk wordt dit onderscheid vaak vergeten hetgeen tot verwarrende en moeilijk te interpreteren resultaten aanleiding geeft.
In de afgelopen 20 jaar is veel onderzoek verricht naar de psychologische aspecten van het keuzegedrag en zijn verschillende modellen ontwikkeld om het keuzegedrag van de respondent te verklaren (Frijters, 1980, 1984; Ennis, 1990, 1991). Hoewel de verschiltests tot de best onderbouwde technieken in het sensorisch onderzoek behoren is de praktische toepassing ervan vrij beperkt. Wanneer is vastgesteld dat er een verschil waarneembaar is volgen er bijna automatisch twee andere vragen: op welke aspecten verschillen de producten en welke wordt geprefereerd. Deze vragen zijn echter niet te beantwoorden binnen de context van de verschiltest, maar vereisen een andere benadering en andere steekproeven.

2.2. Beschrijvende tests

Bij beschrijvende of descriptieve tests wordt de waargenomen sterkte van de sensorische karakteristieken gemeten. Hiervoor wordt gebruik gemaakt van een lijst met pro-

duct-relevante aspecten (bijvoorbeeld kleursterkte, geursterkte, zoete geur, zure geur, smaaksterkte, bittere smaak, enzovoorts) waarop de verschillende producten worden beoordeeld. Hoewel het beschrijvend onderzoek een relatief jonge methodiek is hebben er in de afgelopen 25 jaar een groot aantal veranderingen plaatsgevonden.
De belangrijkste veranderingen hebben betrekking op de manier waarop de relevante kenmerken worden gegenereerd, op de mate van training en op het type respondenten.

2.2.1. Het generen van begrippen

In de zeventiger jaren werd algemeen geaccepteerd dat voor een goed beschrijvend panel langdurig getrainde proevers nodig waren. Met name het ontwikkelen van een gemeenschappelijke begrippenset was een tijdrovende bezigheid. De gangbare procedure was om de panelleden zelf, door middel van groepsdiscussies, een beschrijvende taal te laten ontwikkelen. Alle individuele gewaarwordingen moeten worden genoteerd en bediscussieerd. Dit kan leiden tot zeer veel begrippen (300 tot 400 zijn geen uitzondering). Nadat de begrippen zijn gegenereerd volgt een fase van reductie hetgeen resulteert in een beperkte set begrippen (meestal in de grootte-orde van 60-120). Met behulp van deze begrippenset kan het panel vervolgens gaan oordelen (o.a. Köster, 1990; Van Gemert en Wolters, 1990). De uiteindelijke begrippenset is echter niet altijd interpreteerbaar voor buitenstaanders. Naast inhoudelijke argumenten ten aanzien van de zinvolheid kwamen er steeds meer praktische bezwaren tegen deze techniek: het kostte erg veel tijd en geld.
Een alternatief is om de panelleden aan de hand van vooraf gedefinieerde modelstoffen een set van begrippen aan te leren. Hierbij wordt gebruik gemaakt van bestaande kennis over de producten. Er zijn verschillende standaardlijsten ontwikkeld voor specifieke producten. Voorbeelden zijn bier (Clapperton, 1979; Meilgaard et al., 1979; Meilgaard et al., 1982), thee (Werkhoven, 1973), wijn (Noble, 1978); appels (Williams en Carter, 1977). De laatste 10 jaar is men in de USA bezig met het opzetten van een begrippendatabase (Meilgaard, persoonlijke mededeling).
Dankzij de toepassing van een multivariate techniek, de Gegeneraliseerde Procrustes analyse of GPA, is het ook mogelijk om iedere beoordelaar met zijn of haar eigen begrippenset te laten oordelen (Gower, 1975; Williams and Langron, 1984). Bij deze techniek is de enige eis dat men dezelfde producten heeft beoordeeld. De oordelen van individuele proefpersonen worden samengevoegd tot een groepsresultaat zonder dat er over attributen moet worden gemiddeld (Dijksterhuis en Gower, 1991; Dijksterhuis en Punter, 1991; Dijksterhuis, 1997). Dit heeft geleid tot de ontwikkeling van een nieuwe vorm van beschrijvend onderzoek: Free Choice Profiling of FCP (Cramwinckel en Van Mazijk, 1990; Dijksterhuis en Punter, 1991). Het is van belang om onderscheid te maken tussen de statistische techniek (GPA) en de sensorische methode (FCP). De GPA-techniek kan eveneens zeer goed worden gebruikt voor de analyse van klassieke beschrijvende data. Omdat de techniek niet middelt over individuele beoordelaars kunnen de prestaties van de panelleden op snelle en efficiënte wijze worden gevisualiseerd.
De FCP methode is een vorm van beschrijvend onderzoek waarin de panelleden niet gebruik hoeven te maken van dezelfde begrippenset (alleen de te beoordelen producten zijn gelijk voor iedereen). Er kunnen dus beoordelaars uit verschillende landen worden vergeleken of er kan gebruik worden gemaakt van gegevens van consumenten (zie bijvoorbeeld Steenkamp en Van Trijp, 1988; Steenkamp et al., 1994).

2.2.2. De mate van training

De hoeveelheid training is in sterke mate afhankelijk van het doel van de meting. Naarmate de te detecteren verschillen kleiner worden zal meer tijd moeten worden besteed aan de training. Het aantonen van verschil tussen twee batches pilsner bier vereist meer vaardigheid dan het aantonen van verschil tussen vruchtenyoghurt van het eigen merk en dat van drie concurrenten.

De tijd die wordt besteed aan het trainen van het panel is steeds korter geworden. Dit heeft met name te maken met de kosten, daarnaast speelt de snelheid van productontwikkeling ook een rol. Wanneer producten om de zes maanden moeten worden vernieuwd heeft het weinig zin om voor iedere serie een panel te trainen. Vaak is ook niet duidelijk wat onder training wordt verstaan, dit kan een oefenperiode zijn van meerdere maanden of één groepssessie waarin de attributen worden uitgelegd.

Bij beschrijvend onderzoek met ongetrainde consumenten is per definitie geen sprake van training, in dit geval gaat het er primair om dat de gehanteerde begrippenset begrijpelijk is voor de panelleden.

2.2.3. Type respondenten

In de vijftiger en zestiger jaren werd alleen gebruik gemaakt van zeer ervaren en getrainde experts binnen het bedrijf. De introductie van de QDA techniek maakte het mogelijk om ook leken op te leiden, dit waren echter meestal medewerkers van het bedrijf. Hiervoor zijn twee redenen aanwijsbaar: in de eerste plaats is de training tijdrovend en kostbaar en investeert men liever in eigen personeel dan in consumenten, in de tweede plaats gaat het vaak om vertrouwelijke productinformatie en wil men deze binnen het bedrijf houden. Soms is het echter om praktische redenen niet mogelijk een panel te rekruteren uit het personeelsbestand en moet gewerkt worden met buitenstaanders. In dat geval wordt een groep consumenten geselecteerd die goed kunnen ruiken en proeven en die goed in staat zijn om hun gewaarwordingen te verbaliseren (en die, last but not least, gedurende een voldoende lange periode beschikbaar zijn voor deelname). Deze consumenten worden vervolgens opgeleid tot ervaren proevers.

Dit type panels dient primair als hulpmiddel voor productontwikkeling en kwaliteitscontrole, centraal staat het vermogen om nauwkeurig, herhaalbaar en betrouwbaar te kunnen oordelen. De meeste grote bedrijven in de voedingsindustrie beschikken over dergelijke panels.

Het fundamentele probleem is het vertalen van consumenten-wensen naar concrete, manipuleerbare productspecificaties. Dit kan via multivariate technieken (MacFie and Thomson, 1984; Martens and Naes, 1989) waarbij de consumenten-voorkeur wordt geprojecteerd in de productruimte gebaseerd op de beoordeling door het getrainde of expertpanel. Op deze wijze kunnen de relevante eigenschappen van het geprefereerde product worden bepaald. Dit is echter alleen mogelijk waneer het bedrijf de beschikking heeft over een getraind panel. Vaak is dat echter niet het geval en moet deze informatie op andere wijze worden verkregen.

Het gebruik van ongetrainde consumenten voor beschrijvend onderzoek is lange tijd taboe geweest binnen het sensorisch onderzoek. Nog steeds is een veelgehoorde opvatting dat de consument niet in staat is om een valide, kwantitatief oordeel te geven. Zo stelt Lawless (1998) dat reacties van consumenten op een product vaak gebaseerd zijn op een geïntegreerd patroon van sensorische stimulatie en zich alleen kunnen manifesteren als waardering (lekker of niet lekker). Wanneer consumenten wordt gevraagd om specifieke informatie te geven over producteigenschappen ontstaan er problemen. Ze

redeneren vanuit een niet-analytisch kader en hebben vaak vage concepten over de betekenis van sensorische eigenschappen (ze verwarren bitter met zuur en kunnen grote verschillen vertonen in de interpretatie van een aspect).
Tegen deze stelling zijn een aantal argumenten aan te voeren: in de eerste plaats verschillen consumenten niet van getrainde beoordelaars met betrekking tot de sensorische capaciteiten, in de tweede plaats hebben consumenten jarenlange ervaring op het gebied van ruiken en proeven (dagelijks maken ze een keuze uit het aanbod in voedingsmiddelen etc). De stelling dat ze minder scherp omlijnde definities hebben is correct, het is echter onjuist om te stellen dat ze niet beschikken over een analytisch kader om producten te beoordelen.
Uit jarenlange ervaring blijkt dat ongetrainde consumenten nauwkeurig en reproduceerbaar kunnen oordelen, mits er aan een aantal voorwaarden zijn voldaan. In de eerste plaats moet gebruik worden gemaakt van begrippen (attributen) die ze begrijpen. Dit kan onder meer worden bereikt door vooraf een aantal kwalitatieve sessies uit te voeren waarin consumenten zelf de attributen genereren waarop de producten later moeten worden beoordeeld. Ten tweede zijn er meer consumenten nodig voor een valide resultaat dan getrainde panelleden (getrainde of expert panels bestaan meestal uit 10-15 personen, ongetrainde consumentenpanels uit 80 tot 120 personen). Hoewel consumenten meer variatie in hun oordeel vertonen blijkt deze variatie voornamelijk willekeurig te zijn. Dat betekent dat de gemiddelde waarde nog steeds de beste schatter van de werkelijke waarde is, de grootte van de meetfout (mits deze willekeurig is) doet daar geen afbreuk aan. Vergelijkingen tussen de resultaten van getrainde panels en ongetrainde consumenten laat zien dat er voornamelijk een verschil in spreiding is, bij consumenten is dit circa 50% van het gemiddelde en bij getrainde panelleden circa 35-40%. Hierbij moet wel worden opgemerkt dat de consumenten enige ervaring moeten krijgen in het oordelen, met name het gebruik van de schalen vereist enige oefening.
Een groot voordeel van beschrijvende consumentenpanels is dat er direct een verband kan worden gelegd tussen de waargenomen sterkte van de verschillende aspecten en de waardering. Daarnaast kan de consument tevens worden gevraagd om een oordeel te geven over de gewenste of ideale sterkte van ieder aspect. Het succes van deze benadering staat of valt echter met een strikte controle van de meetomstandigheden.

2.3. Preferentie en voorkeur

Voorkeur en acceptatie verwijzen naar een subjectief oordeel, wat vindt de beoordelaar van het product, is het vies of lekker, acceptabel of onacceptabel. Dit is een subjectieve of hedonische vraag. In dit geval gaat het niet om de zintuiglijke registratie maar om de subjectieve evaluatie ervan. Bij analytische vragen is er sprake van juiste of onjuiste antwoorden (iemand die een zoutoplossing van 10 gram zouter vindt dan een oplossing van 20 gram geeft een fout oordeel). Bij hedonische vragen is nooit sprake van goed of fout. De beoordelaar kan het product op het ene moment acceptabel vinden en op het andere moment niet. Personen kunnen zowel onderling verschillen qua voorkeur als in zichzelf. Bij analytische oordelen staat de kwaliteit en nauwkeurigheid van het meetinstrument centraal, bij hedonische oordelen staat de representativiteit centraal.

In de jaren voor de tweede wereldoorlog werd relatief weinig aandacht besteed aan de voorkeur of acceptatie. Peryam (1990) geeft een aantal voorbeelden van eenvoudige acceptatie of voorkeurstests. Bij Seagram (rond 1940) werden voorkeurstests uitge-

voerd voor het evalueren van ingrediënt-veranderingen en product ontwikkeling. Men maakte gebruik van zowel in- als externe panels. De methode was een simpele paarsgewijze vergelijking (welke prefereert u). De representativiteit was meestal ondergeschikt bij deze onderzoeken, men trachtte met behulp van werknemers uitspraken te krijgen over de acceptatie en voorkeur. Deze werknemerspanels werden als een miniatuur consumentenpanel beschouwd.

Aan het einde van de tweede wereldoorlog en de jaren daarna heeft het voorkeursonderzoek een belangrijke ontwikkeling doorgemaakt. Met name het Amerikaanse leger heeft een belangrijke rol gespeeld in de ontwikkeling van het voorkeursonderzoek. In en na de tweede wereldoorlog begon men te beseffen dat een verantwoorde samenstelling van het voedselpakket voor de militairen op zichzelf geen garantie gaf dat men ook at. Aangezien een goede fysieke conditie essentieel is voor de troepen werd de mate van acceptatie van de voeding een cruciale variabele. Naast een correcte samenstelling moesten de mensen het ook lekker genoeg vinden om te nuttigen.

In principe kunnen twee taken worden onderscheiden bij het bepalen van de voorkeur: keuze (welke van de twee of meerdere varianten prefereert u) en schaling (hoe lekker vindt u dit product).
Voor de tweede wereldoorlog werd voornamelijk gebruik gemaakt van paarsgewijze vergelijking voor het bepalen van de voorkeur. Na de tweede wereldoorlog hebben Peryam en Pilgrim (1957) de 9-punts hedonische referentieschaal geïntroduceerd. Dit is een categorieschaal met negen ankers: like extremely, like very much, like moderately, like slightly, neither like nor dislike, dislike slightly, dislike moderately, dislike very much, dislike extremely. Deze schaal is sinds de introductie overal ter wereld in een zeer groot aantal onderzoeken gebruikt.

Voorkeursgegevens spelen in verschillende stadia van de productie een rol: bij productverbetering of -optimalisatie, bij productontwikkeling en bij het vaststellen van de potentiële vraag in de markt. Daarnaast moeten voorkeursgegevens uiteindelijk worden vertaald naar producteigenschappen: welke producteigenschappen bepalen de voorkeur. Het rendement van dergelijke vertalingen wordt in belangrijke mate bepaald door de nauwkeurigheid waarmee de voorkeur is bepaald. Zoals al eerder is vermeld gaat het hierbij met name om de representativiteit van de steekproef: is de steekproef een correcte afspiegeling van de uiteindelijke doelgroep. Daarnaast speelt de wijze waarop de voorkeur is bepaald eveneens een belangrijke rol, in het algemeen geldt dat gedragsmaten een betere voorspelling geven dan antwoorden op vragenlijsten. Ondanks deze kennis moet worden geconstateerd dat acceptatie en voorkeursmetingen regelmatig (met name in de eerste selectiefase) worden uitgevoerd met analytische panels of medewerkers van het bedrijf. Het komt zelfs nog voor dat de kinderen van de directeur beslissen of een product lekker is of niet.

3. RELATIES MET MARKTONDERZOEK

Binnen het marktonderzoek worden een aantal technieken gehanteerd die een relatie hebben met sensorisch onderzoek, i.e. de concepttesten en producttesten. Hierbij ligt de nadruk echter meestal op de gebruikssituatie, het product wordt in de thuissituatie beoordeeld of consumenten worden uitgenodigd om even te komen proeven in een caravan of supermarkt. Het gaat bij dit type onderzoek primair om de waardering, hoewel daarnaast wordt getracht vast te stellen waarop deze waardering is gebaseerd. Een impliciete aanname bij dit type onderzoek is de gedachte dat de consument slechts in staat is om aan te geven of hij iets wel of niet lekker vindt. Vanuit de optiek van de sensoricus kleven er een aantal bezwaren aan deze procedures. In de eerste plaats is er onvoldoende controle op mogelijke storende variabelen. Daarnaast is het niet mogelijk om meerdere varianten te vergelijken.
In onderstaande tabel zijn de belangrijkste verschillen samengevat (Lawless, 1998).

Sensorisch onderzoek met consumenten	Marktonderzoek (concept- en producttest)
selectie deelnemers uit doelgroep	selectie deelnemers op basis positieve reactie op concept
blinde aanbieding, random codes	aanbieding as marketed, vaak met concept
groot aantal varianten (herhaalde metingen)	slechts enkele varianten, monadisch
nadruk op beleving sensorische aspecten	nadruk op acceptatie en koopintentie
terugkoppeling naar R&D en marketing	terugkoppeling naar marketing
bepaling sensorische aantrekkelijkheid en waargenomen en ideale sterkte op aspecten	bepaling aantrekkelijkheid op basis ‘totale’ product (zowel sensorisch als concept)

Ten aanzien van het laatste punt wordt er de laatste jaren steeds meer onderzoek gedaan naar het effect van merk, verpakking, verwachting, concept en andere niet-sensorische informatie op de beleving van de sensorische aspecten (Cardello,1992; 1994; Deliza and MacFie, 1996).

In feite kan worden geconstateerd dat er een redelijke overlap is tussen (productgericht) marktonderzoek en (consumentgeoriënteerd) sensorisch onderzoek. De sterke kant van marktonderzoek is met name de generaliseerbaarheid van de resultaten naar uiteindelijk koopgedrag (ten koste van sensorische nauwkeurigheid), de sterkte kant van sensorisch onderzoek is de vertaling van de wensen van de consument naar concrete productspecificaties (ten koste van generaliseerbaarheid naar de markt). Vanuit deze optiek zijn beide invalshoeken complementair en aanvullend.

4. EEN VOORBEELD: CONSUMER GUIDED PRODUCT DEVELOPMENT

De probleemstelling luidde als volgt: wat is het effect van verschillende startercultuur (A en B) en toegevoegde cultuur (A+ en B+) op de beoordeling en waardering van kaas.
De eerste stap is selectie van een groep van circa 14 consumenten (frequente gebruikers van kaas) waarmee de attributenlijst wordt gegenereerd door middel van groepsdiscussies.

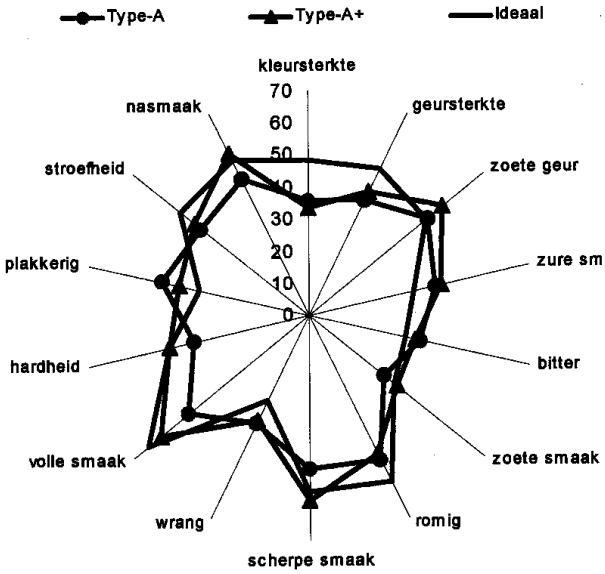
Tabel 1. Voorbeeld van de gebruikte attributen (niet volledig).

attribuut	**linker anker**	**rechter anker**
aantrekkelijke kleur	erg onaantrekkelijk	erg aantrekkelijk
lekkere smaak	niet lekker	erg lekker
lekkere nasmaak	niet lekker	erg lekker
aangenaam mondgevoel	erg onaangenaam	erg aangenaam
totaaloordeel	rapportcijfer 1-10	
kleursterkte	licht	donker
gewenste of ideale kleursterkte	licht	donker
geursterkte	zwak	sterk
gewenste of ideale kleursterkte	zwak	sterk
smaaksterkte	zwak	sterk
gewenste of ideale smaaksterkte	zwak	sterk
zoete smaak	zwak	sterk
gewenste of ideale zoete smaak	zwak	sterk
hardheid	zacht	hard
gewenste of ideale hardheid	zacht	hard
stroefheid in mond	stroef	glad
gewenste of ideale stroefheid	stroef	glad

Met behulp van deze attributen worden vervolgens de vier kazen beoordeeld door een groep van 126 consumenten (regelmatige gebruikers van kaas). In tabel 1 is een aantal attributen vermeld waarop de kazen zijn beoordeeld.
Testafname vindt plaats per computer, de respondenten beantwoorden de vragen op lijnschalen (120 mm) met ankers op 10% en 90% van de schaal.

De respondenten beoordelen de varianten semi-monadisch (één voor één), de attributen worden gevraagd in volgorde van waarneming (eerst uiterlijk en geur, daarna smaak en mondgevoel en tenslotte nasmaak).
De resultaten van de intensiteitbeoordelingen zijn weergegeven in de vorm van spiderplots. In figuur 1 zijn de profielen weergegeven type A en A+ met de ideale kaas.

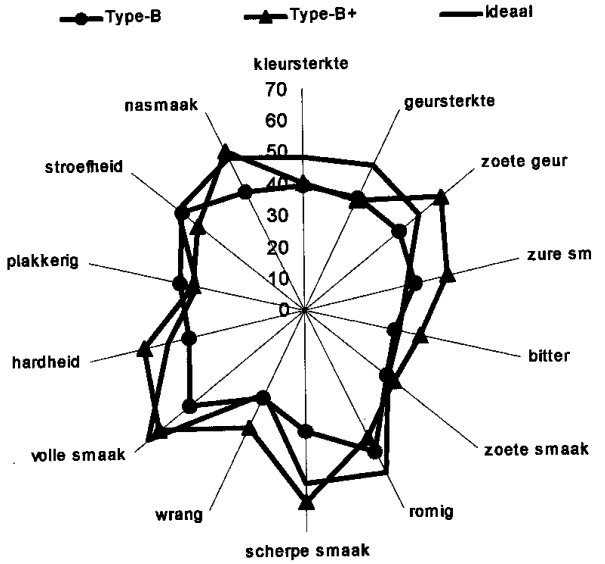
Uit figuur 1 blijkt dat toevoeging van extra cultuur met name effect heeft op de scherpe en volle smaak, op de nasmaak en op de textuur (hardheid en plakkerigheid). Type A+ voldoet beter aan het ideaal dan type A. Dit komt ook tot uitdrukking in het totaaloordeel voor beide varianten: type A scoort 6.2 en type A+ scoort 7.0.



Figuur 1. Profiel van type A en A+ met de ideale kaas.

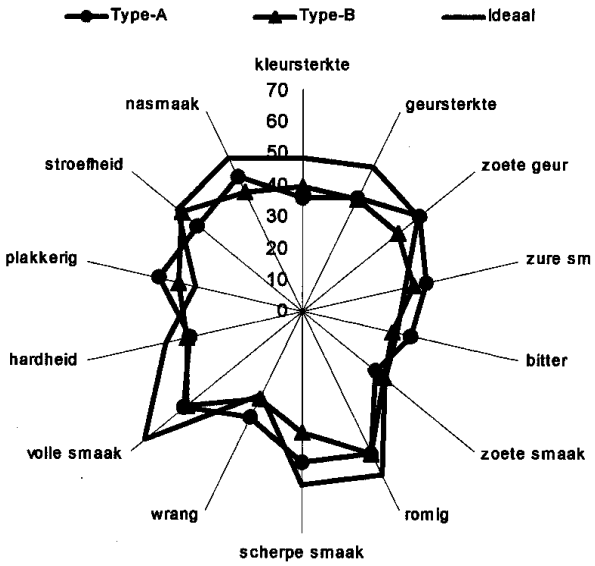
In figuur 2 zijn de resultaten weergegeven voor type B en type B+.

Uit figuur 2 blijkt dat toevoeging van extra cultuur bij type B met name effect heeft op de zoete geur, de zure, bittere smaak, wrangheid, volle smaak, nasmaak en hardheid. Qua scherpe smaak, volle smaak en nasmaak voldoet type B+ beter aan het ideaal, op de meeste andere aspecten voldoet type B beter. Qua totaaloordeel scoort type B 6.5 en type B+ 6.7 (het verschil is niet significant).



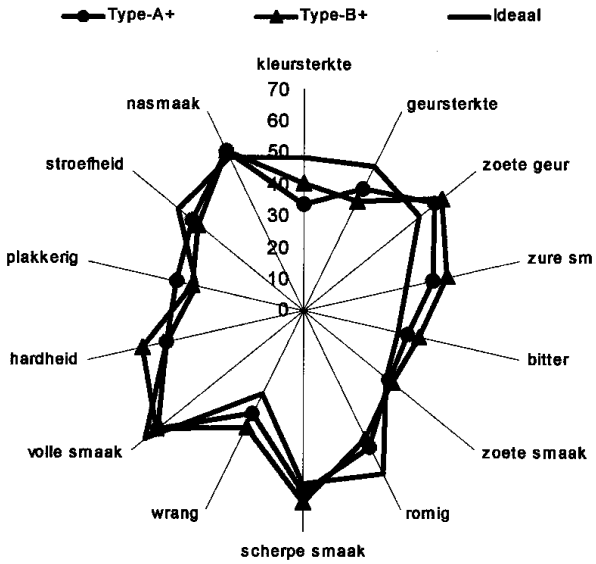
Figuur 2. Profiel van type B en B+ met de ideale kaas.

In figuur 3 zijn de profielen voor type A en type B weergegeven.



Figuur 3. Profiel van type A en B met de ideale kaas.

Uit figuur 3 blijkt dat de twee typen startercultuur met name verschillen qua zoete geur, scherpe smaak en textuur. Type B krijgt een iets hogere waardering (6.5 versus 6.2).

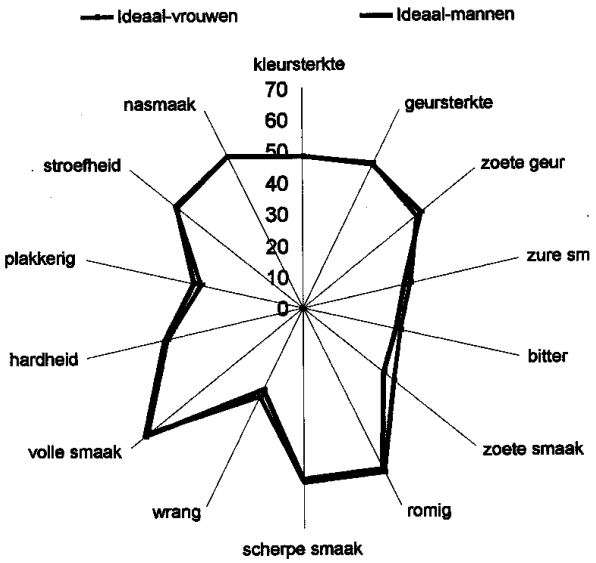


Figuur 4. Profiel van type A+ en B+ met de ideale kaas.

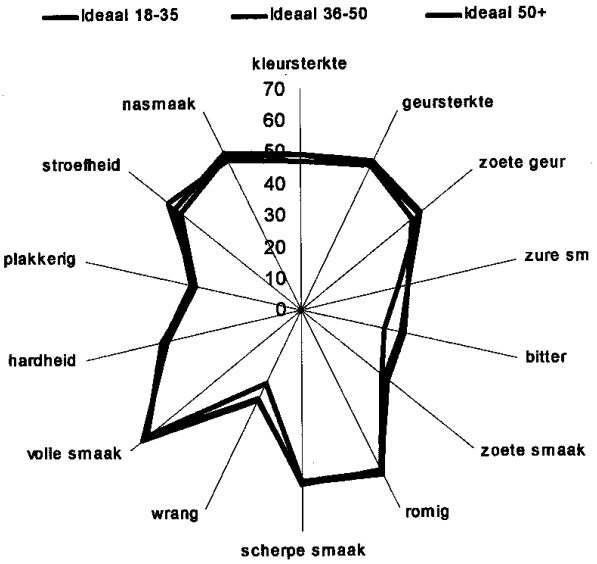
In figuur 4 zijn de profielen voor type A+ en type B+ weergegeven.

Uit figuur 4 blijkt dat het belangrijkste verschil tussen de kazen met toegevoegde cultuur betrekking heeft op textuur (hardheid en plakkerigheid). Type A+ wordt iets meer gewaardeerd dan type B+ (7.0 versus 6.7).

Uit vergelijking van de profielen van de ideale kaas voor mannen en vrouwen of voor verschillende leeftijdsgroepen kan worden nagegaan in hoeverre er verschillen zijn in de ideale kaas. In figuur 5 en 6 zijn de idealen voor mannen en vrouwen weergegeven en voor de drie leeftijdsgroepen.



Figuur 5. Profiel van de ideale kaas voor vrouwen en mannen.



Figuur 6. Profiel van de ideale kaas voor de drie leeftijdsgroepen.

De ideale kaas is identiek voor mannen en vrouwen en voor de drie leeftijdsklassen.

Om na te gaan welke aspecten het meest belangrijk zijn voor de waardering is een PCA uitgevoerd over de sterkte attributen. Er worden vier principale componenten gevonden (percentage verklaarde variantie 63%):

PC1 Scherpe en volle smaak, nasmaak en geursterkte;
PC2 Wrang, bitter, zuur en plakkerig;
PC3 Romigheid, stroefheid (niet stroef), hardheid (niet hard);
PC4 Kleur- en geursterkte, zoete smaak.

Door middel van regressieanalyse van het totaaloordeel is nagegaan in welke mate dit wordt bepaald door de vier factoren. Uit de resultaten van de regressieanalyse blijkt dat PC1 en PC3 een positieve bijdrage aan het totaaloordeel leveren, PC2 draagt negatief bij. De meest gewaardeerde kaas heeft een scherpe en volle smaak, een sterke nasmaak, een sterke geur en is romig, glad en zacht qua textuur. Hij mag niet wrang, bitter, zuur of plakkerig zijn. Kaas type A+ voldoet hieraan het beste.

LITERATUUR

Amerine, M.A., Pangborn, R.M. and Roessler, E.B. (1965). Principles of Sensory Evaluation of Food. London: Academic Press.

Cardello, A.V. and Sawyer, F.M. (1992). Effects of disconfirmed consumer expectations on food acceptability. Journal of Sensory Studies, 7, 255-277.

Cardello, A.V. (1994). Consumer expectations and their role in food acceptance. In: Measurement of
Food Preferences (H.J.H. MacFie and D.H.M. Thomson, eds.), 253-297, Blackie Academic & Professional, London.

Clapperton, J.F. (1979). Sensory characterization of the flavour of beer. In: D.G. Land and H.E. Nursten (eds), Progress in flavour research, London: Elsevier, 1-14.

Cramwinckel, A.B. en van Mazijk-Blokslag, D.M. (1990). De plaats van Free Choice Profiling in het sensorisch onderzoek. VMT, 10, 22-24.

Deliza, R. and MacFie, H.J.H. (1996). The generation of sensory expectation by external cues and its effect on sensory perception and hedonic ratings: a review. Journal of Sensory Studies, 11, 103-128.

Dijksterhuis, G.B. and Gower, J.C. (1991/92). The interpretation of Generalized Procrustes Analysis and allied methods. Food Quality and Preference, 3, 67-87.

Dijksterhuis, G.B. en Punter, P.H. (1990). Interpreting generalized procrustes analysiss of variance' tabels. Food Quality and Preference, 2, 263-270.

Dijksterhuis, G.B. (1997). Multivariate data analysis in sensory and consumer science. Food and Nutrition Press, Trumbull, Ct.

Eilander, G. (1997). Nieuwe producten falen vooral vanwege niet waarmaken belofte. VMT, 23, 35-37.

Ennis, D.M. (1990). Relative power of difference testing methods in sensory evaluation. Food Technol., april, 114-117.

Ennis, D.M. (1991). Probabilistic models of perception. Proefschrift, Landbouwuniversiteit Wageningen

Frijters, J.E.R. (1980). Psychophysical and psychometrical models for the triangular method. Proefschrift, RijksUniversiteit Utrecht.

Frijters, J.E.R. (1984). Sensory difference testing and the measurement of sensory discriminability. In: Piggott, J.R. (ed), Sensory Analysis of Foods. London: Elsevier.

Gemert, L.J. van en Wolters, C.J. (1990). Sensorisch onderzoek met beschrijvende panels. VMT, 26, 9-14.

Gower, J.C. (1975). Generalized Procrustes Analysis. Psychometrica, 40, 33-51.
Köster, E.P. (1990). De dynamiek van de sensorische eigenschappen van levensmiddelen. In: van Dokkum W. en van der Heij D.G. (red). Voedsel in beweging. Wageningen: Pudoc.
Lawless, H.T. and Heymann, H. (1998). Sensory Evaluation of Food-Principles and Practices. Intern. Thomson Publish. Serv., Andover.
MacFie, H.J.H. and Thomson, D.M. (1984). Multidimensional Scaling Methods. In: Piggott, J.R. (ed). Sensory Analysis of Foods. England: Elsevier.
MacFie, H.J.H., Bratchell, N., Greenhoff, K. and Vallis, L.V. (1989). Designs to balance the effect of order of presentation and first order carry-over effects in hall tests. Journal of Sensory Studies, 4, 129-148.
Martens, H and Naes, T. (1989). Multivariate Calibration. England: John Wiley and Sons.
Meilgaard, M.C., Dalgliesh, C.E. and Clapperton, J.F. (1979). Beer flavour terminology. Jn. Am. Soc. Brew. Chem., 37, 47-52.
Meilgaard, M.C., Reid, D.S. and Wyborski, K.A. (1982). Reference standards for beer flavor terminology system. Am. Soc. Brew. Chem., 40, 119-128.
Meilgaard, M.C., Civille, G.V. and Carr, B.T. (1990). Sensory Evaluation Techniques (4th printing). Florida: CRC Press, p 126-140.
Meiselman H.L. (1993). Critical evaluation of sensory techniques. Food Quality and Preference, 4, 33-41.
Moskowitz, H.R. (1997). Base size in product testing: a psychophysical viewpoint and analysis. Food Quality and Preference, 4, 247-257
Muñoz, A.M., Civille C.G. and Carr, B.T. (1992). Sensory Evaluation in Quality Control. Van Nostrand Reinhold, New York.
Noble, A.C. (1978). Sensory and instrumental evaluation of wine aroma. In: Charalambous (ed), Analysis of Food and Beverage. London: Academic Press, p 203-228.
O 'Mahony, M. (1986). Sensory evaluation of Food. New York: Marcel Dekker.
O 'Mahony, M. (1990). Cognitive aspects of difference testing and descriptive analysis: criterion variation and concept variation. In: McBride, R.L. and MacFie, H.J.H. (eds), Psychological basis of sensory evaluation. London: Elsevier, p 117-139.
O' Mahony (1995). Who told you the triangle test was simple?. Food Quality and Preference, 6, 227-238.
Peryam, D.R. (1990). Sensory evaluation-the early days. Food Technol., 44, 86-91.
Peryam, D.R. and Pilgrim, F.J. (1957). Hedonic scale method of measuring food preferences. Food Technology, 11, supp 9-14.
Punter, P.H. en Kolb, C.M. (1994). Verschiltests in sensorisch en consumentenonderzoek. VMT, 13, 29-33
Smit, K. (1998). Trends als kompas. Tijdschrift voor Marketing, 4, 25-28.
Steenkamp, J-B.E.M. and van Trijp, H.C.M. (1988). Free choice profiling in cognitive food acceptance research. In: Thomson, D.M.H. (ed): Food Acceptability. London: Elsevier, p 363-377.
Steenkamp, J-B.E.M., van Trijp, H.C.M. and Ten Berge J.M.F. (1994). Perceptual Mapping Based on Idiosyncratic Sets of Attributes. Journal of Marketing Research, 31, 15-27.
Stone, H. and Sidel, J.L. (1985). Sensory Evaluation Practices. New York: Academic Press.
Sidel, J.L. and Stone, H. (1993). The role of sensory evaluation in the food industry. Food Quality and Preference, 4, 65-75
Wakeling, I.N. and MacFie, H.J.H. (1995). Designing consumer trials balanced for first and higher orders of carry-over effect when only a subset of k samples from t may be tested. Food Quality and Preference, 6, 299-308.
Werkhoven, J. (1974). Tea processing. FAO Agric. Ser. Bull., No 26, FAO, Rome, p 195.
Williams, A.A. and Carter, C.S. (1977). A language and procedure for the sensory assessment of Cox's orange pippin apples. Jn. Sci. Food Agric., 28, 1090-1104.
Williams, A.A. and Langron, S.P. (1984). The use of Free Choice Profiling for the evaluation of commercial ports. J. Sci. Food Agric., 35, 558-568.

6. *Het modelleren van de effecten van assortimentssamenstelling op consumenten-keuzegedrag: Een voorbeeld met betrekking tot bloemenwinkels*

K. KOELEMEIJER en H. OPPEWAL[1]

SAMENVATTING

Dit artikel presenteert een methode om de relaties vast te stellen die bestaan tussen items binnen en tussen produktcategorieën of assortimenten. We modelleren het nut van een artikel in een hypothetisch assortiment als functie van de verkrijgbaarheid en eigenschappen van andere artikelen in dat assortiment met behulp van het extended logit model. Op deze wijze kunnen we substitutie, complementariteit en asymmetrische dominantie-effecten modelleren en schatten als kruiseffecten. De gegevens voor ons model worden verkregen uit keuze-experimenten waarin zowel de samenstelling van assortimenten als de beslissingsomgeving, waaronder winkelsfeer, worden gevarieerd. We illustreren onze benadering met een toepassing in de snijbloemendetailhandel. We vinden dat aankoopbeslissingen door consumenten in een winkel worden beïnvloed door omvang en samenstelling van het assortiment in die winkel en eveneens afhankelijk zijn van de aanwezigheid van een concurrerende winkel. De sfeer in de winkel en de eigenschappen van een eventuele concurrerende winkel, afstand, sfeer en keuzemogelijkheden, blijken echter weinig invloed te hebben op de beslissing of en waar bloemen worden gekocht. Substitutie vindt plaats tussen snijbloemen van hetzelfde soort, maar niet in gelijke mate voor alle soorten.

1. INLEIDING

De samenstelling van assortimenten vormt een complexe taak voor managers in de detailhandel. Gegeven de beperkte schapruimte en de constante stroom nieuwe en vernieuwde produkten die op de markt verschijnt, moeten zij regelmatig beslissen over toevoeging van nieuwe artikelen aan het assortiment en verwijdering van artikelen uit het huidige assortiment. Beslissingen over assortimentssamenstelling zijn ingewikkeld omdat het toevoegen aan of verwijderen van artikelen uit een assortiment gevolgen kan hebben voor de verkopen van andere artikelen uit het assortiment en ook de winkelkeuze kunnen beïnvloeden. Naast kennis van substitutie- en complementariteitsrelaties tus-

[1] De auteurs bedanken het Produktschap Tuinbouw voor hun financiele bijdrage aan het onderzoek en Bloemenveiling Holland voor hulp bij de uitvoering van het onderzoek. Tijdens een deel van de werkzaamheden aan dit onderzoek was de eerste auteur werkzaam aan de vakgroep Marktkunde en Marktonderzoek van de Landbouwuniversiteit Wageningen.

sen de artikelen in een assortiment is daarom een consistente positionering op alle elementen van de detailhandelsmix, zoals winkelsfeer, kwaliteits- en prijsimago, en locatie, van doorslaggevend belang voor het succes van een detailhandelsonderneming.
Ondanks het grote aantal publicaties dat is verschenen over winkelgedrag en keuzeprocessen betreffen slechts enkele publicaties hierbinnen de samenstelling van het assortiment en de effecten daarvan op consumentenkeuze. Het is een bekend gegeven (Oppewal, Louviere, en Timmermans 1997) dat consumenten meestal de voorkeur geven aan grotere assortimenten boven kleine. Er is echter nog weinig onderzoek beschikbaar waarin wordt beschreven welke kenmerken van een assortiment van invloed zijn op de aantrekkelijkheid daarvan. De aanwezigheid van bepaalde artikelen in een assortiment kan tot gevolg hebben dat het assortiment als geheel aantrekkelijker wordt terwijl van het artikel zelf niet eens veel wordt verkocht. Voorbeelden zijn de kledingzaak die een exclusief merk voert naast de meer gangbare collectie en de verkrijgbaarheid van bijzondere delicatessen in een supermarkt. In het algemeen neemt de voorkeur van consumenten voor een assortiment toe naarmate zij de individuele artikelen in het assortiment meer waarderen en het assortiment groter is. Ook werd vastgesteld dat een nieuw toe te voegen artikel de waardering voor het assortiment meer doet toenemen naarmate dat artikel minder lijkt op de artikelen in het bestaande assortiment (Kahn en Lehmann 1991).
Daarentegen richt vrijwel de gehele literatuur over keuzeprocessen en keuzemodellen zich op de vraag hoe consumenten kiezen uit assortimenten. In dit verband is met name onderzoek relevant naar de effecten van aanwezige artikelen in een assortiment op de rationaliteit van de keuze (bijv. Simonson en Tversky 1992). Dergelijk onderzoek stelde vast dat wanneer artikelen worden toegevoegd aan of verwijderd uit een assortiment, of als de kenmerken van artikelen worden veranderd de keuzekansen van andere artikelen kunnen veranderen op een manier die niet in overeenstemming is met bestaande preferentiestructuren. Asymmetrische dominantie (Huber, Payne, en Puto 1982) is een voorbeeld van een dergelijk context effect, dat inhoudt dat uitbreiding van het assortiment met een artikel de kans op aankoop van tenminste één ander artikel vergroot. Hoewel de meeste keuzemodellen vooronderstellen dat de keuzekans voor een artikel juist afneemt wanneer het assortiment wordt vergroot werd dit effect voor een variëteit aan consumentengedrag vastgesteld.
De vraag blijft of het toevoegen van meer substituten alleen tot meer kannibalisatie en daardoor afnemende opbrengsten binnen de categorie leidt, of dat de categorie en misschien het assortiment als geheel juist aantrekkelijker worden gevonden waardoor de verkopen als geheel toenemen.
Dit artikel laat zien hoe keuze-experimenten en kruiseffecten modellen kunnen worden gebruikt om de effecten van competitie tussen en binnen assortimenten te onderzoeken en te voorspellen. Keuze-experimenten stellen de onderzoeker (en dus ook de manager) in staat om een aantal aspecten te beheersen of manipuleren die in ander onderzoek vaak niet aan de orde komen. Hierdoor zijn de geschatte parameters in keuze-experimenten goed interpreteerbaar en efficiënt. Daarnaast kunnen kruiseffecten modellen worden gebruikt om de hierboven beschreven asymmetrische dominantie-effecten te signaleren. Keuze-experimenten winnen in het marktonderzoek aan populariteit als methode voor het meten van consumentenvoorkeuren en voorspellen van consumentenkeuze. We illustreren het gebruik van keuze-experimenten en kruiseffecten modellen met een onderzoek naar het effect van assortimentssamenstelling en winkelsfeer op aankoopbeslissingen door consumenten in de winkel met betrekking tot snijbloemen.

2. KEUZE-EXPERIMENTEN EN EXTENDED LOGIT MODELLEN

Keuze-experimenten verlopen meestal als volgt. Een deelnemer wordt geconfronteerd met een aantal verzamelingen, keuzesets, of assortimenten van hypothetische alternatieven (bijv. produkten, vervoermiddelen, of woningen) en kiest uit elke keuzeset het alternatief dat zijn of haar voorkeur geniet. De alternatieven worden opgebouwd gedacht uit attributen, zoals merknaam, kleur, grootte, of prijs. Met behulp van een experimenteel ontwerp worden de niveaus van de verschillende attributen systematisch gevarieerd en worden keuzesets geformuleerd waarin een deel van de zo ontstane alternatieven in wisselende samenstelling aanwezig is (bijv. Oppewal en Timmermans, 1992).

De meeste toepassingen van keuze-experimenten gaan uit van het Multinomiale Logit (MNL) model. Dit model veronderstelt dat consumenten in hun keuzegedrag naar nutsmaximalisatie streven. Het model veronderstelt tevens dat het nut van keuze-alternatieven door allerlei toevalsfactoren wordt beinvloed en daardoor fluctueert volgens een bepaalde kansverdeling. Een belangrijke implicatie van het MNL model is dat het nut van een alternatief alleen afhankelijk is van de eigenschappen van dat alternatief en niet van de eigenschappen van de andere alternatieven in een keuzeset: de *Independence of Irrelevant Alternatives* (IIA) vooronderstelling. Het is duidelijk dat deze eigenschap niet zal gelden met betrekking tot de keuze uit assortimenten. Consumenten hebben immers de neiging complementaire artikelen uit een assortiment te kopen en een keuze te maken door de artikelen die elkaars substituut zijn te vergelijken.

Het 'extended logit' model, ook wel universeel of moeder logit model genoemd, biedt een oplossing voor de IIA beperking (McFadden et al. 1977; Krishnamurthi et al. 1995). Het nut van een alternatief is in dit model niet alleen afhankelijk van de eigen kenmerken maar ook van de aanwezigheid van andere alternatieven in de keuzeset en de kenmerken van deze alternatieven. De termen in de nutsvergelijking die deze effecten weergeven worden kruiseffecten genoemd (bijv. Krishnamurthi et al. 1995). Kruiseffecten geven het effect weer van de verkrijgbaarheid van een alternatief op de keuzekans van een ander alternatief, zgn. verkrijgbaarheid (of availability) kruiseffecten, of het effect van eigenschappen van een alternatief op de keuzekans van een ander alternatief, zgn. attribuut kruiseffecten (Anderson en Wiley 1992; Lazari en Anderson 1994; Louviere en Woodworth 1983; Oppewal en Timmermans 1991). Beide typen kruiseffecten zijn doorgaans negatief vanwege substitutie tussen alternatieven. Kruiseffecten geven aan in hoeverre keuze-alternatieven substituten of complementen zijn. Stel, we willen weten hoe de marktaandelen zullen veranderen na toetreding van merk C in een markt waar tot nu toe slechts twee merken, zeg A en B, beschikbaar waren. Als we in het model geen signifikant kruiseffect van C op A hebben gevonden en een negatief kruiseffect van C op B, dan betekent dit dat de toetreding van C leidt tot een vermindering van het nut van merk B, zodanig dat B relatief meer marktaandeel verliest dan A. We interpreteren dit als dat B en C in sterkere mate elkaars substituten zijn dan A en C. Omgekeerd, als het kruiseffect positief is leidt dit tot een verhoging van het nut van B en dus een minder sterke daling, of zelfs stijging, van het marktaandeel van B. Een positief kruiseffect van B op C duidt dus op complementariteit tussen B en C. Kruiseffecten zijn als het ware correcties op fouten die zouden worden gemaakt bij gebruikmaking van het gewone MNL model.

Een van de belangrijkste problemen waar assortimentsmanagers zich voor zien gesteld is het vaststellen welke artikelen goede complementen van elkaar zijn en welke artike-

len met elkaar concurreren. Kruiseffecten modellen en keuze-experimenten ondersteunen beslissingen omtrent de samenstelling en omvang van het assortiment. In het vervolg van dit artikel gebruiken we kruiseffecten om substitutie-effecten in en tussen snijbloemenassortimenten te kwantificeren.

3. SUBSTITUTIE-EFFECTEN IN SNIJBLOEMENASSORTIMENTEN

3.1. Een toepassing: de snijbloemendetailhandel

De snijbloemendetailhandel is bij uitstek geschikt voor onderzoek naar de effecten van assortimentssamenstelling en bijbehorende dienstverlening door de detailhandel. Snijbloemen zijn een fast-moving, heterogeen, bederfelijk, merkloos, en tot in de winkel vrijwel onverpakt produkt waarvan de verkrijgbaarheid varieert met het seizoen en de aanvoer door gespecialiseerde telers, waarvan de prijzen dagelijks variëren en waarvoor vrijwel uitsluitend collectief wordt geadverteerd. Daarnaast bestaat het afzetkanaal van snijbloemen voornamelijk uit gespecialiseerde distributeurs en is derhalve in beperkte mate afhankelijk van andere goederen. Vanwege de heterogeniteit met betrekking tot produkt, prijs, en het aangeboden assortiment in de werkelijke snijbloemenmarkt, alsmede het onafhankelijke distributiekanaal, kan een experimentele taak worden ontwikkeld en aan consumenten voorgelegd die zeer representatief is voor de keuze in de werkelijke markt.

3.2. De experimentele assortimenten en keuze-scenario's

Voor het experiment werden vijftien verschillende, frequent verkochte, bossen snijbloemen gebruikt. De aan- of afwezigheid van elk van deze bossen werd systematisch gevarieerd, resulterend in 64 assortimenten bestaande uit 5 tot 12 bossen snijbloemen van verschillende kleur en/of soort.[2] Elk aanwezig alternatief heeft als prijs de werkelijke marktprijs plus of minus f. 1,50. Tevens werd de aan- of afwezigheid van een concurrerende winkel gemanipuleerd. Bij 48 van de 64 assortimenten was een optie om een concurrerende winkel te bezoeken opgenomen, die zich op 1, 5, of 10 minuten reistijd bevond. Het hier beschreven "verkrijgbaarheid x prijs x reistijd" ontwerp was genest onder een 2^3 volledig factorieel ontwerp dat de volgende factoren systematisch varieerde tussen deelnemers aan het experiment: (1) sfeer in de huidige winkel (veel of weinig, aangegeven door middel van foto's van de binnen- en buitenkant van de winkel), (2) sfeer in de concurrerende winkel (veel of weinig, volgens beschrijving van de winkel in de vragenlijst), en (3) keuzemogelijkheden in de concurrerende winkel (groot of klein, volgens beschrijving van de winkel in de vragenlijst).
De deelnemers aan ons experiment waren 323 frequente kopers van snijbloemen die

[2] De selectie van de juiste experimentele ontwerpen voor keuze-experimenten vereist gespecialiseerde kennis en vaardigheden, zeker in geval men kruiseffecten wil schatten. Er bestaan nauwelijks algemene procedures of richtlijnen om optimale ontwerpen voor keuze-experimenten te genereren. Slechts voor specifieke typen keuzeproblemen zijn efficiënte ontwerpen geformuleerd (zie A.G. Lazari en D.A. Anderson, 1994, "Designs of discrete choice set experiments for estimating both attribute and availability cross effects," *Journal of Marketing Research*, 31, 375-383). De experimentele ontwerpen die worden gegenereerd met behulp van recente commerciële software zijn vaak niet het meest efficiënt en in het algemeen niet geschikt voor het schatten van kruiseffecten.

deel uitmaakten van een groot Nederlands consumentenpanel. Zij ontvingen een schriftelijke vragenlijst met foto's van het interieur van één winkel en foto's van acht verschillende keuzesets (assortimenten) bestaande uit verschillende bossen snijbloemen. Alle snijbloemen op de foto's waren even vers en gefotografeerd in gestandaardiseerd veilingfust, vanuit dezelfde hoek van waaruit consumenten in winkels snijbloemen zien staan. De deelnemers beantwoordden vragen over vorige snijbloemenaankopen, hun indrukken van de beschreven hypothetische winkel, en hun keuzen uit elk van de acht keuzesets.

4. RESULTATEN

4.1. Effecten van kleur, variëteit en prijs

We schatten eerst het basisnut van elk van de 15 snijbloemenvarianten, en de effecten van het prijsniveau en de concurrerende winkel op de kans dat een aankoop plaatsvindt in de winkel (zie tabel 1). Omdat we een volledig gebalanceerd experimenteel ontwerp gebruikten zijn de geschatte parameters voor de 15 snijbloemen een weergave van de voorkeuren (het 'basisnut') voor elke kleur-soort combinatie. We zien bijvoorbeeld dat witte fresia (nr. 16) het meest geprefereerd wordt en witte iris (nr. 11) het minst. Er kon ook een basisnut van de concurrerende winkel worden geschat; het alternatief 'geen aankoop' heeft als basisnut de waarde nul. Het negatieve basisnut van de concurrerende winkel geeft aan dat de respondenten er vaker voor kozen geen aankoop te doen dan dat ze naar de andere winkel gingen, echter het verschil is niet significant. Ook de reistijd van de huidige winkel naar de andere winkel blijkt niet significant van invloed op de kans dat men deze winkel kiest.

Tabel 1. Geschatte parameters voor het basismodel.

Variabele	Coëfficiënt	p-waarde	
PRIJS	-0.274	0.000	
Basisnutten van de verschillende snijbloemen:			
V2	0.707	0.006	Grootbloemige Rozen, rood
V3	0.523	0.050	Grootbloemige Rozen, wit
V4	0.982	0.000	Kleinbloemige Rozen, rood
V5	0.780	0.001	Kleinbloemige Rozen, wit
V6	1.137	0.000	Tulpen,rood
V7	0.931	0.000	Tulpen, geel
V8	0.399	0.150	Chrysanten, roze
V9	0.337	0.252	Chrysanten, geel
V10	0.072	0.810	Irissen, blauw
V11	-0.163	0.631	Irissen, wit
V12	1.088	0.000	Anjers, rood
V13	0.508	0.073	Anjers, wit
V14	0.465	0.082	Lelies, rood
V15	0.200	0.521	Lelies, geel
V16	1.193	0.000	Fresia's, geel

Tabel 1. (Vervolg)

Effect (VA_B) van aanwezigheid van bloem A op nut van bloem B:		
V3_2	-0.136	0.446
V2_3	-0.112	0.561
V5_4	-0.052	0.739
V4_5	-0.218	0.171
V7_6	-0.183	0.279
V6_7	-0.327	0.069
V9_8	-0.367	0.032
V8_9	-0.183	0.305
V11_10	0.007	0.973
V10_11	-0.455	0.111
V13_12	-0.176	0.241
V12_13	-0.588	0.008
V15_14	-0.062	0.752
V14_15	-0.536	0.039
Effecten van de concurrerende winkel:		
V1 (basisnut)	-0.879	0.304
Reistijd	-0.132	0.488

Tabel 1 laat verder zien dat het prijsniveau een signifikant effect heeft op de aankoopkans; hogere prijzen verlagen de kans op aankoop. De prijsgevoeligheid blijkt niet te verschillen tussen de verschillende bloemtypen; het opnemen van een afzonderlijke prijsparameter voor elk bloemtype veroorzaakt geen signifikante toename van de verklarende kracht van het model (Likelihood Ratio Chi^2=15.6, d.f.=14, n.s.). De voorkeuren voor de snijbloemenvarianten zijn wel specifiek voor kleur-soort combinaties. Het vervangen van de dummies voor de 15 varianten door een kleiner aantal dummies voor kleuren en soorten verslechterde de fit van het model (L.R. Chi^2=-36.8, d.f.=5, $p<.001$).

4.2. Winkelsfeer en concurrerende winkel

We veronderstelden dat de sfeer in de hypothetische winkel de aankoopkans positief zou beïnvloeden. Daarnaast verwachtten we een effect van de winkelsfeer op de prijsgevoeligheid. Om dit te testen namen we in de nutsfuncties voor elke snijbloemenvariant een parameter op voor de winkelsfeer en een parameter voor de interactie tussen sfeer en prijs. Dit type interactie bleek slechts signifikant (en negatief) voor rode tulpen. De fit van het totale model verbeterde niet door opname van genoemde effecten. We concluderen dus dat winkelsfeer geen invloed uitoefent op zowel de kans op aankoop in de winkel als de keuzekansen van de individuele snijbloemenvarianten. We vinden eveneens geen signifikant effect van de kenmerken van de concurrerende winkel op de aankoopkans.

4.3. Substitutie-effecten

Om meer inzicht te krijgen in substitutie-effecten voegen we termen aan de individuele nutsfuncties toe die substitutie-effecten tussen paren van alternatieven weergeven. We veronderstellen dat wanneer voor een bepaalde bloemsoort twee verschillende kleuren verkrijgbaar zijn deze sterker met elkaar concurreren dan met andere kleur-soort com-

binaties. In de nutsfunctie van elk alternatief nemen we een daarom een kruiseffect op voor de verkrijgbaarheid van het concurrerende alternatief. Dit is slechts mogelijk voor 14 van de 15 alternatieven aangezien fresia's slechts in één kleur in het experiment aanwezig zijn. Daarnaast formuleren we een kruiseffect voor de aanwezigheid van de concurrerende winkel in alle 15 nutsfuncties. Tenslotte schatten we het effect van de omvang van het assortiment voor de kans dat men naar de concurrerende winkel gaat door middel van een kruiseffect in de nutsfunctie van de concurrerende winkel.
De uitbreiding van het basismodel met kruiseffecten veroorzaakt een toename van de verklarende kracht van het model (L.R. $Chi^2 = 74.62$, 34 d.f., $p<.001$). Uit tabel 2 blijkt dat de onderzochte kleurvariëteiten het meest als substituut dienen voor de volgende snijbloemen (in volgorde van afnemend effect): "rode anjer (12) op witte anjer (13)", "rode lelie (14) op gele lelie (15)", "gele chrysanten (9) op roze chrysanten (8)", en "rode tulpen (6) op gele tulpen (7)". De substitutie-effecten zijn niet symmetrisch, we zien de hierboven besproken *asymmetrische dominantie effecten* optreden (Huber et al. 1982). In deze toepassing betekent dit dat de aanwezigheid van rode tulpen (6) de keuzekans voor gele tulpen (7) meer dan proportioneel verlaagt, gele chrysant (9) domineert roze chrysant (8), rode anjer overheerst witte anjer (13) en rode lelie (14) overheerst gele lelie (15). Rood neigt ertoe geel te overheersen, maar niet voor alle soorten bloemen.
Wanneer de concurrerende winkel aanwezig is worden witte rozen (5), roze chrysanten (8) en gele chrysanten (9) vaker gekocht dan bij afwezigheid van een concurrerende winkel. Dit betekent dat consumenten voor de andere alternatieven meer bereid zijn naar de concurrerende winkel te gaan. Chrysanten ondervinden beide typen kruiseffecten, hetgeen impliceert dat consumenten een weinig uitgesproken voorkeur voor deze bloemsoort hebben.
Tenslotte vinden we dat vergroting van het assortiment leidt tot een evenredige afname van de kans dat men naar de concurrerende winkel gaat; het bijbehorende kruiseffect is niet signifikant.

Tabel 2. Geschatte parameters voor basismodel uitgebreid met kruiseffecten.

Variabele	Coëfficiënt	p-waarde
PRIJS	-0.274	0.000
Basisnutten van de verschillende snijbloemen:		
V2	0.707	0.006
V3	0.523	0.050
V4	0.982	0.000
V5	0.800	0.001
V6	1.137	0.000
V7	0.931	0.000
V8	0.399	0.150
V9	0.337	0.252
V10	0.072	0.810
V11	-0.163	0.631
V12	1.088	0.000
V13	0.508	0.073
V14	0.465	0.082
V15	0.200	0.521
V16	1.193	0.000

Tabel 2. Vervolg

Effect (VA_B) van aanwezigheid van bloem A op nut van bloem B:		
V3_2	-0.136	0.446
V2_3	-0.112	0.561
V5_4	-0.052	0.739
V4_5	-0.218	0.171
V7_6	-0.183	0.279
V6_7	-0.327	0.069
V9_8	-0.367	0.032
V8_9	-0.183	0.305
V11_10	0.007	0.973
V10_11	-0.455	0.111
V13_12	-0.176	0.241
V12_13	-0.588	0.008
V15_14	-0.062	0.752
V14_15	-0.536	0.039
Effect (V1_B) van aanwezigheid concurrerende winkel op bloem B:		
V1_2	0.323	0.256
V1_3	0.268	0.362
V1_4	0.393	0.148
V1_5	0.695	0.011
V1_6	-0.266	0.922
V1_7	0.153	0.582
V1_8	1.014	0.001
V1_9	0.840	0.008
V1_10	0.447	0.168
V1_11	0.277	0.462
V1_12	0.435	0.103
V1_13	0.288	0.356
V1_14	0.222	0.458
V1_15	0.092	0.792
V1_16	0.295	0.260
Effecten van eigenschappen concurrerende winkel:		
V1 (basisnut)	-0.879	0.305
Afstand	-0.132	0.488
Keus	-0.108	0.465
Sfeer	0.217	0.121
Interacties van sfeer in huidige winkel (Sfh) met eigenschappen van concurrerende winkel:		
Sfeer:		
Sfh x V1 (basisnut)	0.170	0.311
Sfh x Afstand	0.612	0.001
Sfh x Keus	0.277	0.061
Sfh x Sfeer	0.040	0.774
Kruiseffect van aantal artikelen in assortiment huidige winkel op concurrerende winkel:		
	0.028	0.802

5. CONCLUSIE

In dit artikel lieten we zien hoe keuze-experimenten en kruiseffecten-modellen kunnen worden gebruikt voor het analyseren en voorspellen van substitutie-effecten tussen en binnen assortimenten. Keuze-experimenten zijn bij uitstek geschikt voor het modelleren en voorspellen van substitutie- en complementariteitseffecten in assortimenten. Zoals hier met een vereenvoudigd voorbeeld werd gedemonstreerd stellen ze de onderzoeker, en dus ook managers in de detailhandel, in staat om verschillende factoren die van invloed zijn op de consumentenkeuze uit assortimenten te beheersen. Dit betreft zowel kenmerken van assortimenten, individuele artikelen en overige winkelkenmerken als de fase van het beslissingsproces van de consument. Het extended logit model geeft inzicht in substitutie-patronen binnen en tussen produktcategorieën of winkels. In dit onderzoek vonden we bijvoorbeeld dat substitutiepatronen niet perse symmetrisch zijn en dat een vergroting van het assortiment de kans dat een winkelbezoeker een aankoop doet vergroot.

In verband met de hevige concurrentie in de detailhandel is het voor winkels van levensbelang om brede en soms ook diepe assortimenten te voeren en tegelijk de kosten van voorraad houden en ruimtebeslag te verlagen. Dit betekent dat moet worden beslist over verwijdering van artikelen uit het huidige assortiment en eventuele toevoeging van nieuwe artikelen. Dergelijke beslissingen vereisen inzicht in de bijdrage van individuele artikelen aan de totale omzet van een produktgroep of een winkel. Modellen zoals hier gedemonstreerd kunnen gebruikt worden om de gevolgen van verschillende scenario's van assortimentssamenstelling door te rekenen en op basis van de resultaten van dergelijke scenario's de manager te helpen goede beslissingen voor het assortiment te nemen.

LITERATUUR

Anderson, D.A. en J.B. Wiley (1992), "Efficient Choice Set Designs for Estimating Availability Cross-Effects models," *Marketing Letters*, 3, 357-370.

Huber, J., J.W. Payne en Chr. Puto (1982), "Adding Asymmetrically Dominated Alternatives: Violations of Regularity and the Similarity Hypothesis," *Journal of Consumer Research* 9, 90-98.

Kahn, B.E. en D.R. Lehmann (1991), "Modeling Choice Among Assortments," *Journal of Retailing*, 67, 274-299.

Krishnamurthi, L., S.P. Raj en K. Sivakumar (1995), "Unique Inter-Brand Effects of Price on Brand Choice," *Journal of Business Research*, 34, 47-56.

Lazari, A.G. en D.A. Anderson (1994), "Designs of discrete choice set experiments for estimating both attribute and availability cross effects," *Journal of Marketing Research*, 31, 375-383.

Louviere, J.J. en G. Woodworth (1983), "Design and Analysis of Simulated Consumer Choice or Allocation Experiments: An Approach Based on Aggregate Data," *Journal of Marketing Research*, 20 (November), 350-367.

McFadden, D., K. Train, en W.B. Tye (1977), "An Application of Diagnostic Tests for the Independence of Irrelevant Alternatives Property of the Multinomial Logit Model," *Transportation Research Record*, 637, 39-46.

Oppewal, H., en H. Timmermans (1991), "Context Effects and Decompositional Choice Modeling," *Papers in Regional Science,* 70 (2), 113-131.

Oppewal, H. en H.J.P. Timmermans (1992), "Conjuncte keuzeexperimenten: achtergronden, theorie, toepassingen en ontwikkelingen", *Jaarboek '9293 van de Nederlandse Vereniging van Marktonderzoekers*, 3358.

Oppewal, H., J.J. Louviere, en H.J.P. Timmermans (1997), “Modelling the Effects of Shopping Centre Size and Store Variety on Consumer Choice Behaviour,” *Environment and Planning A*, 29, 1073-1090.
Simonson, I. en A. Tversky (1992), “Choice in Context: Trade-off Contrast and Extremeness Aversion,” *Journal of Marketing Research*, 29 (August), 281-296.

7. Reclamegeheugen en merkattitude: Een meta-analyse

R. PIETERS, T.H.A. BIJMOLT en H. GRUIJTERS

SAMENVATTING

Er heerst een felle discussie binnen de wereld van reclame-onderzoek of reclamegeheugen al dan niet een goede indicatie is voor reclame-effectiviteit. In dit artikel gaan we in op de achtergrond van deze discussie. We presenteren een meta-analyse, een kwantitatieve analyse van de literatuur, over de relatie tussen reclamegeheugen en merkattitude. Daaruit blijkt dat er een kleine, positieve, significante relatie bestaat tussen deze twee grootheden. Bovendien hangt de sterkte van de relatie af van een aantal moderatorvariabelen. We gaan in op de implicaties voor de praktijk en mogelijkheden voor verder onderzoek.

1. INLEIDING

> "Voordat U verder leest, probeert U zich eens een reclame-uiting te herinneren waarmee u recent geconfronteerd bent. Begint u maar.
>
> Welke reclame-uiting komt in U op?
>
> Is het de televisie-commercial met de zijspannende koe van Melk Unie Milkshake? Is het een advertentie voor gevleugelde inlegkruisjes? Of is het misschien de commercial van de dealende Daewoo dealer?"

Welke reclame-uiting ook in U opkwam, de vraag is of het, behalve om onderhoudende redenen, wat uitmaakt of consumenten de reclame-uitingen waarmee ze geconfronteerd worden onthouden. Is reclame die goed herinnerd wordt effectiever dan reclame die slecht herinnerd wordt? Of algemener: "Wat is de relatie tussen het geheugen voor reclame en de effectiviteit van reclame?" Deze vraag houdt de marketingtheorie en –praktijk al geruime tijd bezig, en niet zonder emotie. Er zijn felle voor- en tegenstanders van het gebruik van geheugen als indicator van reclame-effectiviteit, die elkaar te vuur en te zwaard bestrijden, niet altijd met redelijke argumenten of op basis van empirische evidentie, maar soms wel tot bijna in het graf.

In dit hoofdstuk geven we eerst een overzicht van onderzoek naar de relatie tussen geheugen voor en effectiviteit van reclame. We introduceren argumenten voor en tegen, en de resultaten van eerder onderzoek naar de relatie. Daarna presenteren we de resultaten van een meta-analyse naar de relatie tussen reclamegeheugen en merkattitu-

de. Daartoe gaan we nader in op de functie van meta-analyse in marketing en op de verschillende technieken om meta-analyses uit te voeren. Ten slotte formuleren we conclusies voor toegepast en fundamenteel onderzoek naar de relatie tussen reclamegeheugen en merkattitude, en voor verdere toepassingen van meta-analyses.

2. GEHEUGEN VOOR RECLAME

Sinds het begin van deze eeuw bestaan er methoden van reclame-onderzoek die gebaseerd zijn op het geheugen van consumenten. Daniel Starch was in 1922 een van de eersten die in media-onderzoek het geheugen van consumenten onderzocht. Hij liet consumenten vrij bladeren door een tijdschrift en stelde ze er daarna vragen over, terwijl hij met de consumenten opnieuw door het tijdschrift bladerde. Hij vroeg of consumenten zich konden herinneren dat ze een bepaalde advertentie opgemerkt hadden toen ze zelf de eerste keer door het tijdschrift bladerden. Het percentage van de consumenten dat bevestigend antwoordde noemde hij het percentage "Noted." Hij vroeg consumenten ook of zij het merk (naam, logo en dergelijke) in de advertentie hadden opgemerkt de eerste keer (percentage "Seen associated"), en ten slotte vroeg hij of consumenten de tekst van de advertentie de eerste keer voor de helft of meer gelezen hadden (percentage "Read most"). De wijze waarop Daniel Starch vaststelde of consumenten "aandacht" hadden besteed aan gedrukte reclame is nog steeds populair. De organisatie die het Starch systeem uitvoert (Starch INRA Hooper) publiceert de resultaten van analyses op het door haar verzamelde materiaal in een tijdschrift "Tested Copy."

Veel andere vormen van reclame-onderzoek, zowel in de pretest fase als in de posttest fase, maken gebruik van geheugenmetingen. In het McCollum-Spielman systeem om televisie-commercials te pretesten wordt onder meer onderzocht hoe goed consumenten zich nieuwe commercials kunnen herinneren die ze enkele keren in een normale programma context hebben gezien. Het McCollum-Spielman systeem wordt wereldwijd toegepast, en in Nederland aangeboden door Intomart-Qualitatief in Hilversum. Het NIPO in Amsterdam meet onder de naam Impact onderzoek al jaren hoe goed televisiecommercials worden herinnerd, direct nadat ze zijn uitgezonden. Het DAR (Day-After-Recall) systeem van de Burke ondernemingen meet al meer dan 30 jaar reclameherinnering vanuit de visie dat reclame-uitingen een geheugenspoor moet achterlaten om effectief te kunnen zijn.

Niet alleen is geheugen voor reclame een onderdeel van systemen van toegepast reclame-onderzoek, geheugen voor reclame staat ook frequent centraal in fundamenteel onderzoek naar de werking en effectiviteit van reclame. Onderzocht wordt welke kenmerken van boodschappen, media, merken, producten en consumenten van invloed zijn op het geheugen voor reclame. Singh, Misra, Bendapudi en Linville (1994) onderzoeken bijvoorbeeld wat het effect is van herhaling van televisie commercials op het geheugen voor de televisie commercials. Wijzelf (Pieters en Bijmolt 1997) onderzochten het effect van het aantal concurrerende commercials in het reclameblok en de lengte en plaatsing van de commercial op de herinnering en herkenning van de merknaam, voor een database van bijna 3000 door het NIPO onderzochte commercials. Starch INRA Hooper stelt de gegevens van haar onderzoek regelmatig ter beschikking aan universitaire onderzoekers voor verdere analyses. John Rossiter (1981) onderzocht bij-

voorbeeld het effect van het aantal en soort woorden in de kopregel op Starch scores. De resultaten van dergelijke onderzoeken worden in tekstboeken opgenomen die in het marketing-communicatie onderwijs op hogescholen en universiteiten worden gebruikt. In het beroemde boek van John Rossiter en Larry Percy (1997) worden heel specifieke aanbevelingen gedaan over de inhoud, uitvoering en plaatsing van reclame-boodschappen om daarmee de effectiviteit van reclame te verhogen. Veel van deze aanbevelingen zijn gebaseerd op onderzoek waarin het geheugen voor reclame centraal stond.

Het is duidelijk dat geheugen voor reclame een vooraanstaande plaats in marketingonderzoek en -praktijk inneemt. Zolang geheugen voor reclame wordt gemeten om "het geheugen voor reclame te meten" is er weinig aan de hand. Echter, de interesse in marketingonderzoek en -praktijk gaat veelal niet naar het geheugen zelf uit. Geheugen voor reclame wordt vaak gebruikt als een indicator of proxy voor (a) aandacht voor de reclame-uiting, (b) begripsvorming of leren van de reclame-boodschap, of (c) attitude ten aanzien van het geadverteerde merk. De drie toepassingen die oplopen in de sterkte van hun aanname bespreken we hieronder.

In Starch metingen wordt geheugen vaak gemeten als een indicator voor de aandacht die consumenten geschonken hebben aan reclame. Men neemt hierbij aan dat hoe beter een advertentie onthouden wordt, hoe groter de kans dat er veel aandacht aan geschonken is tijdens de eerste confrontatie. Onderzoek toont aan dat deze aanname vaak niet realistisch is (bijvoorbeeld Pieters 1998a; Verstegen 1998). Bovendien zijn er recente ontwikkelingen die het mogelijk maken om aandacht voor reclame grootschalig en frequent, direct te meten met behulp van oogbewegingsregistratie (zie bijvoorbeeld Pieters 1998b). Daardoor vervallen voor een groot deel de voordelen van Starch metingen, zonder de nadelen ervan. Ook in onderzoek naar televisie-commercials wordt geheugen vaak als indicator voor aandacht gebruikt.

Geheugen voor reclame wordt ook vaak geinterpreteerd als een indicator voor begripsvorming of het leren van de reclame-boodschap. In navolging van Herbert Krugman stellen Singh en Rothschild (1983) dat herkenning van de merknaam in televisie-reclame een maat voor begripsvorming zou kunnen zijn. In onderzoek met behulp van Starch scores stelt Adam Finn (1988) iets vergelijkbaars voor print advertenties. Als consumenten beweren dat ze een relatie tussen het merk en de boodschap gelegd hebben en, als ze beweren dat ze de helft of meer van de bodytekst gelezen hebben, dan wordt dat opgevat als een bewijs dat er begripsvorming en leren heeft plaatsgevonden. Veel kritiek op deze beweringen is niet gekomen, veel navolging en impact hebben ze evenmin gehad.

De sterkste bewering is uiteraard dat geheugen voor reclame een indicator is voor de effectiviteit van reclame. Als dat waar zou zijn, zou er een positieve relatie tussen het geheugen voor reclame en de attitude ten aanzien van het geadverteerde merk bestaan (op gedesaggregeerd niveau), en een positieve relatie tussen geheugen voor reclame en marktaandeel (op geaggregeerd niveau, aannemende dat consumenten veelal attitude-consistent gedrag vertonen). De sterkste bewering over de relatie heeft ook tot het meest verhitte debat geleid. In de Verenigde Staten werd de vraag vertaald als, moet herinnering (of herkenning) als maat voor reclame-effectiviteit gebruikt wordt, of moet attitude-verandering als maat gebruikt worden: "Recall or Persuasion."

3. RECLAMEGEHEUGEN EN MERKATTITUDE

Leo Bogart presenteerde in de zestiger jaren de resultaten van onderzoek waarin de correlatie tussen reclameherinnering en omzet een bedroevend lage .03 was, dat wil zeggen praktisch nul. In 1983 verwoordde Larry Gibson (1983) zijn standpunt in een beroemd artikel in het Journal of Advertising Research als volgt: "The 'persuasion versus recall' argument must be the longest playing controversy in all of marketing research. It has been researched, debated, and argued for over 30 years to my personal knowledge. Even today we may not know the final answer to this ancient question, but we do know what the answer is not – and it's clearly not recall!" Deze laatste conclusie was gebaseerd op een grootschalig onderzoek waarin 142 commercials in 55 product categoriën onderzocht werden, en waarin de test resultaten van de commercials met een maat voor marktaandeel werden vergeleken.

Tegenover Gibson en Bogart staan anderen zoals Jan Stapel (1987) grondlegger van het NIPO die zijn stellingname in een artikel met de titel "Viva Recall! Viva Persuasion!" nogmaals benadrukte. Eerder rapporteerde hij (Stapel 1971) al dat niet-gebruikers die de reclame-boodschap kunnen herinneren een twee tot driemaal zo grote aankoopintentie tonen dan degenen die de boodschap niet herinneren. De Zuidafrikaanse onderzoeker Eric du Plessis (1994) concludeert op basis van de onderzoeken die zijn marktonderzoekbureau uitvoerde dat "There remains one variable that is still showing a bigger effect on recall than ad length, and that is likability." Met andere woorden: advertenties die plezierig gevonden worden, worden ook beter onthouden. Immers, mensen zijn bereid hun activiteiten (bijvoorbeeld het lezen van een tijdschrift) te onderbreken voor een plezierige reclame-boodschap, zegt hij. Du Plessis raakt de kern van de redenering waarom geheugen voor reclame een positieve relatie met attitude ten aanzien van het merk zou kunnen vertonen.

De redenering is als volgt: plezierige reclame-uitingen leiden tot verhoogde aandacht. Deze verhoogde aandacht leidt tot een beter geheugen voor de reclame-uiting. Bovendien stimuleren plezierige ("likable"[1]) reclame-uitingen een positieve attitude ten aanzien van het merk. Dat resulteert in een positieve relatie tussen geheugen voor reclame en merkattitude. De analyse is nog wel iets complexer, omdat wanneer de bovenstaande redenering juist zou zijn, daaruit de aanbeveling zou volgen om slechts de "likability" van reclame-uitingen te meten. Immers likability stimuleert zowel het geheugen voor de reclame-uiting als de attitude ten aanzien van het merk. Echter, vaak wordt aangenomen dat geheugen voor reclame-uitingen meer meet dan likability alleen, bijvoorbeeld begripsvorming. Bovendien wordt vaak aangenomen dat geheugen voor reclame-uitingen gemakkelijker te meten is dan attitudes en andere subjectieve oordelen van consumenten.

Dat is de situatie op dit moment. Er zijn twee kampen, de geheugen fans en de persuasion adepten, en veel mensen die een beetje bij beiden horen. Het is moeilijk een conclusie te trekken over de relatie tussen reclamegeheugen en merkattitude. Du Plessis

[1] We bedoelen feitelijk "attitude ten aanzien van de advertentie," afgekort als Aad. Likability is een onderdeel van Aad, maar niet helemaal hetzelfde. In Nederland worden "likability" en Aad veelal aan elkaar gelijk gesteld. We doen dat hier ook.

(1994) merkte terecht al op dat de argumenten voor of tegen geheugenmetingen vaak vanuit politieke en economische visies worden gevoerd. Immers, veel marktonderzoekbureaus hebben grote, gevestigde belangen bij een bepaalde meting van reclame-effectiviteit. Bovendien zijn gepubliceerde onderzoeken niet altijd gemakkelijk met elkaar te vergelijken in opzet en uitwerking. Om dit probleem op te lossen, presenteren wij een meta-analyse van onderzoek naar de relatie tussen reclamegeheugen en merkattitude. Het onderzoek heeft tot doel vast te stellen of er een positieve relatie tussen reclamegeheugen en merkattitude bestaat, wat de sterkte van de relatie is, en welke factoren de omvang van de relatie beinvloeden. We merken op dat dit onderzoek een onderdeel is van een lopend onderzoeksprogramma met meta-analyses, en dat de resultaten daarom eerder voorlopig dan definitief zijn. Allereerst introduceren we doelstellingen en technieken van meta-analyse. Daarna presenteren we de resultaten van onze meta-analyse.

4. META-ANALYSE IN MARKETING

Meta-analyse is een vorm van kwantitatieve literatuuranalyse met als doel om generaliseerbare uitspraken te doen over belangrijke marketing-fenomenen (zie bijvoorbeeld Farley, Lehmann en Sawyer, 1995). Meta-analyse is een verzameling methoden om de resultaten van eerdere onderzoeken op een specifiek terrein samen te vatten, en om daar nieuwe hypothesen uit af te leiden. De opzet van een meta-analyse is vergelijkbaar met de opzet van andere empirische onderzoeken. Informatie uit eerdere onderzoeken wordt gecodeerd en in een gegevensbestand opgeslagen. In veel ander marktonderzoek is de onderzoekseenheid een consument of bedrijf. In meta-analyse is de onderzoekseenheid een eerdere meting van de relatie die centraal staat. Een meta-analyse bevat altijd een statistische analyse van eerdere onderzoeken. Gibson die in de jaren tachtig vaststelde "Not recall" voerde een *kwalitatieve* literatuuranalyse uit. Als iemand een verzameling eerdere onderzoeken kwalitatief analyseert, en samenhangen tussen de onderzoeken probeert te ontdekken zonder daarbij een nieuwe statistische analyse op het materiaal uit te voeren, kan dat heel nuttig zijn, maar dan is het geen meta-analyse.

Een meta-analyse heeft meestal twee doelen. Ten eerste dient de overall grootte van de relatie die centraal staat, te worden vastgesteld. Wat is bijvoorbeeld over het algemeen de prijselasticiteit van de vraag naar een produkt? Of hoe groot is de correlatie tussen tevredenheid van het personeel en de tevredenheid van de klant? Het gaat dus om de 'gemiddelde' grootte van een dergelijke relatie. Vaak zal de omvang van de relatie echter afhangen van andere factoren. Het bepalen van het effect van deze factoren, zogenaamde moderatoren, op de omvang van de relatie die wordt bestudeerd, is het tweede doel van meta-analyse.

Bij het toepassen van meta-analyse stuit men vaak op een aantal methodologische problemen. De wijze van meting van de relatie kan bijvoorbeeld nogal verschillen tussen de publikaties die zijn gevonden. Het ene artikel rapporteert correlaties, de ander regressie-coefficienten, een derde een F-toets om verschillen tussen experimentele groepen na te gaan en een vierde parameters uit een LISREL model. Dergelijke metingen van de effectgrootte kunnen gelukkig naar een standaard worden omgerekend (zie Rosenthal, 1991). In onze meta-analyse hebben we gekozen om te werken met correlaties, omdat dat het meest gangbare grootheid is in de reclame-literatuur.

Ook het "file drawer" of *"publication bias"* probleem (Rust, Lehmann, en Farley, 1990) kan invloed hebben op de resultaten van een meta-analyse. Doordat zowel auteurs als beoordelaars van artikelen graag significante resultaten willen zien, bestaat er een vertekening naar dergelijke resultaten in de literatuur. Dit is met name het geval wanneer er sprake is van concensus of een overheersende stroming. Een studie waarin de correlatie tussen prijs en kwaliteit positief en significant is, is makkelijker te publiceren dan een studie waarin dit niet het geval is. Wanneer er geen duidelijke mening heerst binnen een vakgebied, is de "publication bias" waarschijnlijk gering.

Een onderzoeker die meta-analyse toepast verzamelt studies uit allerlei boeken, tijdschriften, congresbundels, etc. en wel die studies die metingen van het desbetreffende effect rapporteren. Het blijkt dat sommige van deze studies meer dan één meting bevatten! Een studie zou bijvoorbeeld alle vier de correlaties tussen reclameherinnering en -herkenning enerzijds en merkattitude en aankoopkans anderzijds kunnen rapporteren. Alle meta-analyses binnen marketing-onderzoek die wij kennen, bevatten enkele studies die meerdere metingen van het effect rapporteren. Meestal bevatten vrijwel alle studies één effectmeting, een paar studies twee tot vijf effectmetingen en een enkele studie tientallen of zelfs honderden effectmetingen. Het gewoon opnemen van al deze metingen, weegt de studies die meerdere metingen van het effect bevatten natuurlijk zwaarder mee. De vraag is of je dat wil. Bovendien zijn metingen die uit dezelfde studie komen waarschijnlijk niet onafhankelijk; ze zijn met een bepaalde reden in dezelfde studie opgenomen, ze zijn verzameld onder dezelfde steekproef, etc. Met deze meerdere effectmetingen per studie kan op verschillende manieren worden omgegaan. We hebben deze methoden op een rijtje gezet en met elkaar vergeleken (Bijmolt en Pieters, 1998). Uit onze vergelijking kwam de methode gebaseerd op multi-level regressie-analyse (zie bijvoorbeeld Bryk en Raudenbush, 1992 en Goldstein, 1995) als het meest geschikt naar voren. Bij dergelijke multi-level modellen wordt expliciet rekening gehouden met het feit dat er sprake is van een geneste gegevensstruktuur, dat wil zeggen dat de effectmetingen uit een en dezelfde studie als zijnde afhankelijk gemodelleerd worden. Het model ziet er als volgt uit:

$$y_{ms} = b_0 + \Sigma_k\, b_k\, x_{k,ms} + e_{ms} + u_s\,, \qquad (1)$$

waarbij y_{ms} de effectmeting m uit studie s is, $x_{k,ms}$ voor de K moderatorvariabelen staat, b_0 en de b_k's de K+1 regressiecoefficienten zijn en e_{ms} en u_s fouttermen op het meting- en studieniveau zijn. De modelvergelijking uit (1) kan worden geschat met speciale statistische programma's, bijvoorbeeld HLM (Bryk, Raudenbusch en Congdon, 1996).

We voerden een meta-analyse uit naar de relatie tussen reclamegeheugen en merkattitude, waarbij we gebruik maakten van multi-level regressie-analyse. De methode en resultaten van deze meta-analyse worden in de volgende secties besproken.

5. RECLAMEGEHEUGEN EN MERKATTITUDE: METHODE

Steekproef

Om de steekproef samen te stellen werd gezocht naar alle studies naar de relatie tussen het geheugen voor beïnvloedende boodschappen (commercieel en niet-commercieel)

en de attitude of het gedrag van consumenten ten aanzien van de propositie die in de boodschap centraal stond. Binnen deze brede definitie vallen bijvoorbeeld ook studies waarin boodschappen gebruikt werden om studenten te overtuigen van het nut van collegegeldverhoging. De strategie om relevante studies te traceren bestond uit drie fasen. Allereerst werd naar studies gezocht in drie geautomatiseerde literatuur-bestanden: ABI Inform (1/86-11/95), Psyclit (1/74-12/95) en Econlit (1/69-9/1995). Daarnaast werden alle jaargangen doorgenomen van marketing en psychologie tijdschriften waarvan aangenomen werd dat die over het onderwerp studies gepubliceerd zouden kunnen hebben. Het betrof dertien tijdschriften, waaronder Journal of Marketing, Journal of Consumer Research, Journal of Experimental Social Psychology. Ten slotte werden de literatuuropgaven onderzocht van de studies die in de eerste twee fasen werden gevonden.

Studies die uiteindelijk in de meta-analyse betrokken worden, dienen aan een aantal criteria te voldoen. Allereerst dienen de studies gegevens te bevatten over de relatie waarin we geïnteresseerd zijn (in de praktijk hebben nogal wat studies een titel die de inhoud niet voldoende dekt), en dienen de gegevens zo in de studie gepresenteerd te worden dat ze ook gebruikt kunnen worden. Twaalf studies die in de steekproef zaten vielen door dit criterium af. Ten tweede dienen de studies de relatie tussen reclamegeheugen en merkattitude zo te meten dat geheugen de attitude kan beïnvloeden en niet omgekeerd. Daardoor vielen zes studies af. Ten derde dient de attitude die gemeten wordt ook betrekking te hebben op de inhoud van de reclameboodschap en niet op iets anders. Drie studies vielen daardoor af. Daardoor bleven uiteindelijk 45 studies over die in de meta-analyse gebruikt konden worden. In deze 45 studies werd in totaal 307 maal de relatie tussen reclamegeheugen en merkattitude gerapporteerd.

Codeboek

In een codeboek werd aangegeven welke informatie uit de studies werd vastgelegd. Informatie met betrekking tot de studie, de respondenten, de stimuli die gehanteerd werden, en de meting van de relatie tussen geheugen en attitude, werd vastgelegd. Bijvoorbeeld werd informatie opgenomen over de vorm van de publicatie, boek of artikel, over het jaar van publicatie, over steekproefsamenstelling en dergelijke. Daarnaast werd bijvoorbeeld het soort geheugenmaat gecodeerd (ongeholpen herinnering, geholpen herinnering of herkenning) en het soort attitudemaat (merkattitude, attitude ten aanzien van de advertentie, koopintentie, gedrag, enzovoorts). Een volledig codeboek is voor belangstellenden beschikbaar bij de auteurs.

Met behulp van het codeboek werd een inhoudsanalyse uitgevoerd op alle 307 relaties in de 45 studies door twee beoordelaars onafhankelijk van elkaar (gemiddelde betrouwbaarheid .87).

Steekproefkenmerken

De oudste publicatie in onze meta-analyse was uit 1959, de meest recente uit 1994. Meer dan de helft van de studies is afkomstig uit het Journal of Personality and Social Psychology, Journal of Advertising Research en Journal of Consumer Research. Drie studies waren afkomstig uit boeken. Meestal (88%) van de gevallen wordt gebruik gemaakt van een steekproef van studenten; of dat wat uitmaakt onderzoeken we later. In ongeveer 51% van de gevallen hebben de stimuli betrekking op consumentenproducten. In 47% van de gevallen worden de stimuli in omgevingsmateriaal ("in-context")

aangeboden. Meestal worden de respondenten een enkele keer blootgesteld aan de stimulus, voordat hun geheugen ervoor wordt vastgesteld (90%). In 45% van de gevallen betreft het print stimuli, de rest is audio of audio en video.

6. RESULTATEN VAN DE META-ANALYSE

De resultaten van de multilevel regressie analyse op onze database worden nu besproken. Eerst presenteren we de correlatie tussen reclamegeheugen en merkattitude over alle 307 relaties in de 45 studies heen. Daartoe voeren we een random-ANOVA op de database uit. Daarna presenteren gaan we na welke factoren van invloed zijn op de sterkte van de correlatie. Alle gepresenteerde analyses werden uitgevoerd met het programma HLM (Bryk, Raudenbusch en Congdon 1996). In Tabel 1 zijn de resultaten van de random-ANOVA weergegeven.

Tabel 1. Correlatie tussen reclamegeheugen en merkattitude.

Fixed Effect	Coefficient	Standaard fout	t-waarde	p-waarde
Geschatte correlatie	.181	.027	6.788	<.001
Random effect	Variantie component	vr.gr.	Chi-kwadraat	p-waarde
Level-2 effect (studie)	.017	44	126.159	<.001
Level-1 effect (relatie)	.057			

Er blijkt een kleine, maar positieve, significante relatie te bestaan tussen reclamegeheugen en merkattititude. De correlatie bedraagt .181, en is significant op $p < .0001$. Deze resultaten ondersteunen de claims van Stapel, Du Plessis en anderen dat geheugen voor reclame-uitingen wel degelijk positief samenhangt met attitude ten aanzien van het geadverteerde merk. In absolute zin is de correlatie natuurlijk niet groot. Inspectie van Tabel 1 geeft ook aan dat de variantie van de correlatie .057 op het niveau van de relatie is, en .017 op het niveau van de studie. Dit betekent dat verschillen in de correlatie tussen reclamegeheugen en merkattitude vooral (voor meer dan 70%) vooral veroorzaakt worden door verschillen tussen de relatie die onderzocht worden, en slecht voor een klein deel door verschillen tussen studies.

In een vervolganalyse onderzochten we waardoor de verschillen in de correlatie tussen reclamegeheugen en merkattitude precies veroorzaakt werden. De resultaten zijn in Tabel 2 samengevat.

Inspectie van Tabel 2 toont aan dat de correlatie tussen reclamegeheugen en merkattitude *groter* is in studies die in boeken gepubliceerd zijn (publicatievorm), in studies die recent gepubliceerd zijn (publicatiedatum), en in studies die gebruik maken van printadvertenties (medium). De correlatie tussen reclamegeheugen en merkattitude is *kleiner*

in studies die de relatie tussen reclamegeheugen en merkattitude centraal stellen (focus op relatie), waarbij de stimuli verbaal worden aangeboden, en waarbij blootstelling in een laboratorium situatie plaatsvindt.

Tabel 2. Factoren die de relatie tussen reclamegeheugen en merkattitude beïnvloeden.

Factor	Coefficient	t-waarde	p-waarde
Studieniveau:			
Intercept	-16.654	-2.367	.024
Focus op relatie	-.145	-2.220	.033
Medium	-.183	-2.242	.031
Publicatiedatum	.008	2.351	.024
Publicatievorm	.232	2.240	.031
Relatieniveau:			
Geheugenmeting	.152	2.423	.021
Instructie	-.189	-3.957	<.001
Attitudemeting	.159	2.607	.014
Random Effect:	Variantie Component:	Chi-kwadraat (vr.gr.):	p-waarde:
Level-2 effect (studie)	.013	78.046 (36)	<.001
Level-1 effect (relatie)	.052		

We gaan op een specifiek soort resultaten in. Vanaf de jaren 60-70 is de heersende opvatting in marketing en marktonderzoek dat er *geen* significante relatie is tussen reclame-geheugen en merkattitude. In een dergelijk klimaat dienen bureaus en onderzoekers die toch gebruik blijven maken van metingen van reclamegeheugen zich vaak te verdedigen. In een dergelijk klimaat was het daarom "*Comme il faut*" om voortdurend aan te tonen, dat er *geen* relatie tussen de genoemde twee variabelen was. Hoewel onze resultaten niet als bewijs kunnen gelden, lijkt het erop dat dit klimaat tegen reclamegeheugen als effectiviteitsmaat heeft geleid tot een publicatie-bias waarbij positieve relaties moeilijker publiceerbaar waren. Er zijn verschillende aanwijzingen in onze meta-analyse.

Ten eerste, veelal worden de "beste" wetenschappelijke studies in tijdschriften gepubliceerd. We zien dat de relatie tussen geheugen en attitude in deze meta-analyse het laagst is in tijdschriften, en significant hoger in boeken. Ten tweede, het blijkt dat de relatie tussen reclamegeheugen en merkattitude het laagst was in publicaties die hoofdzakelijk deze relatie als onderwerp hadden (wat bijvoorbeeld blijkt uit de titel). In publicaties die over iets anders gingen, maar "toevallig" ook naar de relatie keken was de correlatie tussen reclamegeheugen en merkattitude hoger. Ten derde, hoe ouder de publicatie was (de oudste was van 1954, de meeste zijn uit de jaren 60-70) hoe lager de correlatie die gevonden werd tussen reclamegeheugen en merkattitude. Uiteraard zijn dit geen van allen bewijzen. Echter, het significante positieve verband dat wij in deze meta-analyse vonden, en de factoren die de sterkte van de relatie beïnvloeden rechtvaardigen een hernieuwde aandacht voor geheugen als maat voor reclame-effectiviteit.

7. CONCLUSIE

De relatie tussen geheugen voor reclame en attitudes ten aanzien van de geadverteerde merken is aan een her-evaluatie toe. Uit onze meta-analyse over 45 studies met 307 relaties blijkt een kleine, maar zeer betrouwbare, positieve relatie tussen reclamegeheugen en merkattitude. Deze relatie wordt significant beïnvloed door factoren die met de onderzoekscultuur in marketing te maken lijken te hebben. Immers, sinds Gibson's hartekreet "Not Recall" is het veiliger om met een Ajax T-shirt in een trein Feyenoorders plaats te nemen, dan om aan collega's toe te geven dat je onderzoek doet naar reclamegeheugen. Onze resultaten rechtvaardigen de negatieve reacties op reclamegeheugen niet. Uiteraard ondersteunen ze een juichstemming over de waarde van reclamegeheugen als effectiviteitsmaat evenmin. Wel geven ze aan dat de stelling "Not recall" prematuur was, en dat meer aandacht nodig is voor de condities waaronder reclamegeheugen wel een groot effect kan hebben. Veelal is de relatie tussen reclamegeheugen en merkattitude onderzocht. We kennen geen onderzoek naar de relatie tussen reclamegeheugen en merkbekendheid bijvoorbeeld. Niet alleen Rossiter en Percy (1997) benadrukken keer-op-keer dat naast merkattitude, merkbekendheid een centrale communicatie-doelstelling is. Het is heel wel denkbaar dat reclamegeheugen een groter effect heeft op merkbekendheid dan op merkattitude. Immers beide zijn gebaseerd op geheugen. Ook is het wel mogelijk dat reclamegeheugen vooral een positief effect heeft bij nieuwe advertenties en nieuwe merken (waar er een reden moet zijn om ze te onthouden) dan bij bestaande advertenties en merken. Andere hypothesen kunnen uit de resultaten van onze meta-analyse worden afgeleid, en kunnen in vervolgonderzoek getoetst worden.

Meta-analyse maakt het mogelijk tot generaliseerbare uitspraken over marketing-effecten en processen te komen. Hoewel de techniek reeds geruime tijd bestaat, wordt meta-analyse eigenlijk nog maar weinig toegepast. Dat is jammer. In toenemende mate ontstaan er wereldwijd grote databases waarin de resultaten van meerdere onderzoeken zijn opgeslagen. In plaats van weer een nieuw onderzoek naar een bepaalde relatie uit te voeren, kan er veel geleerd worden van meta-analyses op deze databases. Een gecombineerd gebruik van meta-analyse om de bestaande kennis te integreren, en primair onderzoek om de hypothesen te toetsten die uit de meta-analyse kunnen worden afgeleid, kan leiden tot de marketingwetten die ons vak verder helpen.

LITERATUUR

Bijmolt, Tammo H.A. en Rik G.M. Pieters (1998*), Generalizations in Marketing using Meta-Analysis with Multiple Measurements*, Tilburg University, CentER Discussion Paper No. 9893.

Bryk, Anthony S. and Stephen W. Raudenbush (1992*). Hierarchical Linear Models: Applications and Data Analysis Methods*, Newbury Park: Sage.

Bryk, Anthony S., Stephen W. Raudenbusch and Richard Congdon (1996), *HLM Hierarchical Linear and Nonlinear Modeling with the HLM/2L and HLM/3L Programs*, Chicago: Scientific Software International.

Farley, John U., Donald R. Lehmann, and Alan Sawyer (1995), "Empirical Marketing Generalization using Meta-Analysis," *Marketing Science*, 14 (3, part 2), 36-46.

Finn, Adam (1988), "Print Ad Recognition Readership Scores: An Information Processing Perspective," *Journal of Marketing Research*, 25, 168-177.

Gibson, Lawrence D. (1983), "If the Question is Copy Testing, the Answer is ... "Not Recall'," *Journal of Advertising Research*, 23 (1), 39-46.
Goldstein, Harvey (1995), *Multilevel Statistical Models*, Second edition, LondOn: Arnold.
Pieters, Rik G.M. (1998a), "Aandacht voor Reclame Wordt Steeds Moeilijker te Vangen," *Tijdschrift voor Marketing*, 32 (3), 16-19.
Pieters, Rik G.M. (1998b), "Oog voor Reclame," *Tijdschrift voor Marketing*, 32 (4), 21-24.
Pieters, Rik G.M. and Tammo H.A. Bijmolt (1997), "Consumer Memory for Television Advertising: A Field Study of Duration, Serial Position, and Competition Effects," *Journal of Consumer Research*, 23 (March), 362- 372.
du Plessis, Erik (1994), "Recognition versus Recall," *Journal of Advertising Research*, 34 (May/June), 75-91.
Rosenthal, Robert (1991), *Meta-Analytic Procedures for Social Research,* Revised edition, Newbury Park: Sage.
Rossiter, John R. (1981), "Predicting Starch Scores," *Journal of Advertising Research*, 5, 63-68.
Rossiter, John R. and Larry Percy (1997), *Advertising Communications and Promotion Management*, New York: McGraw-Hill, 2nd. Edition.
Rust, Roland T., Donald R. Lehmann, and John U. Farley (1990), "Estimating Publication Bias in Meta-Analysis," *Journal of Marketing Research*, 27 (May), 220-226.
Singh, Surendra N. and Michael L. Rothschild (1983), "Recognition as a Measure of Learning from Television Commercials," *Journal of Marketing Research*, 20 (August), 235-248.
Singh, Surendra N., Sanjay Mishra, Neeli Bendapudi and Denise Linville (1994), "Enhancing Memory of Television Commercials Through Message Spacing," *Journal of Marketing Research*, 31 (August), 384-392.
Stapel, Jan (1971), "Sales Effects of Print Ads," *Journal of Advertising Research*, 11 (3), 32-36.
Stapel, Jan (1987), "Viva Recall! Viva Persuasion!" *European Research*, 15 (4), 222-225.
Verstegen, Bas (1998), *The Catcher in the Eye: A Study of Visual Attention to Print Advertising by Means of Eye-Tracking Methodologies*, Doctoraal Scriptie, Sectie Marketing en Marktonderzoek, Vakgroep Bedrijfseconomie, Katholieke Universiteit Brabant.

8. De agendering van aankopen: Winkelen als tijdsbesteding

A. VAN DEN BROEK

SAMENVATTING

De hoeveelheid tijd die Nederlanders in een doorsnee werkweek besteden aan winkelen en aan de dagelijkse boodschappen was in de periode 1975-1995 opmerkelijk stabiel. Terwijl menig (vrije)tijdsbesteding duidelijk in de lift zat (vb. sporten en tv kijken) of juist een sterke neergang vertoonde (vb. sociale contacten en lezen), veranderde het tijdsbeslag van winkelen en van dagelijkse boodschappen niet of nauwelijks.
Onder de oppervlakte is wel degelijk van beweging sprake. Meer mensen, meer mannen vooral, zijn tot het regelmatig winkelend publiek gaan behoren. Daarmee loopt winkelen in de pas met een ontwikkeling die kenmerkend is voor de (vrije)tijdsbesteding in het algemeen: mensen krijgen meer bezigheden. Voor een groeiend aantal mensen is winkelen één van de in aantal groeiende bezigheden.

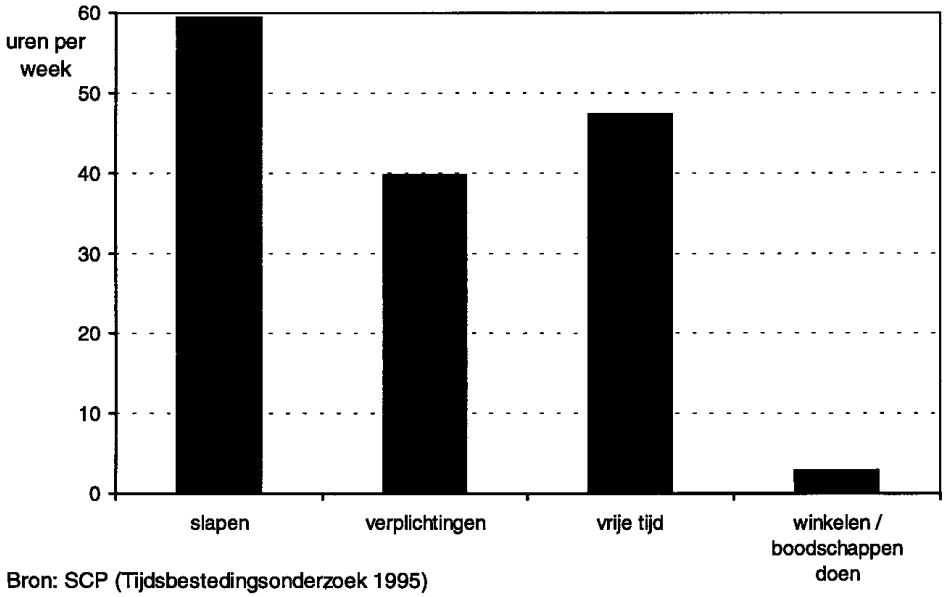
1. INLEIDING

In de toekenning van status en identiteit lag het accent lang op productie, meer in het bijzonder op de plaats in het productieproces. De vraag *"wat doet u?"* beantwoordde men eerder met het noemen van het beroep dat men uitoefende dan met het noemen van de favoriete vrijetijdsbesteding. De positie in het productieproces was bepalend voor iemands status en identiteit: men was student, kostwinner, huisvrouw of gepensioneerd. Daarbij gold voor de aard van het betaalde werk een statushiërarchie, uiteenlopend van jongste bediende of boerenknecht tot notaris of arts. In de verzuilde Nederlandse situatie was ook godsdienst een belangrijk identiteitsverlener.
Er zijn veel publicaties verschenen over de relatie tussen beroepspositie en status, een relatie die geen één-op-één karakter kent, maar die verloopt via bepaalde vormen van bezit, consumptie en gedrag. Hier wordt grotendeels aan deze literatuur voorbij gegaan (de klassieke studie in dit verband is Weber, 1922; een moderne klassieker is Bourdieu, 1984; en een handzame Nederlandse studie is Ganzeboom, 1988).

Exact een eeuw geleden vestigde de Amerikaanse socioloog Veblen de aandacht op 'conspicuous consumption' (opvallende consumptie). Hij lette daarbij niet alleen op de besteding van geld, maar ook op de besteding van tijd. Degenen die min of meer onbeperkt geld en tijd konden verspillen stonden in hoog aanzien. De uitdrukking 'tijd verspillen' dient overigens te worden genuanceerd: het ging Veblen er vooral om dat men in een positie was om straffeloos tijd door te brengen met het aanleren van welbe-

spraaktheid, belezenheid, goede smaak en andere niet-productieve eigenschappen die van grote welstand getuigden.
Veblen onderkende dat iets niet-productiefs doen, of consumeren, statusverhogend kon zijn. Echo's van die gedachte zijn te beluisteren in recente betogen over het belang van consumptie in het algemeen, en dat van winkelen in het bijzonder (o.a.: Bocock, 1993; Falk & Campbell, 1997; Featherstone, 1991; Maffesoli, 1996; Miller, 1995 en 1997; Shields, 1992). Zo is volgens Bocock niet productie maar consumptie typerend voor het moderne kapitalisme, waarin het besef van identiteit van mensen sterker verbonden is met hun consumptiepatroon dan met hun rol in het arbeidsproces (1993: 109). Hij signaleert dat er groepen zijn ontstaan waarbinnen consumptie een centrale rol in het leven speelt, en dat kleding, schoeisel, popmuziek en sporten gebruikt worden om zich een identiteit aan te meten (1993: 28). In dezelfde geest beargumenteren de redacteuren van een bundel artikelen onder de titel "*The shopping experience*" de keus van winkelen als onderwerp door er op te wijzen dat winkelen bij uitstek de maatschappelijke herschikking van productie naar consumptie illustreert (Falk & Campbell, 1997: 1-2).
Bij dergelijke krachtige pleidooien ter onderstreping van het belang van winkelen dient in het achterhoofd te worden gehouden dat ze (mede) dienen ter legitimering van een nieuw onderzoeksveld, waarvoor men met enig verbaal geweld een plaatsje onder de zon tracht te veroveren. Niettemin is duidelijk dat het winkelen zich momenteel in de warme belangstelling van sociologen mag verheugen.

Een relativering van het winkelen lijkt niettemin op z'n plaats. In 1995 werd door de Nederlandse bevolking van 12 jaar en ouder per week gemiddeld nog geen 3 uur (om precies te zijn: 2,8 uur) aan winkelen en boodschappen doen besteed.



Figuur 1. Tijdsbeslag van slapen, verplichtingen, vrije tijd en winkelen/boodschappen doen, in uren per doorsnee werkweek in 1995, bevolking van 12 jaar en ouder.

Een week telt 168 uur. Uitgedrukt in de totale duur van de week bedraagt het aandeel winkelen en boodschappen doen dus niet meer dan zo'n 1.7 procent. Na aftrek van slaap is de gemiddelde Nederlander per week zo'n 110 uur wakker, van die tijd wordt 2.6 procent winkelend en boodschappen doend doorgebracht. In verhouding tot drie algemene categorieën van tijdsbesteding steekt de activiteit winkelen en boodschappen doen schril af (figuur 1).

Resumerend: het belang van winkelen kan onder verwijzing naar enkele welgekozen sociologische auteurs naar gelieve enorm worden uitvergroot, terwijl het onder verwijzing naar de bescheiden positie in de week-agenda van de doorsnee Nederlander ook sterk kan worden gerelativeerd. In het nu volgende wordt geprobeerd de gulden midden weg te bewandelen. Beoogd wordt te komen tot een afgewogen plaatsbepaling van winkelen en boodschappen doen temidden van de vele andere bezigheden die in het leven van alle dag de aandacht vragen.
Daartoe zal allereerst op hoofdlijnen een schets gegeven worden van ontwikkelingen in de tijdsbesteding in de periode 1975-1995. Daarna wordt in meer detail ingegaan op winkelen en boodschappen doen als tijdsbesteding. Het tijdsbeslag van winkelen en boodschappen doen zal naar enige bevolkingscategorieën worden uitgesplitst. Tevens wordt een vergelijking gemaakt tussen het tijdsbeslag van winkelen en boodschappen doen en dat van enkele andere vormen van (vrije)tijdsbesteding. Tot besluit worden de voornaamste bevindingen gerecapituleerd.

De hier gepresenteerde empirische gegevens zijn afkomstig uit het Tijdsbestedingsonderzoek (TBO). Dit vijfjaarlijkse onderzoek, dat in 1975 de eerste editie beleefde, steunt op twee vormen van dataverzameling. Karakteristiek voor het TBO is dat respondenten gedurende een week in oktober hun bezigheden in een hen daartoe ter hand gesteld dagboek noteren. Aan de hand van een lijst van bezigheden geven zij per kwartier hun voornaamste bezigheid aan. Daarnaast wordt informatie ingewonnen door middel van een vragenlijst (dagboeken en vragenlijst zijn afgedrukt in Knulst & Van Beek, 1990). De kracht van dit onderzoek is dat de informatie over bezigheden niet verwijst naar meningen of ambities, maar naar de in het dagelijks leven feitelijk gerealiseerde (vrije)tijdsbesteding.

2. ONTWIKKELINGEN IN (VRIJE)TIJDSBESTEDING

2.1. Tijdsbesteding

Sinds 1975 is het tijdsbeslag van het totaal aan verplichtingen (werk, studie en zorgtaken) van de bevolking toegenomen. Anno 1995 werden Nederlanders in een doorsnee werkweek gemiddeld bijna twee uur langer door hun verplichtingen in beslag genomen dan in 1975. Men bezuinigde op slapen, maar ook de hoeveelheid vrije tijd nam iets af (tabel 1).
Het grotere tijdsbeslag van verplichtingen komt voor rekening van een grotere arbeidsinspanning. Vooral vanaf 1985 nam de gemiddelde arbeidsduur toe. Dit werd overigens niet veroorzaakt door een langere werkweek, maar door een grotere arbeidsdeelname. Volgens de TBO-gegevens is de feitelijk gerealiseerde arbeidsduur van werkenden in een doorsnee werkweek nauwelijks veranderd. Door de jaren heen telde hun werkweek

Tabel 1. Tijd besteed aan slapen, verplichtingen en vrijetijdsactiviteiten, 1975-1995, in uren per doorsnee werkweek, bevolking van 12 jaar en ouder.

	1975	1985	1995
slapen	60,5	59,4	59,4
verplichtingen	38,0	38,1	39,8
vrije tijd	47,9	49,0	47,3

Bron: SCP (Tijdsbestedingsonderzoek 1975-1995).

door de bank genomen iets meer dan 35 uur. Tussen 1985 en 1995 groeide het deel van de bevolking van 18 tot 65 jaar dat werkt (een feitelijke arbeidsduur van 12 uur of meer in de onderzoeksweek) echter van 47% tot 55% (SCP, 1998).
Deze groei komt voor rekening van zowel mannen als vrouwen. Onder vrouwen versnelde tussen 1985 en 1995 de groei van de arbeidsparticipatie die zich ook reeds in de periode 1975-1985 voordeed. Onder mannen compenseerde die groeiende arbeidsdeelname de uitstoot uit het arbeidsproces in het voorgaande decennium.

Omdat meer vrouwen zijn gaan werken en omdat veel volwassen Nederlanders met een partner een huishouden vormen, zijn er steeds meer (echt)paren waarvan beide partners werken. Het aandeel kostwinnershuishoudens daalde, het tweeverdienershuishouden maakte een opkomst door.
Doordat jongeren het ouderlijk huis verlaten zonder direct zelf een gezin stichten, groeide ook het aandeel éénpersoonshuishoudens. Niet langer leeft een meerderheid van de bevolking tussen 18 en 65 jaar in een kostwinnershuishouden (tabel 2).

Tabel 2. Huishoudensvormen waarin Nederlanders leven, 1975-1995, in procenten van de bevolking van 18 tot 65 jaar.

	1975	1985	1995
eenpersoonshuishouden	2,1	6,4	12,1
kostwinnershuishouden	59,2	47,9	33,5
tweeverdienershuishouden	13,4	22,0	32,2
anders[a]	25,3	23,7	22,3

[a] mensen die bij ouders inwonen en samenwonenden in huishoudens waarin niemand betaald werk verricht.
Bron: SCP (Tijdsbestedingsonderzoek 1975-1995).

Het kostwinnershuishouden kent een hoge graad van taakspecialisatie. De huisvrouw concentreert zich op de zorg voor het huishouden en het gezin, de man verdient de kost en de kinderen volgen onderwijs of werken, maar wonen in afwachting van een eigen gezinsleven nog thuis. De neergang van het kostwinnershuishouden deed afbreuk aan het voorheen hoge niveau van taakspecialisatie binnen huishoudens. In plaats daarvan maakte het fenomeen taakcombinatie opgang (Knulst & Van Beek, 1990).
Een eeuwenlang proces van functionele differentiatie kwam hiermee ten einde, in de privé-sfeer althans. Het deel van de bevolking tussen 18 en 65 jaar dat substantiële ver-

plichtingen voor studie en/of werk (≥ 12 uur per week) met substantiële huishoudelijke verplichtingen (≥ 8 uur per week) combineert, groeide van 22% in 1975 tot 40% in 1995. Taakcombinatie betekent dat degenen met het huishouden als hoofdtaak ook betaalde arbeid verrichten, terwijl degenen die primair professioneel bezig zijn tevens huishoudelijk werk doen.

De groeiende deelname van mannen aan onbetaalde arbeid en van vrouwen aan betaalde arbeid – uit oogpunt van een gelijkere bedeling der seksen veelal toegejuicht – heeft de tijdsdruk in gezinnen vergroot. In tweeverdienershuishoudens, de leefvorm van steeds meer Nederlanders, ligt de wekelijkse arbeidsinspanning met 56 uur bijna 16 uur hoger dan de wekelijkse arbeidsinspanning in kostwinnershuishoudens. Het totale tijdsbeslag van verplichtingen (werk, studie en zorgtaken) van beide partners bedraagt 105 uur per week, tegen 91 uur in kostwinnershuishoudens (tabel 3).

Tabel 3. Arbeidsverplichting en totaal aan verplichtingen van paren in kostwinners- en tweeverdienershuishoudens, in uren per doorsnee werkweek, 1975-1995, bevolking van 18 jaar tot 65 jaar.

	arbeidsverplichting	totaal aan verplichtingen
kostwinnerspaar	40	91
tweeverdienerspaar	56	105

Bron: SCP (Tijdsbestedingsonderzoek 1995).

De drukste fase uit de gezinscyclus is die waarin ouders de zorg voor jonge kinderen hebben. Het totaal aan verplichtingen van een ouderpaar met jonge kinderen (jongste kind jonger dan zes jaar) neemt in kostwinnershuishoudens 109 uur per week in beslag, in tweeverdienershuishoudens 119 uur.

2.2. Vrijetijdsbesteding

Om diverse redenen zou mogen worden verwacht dat de verscheidenheid aan activiteiten die mensen in hun vrije tijd ontplooien in de afgelopen decennia is toegenomen. De welvaartsgroei bracht meer activiteiten binnen het financiële bereik van meer mensen, het hogere opleidingsniveau bracht meer cultuurgoederen binnen het culturele bereik van meer mensen en de vrijetijdsindustrie bracht alsmaar meer vrijetijdsattributen en -activiteiten op de markt, terwijl het proces van ontkerkelijking de levensbeschouwelijke restricties op de vrijetijdsbesteding verminderde.
In het licht daarvan is het opmerkelijk dat de diversiteit van de vrijetijdsbesteding van de Nederlandse bevolking in de loop der jaren weinig verandering heeft ondergaan. Het aantal vrijetijdsactiviteiten waar mensen aan deelnemen is op tweeërlei wijze gemeten. Afgemeten aan het aantal activiteiten dat men gedurende de onderzoeksweek in de vrije tijd feitelijk ontplooide is sprake van een geringe daling van de diversiteit van de vrijetijdsbesteding. Afgemeten aan het totaal aan vrijetijdsactiviteiten waar men desgevraagd aan zegt deel te nemen is daarentegen van een lichte toename sprake (tabel 4). De tegenstelling tussen deze twee trends duidt op een zekere vervluchtiging van de vrijetijdsbesteding (SCP, 1998), hetgeen de gedachte dat mensen meer belevenissen zijn gaan najagen (Schulze, 1992) lijkt te ondersteunen.

Tabel 4. Enkele kenmerken van de vrije tijd (indexcijfers, bevolkingsgemiddelde van 1975 = 100), 1975-1995, bevolking van 12 jaar en ouder.

	1975	1985	1995
hoeveelheid vrije tijd in onderzoeksweek	100	102	99
aantal vrijetijdsactiviteiten in onderzoeksweek[a]	100	99	95
totaal aantal vrijetijdsactiviteiten[b]	100	100	103
uithuizigheid van de vrijetijdsbesteding[c]	100	102	110
autogebruik t.b.v. vrijetijdsbesteding[d]	100	118	129
outillage t.b.v. huiselijke vrijetijdsbesteding[e]	100	192	428
consumptieve bestedingen t.b.v. vrijetijdsbesteding[f]	100	117	151

[a] het aantal activiteiten dat men in de onderzoeksweek heeft ontplooid
[b] het aantal activiteiten waaraan men desgevraagd wel eens zegt deel te nemen
[c] het percentage van de vrije tijd dat niet thuis werd doorgebracht
[d] de wekelijkse tijdsduur van het autogebruik t.b.v. vrijetijdsactiviteiten
[e] de aanwezigheid van een band/cassetterecorder, platenspeler, cd-speler en videorecorder in huishoudens
[f] de uitgaven aan apparatuur, lectuur, lidmaatschappen, entreegelden, sport, vakanties, uitstapjes en verteringen in horecagelegenheden, per hoofd van de bevolking, in constante prijzen (CBS-gegevens, SCP-bewerking)

Bronnen: SCP (Tijdsbestedingsonderzoek 1975-1995) en CBS (Nationale rekeningen, diverse jaren).

Met name sinds 1990 brengen Nederlanders meer van hun vrije tijd elders dan thuis door. Daartoe bevindt men zich langer dan voorheen ook in de vrije tijd in de auto. De outillage van huishoudens ten behoeve van huiselijke vrijetijdsbesteding en de (voor inflatie gecorrigeerde) consumptieve bestedingen ten behoeve van de vrije tijd per hoofd van de bevolking namen in twintig jaar sterk toe. Doordat de hoeveelheid vrije tijd en het aantal verschillende vrijetijdsactiviteiten minder veranderden, deed zich een aanzienlijke stijging voor van de automobiliteit en de kapitaalintensiteit van de vrijetijdsbesteding, zowel per tijdseenheid als per activiteit (SCP, 1998).

Achter deze verschijnselen gaan uiteenlopende ontwikkelingen binnen specifieke bevolkingsgroepen schuil. Met name leeftijd is een belangwekkend onderscheid. De hoeveelheid vrije tijd is in de loop der jaren ongelijker over de verschillende leeftijdsgroepen verdeeld geraakt. Mensen tot 50 jaar kregen minder vrije tijd, terwijl het oudere bevolkingsdeel royaler in de vrije tijd kwam te zitten (tabel 5).

Tabel 5. Hoeveelheid vrije tijd in doorsnee werkweek naar leeftijd (indexcijfers, bevolkingsgemiddelde van 1975 = 100), 1975-1995, bevolking van 12 jaar en ouder.

	1975	1985	1995
12 tot 20 jaar	92	89	85
20 tot 35 jaar	94	94	89
35 tot 50 jaar	98	97	92
50 tot 65 jaar	105	114	113
65 jaar en ouder	119	129	127

Bron: SCP (Tijdsbestedingsonderzoek 1975-1995).

Deze polarisatie in de verdeling van de vrije tijd is des te belangwekkender aangezien het karakter van de vrijetijdsbesteding van het oudere bevolkingsdeel relatief sterk is veranderd. De uithuizigheid van hun vrijetijdsbesteding vertoonde een inhaalbeweging (De Haan, 1997), evenals de diversiteit van hun vrijetijdsrepertoire (tabel 6).

Tabel 6. Uithuizigheid[a] en diversiteit[b] van de vrijetijdsbesteding in een doorsnee werkweek naar leeftijd (indexcijfers, bevolkingsgemiddelde van 1975 = 100), 1975-1995, bevolking van 12 jaar en ouder.

	uithuizigheid			diversiteit		
	1975	1985	1995	1975	1985	1995
12 tot 20 jaar	127	122	124	116	107	101
20 tot 35 jaar	113	116	127	114	112	110
35 tot 50 jaar	89	95	100	102	105	108
50 tot 65 jaar	83	88	104	89	91	100
65 jaar en ouder	76	79	86	64	74	87

[a] het percentage van de vrije tijd dat niet thuis werd doorgebracht
[b] het aantal activiteiten waaraan men desgevraagd wel eens zegt deel te nemen
Bron: SCP (Tijdsbestedingsonderzoek 1975-1995).

Sinds 1975 groeide de uithuizigheid van de vrijetijdsbesteding van alle leeftijdsgroepen behalve de jeugd. Ook wat betreft de diversiteit van de vrijetijdsbesteding neemt de jeugd een uitzonderingspositie in, door als enige leeftijdsgroep een aanzienlijke afname van die diversiteit te zien te geven.
Daarentegen maakten mensen ouder dan 50 jaar zowel qua uithuizigheid als qua diversiteit van hun vrijetijdsbesteding een opmerkelijke inhaalbeweging. De aanvankelijke verschillen naar leeftijd zijn hierdoor aanzienlijk genivelleerd.

3. WINKELEN ALS (VRIJE)TIJDSBESTEDING

Gegeven de eerder gebruikte indeling van activiteiten in verplichtingen respectievelijk vrije tijd dringt zich de vraag op onder welke categorie winkelen en boodschappen doen gerangschikt dienen te worden. In onderzoek van het Sociaal en Cultureel Planbureau worden winkelen en boodschappen met huishoudelijk werk en kinderverzorging tot de zorgtaken gerekend. Samen met arbeid en studie vormen die zorgtaken het totaal aan verplichtingen. In deze bijdrage is de aan winkelen en boodschappen doen bestede tijd noch tot de vrije tijd noch tot de verplichtingen gerekend, maar apart beschouwd.
Het is zeer wel denkbaar dat mensen het doen van de dagelijkse boodschappen als huishoudelijke taak beschouwen en winkelen als een vrijetijdsbesteding beleven. Zo onderscheiden Leunissen en Prevo (1998) ‘het doen van dagelijkse boodschappen’ van ‘winkelen’, helaas zonder dat dit onderscheid de status bereikt van een conceptueel verschil tussen een activiteit in de sfeer van huishoudelijke verplichtingen enerzijds en een activiteit ter verpozing in de vrije tijd anderzijds.
In de literatuur over winkelen blijkt een dergelijke conceptualisering niet ongebruikelijk. Zo onderscheiden de eerder genoemde Falk en Campbell “two orientations of

shopping: shopping for and shopping around, the former economic or instrumental, the latter recreational". Ze wijzen er echter tevens op dat dit onderscheid in de praktijk vaak niet helder te maken is. Volgens hen zijn beide facetten veelal in één en dezelfde winkel-episode present (Falk & Campbell, 1997: 6).
Welk deel van het winkelen tot het 'funshoppen' als vrijetijdspassering gerekend kan worden is helaas ook in het hier gebruikte Tijdsbestedingsonderzoek niet duidelijk. Dit is spijtig, want waar 'funshoppen' vooral met andere vrijetijdsactiviteiten in concurrentie is, wordt het doen van de dagelijkse huishoudelijke boodschappen veeleer bepaald door de verdeling van verplichtingen in het huishoudenstype waartoe iemand behoort.

Op basis van het Tijdsbestedingsonderzoek kan echter wel worden onderscheiden wat primair tot de dagelijkse huishoudelijke boodschappen gerekend kan worden. In deze bijdrage worden dan ook telkens twee metingen vermeld: 'winkelen totaal' en 'dagelijkse boodschappen', waarbij het laatste een verbijzondering is van het eerste. De metingen betreffen het tijdsbeslag van het bezoek aan de in tabel 7 genoemde typen winkels, ongeacht of men er iets gekocht heeft. De vervoerstijd ten behoeve van winkelen en boodschappen doen is daarbij buiten beschouwing gelaten (met vervoerstijd is het tijdsbeslag daarvan ongeveer 1½ keer zo groot).

Tabel 7. Gebruikte scoringsmogelijkheden uit tijdsbestedingsregistratie.

grootwinkelbedrijf
andere levensmiddelenwinkels (bakker, slager, kruidenier, melkboer, groenteboer, slijter e.d.)
eten halen (snackbar, chinees e.d.)
markt
warenhuis
kledingzaak of schoenenzaak
doe-het-zelf-zaak (hout, verf, behang, ijzerwaren e.d.)
winkel van huishoudelijke of elektrische artikelen of woninginrichtingzaak
ander type winkel of boodschappen

Bron: SCP (Tijdsbestedingsonderzoek).

Zowel het tijdsbeslag van winkelen in het algemeen als dat van de dagelijkse boodschappen in het bijzonder geven in de periode 1975-1995 een hoge mate van stabiliteit te zien (tabel 8). De hoeveelheid tijd die de doorsnee Nederlander in een doorsnee werkweek in oktober aan winkelen en aan dagelijkse boodschappen besteedt was zelfs nagenoeg constant.
De enige noemenswaardige verandering is de groei van het deel van de bevolking dat binnen het bestek van één week gewinkeld heeft c.q. dagelijkse boodschappen gedaan heeft. Deze volumegroei van het regelmatig winkelend publiek komt vrijwel geheel voor rekening van een gedragsverandering onder mannen, die vaker gingen winkelen en vooral vaker boodschappen gingen doen. Dit neemt overigens niet weg dat vrouwen nog immer vaker winkelen en boodschappen doen dan mannen.

De verklaring voor de toegenomen winkelgeneigdheid van mannen moet gezocht worden bij de eerder beschreven verschuivingen in huishoudensvormen in Nederland. Het kostwinnershuishouden is op haar retour, en daarmee ook het voorheen hoge niveau

Tabel 8. Winkelen en dagelijkse boodschappen in een doorsnee werkweek naar sekse, 1975-1995, bevolking van 12 jaar en ouder.

	winkelen totaal			dagelijkse boodschappen		
	allen	mannen	vrouwen	allen	mannen	vrouwen
tijdsduur[a]						
1975	2,7	1,8	3,6	1,3	0,7	1,9
1985	2,7	1,9	3,5	1,3	0,7	1,8
1995	2,8	2,1	3,5	1,3	0,9	1,7
percentage deelnemers[b]						
1975	85	75	95	69	50	88
1985	88	80	96	74	59	88
1995	90	83	96	76	62	90
tijdsduur onder deelnemers[c]						
1975	3,1	2,4	3,7	1,8	1,3	2,1
1985	3,0	2,3	3,6	1,8	1,5	2,0
1995	3,1	2,5	3,6	1,7	1,4	1.9

[a] het aantal uren dat men gemiddeld per week aan winkelen respectievelijk aan het doen van de dagelijkse bood schappen besteedt
[b] het percentage dat minstens 1 kwartier per week aan winkelen respectievelijk aan het doen van de dagelijkse boodschappen besteedt
[c] het aantal uren dat degenen die in de onderzoeksweek minstens 1 kwartier winkelden in die week aan winkelen hebben besteed, idem voor de dagelijkse boodschappen
Bron: SCP (Tijdsbestedingsonderzoek 1975-1995).

van taakspecialisatie tussen de werkende huisvader en de zorgende huisvrouw. Of hij nu een éénpersoonshuishouden vormt of in een tweeverdienershuishouden leeft, de noodzaak tot taakcombinatie brengt met zich mee dat de man vaker dan voorheen naar de winkel moet.

Dit komt tot uiting in het verschil tussen de alleenverdienende man met partner enerzijds versus de alleenwonende jongere of de tweeverdienende man anderzijds (tabel 9). Vooral bij de dagelijkse boodschappen zijn de verschillen pregnant. Van de traditionele kostwinners doet slechts 50% eens per week of vaker boodschappen, terwijl dat onder alleenstaanden en tweeverdienende mannen 93% respectievelijk 65% bedraagt.

In zijn rapportage na het eerste tijdsbestedingsonderzoek gebruikte Knulst een aan de antropoloog Linton (1937) ontleend begrippenpaar om weer te geven of een bezigheid specifiek tot een welbepaalde sociale rol behoort dan wel of die bezigheid door vrijwel iedereen ondernomen wordt. Het eerste type bezigheid kan als een ‘specialty’ worden gekenschetst, het tweede als een ‘universal’ (Knulst, 1977).
Uiteraard bestaat er geen gouden standaard om te beslissen of een bezigheid tot de ene dan wel tot de andere categorie behoort. Wel kan met dit begrippenpaar in het achterhoofd nog eens met andere ogen naar de tabellen 8 en 9 gekeken worden. Wat dan in het oog springt is dat winkelen en boodschappen doen universeler geworden zijn. De percentages vrouwen die winkelen en dagelijkse boodschappen doen zijn onveranderd hoog, maar die activiteiten worden steeds vaker ook door mannen ondernomen.

Tabel 9. Winkelen en dagelijkse boodschappen (het aantal uren dat men daar per doorsnee werkweek aan besteedt respectievelijk het percentage van de bevolking dat dat minstens 1 kwartier per doorsnee werkweek doet) naar huishoudensvorm en positie in het huishouden, 1995, bevolking van 12 jaar en ouder.

	winkelen totaal		dagelijkse boodschappen	
	uren per week	percentage	uren per week	percentage
gehele bevolking	2,8	90	1,3	76
thuiswonende jongere	1,6	71	0,5	43
alleenwonende jongere	2,5	97	1,2	93
alleenverdienende man met partner	1,6	79	0,6	50
tweeverdienende man	2,0	86	0,8	65
tweeverdienende vrouw	3,7	99	1,9	95
huisvrouw met partner	4,2	100	2,2	98
65+	3,3	96	1,8	89
overig	2,9	93	1,4	81

Bron: SCP (Tijdsbestedingsonderzoek 1995).

Overigens heeft de verandering van kostwinners- naar tweeverdienershuishoudens geen noemenswaardige consequenties voor de totale hoeveelheid tijd die binnen een huishouden aan winkelen en boodschappen doen wordt besteed. Sommatie over de partners in die huishoudens leidt per saldo tot een nagenoeg gelijk aantal winkeluren in beide huishoudensvormen (tabel 10).

Tabel 10. Winkelen en dagelijkse boodschappen van paren in kostwinners- en tweeverdienershuishoudens (in uren per doorsnee werkweek), 1995, samenwonenden van 18 tot 65 jaar.

	winkelen totaal	dagelijkse boodschappen
kostwinnerspaar	5,8	2,8
tweeverdienerspaar	5,7	2,7

Bron: SCP (Tijdsbestedingsonderzoek 1995).

Voorzover men winkelen tot de vrije tijd mag rekenen, duidt dit er op dat winkelen als vrijetijdspassering niet beïnvloed wordt door een grotere tijdsdruk, maar kennelijk een hoog gewaardeerd onderdeel van de vrijetijdsbesteding vormt, waarop men niet snel bezuinigt. Voorzover winkelen tot de verplichtingen behoort, duidt dit er op dat de verminderde taakspecialisatie in tweeverdienershuishoudens kennelijk niet tot inefficiëntie leidt.

Eerder werd uitvoerig stilgestaan bij de vrijetijdsbesteding van ouderen, die de meeste vrije tijd hebben. Hun vrijetijdsbesteding werd uithuiziger en diverser. Ouderen zijn ook trouwe bezoekers van winkels, zowel in termen van het aantal uren dat ze daar besteden als in termen van het percentage deelnemers aan deze activiteit (tabel 11). Het

gegeven dat het vrijetijdsgedrag van ouderen meer op dat van jongeren is gaan lijken, impliceert echter dat ook voor hen winkelen maar één optie uit een breder repertoire van (vrije)tijdsbestedingen is.

Tabel 11. Winkelen (het aantal uren dat men daar per doorsnee werkweek aan besteedt respectievelijk het percentage van de bevolking dat dat minstens 1 kwartier per doorsnee werkweek doet) naar leeftijd, 1975-1995, bevolking van 12 jaar en ouder.

	uren per week			percentage deelnemers		
	1975	1985	1995	1975	1985	1995
12 tot 20 jaar	1,4	1,4	1,3	70	74	66
20 tot 35 jaar	2,7	2,7	2,7	90	91	92
35 tot 50 jaar	2,9	3,0	3,0	85	90	93
50 tot 65 jaar	3,2	2,9	3,2	91	90	93
65 jaar en ouder	3,6	3,0	3,3	91	91	96

Bron: SCP (Tijdsbestedingsonderzoek 1975-1995).

Ter afronding van deze verkenning van winkelen wordt nog de spreiding ervan over de week belicht. Door de jaren heen is de spreiding van het winkelen over de dagen van de week nauwelijks veranderd, het beeld van 1975 wijkt nauwelijks af van dat in 1995.
Wel tekent zich het patroon af van een bezigheid die in de loop van de week geïntensiveerd wordt. Op zondag wordt slechts weinig gewinkeld, vanaf maandag groeit de omvang van het winkelend publiek met de dag, met vrijdag en vooral zaterdag als hoogtepunten (tabel 12).
Overigens geldt dit patroon in veel mindere mate voor het doen van de dagelijkse boodschappen, dat veel gelijkmatiger over de dagen van de week gespreid is. Dit duidt er op dat de dagelijkse boodschappen een verplichtender karakter hebben dan andere vormen van winkelen.

Tabel 12. Winkelen en dagelijkse boodschappen per dag van een doorsnee werkweek (het aantal uren dat men daar per dag aan besteedt respectievelijk het percentage van de bevolking dat dat minstens 1 kwartier per dag doet), 1995, bevolking van 12 jaar en ouder.

	winkelen totaal		dagelijkse boodschappen	
	uren per week	percentage	uren per week	percentage
zondag	0,1	9	0,0	1
maandag	0,3	38	0,2	29
dinsdag	0,3	40	0,2	28
woensdag	0,4	42	0,2	29
donderdag	0,4	45	0,2	31
vrijdag	0,5	53	0,3	38
zaterdag	0,8	58	0,3	39

Bron: SCP (Tijdsbestedingsonderzoek 1995).

Hierbij zij aangetekend dat de meest recente peiling nog van voor de liberalisering van de winkelopenstellingstijden dateert. In hoeverre de nieuwe Winkeltijdenwet een verandering in dit patroon tot gevolg heeft gehad, kan op basis van deze informatie niet worden vastgesteld.
De blijvende differentiatie tussen de verschillende dagen van de week suggereert dat het met de komst van de veelbesproken '7-dagen-per-week, 24-uur-per-dag' samenleving zo'n vaart niet loopt (zie voor vergelijkbare relativeringen van de 24-uurs economie met betrekking tot de arbeidsmarkt: Breedveld, 1998 en SCP, 1998).

4. WINKELEN EN ANDERE VORMEN VAN (VRIJE)TIJDSBESTEDING

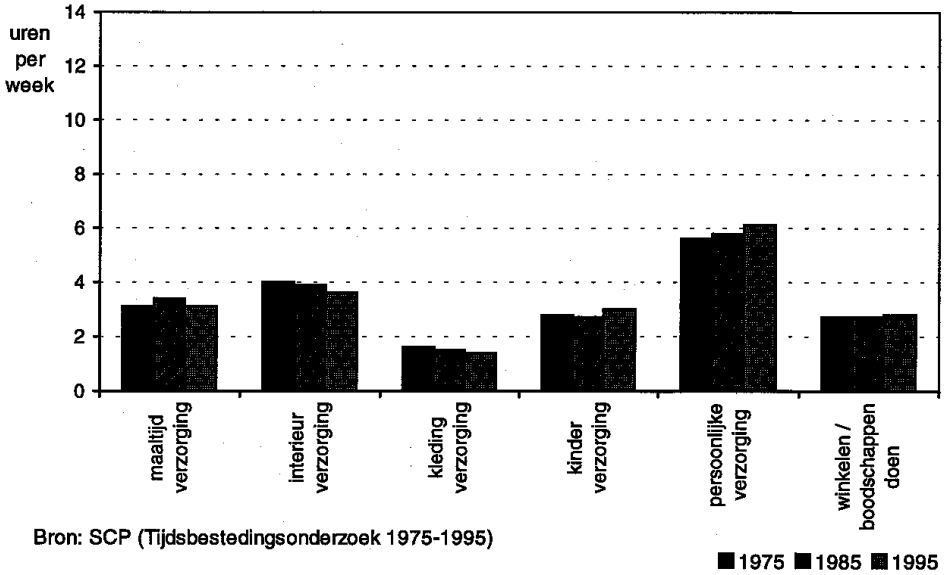
Tot slot van deze inventarisatie van winkelen als tijdsbesteding volgt een genuanceerder plaatsbepaling van het tijdsbeslag van winkelen temidden van andere vormen van (vrije)tijdsbesteding dan de eerdere vergelijking met slapen, verplichtingen en vrije tijd in figuur 1 (zie pag. 98).

Temidden van enkele andere alledaagse verplichtingen is het gemiddelde tijdsbeslag van winkelen van dezelfde orde van grootte als het gemiddelde tijdsbeslag van maaltijdverzorging, interieurverzorging, kledingverzorging en kinderverzorging. Voor persoonlijke verzorging wordt meer tijd ingeruimd (figuur 2, pagina 109).
Belangrijker dan het niveau van het tijdsbeslag is echter de ontwikkeling daarin. Door de jaren heen is het tijdsbeslag van winkelen aanzienlijk constanter van het tijdsbeslag van genoemde zorgtaken. Waar huishoudelijke hygiëne dankzij verbeterde technische toerusting minder tijd vroeg en de persoonlijke hygiëne meer aandacht kreeg, was winkelen een toonbeeld van stabiliteit.

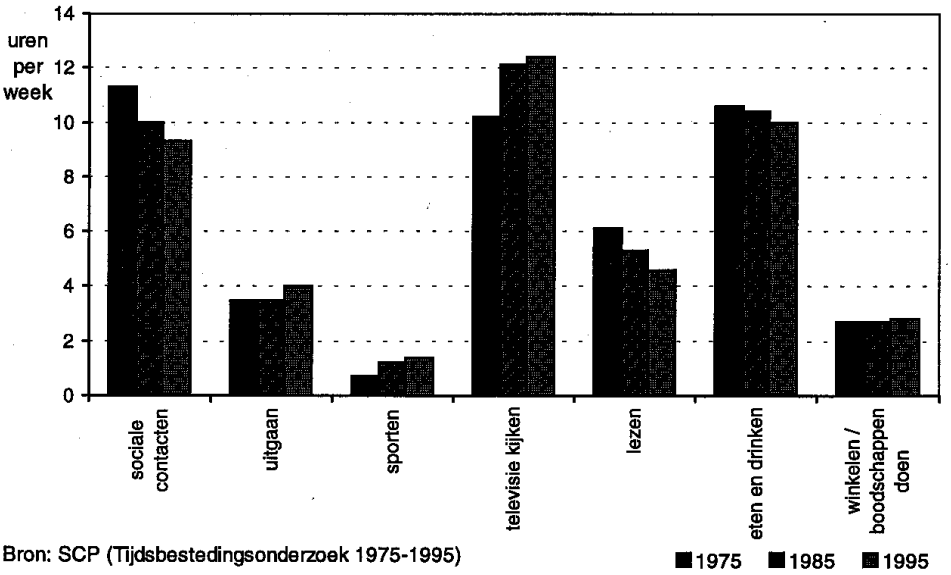
Temidden van enkele vrijetijdsactiviteiten dringt het beeld van winkelen als een stabiele factor in een veranderend patroon van tijdsbesteding zich nog sterker op. Waar sociale contacten, tafelen en lezen terrein verloren, terwijl sporten, televisie kijken – en recent ook uitgaan – terrein wonnen, was winkelen qua gemiddeld tijdsbeslag een toonbeeld van stabiliteit (figuur 3).

5. SLOTBESCHOUWING

De hoeveelheid tijd die Nederlanders in een doorsnee werkweek besteden aan winkelen en aan de dagelijkse boodschappen was in de periode 1975-1995 opmerkelijk stabiel. Terwijl menig (vrije)tijdsbesteding duidelijk in de lift zat (vb. sporten, televisie kijken en persoonlijke verzorging) of juist een sterke neergang vertoonde (vb. sociale contacten en lezen), veranderde het tijdsbeslag van winkelen en van dagelijkse boodschappen niet of nauwelijks.
Onder de oppervlakte is wel degelijk van beweging sprake. Meer mensen, meer mannen vooral, zijn tot het regelmatig winkelend publiek gaan behoren. Daarmee loopt winkelen in de pas met een ontwikkeling die kenmerkend is voor de (vrije)tijdsbesteding in het algemeen: velen doen veel dingen een beetje, waardoor een 'parade der passanten' op gang komt.



Figuur 2. Tijdsbeslag van enkele alledaagse verplichtingen, in uren per doorsnee week, 1975-1995, bevolking van 12 jaar en ouder.



Figuur 3. Tijdsbeslag van enkele vormen van vrijetijdsbesteding, in uren per doorsnee werkweek, 1975-1995, bevolking van 12 jaar en ouder.

Hierbij spelen twee verschillende processen een rol, één in het domein van de verplichtingen en één in het domein van de vrijetijdsbesteding. Onder impuls van veranderende samenlevingsvormen vindt een herverdeling van taken over de seksen plaats, het duidelijkst te herkennen in de opkomst van het tweeverdienershuishouden. Vrouwen gingen vaker betaald werk verrichten, schoorvoetend gevolgd door een grotere bijdrage van mannen aan huishoudelijke taken, waaronder boodschappen doen.
In de sfeer van de vrijetijdsbesteding lijkt van een zekere vervluchtiging sprake. De aandacht moet over tal van vrijetijdsactiviteiten worden verdeeld.
Deze twee processen, hoe verschillend ook, sorteren hetzelfde effect: voor een groeiend aantal mensen is winkelen één van de in aantal groeiende bezigheden.

NOOT

Dit is een bewerking van een presentatie op het congres *Consument en Detailhandel 1998* van het Hoofdbedrijfschap Detailhandel, Houten, 26 maart 1998.

LITERATUUR

Bocock, R. 1993. *Consumption*. London: Routledge.
Bourdieu, P. 1984 (oorspronkelijke Franse uitgave: 1979). *Distinction*. London: Routledge.
Breedveld, K. 1998. Illusies van een 24-uurs economie. In: *Tijdschrift voor Arbeidsvraagstukken* 14/1 (23-36).
Campbell, C. 1995. The sociology of consumption. In: D. Miller (red.). *Acknowledging consumption*. London: Routledge.
Falk, P. & Campbell, C. (red.). 1997. *The shopping experience*. London: Sage.
Featherstone, M. 1991. *Consumer culture and postmodernism*. London: Sage.
Ganzeboom, H. 1988. *Leefstijlen in Nederland*. Rijswijk: SCP; Alpen a/d Rijn: Samson.
De Haan, J. 1997. *Het gedeelde erfgoed*. Rijswijk: SCP; Den Haag: VUGA.
Knulst, W. 1977. *Een week tijd*. Den Haag: Staatsuitgeverij.
Knulst, W. & Van Beek, P. 1990. *Tijd komt met de jaren*. Rijswijk: SCP; Den Haag: VUGA.
Leunissen, P. & Prevo, O. 1998. *Consument en detailhandel 1998. Gevolgen van veranderende tijdsbesteding voor de detailhandel*. Zoetermeer: EIM.
Linton, R. 1937. *The study of men*. New York: Appleton.
Maffesoli, M. 1996. *The time of the tribes*. London: Sage.
Miller, D. 1995. *Acknowledging consumption*. London: Routledge.
Miller, D. 1997. Could shopping ever matter? In: P. Falk & C. Campbell (red.). *The shopping experience*. London: Sage.
SCP (Sociaal en Cultureel Planbureau). 1998. *Sociaal en Cultureel Rapport 1998*. Rijswijk: SCP; Den Haag: VUGA.
Schulze, G. 1992. *Die Erlebnisgeschellschaft: Kultuursoziologie der Gegenwart*. Frankfurt am Main: Campus Verlag.
Shields, R. 1992. *Lifestyle shopping. The subject of consumption*. London: Routledge.
Veblen, T. 1979 (oorspronkelijke uitgave: 1899). *The theory of the leisure class*. New York: Penguin Books.
Weber, M. 1987 (oorspronkelijke Duitse uitgave: 1922). *Over klassen, standen en partijen*. Kampen: Kok Agora.

9. Bellen & gebeld worden: Fabels en feiten

C.VAN ROOY en J.C.VAN STEENIS

SAMENVATTING

Parallel met de ontwikkeling van de telefoon als commercieel fenomeen zwellen ook de kritische geluiden aan. Telefoonmoeheid en belirritatie zijn onderwerpen die de gemoederen meer en meer bezighouden. Wat opvalt is dat de discussies hierover, op verjaardagsfeestjes bijvoorbeeld en in de publieksmedia, veelal gevoerd worden op basis van vage noties en vooroordelen. Hier wreekt zich het ontbreken van adequaat cijfermateriaal. Echter, die situatie behoort inmiddels tot het verleden. In opdracht van de DMSA, De Nederlandse Associatie voor Direct Marketing, Distance Selling en Sales Promotion, en het Voorlichtingscentrum Uw Mening Telt heeft Veldkamp/ Marktonderzoek de problematiek in kaart gebracht.
Uit het onderzoek komt naar voren dat de telefoondruk in de volksmond behoorlijk wordt overschat. Van alle bij huishoudens inkomende telefoontjes heeft 1,5% betrekking op een benadering voor marktonderzoek of telebusiness. Afgezet tegen alle uitgaande telefoontjes is het aandeel voor telebusiness (marktonderzoek is hier niet van toepassing) 2%. Een ander opmerkelijk resultaat is dat consumenten gemiddeld nauwelijks vaker benaderd worden voor telebusiness dan voor marktonderzoek. Daarmee is de aanklacht van onderzoekend Nederland dat een gemiddeld huishouden wordt 'platgebeld' voor telemarketing wel ontkracht. Wel is het zo dat vooral het welstandiger deel van de Nederlandse bevolking relatief vaak wordt benaderd voor telefonische verkoop.

1. INLEIDING

Ondanks de elkaar in rap tempo opvolgende ontwikkelingen op het gebied van nieuwe media, en ondanks voortschrijdende doelgroepsegmentatie bij de massamedia, lijkt aan het eind van de twintigste eeuw de aloude telefoon hèt marketinginstrument bij uitstek. Illustratief is de stormachtige ontwikkeling die de telebusiness industrie doormaakt. De laatste vijf jaar groeide de markt ieder jaar met zo'n 30%, wat heeft geleid tot een sector waarin nu zo'n honderdduizend mensen emplooi vinden. Maar ook binnen het marktonderzoek heeft *information mining* via de telefoon een hoge vlucht genomen. De VMO-statistieken tonen aan dat de groei van de bestedingen aan telefonisch onderzoek zich onverminderd voortzet ten koste van face to face onderzoek. Met een aandeel van 26% voor telefonisch onderzoek onderscheidt Nederland zich op dit punt nadrukkelijk van andere landen in West-Europa, de Scandinavische landen (geringe bevolkingsdichtheid) waar veel minder onderzoek via de telefoon wordt uitgevoerd, uitgezonderd.
Al blijven veel marktonderzoekers zich er hardnekkig en soms zelfs dogmatisch tegen verzetten, marktonderzoek en telebusiness vloeien steeds meer in elkaar over. Kernverschil is dat het bij marktonderzoek alleen om iemands mening gaat en bij telebusiness veelal om verkoop. Bij het eerste is de anonimiteit volledig gegarandeerd, bij het twee-

de niet. Dit alles neemt niet weg dat het onderscheid tussen verkopen en enquêteren vervaagt, al was het alleen maar omdat de consument beide fenomenen over één kam scheert. Onderzoek heeft uitgewezen dat menig consument een interview niet van een verkoopgesprek kan onderscheiden. Dat ligt uiteraard niet zozeer aan de intelligentie van de consument, als wel aan de wijze waarop het gesprek wordt geïntroduceerd.

Lange tijd werd door de VMO nogal geforceerd vastgehouden aan het uitgangspunt dat het gaat over twee wezenlijk verschillende branches. Het liefste wilde men de telemarketingbureaus dwingen om zich niet met marktonderzoek in te laten. De onderzoekswereld weet de problematiek van een teruglopende respons op telefonisch uitgevoerd onderzoek voor een belangrijk deel aan de activiteiten van de telemarketeers, terwijl omgekeerd de telemarketeers de onderzoekers betichtten van onvoldoende professionalisme. Maar deze kunstmatig in stand gehouden waterscheiding loopt op zijn einde. Terecht wierp Peter van Westendorp in zijn Dossier Marktonderzoek de vraag op of een telefonisch verkooppraatje van twee minuten echt zo veel storender is dan een vragenlijst van een half uur of nog meer.
Bemoedigend is dat men de laatste jaren meer begrip heeft gekregen voor elkaars standpunten en, mede onder druk van externe omstandigheden (regelgeving vanuit Brussel, privacy-aspecten), met elkaar in gesprek is geraakt. En dat is des te gewichtiger nu Brussel, en in het verlengde daarvan de Europese lidstaten, er steeds minder toe neigen om in de regelgeving onderscheid te maken tussen marktonderzoek en direct marketing, i.c. telebusiness.
Hoe men het ook wendt of keert, een bindend element is dat beide sectoren via de telefoon informatie verzamelen die uiteindelijk bij dezelfde bron (de consument) vandaan moet worden gehaald. Het is mede vanuit die optiek dat de DMSA (de Nederlandse Associatie voor Direct Marketing, Distance Selling en Sales Promotion) en het Voorlichtingscentrum Uw Mening Telt (destijds opgericht om het draagvlak voor markt- en opinieonderzoek te verbreden) daadwerkelijk de handen ineen hebben geslagen. Op gezamenlijk initiatief is een onderzoek uitgevoerd naar telefoondruk als gevolg van marktonderzoek en telebusiness activiteiten.

2. ONDERZOEKSVERANTWOORDING

De doelstelling van het onderzoek was vast te stellen hoe vaak de consument telefonisch wordt benaderd in het kader van marketingactiviteiten, zoals marktonderzoek of telemarketing, danwel zelf gebruik maakt van bepaalde telebusiness activiteiten. Telebusiness is daarbij gedefinieerd als het systematisch en planmatig gebruik van de telefoon voor marketingdoeleinden.

Omdat verondersteld werd dat een retrospectieve vaststelling van telefoondruk tot vertekening zou leiden, is gebruik gemaakt van een dagboekmethode. Gedurende een vooraf vastgestelde periode hebben respondenten hun belgedrag bijgehouden. Uit oogpunt van efficiëntie is gebruik gemaakt van een bestaand panel, te weten het Telepanel van het Nipo. Het onderzoek kende, om toevalligheden te vermijden, twee registratiegolven: een in het voorjaar van 1997 (mei/juni) en een in het najaar van 1997 (september/oktober). De uitvoering van het onderzoek lag in handen van Veldkamp/Marktonderzoek.

De doelgroep van het onderzoek wordt gevormd door alle huishoudens in Nederland, en daarbinnen alle personen van 18 jaar of ouder. In totaal zijn, verspreid over 4.154 meetweken, meer dan 55.000 telefoontjes geregistreerd.

Bij de registratie is onderscheid gemaakt tussen twee soorten telefoontjes, te weten:
* telefoontjes met een sociaal karakter,
* alle andere telefoontjes.

Onder telefoontjes met een sociaal karakter verstaan we telefoontjes met familie, vrienden en kennissen. Van deze telefoontjes is alleen gevraagd om bij te houden hoe vaak ze plaatsvinden. Alle andere telefoontjes zijn uitgebreid beschreven: wanneer, hoe laat, hoe lang, met wie, waarom, waarover, met welk effect.
Door de respondenten zelf is aangegeven of het een uitgaand of inkomend telefoontje betrof. In het laatste geval is tevens geïndiceerd of men om een mening werd gevraagd (marktonderzoek) of dat het om iets anders ging. Van die laatste cases is door de onderzoekers vastgesteld of het wel of niet om telebusiness gaat, evenals van alle uitgaande telefoontjes.
In vakjargon heet het opbellen van doelgroepen *outbound* telebusiness. Voor de consument is dit een inkomend gesprek. Vanuit consumenten perspectief uitgaande commerciële telefoontjes worden wel aangeduid met de term *inbound* telebusiness. Voor marktonderzoek is uiteraard alleen interessant hoe vaak men gebeld wordt.

3. DE COMMERCIËLE TELEFOONDRUK

Op jaarbasis wordt een Nederlander van 18 jaar of ouder gemiddeld 364,8 keer gebeld. Het overgrote merendeel van deze telefoontjes (95%) heeft een sociaal karakter. Jaarlijks wordt men 2,2 keer gebeld voor medewerking aan een marktonderzoek en 3,1 keer voor telebusiness. Van de 332,9 telefoontjes die men gemiddeld zelf per jaar pleegt, hebben er 6,7 betrekking op telebusiness faciliteiten waarvan de beller gebruik wil maken.

Tabel 1. De commerciële telefoondruk.

	per persoon (18-plus) per jaar	
	inkomend	uitgaand
	gem.	gem.
• totaal	364,8	332,9
• sociaal	345,3	298,5
• niet sociaal	19,5	34,4
marktonderzoek	2,2	nvt
telebusiness	3,1	6,7
niet sociaal, anders	14,2	27,6

Bij deze cijfers moet in de eerste plaats benadrukt worden dat het gaat om bruto benaderingen. Vooral bij marktonderzoek geldt dat lang niet alle benaderingen kunnen wor-

den omgezet in een geslaagd vraaggesprek. In de tweede plaats dient vermeld te worden dat weergave geschiedt op persoonsniveau, terwijl vooral telebusiness plaatsvindt op huishoudensniveau.
Verder is geen rekening gehouden met enkele factoren die een dempende werking hebben op het volume, zoals:
• telefoonbezit – niet iedereen beschikt over telefoon thuis,
• geheime telefoonnummers – men is in principe onbereikbaar,
• bezit antwoordapparaat – men is selectief bereikbaar.
Ook afwezigheid door vakantie is buiten beschouwing gebleven. De voornoemde cijfers moeten derhalve geïnterpreteerd worden als een bovengrens.

Er zijn geen respondentkenmerken die heel sterk samenhangen met enquêtedruk. Dat lijkt ook plausibel, omdat het streven bij markt- en opinieonderzoek er veelal op is gericht om een representatief beeld te krijgen van de bevolking. Er zijn wel enkele zwakke invloeden, maar die zijn veelal terug te voeren op het aspect van de mobiliteit.
Bij telebusiness zien we op diverse achtergrondkenmerken wel een sterke samenhang met telefoondruk. De hoogst opgeleiden, bijvoorbeeld, krijgen twee keer zo veel verkooptelefoontjes als de laagst opgeleiden, en vrouwen worden vaker gebeld dan mannen. Ook leeftijd speelt een rol; vooral de categorie van 25 tot en met 34 jaar vormt een geliefde telefoondoelgroep.
Een vergelijkbare differentiatie doet zich voor bij telebusiness op initiatief van de consument zelf (inbound): hoger opgeleiden en welstandigen nemen het voortouw en bellen vaker, hetgeen naar leeftijd uitgewerkt ook geldt voor de categorie van 25 tot en met 44 jaar.
Voor beide vormen van telebusiness zien we een relatie met omvang van het huishouden, hetgeen ons sterkt in de veronderstelling dat telebusiness zich, anders dan marktonderzoek, vaak nadrukkelijker richt op huishoudens dan op personen.

4. TIMING

Voor marktonderzoekdoeleinden wordt het meest gebeld op maandag en woensdag; voor outbound telebusiness (= inkomend) ligt het accent vooral op woensdag en donderdag. Bij inbound telebusiness (= uitgaand) ligt het zwaartepunt vooral op maandag.
Met gemiddeld 6,1 minuten kost een benadering voor marktonderzoek aanzienlijk meer tijd dan een benadering voor telebusiness (gemiddeld 2,9 minuten). Gemiddeld worden 3,3 minuten besteed aan een uitgaand telebusiness telefoontje.
Bij marktonderzoek duurt 30% van de telefoontjes langer dan tien minuten, terwijl bij telebusiness 80% van de telefoontjes hooguit vier minuten duurt.
In vergelijking tot telebusiness wordt er voor marktonderzoek wat meer in de ochtend en later op de avond gebeld. Voor outbound telebusiness (= inkomend) ligt het spitsuur tussen 17.00 en 20.00 uur. Bij inbound telebusiness (= uitgaand) wordt het merendeel van de gesprekken overdag afgewikkeld; na 17.00 uur wordt betrekkelijk weinig gebeld.

Tabel 2. Tijdstip telefoontjes.

	markt-onderzoek	telebusiness	
	inkomend %	inkomend (outbound) %	uitgaand (inbound) %
• 09.00 - 12.00 uur	23	19	35
• 12.00 - 17.00 uur	26	27	38
• 17.00 - 20.00 uur	33	39	25
• 20.00 uur en later	17	13	

5. THEMATIEK

Marktonderzoek heeft vaak betrekking op media, dienstverlening, eet-/drink-/rookwaren en huishoudelijke producten. Inkomende telebusiness is vaak ten behoeve van de financiële sector of voor kranten. Als reden voor uitgaande telebusiness wordt heel vaak de informatiefunctie genoemd. Vooral financiële informatie die men verlangt vormt een stimulans.

Tabel 3. Onderwerpen telefoontjes.

marktonderzoek (inkomend)		telebusiness inkomend (outbound)		telebusiness uitgaand (inbound)	
• media	35%	• financiële diensten	41%	• financiële informatie	27%
• dienstverlening	33%	• abonnement kranten	22%	• aanschaf diensten/ producten	11%
• eet-/drink-/rook-waren	28%			• productinformatie	10%
• huishoudelijke producten	26%			• reizen, telefoon info	10%

Bijna de helft van de telefoonnummers die gedraaid wordt bij uitgaande telebusiness *calls* (46%), is voorzien van een speciale prefix. Daarvan kan worden afgeleid dat 17% van de nummers gratis kan worden gedraaid (0800) en 29% niet (0900, 0906, 0909). Bij ongeveer de helft van de gevallen komt de eerste verbinding tot stand met een spraakcomputer.

6. EPILOOG

Steeds meer mensen geven aan zich te ergeren aan de ‘telefoonterreur’ vanuit de telebusiness sector. Luider en luider weerspiegelen de publieksmedia een negatieve connotatie. Illustratief is dat het aantal mensen dat zich op de niet-meer-bellen lijst heeft laten zetten in korte tijd meer dan verdubbeld is. Maar niet alleen de telebusiness industrie

ondervindt hinder van een groeiend onbehagen. Ook in de marktonderzoekwereld stuit men op problemen en groeien de zorgen omtrent de problematiek van de non-respons.
Het aanpakken en zo mogelijk verminderen van telefoonmoeheid en irritatie bij de consument kan dan ook worden beschouwd als een van de grootste uitdagingen voor beide branches. Beide groeperingen hebben er baat bij om het vak te ontdoen van kwalificaties die het een slechte naam en faam bezorgen. Het is dan ook een goede zaak dat de beide sectoren, ondanks een zekere animositeit, meer oog hebben voor het gemeenschappelijk belang en de handen ineen hebben geslagen.
Met het hiervoor beschreven onderzoek zijn representatieve kengetallen beschikbaar gekomen, die kunnen dienen als weerwoord voor de veelal ongefundeerde bezwaren. Met de uitkomsten in de hand kan de aanzwellende kritiek vanuit de maatschappij op zijn minst worden genuanceerd. Tevens bieden de onderzoeksresultaten handvatten voor een pro-actieve opstelling (bijvoorbeeld inzake de argumentatie richting Brussel en 's Gravenhage) en kunnen deze worden benut als *tools* voor verbetering van de dienstverlening.
Zo het onderzoek in de toekomst herhaald zou worden, hetgeen in het voornemen ligt, ontstaat ten slotte de mogelijkheid tot het zichtbaar maken van trends.

NOOT

Het Veldkamp-rapport *Bellen & Gebeld Worden* kan worden besteld bij de DMSA, bij het Voorlichtingscentrum Uw Mening Telt of bij Veldkamp. De kosten bedragen hfl 295 voor leden resp. hfl 495 voor niet-leden (beide bedragen exclusief btw).

LITERATUUR

Kuin, H., 1998. Telebusiness blijft plaag voor onderzoek, *Adformatie*, 35, p.48.
Rooy, C. van, 1998. Marktonderzoek, Deventer: Kluwer Bedrijfsinformatie.
Rooy, C. van en D. Verzijden, 1998. Bellen & gebeld worden, Amsterdam: Veldkamp.
Westendorp, P. van, 1996/1997. Dossier marktonderzoek, *Tijdschrift voor Marketing*, div. nummers.

10. Marktonderzoeksmethoden t.b.v. het mededingingsrecht

A.J. OLIVIER

SAMENVATTING

In dit artikel wordt een overzicht gegeven van een aantal uitgangspunten bij het juridisch marktonderzoek en wordt tevens een overzicht gegeven van de beschikbare methoden en vraagstellingen.
Bij juridische procedures kunnen betrokkenen: de eisende, de gewraakte partij en de rechter geïnteresseerd zijn in een aantal zaken:

- Vaststellen van de mate van Inburgering
- Registreren van Verwarringsgevaar
- Registreren van Feitelijke Verwarring
- Registreren van de omvang van Misleidingen
- Registreren van Onderscheidend Vermogen

Het zou goed zijn indien betrokkenen in de toekomst minder op het eigen gevoel zou afgaan en meer gebruik zou maken van marktonderzoek onder consumenten.
Wat men in ieder geval niet moet doen is het marktonderzoek gebruiken als middel om de zaak naar eigen hand te zetten.De kennis bij rechter en advocatuur dient fors vergroot te worden, voordat het marktonderzoek de plaats in de rechtszaal krijgt die het feitelijk verdient

1. INLEIDING

Reeds een groot aantal jaren geleden kwam Peter van Westendorp tot de conclusie dat in de relatie tussen het mededingingsrecht en het marktonderzoek iets fundamenteel fout zat. Hij sprak in die jaren over “a troubled and ill defined relationship” en wij kunnen vaststellen dat het er hierna zeker niet beter op is geworden.

De relatie tussen het mededingingsrecht en het marktonderzoek kenmerkt zich vooral door snel uitgevoerde spoedonderzoeken ter voorbereiding van kort geding-procedures, slecht onderbouwde onderzoeksresultaten en in allerijl ingehuurde contra-experts die de vloer aanvegen met de validiteit van de onderzoeksresultaten zonder aan te geven welke onderzoeksmethode dan wel uitgevoerd had moeten worden.
Het is begrijpelijk dat rechters, die in een kort geding uitspraak moeten doen over een geschil uitermate terughoudend zijn met betrekking tot het gebruik van marktonderzoek-resultaten in hun oordeelsvorming.
Dit heeft overigens wel als consequentie dat marktonderzoek-resultaten in landen zoals

de Verenigde Staten en Duitsland, waar meer geformaliseerd met marktonderzoek wordt omgegaan, een veel belangrijker rol spelen dan in Nederland.

In principe zou men hier vrede mee kunnen hebben omdat marktonderzoek er niet bij gebaat is om voortdurend in diskrediet gebracht te worden door ondeskundige interpretatie van onderzoeksresultaten door buitenstaanders. Zeker indien deze niet op zoek zijn naar de waarheid, maar in de onderzoeksresultaten slechts gegevens zoeken om hun argumentatie te versterken en daarnaast niet aarzelen om onderzoek te bekritiseren als de onderzoeksuitkomsten ongewenste resultaten opleveren.

Een partij die tijdens een gerechtelijke procedure alles in het werk stelt om zijn vermeend gelijk aan te tonen kan nauwelijks iets verweten worden wanneer de integriteit van het marktonderzoek wordt geschonden. Wel kan worden geconcludeerd dat marktonderzoek pas een juiste rol in het mededingingsrecht kan opeisen wanneer veel duidelijker vastligt wat men wèl en wat men niet met marktonderzoek kan aantonen en hieraan heeft het in het verkleden ten ene male ontbroken.
In dit artikel geven wij een beknopt overzicht van een mogelijke standarisatie binnen de gehanteerde vraagstellingen in de overtuiging hiermee een bijdrage te leveren aan de structurering van het juridisch marktonderzoek.

2. DE PLAATS EN BETEKENIS VAN ATP-PROFIELEN EN CONTROLEGROEPEN

2.1. Het ATP-profiel

Alvorens in te gaan op marktonderzoekmethoden en vraagstellingen willen wij allereerst stilstaan bij het begrip ATP-Profiel. (Awareness-Trial-Purchase-Profiel).
Het ATP-profiel is de middels onderzoek bij het relevante publiek verkregen informatie met betrekking tot de voorkennis die aanwezig is omtrent tekens en merken. Het gaat hierbij niet alleen om feitelijke kennis, maar ook om gedrag zoals het verrichten van proefaankopen en het blijvend produktgebruik.

De volgende belangrijke dimensies vormen het ATP-Profiel:

a. Geholpen bekendheid
b. Spontane bekendheid
c. Produktgebruik

Er is een aantal deelterreinen binnen het juridisch marktonderzoek waarbij de registratie van het ATP-profiel nuttig dan wel noodzakelijk is. Wij denken hierbij aan:

a. Bepaling van de beschermingsomvang.
   Voor de bepaling van de mogelijkheid van verwarring of associatie, is het ATP-profiel van de respondent in zoverre van belang, dat daaruit kan blijken in welke mate hij het bewuste merk/teken kent. In dit opzicht is het ATP-beeld van de respondent voor het verwarrings- of associatiegevaar van belang omdat de rechtspraak van Het

Benelux Gerechtshof leert, dat de beschermingsomvang van een merk groter is naarmate de onderscheidende kracht van het teken groter is – andere tekens zullen daarmee eerder in verband gebracht worden – en de onderscheidende kracht van een teken op haar beurt weer afhankelijk is van de mate van haar bekendheid.

b. Registratie van de mate van inburgering.
Wanneer met behulp van de ATP-Profiel spontane produktbekendheden kunnen worden bepaald, kan worden vastgesteld in welke mate reeds beschikbare kennis van invloed is op een vermeende verwarring en verwarringsgevaar. Indien bijvoorbeeld de primaire associatie tussen produktkategorie en merk zonder het tonen van tekens reeds 20% is, kan (indien een verwarring van 50% gemeten wordt), worden vastgesteld, dat het de omvang van feitelijke verwarring op basis van de gewraakte tekens 30% is.

c. Bepaling van de representativiteit.
Natuurlijk dient een steekproef in afdoende mate representatief te zijn.
Wij willen hier echter waarschuwen voor het gevaar dat men zich uitsluitend laat leiden door de socio-economische representativiteit.
Feitelijk is representativiteit van het ATP-Profiel belangrijker omdat deze de associaties van respondenten, en daarmee de registratie van verwarring en verwarringsgevaar, het meest nadrukkelijk kunnen beïnvloeden.

Een rapportage, waarin een vergelijking is opgenomen van het ATP-profiel uit het onderzoek met het landelijke ATP-profiel, verdient verreweg de voorkeur. Marktpartijen beschikken meestal over deze gegevens op basis van landelijk representatief onderzoek en dienen deze ter beschikking te stellen gedurende een procedure teneinde de rechter te helpen bij het doen van een uitspraak en bij het interpreteren van het, ten behoeve van het onderzoek uitgevoerde, marktonderzoek.

d. Matchen van steekproeven.
Het gaat bij onderzoek ten behoeve van het mededingingsrecht niet alleen om de representativiteit van het onderzoek t.o.v. het landelijk universum.
Het is daarnaast van belang, daar waar een controlegroep wordt gehanteerd, vast te stellen dat de samenstelling van beide steekproeven vergelijkbaar is. Met name hier gaat het niet alleen niet om socio-economische kenmerken, maar veel meer om ATP-profiel kenmerken.

2.2. De funktie van een controlegroep

In het wetenschappelijk onderzoek maakt men gebruik van controlegroepen om de effekten van een aktie (= stimulus) zo zuiver mogelijk te registreren. Testopzetten waarbij de verwarring (= effekt) ten gevolge van het hanteren van een teken (= stimulus) gemeten moet worden, maken derhalve bij voorkeur gebruik van controlegroepen.

In een aantal gevallen zal men er, ook in het juridisch marktonderzoek , toe over moeten gaan om controlegroepen te hanteren. Dit is de geëigende methode om de validiteit van een onderzoeksopzet te waarborgen.
De controlegroep onderscheidt zich slechts in één opzicht van de testgroep en wel door

het feit dat men een bepaalde stimulus niet ontvangt. Het gaat er alleen om om vast te stellen of een verschijnsel in de testgroep significant vaker voorkomt dan in de controlegroep.

Dit betekent dan overigens wèl dat aan een tweetal basisvoorwaarden moet zijn voldaan:
1. De steekproef moet voldoende groot zijn om een uitspraak over de significantie te kunnen doen;
2. De controlegroep en de testgroepen moeten gematched zijn, d.w.z. dat alle belangrijke basiskenmerken, voor wat betreft het ATP-profiel, afdoende vergelijkbaar moeten zijn.

Indien de steekproeven voor wat betreft het ATP-profiel niet vergelijkbaar zijn, zullen de onderzoekresultaten moeten worden herwogen, en wel zodanig dat de groepen wel vergelijkbaar zijn voor wat betreft de kenmerken van het ATP-profiel.

Een eisend teken of senior-teken dient niet alleen de verwarring bij het gewraakte teken of junior-teken aan te tonen, maar dient dit ook te vergelijken met de (afwezige) verwarring bij het door henzelf gevoerde teken. Onderzoek waarbij beide tekens (het gewraakte teken en het eisende teken) niet met elkaar vergeleken kunnen worden, is onvolledig.

Hoewel er altijd uitzonderingen zijn die de regel bevestigen, kan in zijn algemeenheid gesteld worden dat marktaandelen en daarmee de penetratieniveau's en de spontane produktbekendheden aanmerkelijk kunnen verschillen per regio. Dit impliceert dat in ieder onderzoek dat gebruikt wordt voor juridische procedures, dus ook het onderzoek dat uitgevoerd wordt in winkelcentra, het ATP-profiel van een consument bepaald moet worden. Dit geldt niet alleen voor het feitelijke onderzoek, maar ook voor eventueel gehanteerde controlegroepen.

3. EEN OVERZICHT VAN DE MOGELIJKE TESTOPZETTEN

Veel onderzoeken binnen het merkenrecht hebben betrekking op het meten van verwarringsgevaar. Hiertoe kunnen een achttal verschillende methoden worden onderscheiden:

1. het tachistoscopisch onderzoek;
2. de product find time test;
3. de simulated shop test;
4. de afplaktest;
5. de modulaire test;
6. de associatietest;

In dit hoofdstuk willen wij deze methoden in het kort behandelen. Van iedere methode geven wij een korte beschrijving en geven tevens aan wanneer en waarom deze methode aanbevelingswaardig is.

3.1 Het tachistoscopisch onderzoek

Een tachistoscoop is een apparaat waarin een object aan de respondent kan worden getoond met de hulp van steeds langer durende exposities. Veelal wordt uitgegaan van 11 expositietijden. De kortste expositietijd bedraagt 0,01 seconde, waarbij vooraf is duidelijk gemaakt dat het gaat om een naam. De maximale expositietijd bedraagt meestal 2 seconden. D.m.v. dit apparaat kan worden nagegaan hoe snel een respondent het produkt kan identificeren en in welke mate hierbij verwarring optreedt.

Bij de interpretatie van de cijfers dient slechts in beperkte mate naar de absolute hoogte van de uitkomsten te worden gekeken. Het gaat hierbij immers vooral om de vergelijking tussen twee merken. Het eisende senior-teken/merk en het inbreuk makende junior-teken. Daarnaast dient er een vergelijking gemaakt te worden tussen degenen die een koopervaring met het bestaande merk gehad hebben.

De opdracht aan de respondent zal simpel en eenduidig moeten zijn: Kunt u voor mij zo goed mogelijk omschrijven wat u op dit moment ziet?
Iedere vorm van nadere concretisering in de vraagstelling brengt een subjectief manupuleerbaar element in en dient derhalve te worden vermeden.

Een tweede vraag die vast onderdeel moet zijn van de testprocedure is er vraag naar de motivering van de uitspraak: Kunt u mij zeggen op basis waarvan u dit zegt?
Bij deze vraag mogen geen voorcoderingen worden gebruikt en de antwoorden zullen op basis van protocol-analyse moeten worden gerubriceerd.
Het basismateriaal dient beschikbaar te blijven t.b.v. verificatie en eventueel aan de rechter of de tegenpartij ter beschikking te worden gesteld.

Het nut van tachistoscopisch onderzoek is vooral daarin gelegen, dat benaderd kan worden wat de indruk is die het getoonde teken op het publiek maakt wanneer men daar slechts vluchtig mee in aanraking komt. Omdat de grens tussen niet-herkenning, verwarring en herkenning uitermate sterk beïnvloed wordt door de vluchtigheid van de blik, worden kijktijden essentieel. Dit gegeven is zo gevoelig, dat er geen eenduidige vaststelling is van optimale kijktijd en dat er geen andere keuze resteert dan kijktijden verfijnd uit te splitsen.
Aldus kan worden vastgesteld, welke elementen van het teken/merk bij een nog steeds vluchtige, dan wel iets langere beschouwing, tot de aandacht van de respondent doordringen. Vervolgens kan worden vastgesteld, welke elementen van het teken aan de herkenning van het merk het meeste bijdragen.

Dit is juridisch van eminent belang omdat algemeen wordt aangenomen dat voor het verwarrings- of associatiegevaar de totaalindruk van het eisende teken/merk en de imitatie daarvan beslissend zijn, terwijl niet alle bestanddelen van het teken aan de vorming daarvan gelijkelijk bijdragen.
Sommige delen van het teken zullen een zwak onderscheidend vermogen hebben en sommige delen van het teken een sterk onderscheidend vermogen. Dit laatste is weer van belang voor het verwarrings- of associatiegevaar in die gevallen waarin de eisende of de gedaagde partij bepaalde elementen binnen tekens of merken als inbreuk bestempelt.

3.2. De product find time test

De Product Find Time test is grotendeels vergelijkbaar met de tachistoscoop test. Bij een product find time test worden zowel het benadeelde als het benadelende merk in hun conkurrerende omgeving geplaatst. Een respondent ontvangt een zoekopdracht en moet het eisende teken/merk aanwijzen waarbij vastgesteld wordt hoe lang men hierover doet en in welke mate verwarring tussen het eisende en het gewraakte merk ontstaat.

In de tachistoscoop test is de zoektijd een gegeven en de herkenning een uitkomst. Bij een product find time test is het zoeken de stimulus en de zoektijd een uitkomst. Ook bij deze test geldt dat de vergelijking tussen de experimenten onderling belangrijker is dan de feitelijke uitkomsten van de individuele experimenten. De product find time uitslagen van het benadelende merk zullen moeten worden vergeleken met de product find time test uitslagen van het benadeelde merk.

Er dient speciaal gekeken te worden naar het oordeel van degenen die feitelijke aankoop- ervaring hebben met het produkt. De tachistoscoop test en de product find time test meten in belangrijke mate hetzelfde, namelijk het verwarringsgevaar. De tachistoscoop test maximeert echter de kans op het meten van verwarring terwijl de product find time test de kans op het meten van verwarring juist minimaliseert. De keuze voor de test wordt veelal door deze overweging bepaald.

Wanneer de respondent de opdracht krijgt het eisend teken dan wel het gewraakte teken/merk aan te wijzen, dan zou een respondent, als hij de opdracht goed vervult, altijd aankomen met het produkt dat voldoet aan de opdracht. Aangenomen uiteraard dat de opdracht zo duidelijk geformuleerd is, dat de respondent het produkt in de vrijheid van getoonde produkten kan terugvinden en hij voldoende tijd krijgt en neemt om de opdracht uit te voeren.
Wanneer de respondent niettemin met het verkeerde produkt op de proppen komt, kan daaruit niet meer worden afgeleid, dan dat hij kennelijk het ene produkt voor het andere heeft aangezien en dat zich dus feitelijke verwarring heeft voorgedaan.

De vraag blijft echter bestaan in hoeverre dit resultaat afhankelijk is van of samenhangt met de tijd die de respondent gebruikt om te zoeken. Het is dus sterk de vraag of kortere of langere zoektijden indikaties vormen voor verwarring of verwarringsgevaar. Met andere woorden, hoe vergelijken wij de groep die na korte tijd met het verkeerde produkt aankomt met de groep die na extreem lange tijd met het goede produkt aankomt.

Hiermee wordt tevens een belangrijk bezwaar van de methode aangegeven. De onderzoeksmethode is zeer geschikt om na te gaan of een nieuw verpakkingsontwerp net zo opvallend is in het schap als het oude verpakkingsontwerp. Voor onderzoek in het kader van het mededingingsrecht is deze methode feitelijk niet geschikt.

3.3. De simulated shop test

Degenen die onderzoek wensen te doen naar feitelijke verwarring i.p.v. verwarringsgevaar of verwarringsgevaar willen ontkennen zullen de voorkeur geven aan de methode

die zo nauwkeurig mogelijk aansluit bij de praktijk. In een dergelijke situatie verdient de “simulated shop” methode de voorkeur.

Hierbij wordt gebruik gemaakt van een winkelsituatie alsmede een met de werkelijkheid overeenkomend schappenplan.
Het beste is om gebruik te maken van een echte winkel, maar ook het werken met nagebouwde schappen behoort tot de mogelijkheden.
De respondenten krijgen de opdracht een aantal produkten aan te kopen en een van de produkten is het benadeelde merk. Door na het experiment vast te stellen welke respondenten het benadelende merk i.p.v. het benadeelde merk gekocht te hebben, kan aangetoond worden in welke mate de verwarring tot het daadwerkelijk aankopen van een foutief produkt heeft geleid.

Deze methode minimaliseert de kans op het meten van verwarring. Omdat de respondent in de winkel alle tijd heeft om produkten op volstrekt natuurlijke wijze met elkaar te vergelijken, wordt de kans op verwarring immers geminimaliseerd. Hier staat tegenover dat indien in deze testopzet feitelijke verwarring gemeten wordt, er inderdaad sprake is van daadwerkelijke schade die toegebracht wordt aan het benadeelde merk. Er worden oneigenlijke aankopen geforceerd en het reeds op de markt aanwezige merk lijdt omzetverlies.

Deze methode is derhalve geschikt wanneer een partij wil aantonen dat van commerciële schade in de vorm van aandeelverlies bij de eigenaar van het reeds bestaande merk (Wij spreken hier van senior teken/merk) geen sprake is.

3.4. De afplaktest

Bij de afplaktest worden de merktekens van de verpakking verwijderd (afgeplakt). Nadat alle tekens die indikatie geven m.b.t. de herkomst van de verpakking zijn verwijderd, wordt aan de consument gevraagd desondanks een fabrikant/merk aan te geven. Voor deze methode pleit het feit, dat veel inbreuken op het merkenrecht juist erop gericht zijn het basisontwerp van het senior teken, alsmede de kleurstelling, op zodanige wijze te imiteren, dat het positieve image van het benadeelde merk ook toegedicht wordt aan het junior teken.
Hierbij gaat het niet alleen om de gehanteerde kleurstellingen maar ook om de opbouw en de indeling van het verpakkingsontwerp. Deze methode meet dus zo zuiver mogelijk of de fabrikant daarin, in de ogen van de consument, is geslaagd.

Tegenstanders van deze methode zijn van mening dat er te veel subjectieve elementen in de testopzet worden ingebracht. Immers, wat plakt men nu precies af en wat niet? Daarnaast stelt men dat de verpakking in deze vorm nooit in de winkel aangeboden zal worden en dat het toevoegen van een afwijkende merknaam de associaties die wellicht in een afplaksituatie zouden kunnen worden opgeroepen weer te niet doen.
Ook het feit dat uit jurisprudentie blijkt dat gehanteerde kleurstellingen in zijn algemeenheid niet beschermd kunnen worden op basis van het merkenrecht, heeft de populariteit van deze methode verder doen afbrokkelen.

Dit neemt niet weg dat, afgezien van de bovengenoemde bezwaren de onderzoeksaan-

pak nauw aansluit bij de wijze waarop er binnen het mededingingsrecht wordt gedacht. Des te meer reden om stil te staan bij de modulaire test die een aantal nadelen van de afplaktest niet bezit.

3.5. De modulaire test

Bij de modulaire test wordt de verpakking teruggebracht tot zijn modulaire componenten.
Zo kan de verpakking van een merkartikel teruggebracht worden tot de volgende basiscomponenten:

Douwe Egberts	*Heineken*
1. een rode achtergrond	1. de groene cirkel rondom
2. een witte topbalk	2. de zwarte horizontale balk
3. een rond front in het midden	3. de rode ster
4. lauwerkransen rondom het fond	4. premium quality opdruk
5. een Fries vrouwtje in het midden	5. het halsetiket

Middels onderzoek kan de vraag beantwoord worden op basis van welke onderdelen of combinatie van onderdelen van de Douwe Egberts verpakking, een consument dit merk herkent. Bij deze testopzet ontbreekt de subjectiviteit omdat niet één maar alle onderdelen van de verpakking in een geïsoleerde vorm beschikbaar komen.

Een belangrijk voordeel van deze methode is dat kan worden vastgesteld welke basiscomponent nu het meeste bijdraagt aan de identiteit van het merk, terwijl tevens kan worden vastgesteld welke kombinatie van komponenten het onderscheidend vermogen van het merk doet ontstaan. De kracht van de methode wordt gevormd door het feit, dat eisende en gedaagde partijen hun pleidooien voeren op basis van argumentaties die betrekking hebben op onderdelen van het merk of het teken.

Middels deze testmethode kan per onderdeel van het teken vastgesteld worden wat de realiteitsgehalte van de argumentatie is. Nadeel van deze methode wordt gekenmerkt door het feit dat er een groot aantal steekproeven nodig zijn omdat iedere kombinatie van komponenten onderdeel dient te zijn van een separate beoordeling.
Ook indien men d.m.v. een trial and error methode werkt zijn er aanzienlijke onderzoeksbudgetten mee gemoeid, die alleen haalbaar zijn voor de echt grote merken.

3.6. De associatietest

De methode van de associatietest wordt veelvuldig toegepast binnen het juridisch marktonderzoek. Meestal wordt er gevraagd naar de herkomst van een merk of een teken. Het produkt, het merk, het teken wordt, voor onbepaalde tijd, in zijn isolement getoond. Een belangrijk verschil t.o.v. de tachistoscoop test is dat de zoektijd/kijktijd/denktijd onbeperkt is en een belangrijk verschil met de product find time test is het feit dat er niet gezocht behoeft te worden.

De methode is met name geschikt om de bekendheid van commerciële tekens, verpakkingsvormen of slogans te onderzoeken. Op het moment dat de merknaam zelf op de

verpakking staat, worden de onderzoeksresultaten minder relevant omdat men uiteindelijk slechts onderzoekt of de respondent kan lezen.

De methode wordt wel gebruikt door een junior teken dat wil bewijzen dat er geen sprake is van feitelijke merkverwarring. Of er sprake is van verwarringsgevaar is niet aan te tonen met deze vorm van onderzoek.

Een belangrijk bezwaar van de methode wordt gevormd door het feit dat de vraagstelling een gokelement uitlokt. Men wordt gedwongen te associëren en noemt men dan het merk dat primair geassocieerd wordt met de desbetreffende produktkategorie of het merk dat men zelf meestal gebruikt.
De richting waarin men raadt, wordt bepaald door het kennisniveau en de productervaring van de respondent. Hiervoor kan alleen afdoende gekorrigeerd worden indien er met controlegroepen wordt gewerkt, waarbij de resultaten van het senior merk/teken vergeleken worden met de resultaten van het junior merk/teken

De volgende vraagformuleringen komen voor:

- Welke producent denkt u dat dit produkt maakt;
- Aan welke producent denkt u als u deze verpakking ziet;

Aan dit type vraagstelling kleeft een belangrijk nadeel. Indien deze vraagstelling aan de respondent wordt voorgelegd, moet men rekening houden met het feit dat beoordelaars van het van het teken voorziene produkt veelal niet weten wie de producent van dat produkt is. Ook wanneer dat teken een bekend teken of een bekende verpakking omvat.

Zelfs wanneer, zoals niet ongebruikelijk is, op een verpakking de naam van de producent voorkomt, hoeft die producent nog niet aan het publiek bekend te zijn. Vooral niet in die gevallen waarin de vermelding van de naam van de producent van het geheel der verpakking slechts een visueel onderdeel uitmaakt.
Men kan derhalve beter aan de respondent vragen, waaraan hij bij het zien van het produkt denkt en en er, met daarop gerichte maar niet leading questions, achter zien te komen of de respondent daarbij ook aan een ander teken denkt.
In alle gevallen is deze associatie methode minder zuiver dan de resultaten van tachistoscopisch onderzoek.

Het minimaliseren van de aangeboden stimuli en het maximeren van de vrijheidsgraden binnen het te kiezen antwoord, verhoogt de objectiviteit van de marktonderzoek resultaten als het gaat om onderzoek t.b.v. het mededingingsrecht.
Met name in de Amerikaanse literatuur is veel gepubliceerd over de interpretatie van de uitkomsten van de diverse assiociatietechnieken. Hoewel het Amerikaanse mededingingsrecht volstrekt afwijkt van het Nederlandse mededingingsrecht is het toch zinvol kennis te nemen van de vraagstellingen (protocollen) die aldaar worden gehanteerd.

4. DE PROTOCOLLEN BIJ ASSOCIATIETESTEN

4.1. Inleiding

In het voorafgaande werd besproken, wat de bezwaren zijn, die kleven aan het leggen van rechtstreekse associaties tussen tekens/merken enerzijds, en fabrikanten anderzijds. Wij moeten ons afvragen wat nu de juiste vraagstelling is.
Iedere ervaren marktonderzoeker weet dat kleine veranderingen in de vraagstelling kunnen leiden tot andere onderzoeksresultaten.Feitelijk dient iedere vraagstelling te zijn gevalideerd alvorens deze te gebruiken als onderdeel van een merkenrechtelijke procedure.

Overigens dient op voorhand te worden vastgesteld dat er niet één vraagstelling te formuleren is, die optimaal is voor alle situaties.
De juistheid van de vraagstelling is afhankelijk van datgene wat men wenst aan te tonen.
Toch blijkt het mogelijk een aantal standaardisaties te ontwikkelen dat in ieder geval afrekent met het probleem dat kleine veranderingen in de vraagstelling tot wezenlijk andere resultaten kunnen leiden.

Wij willen in dit hoofdstuk de belangrijkste protocollen behandelen en hiervan de voor- en nadelen aangeven.

Wij behandelen achtereenvolgens:

1. Het Exxon protocol;
2. het Eveready protocol;
3. het Squirt protocol;
4. het Zatarains protocol;
5. Simulated Choice;
6. het Teflon protocol.

4.2. Het Exxon protocol

Het Exxon protocol bestaat uit 4 vraagstellingen;

Toon junior merk;

1. Wat is het eerste dat u te binnen schiet wanneer u naar deze afbeelding kijkt?
2. Op basis waarvan komt u tot deze mening?

Toon senior merk;

3. wat is het eerste dat u te binnen schiet wanneer u kijkt?
4. op basis waarvan komt u tot deze mening?

Deze testmethode is bij uitstek geschikt om feitelijke verwarring aan te tonen. De test is gebaseerd op de aanname, dat indien het senior-merk genoemd wordt terwijl het junior-

merk getoond word, verwarring een aan zekerheid grenzende waarschijnlijkheid is.
Juist door te vragen wat het eerst te binnenschiet, zullen respondenten de neiging hebben om de meest gelijkende naam te noemen indien er sprake is van verwarring. Uit onderzoek blijkt ook dat merknamen steeds het eerst genoemd worden. Daarnaast kan worden vastgesteld op basis waarvan de verwarring nu in feite ontstaat in de ogen van de consument.
Het feit dat men het seniorteken noemt bij vraag 1, geeft aan dat men een onderscheid maakt tussen het junior- en het senior-teken en dat er sprake is van associatie.

Hier hoeft van feitelijke verwarring geen enkele sprake te zijn. Overigens wel van aantoonbare associatie. Zeker in situaties waar niet alleen namen maar ook tekens of kombinaties van tekens een hoge mate van parallelliteit vertonen.

4.3. Het Eveready protocol

Bij het Eveready protocol wordt geen feitelijke verwarring maar verwarringsgevaar gemeten.

Bij het Eveready protocol wordt het junior-teken/-merk getoond en wordt gevraagd:
1 Wie denkt u dat dit produkt op de markt brengt?
2 Waarom denkt u dat?
3 Kunt u ook andere produkten noemen die door dit bedrijf op de markt gebracht worden?

Deze vraagstelling is vooral geschikt voor het meten van verwarringsgevaar en niet van feitelijke verwarring. Een mogelijk bezwaar van deze methode is het feit dat respondenten die het senior-merk noemen als bron van herkomst van het junior-merk, deze associatie niet zouden leggen wanneer het om een normale aankoopbeslissing zou hebben gehandeld.
Daarnaast kan de vraag gesteld worden hoe zeker men is van de genoemde herkomst. Onderzoek heeft overigens wel aangetoond dat de genoemde gelijkenissen groter worden naarmate de tekens meer op elkaar lijken. Alles bijeengenomen kan de conclusie getrokken worden dat de betrouwbaarheid van het Eveready protocol voor wat betreft definiëring van het verwarringsgevaar afhankelijk is van de specifieke karakteristieken van de merken en de tekens en (iets eerder) de voorkennis van de consument.

4.4. Het Squirt protocol

Het grote voordeel van deze methode is, dat dit de meest direkte aanpak is voor wat betreft van het meten van verwarring.

Men stelt de respondent de volgende vraag:
Denkt u dat het junior-teken en het senior-teken door het zelfde bedrijf op de markt gebracht worden?

Deze methode zal leiden tot een onderschatting van het verwarringsgevaar omdat het vrij onlogisch is, dat een bedrijf twee vergelijkbare of een zo goed als vergelijkbare tekens of merken op de markt zal brengen.

Verder geldt een belangrijk aantal van de nadelen die beschreven zijn bij het Eveready protocol ook voor deze vraagstellingstechniek. Een verder nadeel van deze methode is het feit dat deze niet aansluit bij de wijze waarop de consument denkt en dat hiermee de bekende slapende honden worden wakker gemaakt.
Soms is de methode echter bij uitstek geschikt. Wij denken hierbij aan een situatie waar bijvoorbeeld New Spice als junior-teken t.o.v. Old Spice als senior-teken met elkaar een juridische procedure zouden voeren.

4.5. Het Zatarains protocol

Deze vraagstellingstechniek zoekt aansluiting bij de wijze waarop de consument denkt en het desbetreffende merk benoemt.

De volgende vragen worden aan de consument gesteld:

Heeft u in de afgelopen periode met enige regelmaat (van geval tot geval te definiëren) dit produkt gekocht en als u morgen dit produkt zou kopen (toon senior-verpakking) hoe zou u dit dan omschrijven?

Deze vraagstellingstechniek is bij uitstek geschikt wanneer junior-tekens inbreuk maken op de rechten van variëteit aanduidingen, bijv. het verkopen van pinda's onder de merknaam Poesta (Calve Borrelnoten); motorolie verkopen onder de merknaam Q.(Q8)

4.6. De Simulated Choice

In deze testopzet worden zowel junior- als senior-merken/tekens getoond en vervolgens wordt gevraagd aan welke van de twee merken/tekens men de voorkeur geeft. Het gaat hierbij altijd om een volledig omschreven produkt waarbij de technische specificaties bekend zijn. Daarnaast wordt een vergelijkbaar prijsniveau aan beide merken/tekens gekoppeld.
Deze methode wordt gebruikt als een uitbreiding t.o.v. de bestaande technieken waarbij verwarringsgevaar gemeten wordt in situaties waarin men de gevolgen van het verwarringsgevaar wenst te inventariseren. Deze methode is slechts gedurende een korte tijd populair geweest omdat er erg veel kritiek op deze methode werd geuit:

1. de produkten worden bijna nooit naast elkaar in zijn isolement in de winkel aangeboden;
2. het toevoegen van technische specificaties en het toevoegen van prijzen brengt een onnodig aantal beïnvloedende variabelen binnen in een testmethodologie die van nature al uitermate gevoelig is voor stimuli.

4.7. Het Teflon protocol

In een belangrijk aantal gerechtelijke procedures gaat het om de vraag of het merk een generiek karakter heeft of een specifieke merknaam is. In deze onderzoeksituatie zal men de ondervraagde eerst goed uit moeten leggen wat het verschil is tussen een merknaam en een generieke naam. Een goed voorbeeld waarin oefenvragen hun nut afwerpen.

Daarna krijgt men een groot aantal voorbeelden waarvan men moet aangeven of het hier om een merknaam of om een produktkategorie gaat. Dit moet men voor ongeveer tien situaties doen. Een van de tien situaties is de situatie waar het feitelijk om gaat.

Sommige van de genoemde voorbeelden zijn duidelijke voorbeelden van generieke namen en sommige zijn duidelijke voorbeelden van merknamen. Ook een aantal notoire twijfelgevallen (Maggi, Coca-Cola) kunnen worden ingebracht. Niet alleen om de zaak in perspectief te plaatsen maar ook om de validiteit van het onderzoek te testen, door na te gaan of de testresultaten hiervan in hoge mate vergelijkbaar zijn met reeds bekende gegevens.

LITERATUUR

Allen, M.J. (1993), *The role of actual confusion Evidence in Trademark Infringement Litigation. The Trademarkreporter,* Vol 83, May-June 1993

Blackwell., R.D. et al. (1970), *Laboratay Equipment for markety Research.,* Kendall/Hurt Publishing Company., Dubuque, IOWA.

Cohen, D. (1991), Trademark Stategy Revisted., *Journal of marketing,* July 1991. vol 55. 46-59.

Evens, E.,and Cunn (1991), TrademarkSurveys. *The Trademarkreporter,* Vol 79.

Hera, A. (1978), *Die identifications geschwindigkeit fur plakative Werbe darbietungen mit emotionalen Blichfangen.* Inaugural dissertetion, Saarbruchen.

Hossinger, H.P. (1968), *Pretest in der Marktforschung Konsum und Verhalten* Band 2 Physica Verslag Wien.

Sunonooy I. (1993), The effect of Survey Method on Likehood of Confusion. Estimates conceptual analyses and empucal Test. *The Trademarkreporter,* Vol 83, May- June 1993

Kaufman, P. (1982), *Het marktonderzoek is de praktijk van het mededingingsrecht en het merkenrecht.*

Noëlle Neuman E. (1961), *Umfrage Forschung in der Rechtspraxis.* Verslag Chemie GMBH, Weinheim.

Van Nieuwenhoven Helbach, E.A., (1983), Het opinieonderzoek als bewijsmiddel in het mededingingsrecht, in: *Een goede procesorde,* Kluwer 1983,.

Van Nieuwenhoven Helbach, RAO-document, unpublished paper.

Thomson,. D.L. (1965), Survey data as Evidence in Trademark Infringement cases., *Journal of marketing Research,* pp64-73.

V. Westendorp, P.H., *A troubled and ill defined relationship.* De rol van het marktonderzoek bij rechtszaken, unpublished paper.

V. Westendorp P.H. (1988),. De rol van het marktonderzoek bij mededingingsrechtelijke procedures. *Publikatie nr. 1 van de vereniging voor mededingingsrecht.*

V. Westendorp, P.H. (1993), Ontwerp van een beroepscode voor de uitvoering van juridisch marktonderzoek.

11. Wat is er aan de hand met ons gezin? Een exploratie van de Europese opvattingen over gezin, huwelijk, ouderschap, ouder-kind relaties en vrouwenrollen

L. HALMAN

SAMENVATTING

Recentelijk is het gezin weer op de politieke agenda geplaatst en onderwerp geworden van het publieke debat. De recente demografische ontwikkelingen lijken erop te wijzen dat het gezin niet langer de hoeksteen van de samenleving is. Individualisering en het behoud van de traditionele familiale solidariteit worden in het algemeen beschouwd als fenomenen die niet langer met elkaar kunnen samengaan.
In dit artikel wordt ingegaan op de vraag in hoeverre het gezin nog populair is in de hedendaagse geïndividualiseerde, gesecularisseerde en vooral ook geëmancipeerde samenleving. Onze empirische analyses laten geen andere conclusie toe dan dat het nog maar zeer de vraag is of het gezin eigenlijk wel van het toneel is weggeweest. Alle verhalen over modernisering en postmodernisering van de samenleving ten spijt blijken de opvattingen over het gezin, het huwelijk, ouder-kind relaties, alsook de mannen en vrouwen rollen nog immer overwegend traditioneel.

1. INLEIDING

Recentelijk is 'het gezin' opnieuw op de politieke agenda geplaatst. Met name CDA politici ijveren zich het belang van het gezin in de hedendaagse samenleving te benadrukken, maar ook binnen andere politieke stromingen en in de media staat het gezin weer volop in de belangstelling. Het gezin is 'weer onderwerp van serieus debat geworden' (Liefbroer, 1998: 20).

De sterk toegenomen individualisering zou er debet aan zijn dat het gezin niet langer als hoeksteen van de samenleving wordt beschouwd. Daarbij wordt dan vaak gewezen op de recente demografische ontwikkelingen die tot de conclusie zouden kunnen leiden dat het traditionele gezinstype op z'n retour is. Allerlei nieuwe varianten zijn opgekomen en in toenemende mate populair geworden. Het huwelijk zou aan betekenis hebben ingeboet hetgeen bijvoorbeeld zou kunnen worden opgemaakt uit de dalende huwelijkscijfers en het toegenomen aantal echtscheidingen. Gezinnen zijn kleiner geworden en de traditionele rolverdeling is verdwenen en heeft plaats moeten maken voor een meer evenwichtige taakverdeling tussen mannen en vrouwen. Volgens het conventionele patroon, zoals dat cultureel en ideologisch door de kerken werd uitgedra-

gen, diende het gezin de seksuele relaties, voortplanting en de opvoeding van kinderen te reguleren. Het huwelijk werd daarbij beschouwd als een onverbrekelijke band van 'women's mothering and men's authority' (Elliot, 1996: 6). Het klassieke gezinspatroon bestond uit een mannelijke hoofdkostwinner en een vrouw die de zorgtaken in het huishouden op zich nam.

Dit patroon zou ten gevolge van modernisering in het algemeen en individualisering, secularisatie, emancipatie en de toenemende participatie van vrouwen op de arbeidsmarkt in het bijzonder, naar de achtergrond zijn verdwenen. Met name de toetreding van vrouwen tot de arbeidsmarkt betekende dat zij niet langer meer alleen werkzaam zijn binnen de eigen huishouding. Deze kentering wordt in het algemeen beschouwd als de definitieve doorbraak van de traditionele vrouwenrol die voornamelijk bestond uit huishoudelijk werk en de zorg voor echtgenoot en kinderen, ouderen en zieken (Fox Harding, 1996: 2).

Hoewel voor veel vrouwen hun toetreding tot de arbeidsmarkt een economische noodzaak mag zijn geweest, bijvoorbeeld om het gezinsinkomen op een acceptabel niveau te krijgen, worden individualisering en emancipatie als belangrijke factoren genoemd die de veranderingen in gang hebben gezet of minstens hebben gestimuleerd. Arbeid betekent voor veel vrouwen onafhankelijkheid of in ieder geval de mogelijkheid daartoe (Scott, Braun & Alwin, 1993).

Met name deze veranderingen in vrouwenrollen zouden belangrijke verschuivingen in oriëntaties ten aanzien van arbeid en het gezin teweeg hebben gebracht, vooral onder vrouwen (Boh, 1989: 266). Maar niet alleen de opvattingen over de traditionele vrouwenrollen, ook de opvattingen over huwelijk en gezin en over relaties in het algemeen zouden erdoor zijn veranderd. Met name de voorstanders van wat wel de Tweede Demografische Transitie Theorie wordt genoemd wijzen erop dat de afname van het traditionele gezinspatroon een resultaat is van toegenomen individualisering en modernisering. De nadruk die moderne mensen leggen op persoonlijke ontwikkeling en eigen geluk is ten koste gegaan van de plichten die traditioneel verbonden zijn aan het gezin. De veranderingen zijn aangeduid in evolutionaire termen 'from altruism to individualism' (Van de Kaa, 1987: 5; Lesthaeghe & Van de Kaa, 1986), daarmee suggererend dat individualisering en het behoud van traditionele familiale solidariteit niet samen kunnen gaan. Immers, als gevolg van individualisering zal men steeds vaker voor zichzelf kiezen en voor persoonlijk gewin in plaats van voor het huwelijk en een traditioneel gezinsleven.

De vraag is nu of en in hoeverre dergelijke wijzigingen in opvattingen inderdaad zijn opgetreden. In hoeverre is het gezin nog populair in de moderne, geïndividualiseerde en gesecularisseerde samenleving? Is afscheid genomen van het traditionele denken over gezin, relaties en mannen- en vrouwenrollen?

In deze bijdrage worden een aantal opvattingen over gezin en gezinsleven gepresenteerd. Daarbij richten we ons op Europa omdat enkele interessante verschillen worden verwacht. Er bestaan immers markante verschillen tussen landen waar het gaat om de mate van modernisering, individualisering, emancipatie, religieuze tradities, het gevoerde gezinsbeleid, de beschikbaarheid en toegang tot kinderzorg (Kamerman, 1991),

en ook in de uiteindelijke arbeidsparticipatie van vrouwen (Boh, 1989; Elliot, 1996: 4). Tevens zullen enkele belangrijke verschillen bestaan tussen Oost- en Westeuropese landen. Veelal wordt geconstateerd dat processen van modernisering en individualisering minder ver zijn voortgeschreden in Oost- en Centraal-Europa dan in West- en vooral Noord-West-Europa (Chirot, 1989; Glenny 1990). Getracht wordt nu een antwoord te krijgen op de vraag in hoeverre landen waar secularisatie, individualisering, emancipatie, arbeidsparticipatie van vrouwen en democratische processen verder zijn voortgeschreden, minder traditioneel zijn waar het opvattingen over gezin, relaties en mannen- en vrouwenrollen betreft, dan landen die minder ver zijn voortgeschreden.

Onze bevindingen zijn gebaseerd op de onderzoeksgegevens van het Europese Waardenonderzoek dat in 1990 heeft plaatsgevonden (zie Ester et al, 1994). Uiteraard kan worden tegengeworpen dat deze data reeds gedateerd zijn en dat meer recente gegevens wellicht tot geheel andere conclusies aanleiding zouden kunnen geven. Dat valt natuurlijk niet te ontkennen, maar daar valt tegen in te brengen dat mensen in het algemeen niet snel geneigd zijn de fundamentele oriëntaties te wijzigen. Opinies kunnen sterk fluctueren mede onder invloed van de omstandigheden en recente gebeurtenissen, maar de dieper liggende waarden blijken over het algemeen redelijk stabiel te zijn, ook al wordt men ouder en veranderen de omstandigheden.

2. HET BELANG VAN HET GEZIN

Het gezin blijkt nog immer een belangrijk gegeven voor de meeste Europeanen. En dit is niet alleen het geval in Zuid- en Oost-Europa, ook in de Noordeuropese landen vindt een (over) grote meerderheid het gezin (nog immer) belangrijk. Dat kan niet alleen worden opgemaakt uit de antwoorden op de vraag hoe belangrijk het gezin is, ook de grote mate van instemming met de uitspraak dat een sterkere nadruk op het gezin in de toekomst een goede zaak is, bevestigt dit beeld.
Het gezin is dus bij lange na niet van het toneel verdwenen en blijft in alle landen een hoeksteen van de samenleving (zie ook Prinz, 1995: 2). Wellicht is dit resultaat niet zo vreemd, want zoals Sgritta (1989: 2) opmerkt, het gezin is de plaats bij uitstek om tegemoet te komen aan individuele behoeften. Uiteraard is hiermee nog niet gezegd dat bij het gezin alleen maar aan de traditionele vorm wordt gedacht. Het begrip 'gezin' wordt heden ten dage veel breder gedefinieerd dan enkele decennia geleden (zie van Nimwegen & Moors, 1998: 17). De beantwoording van de vragen kan dan ook betrekking hebben op een groot scala aan gezinstypen en samenlevingsvormen die allemaal met gezin worden aangeduid. Wat er wel mee lijkt te worden bevestigd is het idee dat het 'gezin' een blijvende waarde lijkt te vertegenwoordigen, niet alleen in de meer traditionele culturen, ook in de ver geïndividualiseerde samenlevingen is dat klaarblijkelijk het geval.

3. HUWELIJK EN ECHTSCHEIDING

Het huwelijk geldt niet langer als enige vorm van relatie tussen een man en een vrouw. Ook andere vormen van partnerschap zijn in toenemende mate geaccepteerd geraakt en gepraktiseerd. Het ongehuwd samenwonen is steeds populairder geworden, met name

Tabel 1. Percentages respondenten die van mening zijn dat het gezin zeer belangrijk is en meer nadruk op het gezin in de toekomst een goede zaak vinden.

Land	Gezin zeer belangrijk	Nadruk op gezin is goed
Frankrijk	82	90
Groot-Britannië	88	88
West-Duitsland	71	87
Oost-Duitsland	84	92
Italië	88	93
Spanje	84	88
Portugal	65	95
Nederland	80	67
België	84	85
Denemarken	88	95
Noorwegen	88	95
Zweden	87	85
Noord-Ierland	94	92
Ierland	91	94
IJsland	91	96
Polen	91	96
Hongarije	88	94
Tsjechië	86	94
Slowakije	88	90
Bulgarije	76	90
Oostenrijk	85	92
Estland	69	86
Letland	73	85
Litouwen	65	96
Roemenië	83	–

onder jongeren, maar ook onder hen die gescheiden zijn. Desondanks heeft het huwelijk nog niet afgedaan. Het huwelijk wordt nog immer gewaardeerd, en zeker niet beschouwd als gedateerd. Uiteraard geldt ook nu weer, net zoals dat het geval is bij het gezin, dat onduidelijk is wat de respondenten precies voor ogen hadden met het huwelijk toen zij de vraag kregen voorgelegd, maar het lijkt het meest waarschijnlijk dat gedacht werd aan de 'levenslange verbintenis' tussen man en vrouw. Dat kan bijvoorbeeld worden afgeleid uit de grote mate waarin Europeanen van mening zijn dat 'trouw' zeer belangrijk is voor het welslagen van een huwelijk, maar ook, en nog meer, uit het gegeven dat de meeste jonge mensen hopen te trouwen en daarin wederzijdse trouw maar ook een 'lifelong union' verwachten (Elliot, 1996: 17). Het blijkt ook uit de antwoorden die jong-volwassenen in een Australisch project hebben gegeven op de vraag wat het huwelijk voor hen betekende. Het meest genoemd zijn aspecten als samen delen, wederzijdse steun, eerlijkheid, vertrouwen, stabiliteit en zekerheid (Millward, 1987: 28).

In tegenstelling tot wat zo vaak wordt gedacht hebben individualisering, modernisering en de emancipatie van alternatieve vormen van samenleven dus niet geleid tot een afwijzing van het huwelijk. Ongehuwd samenwonen wordt vaak beschouwd als een

Tabel 2. Percentages respondenten die van mening zijn dat het huwelijk niet verouderd is, dat trouw belangrijk is in het huwelijk, en die echtscheiding aanvaardbaar vinden (gemiddelde score op 10 punt schaal: 1 = nooit te rechtvaardigen; 10 = altijd te rechtvaardigen).

Land	Huwelijk is niet verouderd	Trouw belangrijk in huwelijk	Acceptatie van echtscheiding
Frankrijk	71	74	5.6
Groot-Brittannië	82	90	5.3
West-Duitsland	85	78	5.7
Oost-Duitsland	85	78	5.1
Italië	87	85	5.2
Spanje	86	81	5.3
Portugal	77	78	5.4
Nederland	79	88	6.1
België	78	85	4.9
Denemarken	82	81	5.9
Noorwegen	90	91	5.3
Zweden	86	89	6.3
Noord-Ierland	86	94	4.1
Ierland	90	93	4.1
IJsland	94	90	6.2
Polen	94	83	4.2
Hongarije	89	84	5.0
Tsjechië	93	72	5.2
Slowakije	93	86	4.8
Bulgarije	89	63	4.5
Oostenrijk	88	83	4.9
Estland	89	61	5.4
Letland	91	74	5.7
Litouwen	91	70	4.0
Roemenië	91	76	4.9

bepaalde fase voorafgaand aan het huwelijk. Als zodanig wordt het dan niet opgevat als een echt alternatief voor de huwelijkse vorm van samenleven. Juist door het huwelijk zijn enkele zaken wettelijk geregeld waarvoor men bij samenwonen allerlei regelingen moet treffen, bijvoorbeeld om de nalatenschap te regelen. Met name wanneer er kinderen komen wordt dan ook vaak (alsnog) voor het huwelijk gekozen. Dit alles betekent dat het huwelijk dus nog lang niet uit de gratie is, al was het maar omdat het vaak het meest praktische is. Ook onder jongeren bestaat er een grote voorkeur voor. Een overgrote meerderheid van de jongeren blijkt een toekomst in gehuwde staat na te streven (Goldscheider & Waite, 1991). Uit de onderzoeksgegevens van het International Social Survey Project blijkt eveneens dat het huwelijk nog immer belangrijk is. Gevraagd wat men jongeren zou adviseren: alleen blijven, ongehuwd samenwonen of trouwen, blijkt een overgrote meerderheid van de ondervraagden ‘advice a young person ultimately to marry, with or without prior cohabitation’ (Scott, Braun & Alwin, 1993: 26).

Het huwelijk wordt dus nog steeds zeer gewaardeerd, zelfs in landen waar alternatieven in groten getale voorkomen. Maar tegelijkertijd kan worden geconstateerd dat echt-

scheiding niet wordt afgewezen. Vooral in Noord-West Europa blijkt echtscheiding zeer geaccepteerd. Dat is wat minder het geval in Zuid- en Oost-Europa, en het minst in Ierland. Mensen die een echtscheiding aangaan worden niet langer gestigmatiseerd in de hedendaagse geïndividualiseerde samenleving. Bovendien is echtscheiding niet langer meer illegaal. Soms lijkt het er zelfs meer op dat er op het gebied van relaties sprake is van consumentisme. In ieder geval wordt echtscheiding heden ten dage niet meer als uitzonderlijk gekwalificeerd, maar als 'normaal' beschouwd (Fox Harding, 1996: 55-56).

Het feit dat het huwelijk niet langer als onlosmakelijk wordt beschouwd legt wel grotere claims op beide huwelijkspartners. Immers, steeds bestaat de mogelijkheid de relatie te beëindigden wanneer deze niet voldoet aan de wensen van (één van) de partners. Daaraan moet overigens direct worden toegevoegd dat van consumentisme in de zin dat wanneer een relatie niet meer voldoet deze wordt beëindigd en een nieuwe wordt aangegaan die wel aan de eisen voldoet (Zijderveld, 1983), geen sprake is. Het feit dat velen hertrouwen of een nieuwe relatie aangaat geeft wel aan dat het huwelijk of de twee-relatie nog immer 'up to date' is.

4. OUDERSCHAP

Het aantal respondenten dat geen kinderen wenst te krijgen blijkt verwaarloosbaar klein te zijn. In vrijwel alle landen bestaat een voorkeur voor twee kinderen. Dat kinderen belangrijk zijn in het leven van de meeste respondenten, blijkt eveneens uit het grote aantal respondenten dat van mening is dat kinderen in belangrijke mate kunnen bijdragen aan het welslagen van het huwelijk. De aanwezigheid van kinderen in een relatie wordt blijkbaar beschouwd als een mogelijkheid de kwaliteit van de relatie te verhogen. In Oost-Europa wordt deze mening vaker gehoord dan in bijvoorbeeld Denemarken en West-Duitsland. Dat het krijgen van kinderen geen noodzaak meer is, maar meer berust op de eigen wens tot kinderen krijgen, blijkt bijvoorbeeld uit het antwoord op de vraag of een vrouw om aan haar bestemming te voldoen kinderen behoort te krijgen. In de meeste Westeuropese landen is hiervoor geen meerderheid te vinden. Daar waar wel meerderheden bestaan in West-Europa, lijkt de vertaling van het item debet aan de hoge mate van instemming.[1] In Oost-Europa bestaat meer instemming met deze opvatting, hetgeen wellicht kan worden uitgelegd als een uiting van een minder ver gevorderd proces van individualisering.

Net zoals het huwelijk heeft ook ouderschap zijn vanzelfsprekendheid verloren in de moderne of postmoderne samenleving. Het impliceert overigens niet dat moderne mensen geen kinderen meer willen krijgen, wel dat de noodzaak voor vrouwen om kinderen te krijgen omdat vrouwen daarvoor nu eenmaal zijn voorbestemd of omdat dat nu eenmaal hun taak is om kinderen te krijgen, is verdwenen. Dit traditionele idee is vervangen door een meer individualistisch standpunt dat vrouwen zelf moeten uitmaken of, hoe en wanneer ze kinderen willen krijgen. Ook het idee dat de kinderwens zelfs hoorbaar is in homoseksuele relaties onderstreept nog maar eens de wijdverspreide kinderwens. Als gevolg van de toenemende individualisering is het idee gerijpt dat mensen vrij en autonoom moeten beslissen onder meer over het krijgen van kinderen. Er bestaat geen maatschappelijke dwang of sociale norm waaraan men moet voldoen. In toenemende mate is het geaccepteerd dat mensen weinig of geen kinderen wensen te

Tabel 3. Percentages respondenten die vinden dat vrouwen kinderen mogen krijgen ook als ze geen vaste relatie met een man hebben, die van mening zijn dat kinderen belangrijk zijn voor het welslagen van een huwelijk, en die van mening zijn dat kinderen noodzakelijk zijn voor vrouwen om aan hun bestemming te voldoen.

Land	Acceptatie dat vrouw kind heeft zonder vaste relatie	Kinderen zijn belangrijk voor goed huwelijk	Kind is noodzaak voor een vrouw
Frankrijk	39	65	75
Groot-Britannë	36	58	21
West-Duitsland	23	44	45
Oost-Duitsland	34	64	67
Italië	39	64	67
Spanje	62	73	51
Portugal	40	65	61
Nederland	38	54	12
België	32	55	46
Denemarken	67	42	82
Noorwegen	27	61	23
Zweden	25	61	21
Noord-Ierland	25	69	32
Ierland	23	62	26
IJsland	84	66	42
Polen	13	78	75
Hongarije	39	85	96
Tsjechië	29	86	67
Slowakije	26	88	73
Bulgarije	46	84	92
Oostenrijk	39	63	50
Estland	32	74	91
Letland	26	78	96
Litouwen	55	66	89
Roemenië	38	68	85

krijgen. Er is dan ook niet langer een economische noodzaak meer om kinderen te krijgen, immers als 'voorziening' voor de oude dag zijn kinderen niet meer noodzakelijk (Caldwell, 1981: 24). De keuze om kinderen te krijgen is niet zozeer gebaseerd op dergelijke economische of financiële motieven (Saporitti, 1989: 201-205), maar meer afhankelijk geworden van individuele voorkeuren.

De noodzaak, alsmede de vanzelfsprekendheid van kinderen mag dan verdwenen zijn, de kinderwens is gebleven. Dat blijkt onder meer uit het feit dat alternatieve medische mogelijkheden (bijv. in vitro fertilisation) worden benut om vrouwen ook op latere leeftijd nog kinderen te laten krijgen. Al met al is het idee dat kinderen meer dan vroeger een bewuste, geïndividualiseerde keus zijn geworden (Fox Harding, 1996: 28).

Toch betekent dit niet dat het individualisme ongelimiteerd is en dat men van mening is dat men kan en mag doen en laten wat men maar wil. Zo is een overgrote meerderheid

van de ondervraagden van mening dat een vrouw een stabiele relatie met een man moet hebben wanneer ze kinderen wil krijgen. Blijkbaar wordt het belang van het kind zwaar gewogen en dat belang weegt kennelijk zwaarder dan het eigen belang van de vrouw. Het blijkt ook uit het feit dat een omvangrijke groep van mening is dat een kind zowel een vader als een moeder nodig heeft om gelukkig te kunnen opgroeien. Het twee ouder-model wordt dus blijkbaar nog immer beschouwd als essentieel voor de normale ontwikkeling van kinderen (Rodger, 1996: 149).

Aan kinderen wordt dus nog steeds een belangrijke plaats in de hedendaagse samenleving gegeven. Hoewel ze niet noodzakelijk meer worden beschouwd voor een betekenisvol leven van individuele personen, vindt men in het algemeen wel dat kinderen cruciale factoren zijn voor het welslagen van een relatie. Tenminste in West-Europa. In Oost-Europa zijn de opvattingen meer traditioneel in de zin dat kinderen krijgen meer een soort sociale norm is. Opvallend is echter dat wanneer er sprake is van kinderen, het welzijn van die kinderen zwaar weegt en kennelijk doorslaggevend is voor de meeste mensen. Het traditionele gezin met zowel een vader als een moeder wordt essentieel geacht voor het geluk van een kind.

5. RELATIES TUSSEN OUDERS EN HUN KINDEREN

In het algemeen wordt aangenomen dat de individualiseringstendens heeft geleid tot meer egalitaire verhoudingen tussen ouders en hun kinderen. Kinderen zijn een steeds centraler plaats in het gezinsleven gaan innemen (Allan, 1985; Walby, 1990). Ouders hebben hun opvoedende taken deels moeten afstaan aan gespecialiseerde instanties en hun relaties met de kinderen hebben zich aangepast aan de nieuwe situatie. In het algemeen, zo wordt gesteld, zijn de relaties tussen ouders en hun kinderen minder autoritair dan voorheen het geval was. In plaats van opvoeder is de ouder nu leraar geworden, en het gezin is meer en meer democratisch geworden tot een plaats waarin ouders en kinderen als meer gelijken tegenover elkaar staan (Slater, 1969).

Toch blijkt, wanneer de antwoorden over vragen naar de relaties tussen ouders en (hun) kinderen worden geanalyseerd, dat hierin van veel democratie geen sprake is. Met uitzondering van de Scandinavische landen en Nederland, wordt de mening dat kinderen hun ouders behoren te gehoorzamen en te respecteren, ongeacht hoe die ouders zijn, bij een meerderheid van de ondervraagden gehoord. Met andere woorden, het traditionele hiërarchische model blijkt te prevaleren in vrijwel alle landen, zowel in Oost- als in West-Europa.

Ook wanneer vanuit het standpunt van de ouders wordt gekeken valt op dat het individualistische standpunt weinig aanhangers kent. Ouders behoren hun uiterste best te doen voor hun kinderen ook al gaat dat ten koste van hun eigen welzijn is het vaakst gehoord. Met andere woorden, men accepteert kennelijk een beperking van de eigen (ouderlijke) vrijheid wanneer het om het kind gaat.

Op basis van de antwoorden op beide vragen onderscheidt Therborn vier typen van wat hij ‘world views on parenting’ noemt (Therborn, 1995: 292-293). Allereerst is daar het type van het *consistent gezins-individualisme*. Kenmerkend hiervoor is dat voor zowel

Tabel 4. Percentages respondenten die vinden dat ouders hun uiterste best moeten doen voor hun kinderen ook al gaat dat ten koste van hun eigen welzijn en die vinden dat kinderen hun ouders moeten respecteren, ongeacht hoe die ouders zijn.

Land	Ouders moeten uiterste best doen	Kinderen moeten ouders respecteren
Frankrijk	81	77
Groot-Brittannië	74	68
West-Duitsland	54	62
Oost-Duitsland	71	77
Italië	79	84
Spanje	78	83
Portugal	84	77
Nederland	68	42
België	64	76
Denemarken	52	47
Noorwegen	73	45
Zweden	63	51
Noord-Ierland	82	80
Ierland	73	78
IJsland	50	61
Polen	67	89
Hongarije	62	81
Tsjechië	65	81
Slowakije	60	86
Bulgarije	58	83
Oostenrijk	60	74
Estland	55	61
Letland	51	72
Litouwen	38	80
Roemenië	85	83

ouders als kinderen geldt dat ze vrij moeten zijn. Het tegenovergestelde type noemt hij *consistent gezins-collectivisme*. Kinderen behoren te gehoorzamen en ouders moeten zich desgewenst opofferen voor hun kinderen. Daarnaast onderscheidt hij nog het *kind-gecentreerde individualisme* waarin kinderen dus een centrale plaats innemen: kinderen moeten onafhankelijk zijn en ouders moeten zich desnoods opofferen. Tenslotte is er dan nog het *ouder-gecentreerde individualisme*: kinderen moeten gehoorzamen, maar ouders hoeven zich niet op te offeren voor hun kinderen.

Het type van consistent gezins-collectivisme komt het vaakst voor, maar van een duidelijk patroon van landen is geen sprake. De meest moderne landen, dat zijn de Scandinavische landen, hebben niet een duidelijk van de andere landen afwijkende voorkeur, net zo min als dat het geval is met de landen waar het moderniseringsproces minder ver is voortgeschreden, bijvoorbeeld de Oost- en Zuideuropese landen. Polen, bijvoorbeeld, komt meer overeen met Frankrijk dan met de andere Oost- en Centraaleuropese landen. Noorwegen wijkt van de overige Scandinavische landen af in die zin dat zij meer kind georiënteerd zijn dan de andere Scandinavische landen.

6. ROLPATRONEN

De toenemende arbeidsparticipatie van vrouwen wordt beschouwd als een van de meest ingrijpende veranderingen van de laatste tijd. Het is het resultaat van een 'complex interplay of economic, technical and demographic developments, which in turn are closely connected to a sociocultural change towards individualism, equality of opportunity and a basic new order of gender roles' (Haller & Höllinger, 1994: 87). Wat ook de drijvende kracht is geweest, vrouwen zijn in toenemende mate toegetreden tot de arbeidsmarkt. Het idee is dat daarmee ook de opvattingen over vrouwenrollen zijn veranderd in de zin dat traditionele rollen hebben afgedaan en weinig ondersteuning meer zullen hebben.

In onze dataset zijn enkele items beschikbaar over hoe mensen denken over de rol van de vrouw in de hedendaagse samenleving.[2] Twee items (1 en 2) verwijzen naar wat Scott, Braun en Alwin hebben aangeduid als 'consequences of women working outside the home' (Scott, Braun & Alwin, 1993: 32), de gevolgen voor de kinderen. Het lijkt erop dat de arbeidsparticipatie van vrouwen geaccepteerd wordt, maar dat het niet ten koste moet gaan van het welzijn van het kind, met name niet wanneer die kinderen nog erg klein zijn en niet naar school gaan. Vooral in Centraal- en Oost-Europa blijkt men van mening dat het kind eronder lijdt wanneer de moeder werkt. In Denemarken is men het hier het minst vaak mee eens, hetgeen het geëmancipeerde beeld van Denemarken bevestigt. Echter de Zweden blijken dan minder geëmancipeerd en als zodanig wordt het Scandinavische patroon niet unaniem gevonden.

Twee uitspraken (3 en 4) verwijzen direct naar de ideologie van de sekse-specifieke rollen. Ze zijn indicatief voor de 'normative position of respondents to the female gender role' (Haller & Höllinger, 1994: 100-101). Centraal staat hetgeen feministen bestrijden, nl. het idee dat vrouwen thuis horen. Desalniettemin is de aanhang voor deze ideologie nog immer groot in Europa. Alleen in Nederland, Groot-Brittannië en Denemarken is de instemming minder. Vooral in Oost-Europa blijkt men sterk voorstander van de traditionele vrouwenrol.

Tenslotte zijn er nog twee uitspraken (5 en 6) die verwijzen naar het belang van werk voor vrouwen. Het betreft dan vooral het materiële aspect en dus de mogelijkheid van vrouwen om niet langer afhankelijk te moeten zijn van het inkomen van de man. In de meeste landen bestaat er een grote mate van instemming met beide uitspraken.

In hoeverre kunnen nu de (niet zo erg grote) variaties in Europa in verband worden gebracht met de verschillen die bestaan tussen landen in arbeidsparticipatie van vrouwen? Immers, hoewel arbeidsparticipatie van vrouwen is toegenomen bestaan er belangrijke verschillen in Europa. De hoogste arbeidsparticipatie wordt gevonden in Scandinavië. De arbeidsparticipatie van vrouwen ten tijde van de dataverzameling in 1990, varieert van 71% in Noorwegen tot 81% in Zweden (OECD, 1993: 192). Daarentegen is de arbeidsdeelname van vrouwen beperkt in Ierland (39%), Spanje (41%) en Italië (45%). De arbeidsparticipatie van vrouwen is in Oosteuropa ongeveer vergelijkbaar met het Scandinavische niveau. In Hongarije bedraagt die 84%; in Tsjechië 82% en in Slowakije 79%. Polen blijft wat achter met ongeveer 55% (ILO, 1995)

Tabel 5. Percentages respondenten die instemmen met de uitspraken over vrouwenrollen.

Land	a	b	c	d	e	f
Frankrijk	73	65	68	60	79	80
Groot-Brittannië	70	55	45	60	68	70
West-Duitsland	41	84	51	54	75	64
Oost-Duitsland	55	80	52	36	74	87
Italië	63	78	72	56	74	80
Spanje	66	57	55	59	78	83
Portugal	76	84	62	49	80	98
Nederland	69	63	41	56	55	30
België	74	61	61	67	70	67
Denemarken	83	32	25	55	81	71
Noorwegen	69	46	51	53	75	74
Zweden	73	74	-*	62	74	87
Noord-Ierland	71	44	53	70	70	83
Ierland	63	53	59	72	61	70
IJsland	82	52	71	71	43	68
Polen	38	94	88	63	64	81
Hongarije	69	70	76	76	48	83
Tsjechië	63	69	84	84	62	91
Slowakije	58	72	85	48	62	82
Bulgarije	61	76	90	87	63	83
Oostenrijk	51	83	62	63	74	74
Estland	82	91	85	71	56	81
Letland	47	92	91	66	62	77
Litouwen	42	90	97	85	41	75
Roemenië	84	58	82	48	68	91

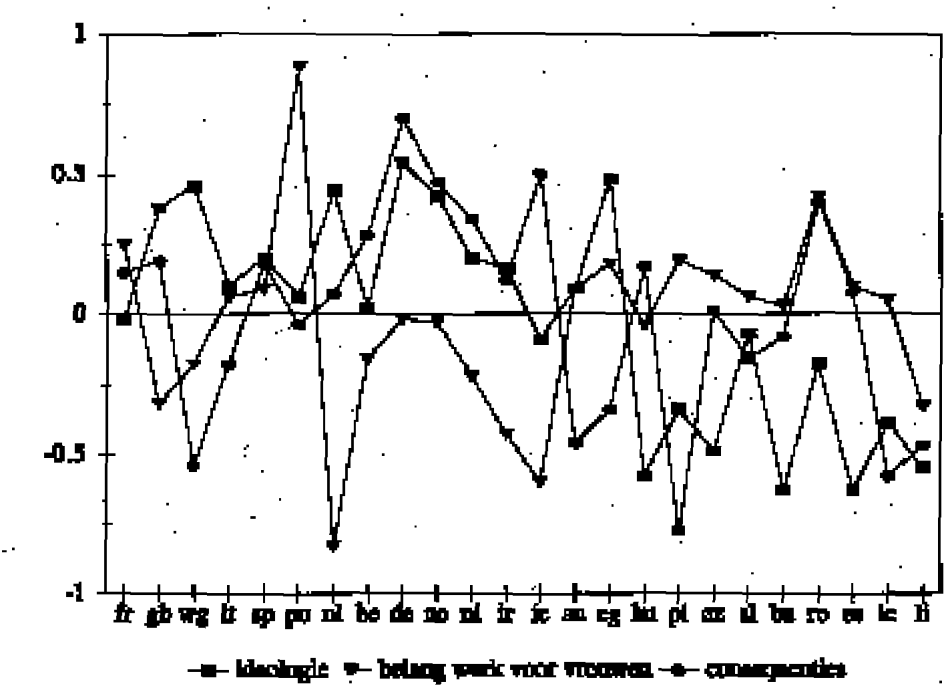
a. Een werkende vrouw kan net zo'n warme relatie met haar kinderen hebben als een moeder die niet werkt
b. Een klein kind lijdt eronder wanneer zijn of haar moeder werkt
c. Een baan is goed, maar wat de meeste vrouwen willen is een thuis en kinderen
d. Huisvrouw zijn geeft net zoveel voldoening als betaald werken
e. Een baan is de beste manier voor een vrouw om onafhankelijk te zijn
f. Zowel man als vrouw moeten bijdragen aan het huishoudinkomen

Hoewel dergelijke cijfers dus al de gevarieerdheid in het Europese patroon van vrouwelijke deelname aan het arbeidsproces weergeven moet, volgens onder andere Boh, ook rekening worden gehouden met de intensiteit van die arbeidsparticipatie. Want, zoals blijkt uit een nadere analyse, vrouwen blijken meer dan mannen werkzaam in part-time banen. Vaak wordt op die manier een baan gecombineerd met de zorgtaken voor de kinderen (Frinking & Willemsen, 1995: 1).

Boh maakt onderscheid in een viertal patronen van vrouwen arbeidsparticipatie. Zo is er het patroon dat wordt gevonden in Spanje, Italië en Ierland waar sprake is van een lage arbeidsparticipatie. Het middelmatige patroon wordt gevonden in Portugal, Frankrijk, West-Duitsland, Oostenrijk en België. Een hoge participatiegraad bestaat in de Scandinavische landen, terwijl met name Nederland een speciale positie inneemt door de hoge mate van part-time werk van vrouwen (Boh, 1989: 270). Men zou toch mogen

verwachten dat daar waar de participatie hoog is meer geëmancipeerde opvattingen domineren. In Figuur 1 hebben we de gemiddelde scores op de drie dimensies weergeven voor de landen in Europa. Een duidelijk patroon valt er niet in te herkennen en dus ook niet een patroon dat overeenkomt met de mate en intensiteit van vrouwen arbeidsparticipatie. Wel kan worden geconstateerd dat de feministische ideologie minder pregnant aanwezig is in Oost-Europa. Volgens Therborn zou dat wel eens te maken kunnen hebben met het feit dat de arbeidsparticipatie van vrouwen in de voormalige Oostblokland en reeds lange tijd zo hoog was (Therborn, 1995: 63).

Ook lijkt de variatie in antwoorden niet samen te vallen met verschillen en overeenkomsten in 'gender welfare state regimes'. Siaroff heeft een typologie van welvaart-staten opgesteld waarin rekening werd gehouden met de positie die vrouwen innemen en het beleid dat wordt gevoerd op dat terrein. Een viertal modellen werd onderscheiden: een Christen-democratisch model (Oostenrijk, België, Frankrijk, Duitsland, Italië, Nederland), een Protestant sociaal-democratisch model (Denemarken, Noorwegen, Zweden, IJsland), een 'late female mobilization' (Ierland, Portugal, Spanje), en een Protestant-liberaal model (Groot-Brittannië). Ook deze indeling loopt niet parallel met de indeling op basis van de attitudes ten aanzien van vrouwenrollen. Wat blijkt is dat er nog immer verschillen bestaan in oordelen over vrouwenrollen, en dat deze blijkbaar niet zonder meer te relateren zijn aan een bepaald beleid dat wordt gevoerd en ook niet aan de actuele participatie van vrouwen aan het arbeidsproces.



Figuur 1. Vrouwenrollen.

Vooralsnog lijkt de traditionele vrouwenrol nog niet verdwenen. Dat blijkt niet alleen uit de opvattingen, ook in de gedragingen komt dit tot uiting. Hoewel meer mannen betrokken zijn bij het huishoudelijke werk en verzorging van de kinderen dan vroeger (Gershuny & Robinson, 1988), worden nog immer de meeste huishoudelijke taken door vrouwen verricht. Met name waar het de verzorging van kinderen betreft blijken mannen weinig actief (Furstenberg & Cherlin, 1991; Van den Akker & Halman, 1994; Elliot, 1996: 33). De 'traditionele man' is dus nog 'far from an endangered species' (Scott, Braun & Alwin, 1993: 25).

7. CONCLUSIE

Het voorgaande lijkt te kunnen worden samengevat met 'de persistentie van de burgerlijkheid', maar wellicht dat zo'n opmerking weinig recht doet aan de realiteit. Inderdaad blijken traditionele opvattingen over gezin en huwelijk tamelijk dominant, niet alleen in de meer traditionele samenlevingen in Europa, ook in de ver geïndividualiseerde samenlevingen is dat het geval. De ideeën van de aanhangers van de theorie van de tweede demografische revolutie lijken dan ook niet te kunnen worden bevestigd. Er zijn geen aanwijzingen dat er een overgang heeft plaatsgevonden van altruïsme naar individualisme. Bovendien, dergelijke ideeën lijken vooral geïnspireerd door een tamelijk negatieve interpretatie van individualisme, namelijk van individuen die zich niets meer gelegen willen laten aan anderen in de samenleving en als volstrekte egoïsten narcistisch door het leven gaan. Inderdaad kan een ver doorgeschoten individualisering daar aanleiding toe geven, maar een noodzakelijke uitkomst van het proces van individualisering is het niet. Individualisering heeft geleid tot meer autonomie en meer vrijheid voor het individu om keuzes te doen, terwijl het aantal keuzemogelijkheden is toegenomen. Het zelfstandig en in alle vrijheid doen van keuzes betekent dat ook traditionele opties kunnen worden verkozen boven allerlei meer hedendaagse varianten. Op het gebied van primaire relaties lijkt dit inderdaad aan de hand. Individualisering heeft niet geleid tot een afwijzing van de traditionele opvattingen over relaties, ondanks de toegenomen vrijheden en mogelijkheden. Het rolgedrag van mannen en vrouwen kan het best worden getypeerd als 'gemankeerde emancipatie' (Van den Akker, Van der Avort & Van den Elzen, 1994). De conclusie kan dan ook niet anders zijn dan dat traditionele opvattingen 'co-exist with values of individual autonomy and gender-equality, allowing people's options to become more and more a matter of choice rather than expectations' (Scott, Braun & Alwin, 1993: 24).

Het gezin mag dan wel weer op de politieke agenda zijn geplaatst en een belangrijk onderwerp in het publieke debat zijn geworden, het is eigenlijk nooit weggeweest. Niet alleen leven de meeste mensen in gezinsverband, ook de opvattingen die ze er op na houden duiden niet op een afwijzing van de traditionele vormen van samenleven, het doorbreken van traditionele rolpatronen en het afscheid van het gezin. Alle omvangrijke en diepingrijpende veranderingen in de samenleving ten spijt, zijn de opvattingen over het gezin, relaties en mannen en vrouwen rollen overwegend opmerkelijk traditioneel gebleven. Als zodanig vormt het gezin een belangrijke uitzondering op de (post-) modernisering van de samenleving.

NOTEN

1. In Denemarken blijkt het item veel minder nadruk te leggen op de noodzaak dat vrouwen kinderen krijgen om aan hun bestemming te voldoen. In het Deens staat er zoiets als: kinderen zijn belangrijk voor vrouwen.
2. De volgende uitspraken zijn voorgelegd met de vraag om aan te geven of men het er sterk mee eens was, mee eens was, mee oneens was of zeer mee oneens was:
   1. Een werkende vrouw kan net zo'n warme relatie met haar kinderen hebben als een moeder die niet werkt;
   2. Een klein kind lijdt eronder wanneer zijn of haar moeder werkt
   3. Een baan hebben is goed, maar wat de meeste vrouwen willen is een thuis en kinderen
   4. Huisvrouw zijn geeft net zo veel voldoening als betaald werken
   5. Een baan hebben is de beste manier voor een vrouw om onafhankelijk te zijn
   6. Zowel man als vrouw moeten bijdragen aan het huishoudinkomen.

LITERATUUR

Allan, G., 1985. Family life, domestic roles and social organisation. Oxford: Basil Blackwell.
Boh, K., 1989. "European family life patterns – a reappraisal". In: K. Boh et al. (eds.), Changing patterns of European family life. A comparative analysis of 14 European countries. London: Routledge, pp. 265-298.
Caldwell, J.C., 1981. The mechanisms of demographic change in historical perspective. Population Studies 35, 5-27.
Chirot, D., 1989. "Causes and consequences of backwardness". In: D. Chirot (ed.), The origins of backwardness in Eastern Europe. Economics and politics from the middle ages until the early twentieth century. Berkeley: University of California Press, pp. 1-14.
Elliot, F.R., 1996. Gender, family and society. Houndmills: Macmillan.
Ester, P., L. Halman en R. de Moor (eds.), 1994. The individualizing society. Tilburg: Tilburg University Press.
Fox Harding, L., 1996. Family, state and social policy. Houndmills: Macmillan.
Frinking, G. en T. Willemsen, 1995. "Introduction: Policies regarding the gender division of paid and unpaid work in Europe". In: T. Willemsen en G. Frinking (eds.), Work and family in Europe: The role of policies. Tilburg: Tilburg University Press, pp. 1-8.
Furstenberg, F. en A. Cherlin, 1991. Divided families: what happens to children when parents part. Cambridge: Harvard University Press.
Gershuny, J. en J. Robinson, 1988. Historical changes in the household division of labor. Demography 25, 537-552.
Glenny, M., 1990. The rebirth of history. Eastern Europe in the age of democracy. Harmondsworth: Penguin Books.
Goldscheider, F. en L. Waite, 1991. New families, no families? Berkeley: University of California Press.
Haller, M. en F. Höllinger, 1994. Female Employment and the Change of Gender Roles: The Conflictual Relationship between Participation and Attitudes in International Comparison. International Sociology 9, 87-112.
ILO, 1995. Yearbook of labour statistics. Geneva: ILO.
Kamerman, S., 1991. Child care policies and programs. An international overview. Journal of Social Issues 47, 179-196.
Lesthaeghe, R. en D. van de Kaa, 1986. Twee demografische transities?, Mens en Maatschappij 61, 9-24.
Liefbroer, A., 1998. Gezinsklimaat: tolerant, maar minder vrijblijvend. Demos 14, 20-23.
Millward, C., 1987. What marriage means to young adults in the 1990's. Family Matters 19, 27-28.

OECD, 1993. Employment Outlook. Paris: OECD.
Prinz, C., 1995. Cohabiting, Married, or Single. Aldershot: Avebury.
Rodger, J.J., 1996. Family Life & Social Control. A Sociological Perspective. Houndmills: MacMillan.
Saporiti, A., 1989. "Historical Changes in the Family's Reproductivity Patterns". In: K. Boh et al. (eds). Changing patterns of European family life. A comparative analysis of 14 European countries. London: Routledge, pp. 191-216.
Scott, J., M. Braun en D. Alwin, 1993. "The family way". In: R. Jowell, L. Brook en L. Dowds (eds.), International Social Attitudes. The 10th BSA report. Aldershot: Dartmouth, pp. 23-47.
Siaroff, A., 1995. "Work, welfare and gender equality". In: D. Sainsbury (ed.), Gendering welfare states. London: Sage, pp. 82-101.
Sgritta, G.B., 1989. "Towards a new paradigm: Family in the welfare state crisis". In: K. Boh et al. (eds.), Changing patterns of European family life: A comparative analysis of 14 countries. London: Routledge, pp. 71-92.
Slater, P., 1969. Parental role differentiation. American Journal of Sociology 67, 296-311.
Therborn, G., 1995. European modernity and beyond. London: Sage.
Van de Kaa, D., 1987. Europe's second demographic transition. Population Bulletin 42, 1-57.
Van den Akker, P., A. Van der Avort en A. Van den Elzen, 1994. "Rolgedrag van mannen en vrouwen: gemankeerde emancipatie". In: P. Ester en L. Halman (red.), De cultuur van de verzorgingsstaat. Tilburg: Tilburg University Press, pp. 135-167.
Van Nimwegen, N. en H. Moors, 1998. Kinderen krijgen.....en dan?. Demos 14, 17-20.
Walby, S., 1990. Theorising patriarchy. Oxford: Basil Blackwell.
Zijderveld, A., 1983. De culturele factor. 's-Gravenhage: Vuga.

12. Sterk en sociaal: Ontwikkeling van de PvdA-boodschap in de campagne van 1998

H. ANKER

SAMENVATTING

Deze bijdrage gaat over de wijze waarop campagneboodschappen ontwikkeld en getoetst kunnen worden. Dit wordt geïllustreerd aan de hand van de 'sterk en sociaal'-boodschap van de Partij van de Arbeid in de Tweede-Kamerverkiezingen van 1998. Daarbij heeft een voor Nederland nieuw soort toetsingsprocedure een belangrijke rol gespeeld. Deze procedure wordt uitgebreid beschreven, alsmede de tekst en inhoudelijke achtergronden van de vier campagneboodschappen die aan de toets zijn onderworpen. Een boodschap bestaande uit zowel 'eigen' en 'niet-eigen' issues leverde, zoals verwacht, de beste toetsingsresultaten op. Deze boodschap is vervolgens op allerlei manieren verfijnd, uiteindelijk resulterend in de boodschap waarmee de PvdA de campagne is ingegaan.

1. INLEIDING

Verkiezingscampagnes in Nederland zijn lange tijd op vrij traditionele wijze gevoerd. Pas vanaf het begin van de jaren negentig worden vernieuwingen zichtbaar. Zo bracht de campagne van 1994 de doorbraak van electoraal onderzoek, 'infotainment' en een forse toename in het gebruik van regionale media. In 1998 werden deze vernieuwingen aangevuld met een sterke nadruk op het uitdragen van de centrale campagneboodschap en de bijbehorende 'boodschap-discipline'. Daarbij trok onder meer de campagne van de Partij van de Arbeid de aandacht. De PvdA had grondig geïnvesteerd in haar centrale boodschap ('sterk en sociaal') en haar politici hielden daar in hun optredens zo veel mogelijk aan vast. Voor een door Perscentrum Nieuwspoort ingestelde jury was dit zelfs reden om de PvdA te belonen met de prijs voor 'de beste campagne', een prijs die in 1998 voor het eerst werd uitgeloofd.

Elk campagnehandboek benadrukt het belang van een heldere, compacte boodschap (zie bijvoorbeeld Faucheux 1995). Die boodschap moet helder en beknopt zijn, een zekere emotionele urgentie uitstralen, voldoende aansluiten bij reeds bestaande cognities van kiezers, contrasteren met de boodschappen van andere partijen, en geloofwaardig zijn (Bradshaw 1995; Anker 1996a). De boodschap bestaat meestal uit een kort verhaal van enkele regels dat in zijn geheel de positie van een partij of kandidaat

weergeeft en kiezers een rationale verschaft om op die partij of kandidaat te stemmen. De elementen van de boodschap moeten verbonden zijn door een aansprekend *thema*, meestal de eerste zin van de boodschap. Het thema fungeert als een soort van lens waarlangs zo veel mogelijk issues tegemoet worden getreden. In de praktijk neemt het thema meestal de vorm aan van een slogan, zoals bijvoorbeeld "Tijd voor verandering". Als het goed is wordt zo niet alleen de positie van de eigen partij gemarkeerd ('verandering'), maar worden tevens de politieke tegenstanders 'gedefinieerd' ('meer van hetzelfde').
Deze bijdrage gaat over de wijze waarop campagneboodschappen ontwikkeld en (vooral) getoetst kunnen worden. Ik doe dit aan de hand van de 'sterk en sociaal'-boodschap van de PvdA in de Tweede-Kamerverkiezingen van 6 mei 1998.[1] Ik ga allereerst dieper in op de achtergrond van de ontwikkeling van de PvdA-campagneboodschap, waarin een nieuwe toetsingsprocedure centraal heeft gestaan (paragraaf 1). Vervolgens bespreek ik het design van deze procedure (paragraaf 2), alsmede de inrichting daarvan voor de PvdA-campagne van 1998 (paragraaf 3) en de bijbehorende resultaten (paragraaf 4). Daarna ga ik dieper in op de wijze waarop de boodschap verder is verfijnd (paragraaf 5). Ik sluit af met enkele korte concluderende opmerkingen (paragraaf 6).

2. ACHTERGROND VAN DE PVDA-BOODSCHAP

De PvdA werd op 6 mei 1998 opnieuw de grootste partij in de Tweede Kamer, boekte de grootste zetelwinst van alle partijen en manoeuvreerde zich in een politieke positie met een zo groot mogelijke kans op voortzetting van regeringsmacht. De boodschap waarmee de Partij van de Arbeid de verkiezingen van 1998 is ingegaan luidde als volgt:

> "**De PvdA wil Nederlandeconomisch en sociaal nog sterkermaken.** Iedereen moet kunnen delen in de welvaartsgroei, in een betere leefomgeving. Meer mensen uit de bijstand naar een baan, een rechtvaardig belastingstelsel, kleinere klassen voor onze kinderen, 30.000 extra banen in de gezondheidszorg, en de politie dichterbij. Dáár hebben we de afgelopen jaren ruimte voor gemaakt. Het Nederland van de 21e eeuw moet **sterk** en **sociaal** zijn."

Deze boodschap voldoet op hoofdlijnen aan alle zojuist gestelde eisen (helder en beknopt, etc.) en kan vrij gemakkelijk door kiezers worden gereproduceerd. Ondanks haar relatieve eenvoud, is deze boodschap het product van een lange reeks van gedachtenwisselingen, discussies, verkenningen, try-outs, toetsing en – opnieuw – gedachtenwisselingen. In dat ontwikkelingsproces is een belangrijke rol weggelegd voor electoraal onderzoek. Eigenlijk is sprake van een voortdurende interactie tussen 'theoretische' inzichten (gebaseerd op wetenschappelijke theorieën, eerdere ervaringen, 'common sense' of intuïtieve veronderstellingen) en empirische toetsing door middel van onderzoek. Daarbij begint men met een zo breed mogelijk perspectief (begrensd door de politieke beginselen van een partij) en zoomt men geleidelijk in op de contouren van de definitieve boodschap.
De ontwikkeling en toetsing van een campagneboodschap is geen 'schoon', technocratisch en geïsoleerd wetenschappelijk proces. De meeste activiteiten op dit terrein spelen zich af in een zeer dynamische omgeving waarbij in ieder geval rekening moet worden

gehouden met al of niet geplande 'try-outs' van de (embryonale) boodschap in de politieke arena, externe (politieke) gebeurtenissen en, *last but not least,* interne machts- en communicatieprocessen. Deze dynamiek verklaart voor een deel waarom de ontwikkeling van campagneboodschappen vaak wordt gezien als een combinatie van kunst en wetenschap ('art and science').
De campagneboodschap van de PvdA is via een reeks tussenstappen ontwikkeld. De meeste van deze stappen werden gezet in het najaar van 1997. De belangrijkste 'kristallisatiepunten' bestonden uit een kwantitatieve toets van de boodschap, een kort daarna verrichte focusgroep en verschillende try-outs in de politieke arena (o.a. interviews met de beoogd lijsttrekker en toespraken op het verkiezingscongres). Terugkijkend op de campagne bleken de resultaten van een kwantitatieve toetsingsprocedure het meest te hebben bijgedragen aan de ontwikkeling van de uiteindelijke campagneboodschap. Deze toets zal daarom in de komende paragrafen gedetailleerd worden besproken.

3. DESIGN VAN DE KWANTITATIEVE TOETSINGSPROCEDURE

De door de PvdA gebruikte kwantitatieve toetsingsprocedrue voor de boodschap is afkomstig uit de Verenigde Staten en is bij mijn weten niet eerder toegepast in Nederland. Het gebruikte design is betrekkelijk eenvoudig en kan het beste worden gezien als een combinatie van experimenteel en survey-onderzoek.[2] Het ontwerp behoort tot een grote familie van designs die sinds het begin van de jaren tachtig door Amerikaanse opiniepeilers worden gebruikt en die dankzij technologische vernieuwingen steeds makkelijker toepasbaar zijn geworden (zie ook Moore 1995; zie Greenberg (1997) voor een toepassing binnen de Amerikaanse context). De door de PvdA gevolgde procedure bestaat uit de volgende vier stappen:

Stap 1: Aan respondenten wordt gevraagd om aan te geven hoe zij zouden stemmen als zij uitsluitend zouden mogen kiezen tussen de PvdA en de VVD. Dit is de *voormeting*.

Stap 2: Er worden twee campagneboodschappen voorgelezen aan de respondenten: één voor de VVD en één voor de PvdA. Voor de VVD is de boodschap voor alle respondenten identiek; voor de PvdA zijn vier verschillende boodschappen onderscheiden waarbij elk aan een kwart van de respondenten is voorgelezen (*split sample*). De letterlijke formulering van deze boodschappen is verderop in deze bijdrage weergegeven.

Stap 3: Aan respondenten wordt opnieuw gevraagd om aan te geven hoe zij zouden stemmen als zij uitsluitend zouden mogen kiezen tussen de PvdA en de VVD. Dit is de *nameting*.

Stap 4: De verschillen tussen de voormeting en de nameting worden in kaart gebracht en onderling vergeleken. De verschillen worden eventueel verder geanalyseerd om te kijken welke groepen (jongeren, vrouwen, etc.) de veranderingen hebben veroorzaakt en welke aspecten van de boodschap daar de belangrijkste rol in hebben gespeeld.

Overeenkomstig de politieke inschattingen van de PvdA in het najaar van 1997 is er binnen de onderzoeksopzet voor gekozen om de verkiezingen te modelleren als een tweestrijd tussen de PvdA (Kok) en de VVD (Bolkestein). Deze inperking tot twee politieke partijen is handig omdat anders voor te veel partijen een boodschap zou moeten worden voorgelezen. De Amerikaanse en Britse praktijk leert overigens dat er

krachtige alternatieve designs voorhanden zijn voor het geval er sprake is van meer dan twee partijen.
Het design wijst eveneens op de noodzaak van een helder omschreven boodschap voor de VVD. Dit wekt wellicht enige bevreemding omdat het weinig consistent lijkt om de VVD-boodschap wèl als bekend te veronderstellen en de boodschap van de PvdA niet. Het bijzondere is echter dat het tijdens campagnes meestal niet zo heel erg moeilijk blijkt om een redelijk adequaat beeld te schetsen van de boodschap van de politieke opponent. De hoofdlijnen van de campagne zijn meestal in een vrij vroeg stadium al bekend – en de Kamerverkiezingen van 1998 vormden daar geen uitzondering op. Overigens kan in gevallen waar men écht in twijfel verkeert altijd besloten worden om het design aan te passen (bijvoorbeeld door te werken met twee PvdA- en twee VVD-boodschappen en alle vier mogelijke combinaties te beschouwen). Daarnaast kan men het aantal respondenten zodanig vergroten dat een groter aantal boodschapcombinaties in het onderzoek kan worden betrokken. De laatste mogelijkheid is uiteraard direct afhankelijk van het beschikbare budget.
Men kan zich op voorhand afvragen in hoeverre Nederlandse respondenten, net als hun Amerikaanse tegenhangers, sowieso bereid zijn om binnen een dergelijke vraagsessie van partijvoorkeur te veranderen. Als die bereidheid door wat voor oorzaak dan ook niet aanwezig zou zijn, dan zou de zojuist beschreven toetsingsprocedure niet goed toepasbaar zijn voor Nederland. Zoals hieronder nog zal blijken bleek dit geenszins het geval en veranderden ook de Nederlandse respondenten in grote aantallen van partijvoorkeur. De vrees voor een mislukking van het onderzoeksproject was daarmee geheel ongegrond.
Ten slotte dient men zich bij de interpretatie van de aantallen partijwisselingen goed te realiseren dat deze veranderingen plaats vinden binnen een kunstmatig scenario, onder constant-houding van alle overige factoren. In werkelijkheid zullen de verschuivingen minder groot zijn doordat burgers blootstaan aan zeer uiteenlopende politieke stimuli. Juist deze gevoeligheid is echter de grote kracht van dit instrument omdat ‘de meters direct uitslaan’ waardoor de electorale aantrekkingskracht van een boodschap (of het gebrek daaraan) eenvoudig kan worden afgelezen.

4. BESCHRIJVING VAN DE TE TOETSEN BOODSCHAPPEN VOOR DE PVDA

De kardinale vraag bij de inrichting van de toetsingsprocedure is welke precieze campagneboodschappen men wil toetsen. Het ligt voor de hand om bij de formulering van de te toetsen boodschappen zo veel mogelijk gebruik te maken van reeds bestaande inzichten. Terugkijkend op de verkiezingscampagne van de Partij van de Arbeid is daartoe, naast talloze andere bronnen, vooral geput uit de volgende vier bronnen van informatie:
1. Heden en verleden van lijsttrekker en partij;
2. Resultaten van eerder verkennend onderzoek;
3. Resultaten van academisch onderzoek, in het bijzonder de theorie van ‘issue-eigenaarschap’;
4. Inzichten uit de Clinton-campagne van 1996.

Ik zal elk van deze inspiratiebronnen nu kort bespreken.

Heden en verleden van lijsttrekker en partij

De identiteit van de beoogd lijsttrekker van de PvdA (Wim Kok) was reeds in een zeer vroeg stadium bekend, waardoor vanaf het prilste begin van de boodschapontwikkeling rekening kon worden gehouden met zijn specifieke politieke opvattingen en gedragingen. Dit alles vond natuurlijk plaats binnen de bredere context van de standpunten van de PvdA zelf, waarbij een centrale rol was weggelegd voor het verkiezingsprogramma 'Een wereld te winnen'. Afgezien van normatieve overwegingen is de overheersende gedachte bij dit alles dat de campagneboodschap voldoende verbonden moet zijn met bestaande cognities van kiezers over waar de partij en de lijsttrekker zich bevinden in het politieke spectrum. In de praktijk komt dit neer op de afbakening van een zekere ideologische bandbreedte (links van het midden) die grenzen stelt aan het soort boodschappen dat de PvdA kan uitstralen. Het zou, om maar een dwarsstraat te noemen, weinig realistisch zijn als de PvdA zich plotseling voorstander zou tonen van afschaffing van het ziekenfonds of van denivellering van de inkomens. Voor een voorbeeld van hoe het in dit opzicht mis kan gaan hoeft men slechts te kijken naar de verkiezingscampagne van 1994. Het CDA stelde toen voor om de AOW-pensioenen voor de komende jaren te bevriezen. De partij leek daarmee buiten haar ideologische bandbreedte te treden, waardoor veel kiezers zich niet meer voldoende in het CDA konden herkennen en zich van de partij afkeerden.

Resultaten van verkennend onderzoek

Reeds eerder verricht verkennend onderzoek had duidelijk gemaakt dat er grote tevredenheid bestond over de resultaten van het kabinet-Kok en dat er zo mogelijk een nog groter vertrouwen bestond in de persoon van Wim Kok. Onder invloed van de gebeurtenissen rondom Meindert Tjoelker – op straat doodgeschopt nadat hij tijdens een avondje uit een opmerking had gemaakt tegen enkele vandalen – maakten kiezers zich ernstig zorgen over criminaliteit en het vervagen van normen en waarden. Daarnaast was zichtbaar dat, hoewel men zeer tevreden was over de economische stand van zaken en de bestrijding van de werkloosheid, kiezers zich in toenemende mate bezorgd toonden over de vraag of de vruchten van de toegenomen welvaart wel voldoende eerlijk werden verdeeld. Ten slotte bestond er een reële voedingsbodem voor angst rondom de Euro, Europa en allochtonen.

Theorie van issue-eigenaarschap

Uit kiezersonderzoek in binnen- en buitenland blijkt steeds weer opnieuw dat het electoraat specifieke partijen associeert met de beste bijdrage op specifieke beleidsterreinen ('issue-eigenaarschap', zie Petrocik 1996). Zo geldt de PvdA als de beste partij voor werk en sociale zekerheid, de VVD als de beste partij voor de bestrijding van de criminaliteit, GroenLinks als de beste partij voor het milieu, enzovoort. Deze associaties zijn bovendien door de tijd en over de grenzen heen opmerkelijk stabiel.
In de afgelopen jaren is duidelijk geworden dat het eigenaarschap van issues een belangrijke predictor is voor zowel de wijze waarop partijen hun campagnes inrichten als de omvang van de aan die campagnes verbonden electorale effecten. Kiezers hanteren sjablonen – schemata – die hen in staat stellen om het gedrag van partijen en politici te duiden. Het links-rechts schema is daarbij nog steeds dominant, maar eigenaar-

schap-schemata spelen een belangrijke aanvullende rol (en laten meer ruimte voor korte-termijn sturing door partijen). Zo blijken de electorale effecten van politiek nieuws direct afhankelijk te zijn van de mate waarin een partij wel of niet als eigenaar van het nieuws-onderwerp wordt gezien. Als de Republikeinen een identieke verkiezingsspot over defensie uitzenden als de Democraten, heeft dat voor de Republikeinen als issue-eigenaar wèl het gewenste electorale effect, en voor de Democraten als niet-eigenaar niet. Het omgekeerde geldt voor onderwerpen waarvan de Democraten wèl eigenaar zijn (Ansolabehere en Iyengar 1994).

De Clinton-campagne van 1996

Bij de aanvang van de voorbereidingen voor de verkiezingscampagne van 1998 is binnen de PvdA steeds de nadruk gelegd op de vergelijkbaarheid met de Clinton-campagne van 1996. Clinton en Kok waren, internationaal bezien, de eersten van een nieuwe lichting van centrum-linkse (minister)-presidenten die geconfronteerd werden met een *her*verkiezingsopdracht. Deze parallellen zijn zeer relevant omdat juist de ervaringen van de Clinton-campagne tot een belangrijke herwaardering hebben geleid van de implicaties van de theorie van issue-eigenaarschap – met uiterst belangrijke consequenties voor de wijze waarop campagnes inhoudelijk worden ingericht. Het betreft hier specifiek de vraag hoe een partij met haar eigen en haar niet-eigen issues om zou moeten gaan. Lange tijd is gedacht dat een partij alleen maar die issues hoefde te benadrukken waarvan zij zelf eigenaar is ("selectieve nadruk op eigen issues", Budge en Farlie 1983a, 1983b). In deze gedachtegang zouden linkse partijen het beste kunnen hameren op eigen thema's als werk en sociale zekerheid en rechtse partijen op zaken als inflatie en defensie. Politici en partijen worden hierbij verondersteld niet mét elkaar over bepaalde vraagstukken te praten, maar vooral langs elkaar heen te praten.
Deze laatste assumptie is echter niet volledig houdbaar gebleken. De herverkiezingscampagne van Clinton heeft duidelijk gemaakt dat partijen de issues waar zij niet de eigenaar van zijn niet straffeloos kunnen negeren. Daarmee is tevens de strategische vraag op tafel gekomen hoe partijen dan wél met hun niet-eigen issues om moeten gaan. Na een fel debat tussen Amerikaanse opiniepeilers (Penn 1996; Greenberg Research 1996) is geleidelijk aan consensus ontstaan over het idee dat linkse partijen hun geprononceerde standpunten op hun eigen issues moeten zien te behouden, maar dat ze zich tegelijkertijd immuun moeten zien te maken voor aanvallen op de niet-eigen issues ('inoculatie'). In de praktijk komt dit laatste meestal neer op een beweging naar het politieke centrum.[3] Bijkomend voordeel is dat een aantal van dergelijke centristische standpunten (bijvoorbeeld de bereidheid om bij criminaliteit niet alleen naar preventie maar ook naar repressie te kijken) helpt om linkse partijen te 'aarden', in de zin dat zij vaste grond onder de voeten houden en niet beticht kunnen worden van enigerlei vorm van soft- of zweverigheid (een klassiek aanvalspunt tegen linkse partijen).

De combinatie van de inzichten uit deze vier informatiebronnen leidt niet automatisch tot de formulering van één enkele campagneboodschap. Integendeel, er zijn vele boodschappen denkbaar die alle geënt zijn op de zojuist besproken inzichten en die bovendien passen binnen het ideologische gedachtegoed van de PvdA. Wat de eindeloze reeks keuzemogelijkheden betreft denke men slechts aan de enorme variatie die kan worden aangebracht door het toevoegen of weglaten van bepaalde issues. Uiteindelijk is een viertal boodschappen geformuleerd:

1. Een *empowerment*-boodschap. Deze boodschap was sterk geïnspireerd door de inhoudelijke thematiek van de Clinton-campagne ("giving people the tools to succeed").
2. Een boodschap met *uitsluitend eigen issues*. Dit is min of meer de klassieke stellingname van sociaal-democratische partijen in West-Europa, door vernieuwers meestal afgeschilderd als 'oud-links' of 'old Labour'.
3 Een boodschap met *uitsluitend niet-eigen issues*. De opname van deze boodschap diende hoofdzakelijk een methodologisch doel (volledigheid en extra contrastmogelijkheid voor het vaststellen van de effectiviteit van de andere boodschappen). De boodschap bevindt zich in velerlei opzicht buiten de ideologische kaders van de PvdA en werd daarom uitdrukkelijk niet gezien als een serieuze suggestie voor de uiteindelijke boodschap.
4 Een boodschap met *zowel eigen als niet-eigen issues*. Dit is het type boodschap waar men op uit zou komen als men de lessen van de Clinton-campagne serieus zou nemen en als deze ook van toepassing zouden zijn op de Nederlandse situatie.

De tekst van de vier getoetste boodschappen luidde als volgt:

PvdA: Empowerment-boodschap **[Variant I: 255 respondenten]**
De PvdA koestert het nieuwe vertrouwen dat met veel inspanningen gezamenlijk is opgebouwd. Nu ons financiële huis weer op orde is gebracht moeten we er alles aan doen om mensen goed toe te rusten voor de uitdagingen van de volgende eeuw. Dat betekent jonge mensen via scholing en onderwijs toerusten voor de toekomst, verder werken aan werk, en veilig stellen van het AOW-pensioen. Alleen DAN kunnen we een samenleving tegemoet zien die voor iedereen de moeite waard is.

PvdA: Eigen issues-boodschap **[Variant II: 252 respondenten]**
De PvdA koestert het nieuwe vertrouwen dat met veel inspanningen gezamenlijk is opgebouwd. Nu ons financiële huis weer op orde is gebracht moeten we er alles aan doen om de mensen die het meeste hebben ingeleverd als eerste te laten profiteren van de economische groei. Dat betekent dat de gezondheidszorg betaalbaar moet blijven, dat de tweedeling tussen arm en rijk moet worden tegengegaan, en dat armoedebestrijding een hoge prioriteit verdient. Alleen DAN kunnen we een samenleving tegemoet zien die voor iedereen de moeite waard is.

PvdA: Niet-eigen issues-boodschap **[Variant III: 245 respondenten]**
De PvdA koestert het nieuwe vertrouwen dat met veel inspanningen gezamenlijk is opgebouwd. Nu ons financiële huis weer op orde is gebracht moeten we er alles aan doen om een veilige leefomgeving te creëren. Criminaliteit moet met harde hand worden bestreden, druggebruik moet worden teruggedrongen en de politie moet actiever optreden. Voor de integratie van minderheden is 'gelijk oversteken' het motto. Kansen en plichten gaan hand in hand. Alleen DAN kunnen we een samenleving tegemoet zien die voor iedereen de moeite waard is.

PvdA: Eigen en niet-eigen issues-boodschap **[Variant IV: 251 respondenten]**
De PvdA koestert het nieuwe vertrouwen dat met veel inspanningen gezamenlijk is opgebouwd. Nu ons financiële huis weer op orde is gebracht moeten we er alles aan doen om iedereen te laten delen in de welvaartsgroei. De AOW moet worden veilig gesteld. Criminaliteit moet met harde hand worden bestreden, druggebruik moet worden teruggedrongen, en de politie moet actiever optreden. Voor de integratie van minderheden is 'gelijk oversteken' het motto. Verder werken aan werk is daarvoor onontbeerlijk. Alleen DAN kunnen we een samenleving tegemoet zien die voor iedereen de moeite waard is.

Deze vier boodschappen werden elk getoetst tegen de volgende vaste VVD-boodschap:

> *VVD-boodschap* **[1002 respondenten]**
> De VVD vindt dat het huidige beleid versneld moet worden doorgezet. Dat betekent niet achterover leunen, maar zorgen dat wat NU nodig is NU ook daadwerkelijk wordt gedaan. De VVD wil daarom verdere lastenverlichting voor burgers en bedrijfsleven, een harde aanpak van de criminaliteit, hard doorwerken aan de integratie van allochtonen, en werken aan een Europese munt die even hard is als onze gulden. Ten slotte spreekt het vanzelf dat de VVD geen schulden voor onze kinderen wil achterlaten.

5. RESULTATEN

De resultaten van de toetsingsprocedure zijn samengevat in Tabel 1. De gegevens in deze tabel maken duidelijk dat er na het voorlezen van de boodschappen forse veranderingen zijn opgetreden in de partijvoorkeuren van de respondenten.

Tabel 1. Resultaten toetsingsprocedure.

	Stemintentie *voor* voorlezen boodschappen			Stemintentie *na* voorlezen boodschappen			
Aard boodschap	PvdA	VVD	Marge	PvdA	VVD	Marge	Verschil in marge
Empowerment (I)	56	37	**+19**	45	51	**–6**	**–25**
Eigen issues (II)	56	39	**+17**	54	42	**+12**	**–5**
Niet-eigen issues (III)	59	34	**+25**	54	42	**+12**	**–13**
Eigen en niet-eigen issues (IV)	51	42	**+9**	59	35	**+24**	**+15**

De gegevens zijn in het najaar van 1997 verzameld door middel van telefonische interviews door het marktonderzoeksbureau Marketresponse. Het totaal aantal respondenten bedraagt 1.002. De gegevens zijn gewogen op basis van geslacht, leeftijd, opleiding en regio. Percentages tellen niet op tot 100, omdat de categorieën 'ander' en 'weet niet' zijn weggelaten.
Leesvoorbeeld: voordat de empowerment-boodschap en de VVD-boodschap aan de respondenten werden voorgelezen gaf 56 procent te kennen bij een geforceerde keuze op de PvdA te zullen stemmen en 37 procent op de VVD. Het verschil tussen deze beide percentages (de 'marge') is positief voor de PvdA en bedraagt +19 procent. Na het voorlezen van de beide boodschappen zegt nog maar 45 procent op de PvdA te gaan stemmen en zegt 51 procent op de VVD te gaan stemmen. Het verschil tussen beide percentages is nu *negatief* (-6 procent). Het verschil tussen de beide marges bedraagt (-6) – (+19) = -25 procentpunten.

Op basis van Tabel 1 kunnen twee belangrijke conclusies worden getrokken.[4] Ten eerste blijken boodschappen die uitsluitend gericht zijn op de eigen issues (variant I en II) of uitsluitend op de niet-eigen issues (variant III) de positieve marge tussen PvdA en VVD te *verkleinen*. Dat betekent dat zij een *on*gunstig effect hebben op de electorale positie van de PvdA. In het eerste scenario (empowerment) verliest de PvdA zelfs de verkiezingen (negatieve marge van 6 procent). In het tweede en derde scenario behoudt de PvdA weliswaar een voorsprong van 12 procent, maar is deze voorsprong kleiner

dan de marge die bestond voordat de boodschappen werden voorgelezen.
Ten tweede blijkt een boodschap die gericht is op zowel de eigen als de niet-eigen issues (variant IV) de positieve marge tussen PvdA en VVD te *vergroten*. De PvdA vaart in electoraal opzicht wel bij een dergelijk type boodschap. Dit blijkt zowel uit het toegenomen stemmenpercentage van de PvdA (van 51 naar 59 procent) als het afgenomen stemmenpercentage van de VVD (van 42 naar 35 procent).
Gezamenlijk vormen deze conclusies de spectaculaire bevestiging van de eerdere vermoedens dat de PvdA (binnen de context van de Tweede-Kamerverkiezingen van 1998 althans) niet eenvoudigweg kon volstaan met het benadrukken van de eigen issues, maar dat zij tevens een inoculatiestrategie moest volgen op de niet-eigen issues.[5]

Analyse voor subgroepen

Voor elk van de boodschappen zijn vergelijkbare analyses uitgevoerd voor een groot aantal subgroepen. Tabel 2 bevat de subgroepen met de grootste netto-verschuivingen van de 'eigen en niet-eigen issues-boodschap', de boodschap waarvan zojuist bleek dat deze in electoraal opzicht het grootste positieve effect sorteerde.

Tabel 2. Grootste effect van 'eigen en niet-eigen issues-boodschap.'

Subgroep	Verschil in Marge	Omvang subgroep (% electoraat)
Laag opgeleide vrouwen (MAVO of lager)	**+41**	**35**
Laag opgeleiden	**+35**	**49**
Ideologisch midden	**+30**	**43**
Jongeren <30	**+27**	**26**
Noord-Brabant en Limburg	**+27**	**19**
Vrouwen	**+26**	**54**
Laag opgeleide mannen	**+22**	**14**
Gelderland, Overijssel en Utrecht	**+19**	**27**

Leesvoorbeeld: na het voorlezen van de 'eigen en niet-eigen issues-boodschap' en de vaste VVD-boodschap is de marge tussen de PvdA en de VVD onder laag opgeleide vrouwen *vergroot* met 41 procentpunten. De groep laag opgeleide vrouwen beslaat ruim eenderde (35 procent) van het gehele electoraat.

De vergroting van de marge in het vierde scenario wordt vrijwel uitsluitend door vrouwen veroorzaakt. Het verschil in marge tussen voor- en nameting bedraagt voor vrouwen maar liefst 26 procentpunten, terwijl voor de mannen dit verschil slechts 2 punten bedraagt (niet in tabel). Deze gegevens indiceren de aanwezigheid van een 'gender gap' in het electoraal gedrag van Nederlandse kiezers (cf. Leijenaar 1989). Dit gender-patroon is sterk gerelateerd aan het opleidingsniveau van de respondenten. Laag opgeleide vrouwen bewegen het sterkst (vergroting van de marge met maar liefst 41 procent). Zij beslaan ruim een-derde van het electoraat (35 procent) en vormen derhalve een cruciale doelgroep voor de PvdA-campagne.[6] Dat naast gender ook opleiding een belangrijke rol speelt wordt zichtbaar in het feit dat ook laag opgeleide mannen sterk door de boodschap worden aangetrokken (+22 procent). De grote aantrekkingskracht van de boodschap op het ideologische midden is geen verrassing als men zich realiseert dat deze categorie disproportioneel bevolkt wordt door (laag opgeleide) vrouwen.

Opvallend zijn tevens de positieve reacties bij twee categorieën kiezers die zich traditioneel in de warme belangstelling van PvdA-politici mogen verheugen omdat zij worden gezien als belangrijke electorale wingewesten: jongeren en kiezers woonachtig in het zuiden. Jongeren onder de dertig – goed voor een kwart van het electoraat – reageren zeer positief op de boodschap (vergroting van de marge met 27 procent). Respondenten uit de zuidelijke provincies reageren eveneens zeer positief (+27 procent), evenals de mensen uit de oostelijke provincies (+19 procent).

Spontane herinnering

Een nog dieper inzicht in de werking van de boodschappen kan worden verkregen door respondenten een open vraag te stellen over welke van de aan hen voorgelezen zaken hen het meeste hebben aangesproken (de 'spontane herinnering'). Op deze wijze wordt zicht verkregen op welke informatie beklijft bij kiezers en welke informatie min of meer langs hen heen gaat. De aldus vergaarde gegevens maken duidelijk dat de meeste politieke zaken die in de boodschappen worden aangeroerd goed doorkomen bij de respondenten, zij het dat sommige elementen (in het bijzonder criminaliteit en gezondheidszorg) duidelijk beter doorkomen dan andere aspecten. Verdere analyse van het datamateriaal laat zien dat deze patronen sterk samenhangen met de urgentie die kiezers aan die vraagstukken toekennen. Minstens zo opvallend was dat het thema van 'nieuw vertrouwen' nauwelijks resoneerde bij de respondenten: een belangrijke aanwijzing dat de boodschap nog lang niet af was en dat extra geïnvesteerd moest worden in het vinden van een voldoende aansprekend thema.

6. VERDERE VERFIJNING VAN DE BOODSCHAP

De via de kwantitatieve toets verkregen resultaten en conclusies hebben zeer veel richting gegeven aan de verdere ontwikkeling van de campagneboodschap. Hoewel inhoudelijk niet onaantrekkelijk, bieden campagneboodschappen op basis van de empowerment-gedachte of alleen eigen issues te weinig perspectief op electoraal succes. Vandaar dat werd besloten om deze hoofdrichtingen (tijdelijk) af te sluiten en de aandacht te concentreren op de boodschap met eigen en niet-eigen issues.
Besloten werd om de passage over drugs weg te laten en de passage over vertrouwen te vervangen door – naar wat men vermoedde – een meer aansprekend thema ("Nederland economisch en sociaal nog sterker maken"). Daarnaast werd de behoefte gevoeld om meer eigen issues toe te voegen: allereerst werk (als alom erkend 'superissue' van de PvdA) en daarnaast meer aandacht voor onderwijs en gezondheidszorg.[7] Voor de passage over criminaliteit werd inspiratie geput uit de succesvolle slogan van de Britse Labour Party ('tough on crime and the causes of crime'). Na verwerking van al deze suggesties en ideeën ontstond de volgende alternatieve versie voor de 'eigen en niet-eigen issues-boodschap':

> De PvdA wil Nederland economisch en sociaal nog sterker maken. Ons financiële huis is op orde gebracht en het gaat beter met de werkgelegenheid. Nu moeten we er alles aan doen om iedereen te laten delen in de welvaartsgroei. De kwaliteit van het bestaan en van de leefomgeving moet verbeteren. De AOW moet worden veilig gesteld. Er moeten meer banen bijkomen in de gezondheidszorg. Criminali-

teit en de oorzaken van criminaliteit moeten effectiever worden bestreden. Voor de integratie van minderheden is ‘gelijk oversteken’ het motto. Verder werken aan werk en beter onderwijs zijn daarvoor onontbeerlijk. Alleen dan gaan we de 21e eeuw in met een samenleving die voor iedereen de moeite waard is.

Na uitgebreide interne besprekingen en brainstorms werd deze boodschap tezamen met de ‘oorspronkelijke’ boodschap (variant IV) bij wijze van laatste toets aan een focusgroep voorgelegd. Deze werkwijze verdiende de voorkeur boven een nieuwe ronde van kwantitatief onderzoek omdat de ervaring leert dat regelmatige afwisseling van kwantitatief en kwalitatief onderzoek meer ‘diepte’ verschaft om de materie te analyseren.
De deelnemers in de focusgroep reageerden enthousiast op de nieuwe boodschap. De boodschap werd goed begrepen (“heel duidelijk”), als voldoende doortastend ervaren en – niet onbelangrijk – beter gewaardeerd dan de oorspronkelijke boodschap. De deelnemers voelden zich zelfs gerustgesteld: “het is lekker om te horen dat we op de goede weg zitten”. Dit was op zijn beurt weer een geruststellende gedachte voor de PvdA omdat één van de hoofddoelstellingen van de campagne er uit bestond om directe aansluiting te vinden bij juist deze klassieke ‘herverkiezingsnotie’.
Na deze focusgroep is een groot deel van de beschikbare tijd en energie gebruikt om de boodschap verder te stroomlijnen. Onderzoek speelde hierbij vrijwel geen rol meer. De meeste wijzigingen kwamen tot stand door middel van verdere interactie met het verkiezingsprogramma, met de strofe over ‘30.000 banen in de gezondheidszorg’ als meest herkenbare punt. Andere wijzigingen kwamen voort uit reacties op ‘try-outs’ van de boodschap via interviews of door simpel logisch nadenken. De uiteindelijke slogan ‘sterk en sociaal’ is bijvoorbeeld niets anders dan een verbastering van de eerste zin van de boodschap (achteraf bleek een variant op deze slogan ook al in het begin van de zomer in een memorandum te hebben gecirculeerd: “Nederland vooraanstaand, sterk en sociaal in de 21e eeuw”). Met dit thema omschreef de Partij van de Arbeid niet alleen haar eigen missie, maar definieerde zij ook de andere partijen. De boodschap gaf impliciet aan dat politieke partijen ter rechterzijde van de PvdA te weinig aandacht hebben voor het sociale en dat men ondanks goede sociale intenties ter linkerzijde niet voldoet aan de economische randvoorwaarden voor duurzaam sociaal beleid. De PvdA, daarentegen, staat voor het ‘juiste’ evenwicht tussen ‘sterk’ en ‘sociaal’.
De volgorde tussen de in de boodschap vervatte ambities en prestaties werd aangepast nadat op het verkiezingscongres van januari 1998 duidelijk was geworden dat het beter was om te beginnen met de ambities voor de toekomst en pas daarna over te gaan op de prestaties van het kabinet-Kok in plaats van andersom. De PvdA trok hierbij mede lering uit de ervaringen van haar Ierse zusterpartij die verwachtte dat zij electoraal zou worden beloond voor de geleverde prestaties en vervolgens bij de stembus bedrogen uitkwam. Na enkele tekstuele wijzigingen resteerde de boodschap waarmee de PvdA uiteindelijk de verkiezingen is ingegaan (zie paragraaf 1).
De oorspronkelijke campagne-opzet was om na de raadsverkiezingen van 4 maart massief met deze boodschap naar buiten te komen en bij wijze van appèl in de laatste tien dagen van de campagne nog een stevig ‘Kies Kok’-accent te leggen. De slogan zou dan overgaan van ‘sterk en sociaal’ naar ‘sterk en sociaal: Kies Kok’. De voor de maand april voorziene tweestrijd tussen Bolkestein en Kok speelde zich echter reeds in februari tijdens de raadscampagne af. De electorale vooruitzichten voor de PvdA zagen er mede daardoor reeds in een vroeg stadium zeer gunstig uit. Bovendien bleek tijdens de campagne voor de Tweede-Kamerverkiezingen een onverwacht grote appellerende

werking uit te gaan van de 'sterk en sociaal'-boodschap. Hierdoor ontstond geleidelijk aan twijfel over de vraag of een 'Kies Kok'-appèl nog wel nodig was. Uiteindelijk resulteerde dit in een enigszins schizofrene situatie waarbij bepaalde groepen kiezers (zoals radioluisteraars) zullen zweren dat de PvdA massief heeft ingezet op 'Kies Kok', terwijl andere groepen kiezers (bijvoorbeeld krantenlezers) met recht kunnen claimen nooit van die leus gehoord te hebben.

7. DISCUSSIE

Ik heb in deze bijdrage laten zien hoe een campagneboodschap wordt ontwikkeld en empirisch kan worden getoetst. Ik deed dat specifiek aan de hand van mijn eigen ervaringen met de campagneboodschap van de PvdA in de Tweede-Kamerverkiezingen van 1998. Daarbij is speciale aandacht uitgegaan naar de toetsing van campagneboodschappen, in het bijzonder de kwantitatieve toetsingsprocedure die daarbij werd gebruikt.
Het is belangrijk om zich goed te realiseren dat de beschikbaarheid van een afdoende getoetste boodschap geen enkele garantie biedt voor een goede verkiezingsuitslag. Een tweede cruciale taak is om de boodschap daadwerkelijk over het voetlicht te krijgen en zo veel als mogelijk bij de mensen thuis te krijgen (in jargon: 'getting the message out', zie Anker 1996a). Dat vereist op zijn minst een goede boodschapper en een gedegen media-strategie. De PvdA heeft daarbij onmiskenbaar zeer veel baat gehad bij de kwaliteiten van haar lijsttrekker Wim Kok.
Onvermijdelijk rijst bij dit soort beschouwingen de vraag of het gebruik van peilingen en onderzoek niet te veel is doorgeslagen en politici zich niet op kunstmatige wijze allerlei boodschappen aan laten meten. Hoewel het te ver gaat om op deze plaats uitvoerig op deze discussie in te gaan (daarvoor verwijs ik naar Anker 1996a en 1996b), denk ik dat de invloed van peilingen op de inhoudelijke stellingnamen van partijen gemakkelijk wordt overschat. Dat blijkt eens te meer als we de 'sterk en sociaal'-boodschap vergelijken met de uitingen van de zusterpartijen van de PvdA. Zo maakt de Britse Labour-partij herhaaldelijk gebruik van variaties op ditzelfde thema (economic prosperity and social fairness), spreekt men binnen de Duitse SPD over het verwante 'Innovation und Gerechtigkeit' (Cuperus en Kandel, 1998), worden in het Amerikaanse Witte Huis sessies belegd om te discussiëren over 'the concept of broadening those who benefit from the improved economy'[8], en stelde Lionel Jospin, de Franse sociaaldemocratische premier, recentelijk dat hij zo veel mogelijk wil handelen vanuit de principes van "economische doelmatigheid en sociale rechtvaardigheid".[9] Deze frappante inhoudelijke overeenkomsten maken duidelijk dat veel boodschap-elementen minder vergezocht of kunstmatig zijn dan ze op het eerste gezicht misschien lijken. Het onderzoeksproces geeft echter de feitelijke empirische steun die nodig is om ideeën aan te scherpen en onzekerheid te reduceren. Dat draagt vervolgens weer bij aan een breder draagvlak binnen de partij en uiteindelijk tot een zelfbewuste, offensieve presentatie van het eigen gedachtegoed.

NOTEN

1. De auteur was verantwoordelijk voor onderzoek en strategie in de PvdA-campagne van 1998 en werkt(e) daarnaast als *pollster* (opiniepeiler) voor diverse campagnes in het buitenland. E-mail: Hanker @ IGC. ORG

2. Opiniepeilers zelf spreken meestal van een ‘simulatie’ (van een verkiezing), hoewel het technisch gezien gaat om een (laboratorium)-experiment met voor- en nameting binnen de context van een survey.
3. De specifieke aard van de te volgen inoculatie-strategie is direct afhankelijk van de aard van het niet-eigen issue. De meeste niet-eigen issues zijn ‘rechtse’ issues (zie Budge en Farlie 1983a en 1983b), waardoor linkse partijen zich immuun kunnen maken door naar het ideologische midden op te schuiven. Maar voor het milieu-vraagstuk, eigendom van GroenLinks, ligt dit weer anders. De PvdA inoculeert zich daar niet door naar het midden op te schuiven. In 1998 koos de partij er voor om het issue van een ander ‘frame’ te voorzien door zo veel mogelijk te wijzen op de noodzaak van economische duurzaamheid.
4. Een allesbehalve irrelevant ‘bijproduct’ van deze exercitie is dat bij een geforceerde tweestrijd tussen de PvdA en de VVD eerstgenoemde partij steeds een ruime meerderheid van de stemmen krijgt. Dit vormde voor de PvdA aanvullend bewijs dat zij een electorale tweestrijd met de VVD met vertrouwen tegemoet kon zien.
5. Daarnaast werd de grote interpreteerbaarheid van de resultaten opgevat als verdere empirische steun voor het denken in termen van eigen en niet-eigen issues. Issues spelen daarbij in allereerste instantie een instrumentele rol als eigen of niet-eigen issue; de intrinsieke betekenis (‘werk’, ‘gezondheidszorg’) is daaraan ondergeschikt.
6. Dit bleek ook op 6 mei toen exit-polls aangaven dat vrouwelijke kiezers waren oververtegenwoordigd onder het electoraat van GroenLinks, D66 en de PvdA, terwijl het electoraat van de VVD en SP disproportioneel uit mannen bestond.
7. Deze neiging om alleen eigen issues in de boodschap op te willen nemen is zeer sterk en past volledig in de theorie van issue-eigenaarschap zoals beschreven en ontwikkeld door Budge en Farlie (1983a en 1983b) en Petrocik (1996).
8. Geciteerd in de Washington Post, 19 juli 1998, p20.
9. Geciteerd in de Washington Post, 2 juni 1998, pA7.

LITERATUUR

Anker, H. 1996a. “De campagne-strategie: ontwikkeling, implementatie en effectiviteit.” In: N.P.G.W.M. Kramer, E.H.T.M. Nijpels, B.M.J. Pauw, en L. Tiddens (red.). *Politieke marketing: winst of verlies?* ’s-Gravenhage: SDU. pp. 89-104.

Anker, H. 1996b. “Voorbij ‘Berlusconi voorbij’. Methoden en technieken van politiek marktonderzoek.” In: *G. Voerman (red.) Jaarboek Documentatiecentrum Nederlandse Politieke Partijen 1995.* Groningen: Rijksuniversiteit Groningen. pp207-231.

Ansolabehere, S. en S. Iyengar, 1994. Riding the wave and claiming ownership of issues: the joint effects of advertising and news coverage in campaigns. *Public Opinion Quarterly.* 58:335-357.

Bradshaw, J., 1995. “Who will vote for you and why.” In: *J.A. Thurber en C. Nelson (red.). Campaigns and Elections American Style.* Boulder: Westview Press.

Budge, I. en D.J. Farlie. 1983a. *Explaining and predicting elections. Issue effects and party strategies in twentythree democracies.* London: George Allen & Unwin.

Budge, I. en D.J. Farlie. 1983b. “Party competition selective emphasis or direct confrontation? An alternative view with data”. In: *Hans Daalder en Peter Mair (red.), Western European party systems. Continuity and change,* London: Sage.

Cuperus, R. en J. Kandel. 1998. “The magical return of social-democracy. An introduction.” *In: R. Cuperus en J. Kandel (red). European social democracy. Transformation in progress. Social democratic think tanks explore the magical return of social democracy in a liberal era.* Amsterdam: Wiardi Beckman stcihting/Friedrich Ebert Stiftung.

Faucheux, R., 1995. *The road to the White House.* The complete guide to winning in politics. Dubuque: Kendall/Hunt.

Greenberg Research, 1996. *The popular mandate of 1996.* The voter concerns that decided the election and that will shape the future agenda. A national post-election survey sponsored by The Campaign for America's Future. November 1996.
Greenberg, S.B., 1997. "Popularizing progressive politics". In: *S.B. Greenberg en Th. Skocpol (red.), The new majority.* Toward a popular progressive politics. New Haven en Londen: Yale University Press, pp.279-298.
Leijenaar, M.H., 1989. *De geschade heerlijkheid. Politiek gedrag van vrouwen en mannen in Nederland,* 1918-1988. Rijksuniversiteit Leiden.
Moore, D.M., 1995. *The superpollsters.* How they measure and manipulate public opinion in America. New York: Four Walls Eight Windows.
Penn, M.J., 1996. *The mandate for the middle.* Unpublished manuscript.
Petrocik, J.R., 1996. Issue ownership in presidential elections, with a 1980 case study. *American Journal of Political Science* 40:825-850.

13. Niveaus van vertrouwen in landen van de Europese Unie

P. DEKKER

SAMENVATTING

Metingen van sociaal vertrouwen staan de laatste jaren sterk in de belangstelling als indicatoren voor 'sociaal kapitaal', dat wil zeggen voor het vermogen tot vrijwillige samenwerking en collectieve actie van een maatschappij. Onderzoek richt zich thans vooral op de individuele correlaten van sociaal vertrouwen en zelden op de mechanismen waardoor dit vertrouwen politiek en economisch werkzaam wordt. In dit hoofdstuk gaat het om nationale niveaus van vertrouwen in volken, vertrouwen in informatie over publieke kwesties en institutioneel vertrouwen. Verschillen in toegekend vertrouwen vertonen overeenkomstige patronen tussen de elf onderzochte EU-landen: Denen worden meer vertrouwd dan Russen, informatie van milieuorganisaties meer dan informatie van vakbonden en er is meer vertrouwen in de media dan in de politiek. Denen en Nederlanders schenken gemakkelijk vertrouwen, maar voor een scherpe scheiding van 'high trust' en 'low trust' landen is weinig reden.

1. INLEIDING: SOCIAAL EN INSTITUTIONEEL VERTROUWEN ALS HYPE VAN DE JAREN NEGENTIG

'Vertrouwen' is in hoog tempo een populair onderwerp geworden in de sociale weten schappen. Het onderwerp raakte echt in de mode door 'Making democracy work' van Putnam (1993) en 'Trust' van Fukuyama (1995). Voor beide auteurs is wederzijds vertrouwen van burgers een centraal element van het 'sociale kapitaal' van een samenleving, dat wil zeggen van het vermogen van een maatschappij om te komen tot vrijwillige samenwerking en gecoördineerde actie. Putnam biedt een diepgaande empirische vergelijking van de Italiaanse regio's om te laten zien hoe vertrouwen is ingebed in veelomvattende 'civic community' tradities en positieve economische, politieke en bestuurlijke effecten genereert. Fukuyama gaat globaler te werk – het gaat bij hem dan ook om de economische wereldorde – en maakt landen- en regiovergelijkend het economisch belang van sociaal vertrouwen aannemelijk. Inmiddels zijn er de afgelopen jaren diverse monografieën verschenen waarin vertrouwen veelomvattend, multidisciplinair en op zeer verschillende schaal en in zeer uiteenlopende contexten, aan de orde wordt gesteld (Misztal 1996, Govier 1997, Seligman 1997, Hollis 1998). Probleemstellingen voor empirische onderzoek en eerste onderzoeksbevindingen zijn nog vooral te vinden in papers en artikelen (b.v. Edwards en Foley 1997, Uslaner 1998, Newton 1999). Overigens leidt de hype ook tot herwaardering van oud werk, zoals de studie van Banfield (1958) naar de belemmerende werking van sociaal wantrouwen voor de

economische vooruitgang in een dorp in Zuid-Italië en de landenvergelijkende studie van Almond en Verba (1989 [1963]) waarin sociaal vertrouwen als fundament van een democratische 'civic culture' werd gepresenteerd.

De recent toegenomen belangstelling voor vertrouwen heeft niet één oorzaak. Het is eerder een samenloop van omstandigheden geweest, waarin het onderwerp zich als gemeenschappelijk terrein ontwikkelt en het ene onderzoek het andere uitlokt. Van belang is de al langer toenemende belangstelling voor culturele factoren binnen diverse sociaal-wetenschappelijke disciplines en stromingen (organisatiesociologie, vergelijkende politicologie), de zich opeenstapelende aanwijzingen voor een verontrustende daling van het vertrouwen van burgers in de politiek en in publieke instituties in de Verenigde Staten en andere landen (Putnam 1995, Pew 1998, Curtis and Jowell 1997), de ontdekking dat wijd verbreid wantrouwen, cynisme en onverschilligheid grote belemmeringen zijn voor de nieuwe democratieën in Oost- en Centraal-Europa, en toenemende onzekerheid in brede kring over de morele grondslagen van het beschaafde samenleven en de gevestigde democratieën in het westen ('sociale cohesie', 'normen en waarden', 'de kloof tussen burgers en politiek').

Een centrale veronderstelling in beschouwingen over sociaal kapitaal is dat wederzijds vertrouwen van burgers doorwerkt in het functioneren van maatschappelijke instellingen en naast een aangenamer sociaal-cultureel klimaat ook tastbare bestuurlijke en economische opbrengsten heeft. Economisch is het basisidee dat men met wederzijds vertrouwen sneller tot transacties komt en deze transacties soepeler verlopen. Er wordt minder gescreend en gecontroleerd, minder schriftelijk vastgelegd en minder geld aan advocaten besteed. Politiek is een verwante overweging dat vertrouwen collectieve actie facilieert. Ik lever mijn vrijwillige bijdrage zolang ik erop kan vertrouwen dat anderen daar geen misbruik van maken als free-riders (vgl. Hollis 1998). Naarmate er meer vertrouwen is, zullen burgers gemakkelijker in staat zijn om gemeenschappelijke problemen op te lossen en dat zonder een beroep te hoeven doen op de overheid met haar controle en sanctiemogelijkheden (Fukuyama 1995). Een andere overweging is dat burgers zich pas politieke zullen engageren als ze erop mogen rekenen dat politieke controverses tot de politiek beperkt blijven (Almond en Verba 1989). Vrije meningsuiting is gebaseerd op sociaal vertrouwen (of anders gezegd: sociaal vertrouwen is de basis voor politieke tolerantie, niet voor politieke eensgezindheid). Vrije meningsuiting en open debat kenmerken een democratie en dragen bovendien bij aan effectief bestuur voor zover door discussie een draagvlak voor beleid wordt gecreëerd en onaangename verrassingen in de beleidsuitvoering worden voorkomen. Met het oog op het empirisch materiaal dat verderop in dit hoofdstuk aan de orde komt, ga ik iets verder in op de politieke wilsvorming in moderne grootschalige representatieve democratieën. In dergelijke maatschappijen vinden politieke discussies veelal niet face-to-face plaats tussen burgers, maar tussen organisaties en vertegenwoordigers van groepen in en via de media. De bevolking is doorgaans toeschouwer en doet mee via het uitbrengen van een stem bij verkiezingen en via de ondersteuning van strijdende partijen in de politieke arena door middel van lidmaatschappen, donaties, handtekeningen et cetera. Van belang is dat het publiek kan uitgaan van de integriteit van de organisaties en personen die de politieke wilsvorming gestalte geven en van de betrouwbaarheid van de informatie over politieke kwesties die over hen wordt uitgestort. Burgers hoeven niet overtuigd te zijn van de betrouwbaarheid van alle strijdende partijen en de geloofwaardigheid van

alle mogelijk relevante informatiebronnen, maar ze moeten toch wel kunnen vertrouwen in een aantal actoren en bronnen van informatie en zeker ook in de basisinstellingen van de democratie en het openbaar bestuur. Misschien is dat vertrouwen tegenwoordig ook wel een voorwaarde voor vertrouwensrelaties tussen elkaar onbekende burgers (Offe 1997). Het is gemakkelijker om andere mensen te vertrouwen als je ervan uit kan gaan dat je sociale zekerheid is gegarandeerd en bedriegers worden gepakt en bestraft. Naast sociaal vertrouwen is 'institutioneel vertrouwen' al langer een onderwerp van empirisch sociaal-wetenschappelijk onderzoek. In dit hoofdstuk voeg ik daar het vertrouwen in de geloofwaardigheid van informatiebronnen over publieke zaken aan toe. In een sterk 'gemedialiseerde' democratie, waarin burgers politiek zelden kunnen afgaan op hun eigen directe ervaringen, verdient de betrouwbaarheid van publieke informatie apart aandacht.

In de literatuur die ter introductie werd vermeld en in de publieke belangstelling voor vertrouwen, gaat het primair om niveaus van vertrouwen als kenmerk van collectiviteiten en om de gevolgen van (gebrek aan) aan vertrouwen voor collectieve verschijnselen: de kwaliteit van de democratie, de economische dynamiek, de omgangsvormen, de integriteit van het bestuur, het aanzien van buurten in de grote steden et cetera. De individuele verscheidenheid is hier van secundair belang. Opvallend is dat veel empirische literatuur zich bij uitstek concentreert op de individuele diversiteit en kenmerken van vertrouwen. Er is inmiddels veel geschreven over individuele correlaten zoals opleidingsniveau, religiositeit (te onderscheiden in waardenoriëntaties en institutionele bindingen), optimisme, tevredenheid, gemeenschapszin en vrijwillige participatie (o.a. Gabriel 1995, Wuthnow 1997, Elchardus en Smits 1998, Pew 1998, Newton 1999). Doorgaans gaat het vanzelfsprekend om oorzaken of achtergronden van vertrouwen; met name voor vrijwillige participatie wordt wel expliciet de vraag gesteld in hoeverre ze niet (ook) een gevolg is van vertrouwen is (Putnam 1995, Dekker en Van den Broek 1998, Uslaner 1998). Voor onderzoek van de collectieve uitkomsten van verschillen in niveaus van vertrouwen ontbreekt het doorgaans aan gegevens. Landenvergelijkende uitzonderingen, gebaseerd op de *World values study* van 1990/'91, zijn Inglehart (1997: 172-174) en Whiteley (1998) met beschrijvingen van positieve verbanden van niveaus van vertrouwen met het aantal jaren dat democratische instellingen ononderbroken hebben gefunctioneerd, respectievelijk economische groeicijfers.

In dit hoofdstuk gaat het niet om de individuele correlaten, maar uitsluitend om het macroniveau van vertrouwen: hoe verschillen de niveaus van vertrouwen tussen Europese landen en wat is de samenhang tussen verschillende soorten vertrouwen? De beschrijving van internationale patronen in de verschillende soorten vertrouwen is op zich van belang voor een beter begrip van andere Europese landen. De samenhang tussen verschillende soorten kan laten zien in hoeverre vertrouwen tussen burgers samengaat met steviger vertrouwensbases voor politieke processen. Aldus kan ook duidelijk worden of vertrouwen het stabiele en veelomvattende kenmerk is dat wordt verondersteld als in het algemeen tussen 'high trust' en 'low trust' samenlevingen wordt onderscheiden.

2. INDICATOREN EN DATABESTANDEN

In het voorafgaande werd vertrouwen in verschillende betekenissen gebruikt. Gangbaar in onderzoek van houdingen van burgers is de onderscheiding naar object van vertrouwen: mensen ('interpersonal' en 'social trust') of instellingen ('institutional trust'). Een populaire variant van institutioneel vertrouwen is politiek vertrouwen. Daarin begrepen zijn vaak ook niet-politieke instellingen van het openbaar bestuur (politie, rechtbanken) en soms wordt het gemeten met vragen over functies (ministers, kamerleden), waarbij niet altijd duidelijk is in hoeverre de instituties of hun momentane belichamingen moeten worden beoordeeld. Dwars op het onderscheid naar object staat de vraag of vertrouwen wel of niet op kennis en ervaring gebaseerd is of kan zijn: tegenover vertrouwen in mensen en organisaties die men kent staat diffuus vertrouwen ('general trust') die niet beperkt is tot specifieke objecten en die door de makers van het onderscheid betiteld is als een 'cognitive bias' (Yamigichi en Yamigichi 1994). Een onderscheid zou ook nog gemaakt kunnen worden naar de aard van het vertrouwen: erop vertrouwen dat iets of iemand het beste met je voor heeft of zich moreel gedraagt is iets heel anders dan erop vertrouwen dat iets of iemand zich volgens een bepaalde regelmaat of berekening gedraagt (Wuthnow 1997).

Diffuus sociaal vertrouwen komt in empirisch onderzoek het meest aan bod. Dé meting van dit type vertrouwen of van vertrouwen *tout court* is de simpele vraag "Vindt u dat over het algemeen de meeste mensen wel te vertrouwen zijn of vindt u dat men niet voorzichtig genoeg kan zijn in de omgang men mensen?" met de antwoordmogelijkheden "wel te vertrouwen" en "je kan niet voorzichtig genoeg zijn" (en "weet niet"). Deze vraag uit Amerikaans onderzoek uit de jaren vijftig heeft via het landenvergelijkende onderzoek van Almond en Verba (1989) de wereld veroverd. De data die ermee werden vergaard in de *World values study* van *1990/'91* zijn terug te vinden in tal van publicaties over sociaal kapitaal en vertrouwen in de jaren negentig (Fukuyama 1995, Uslaner 1998, Whiteley 1998, Newton 1999).

In dit hoofdstuk worden gegevens gebruikt uit diverse edities van de Eurobarometer, een regelmatig opinieonderzoek in opdracht van de Europese Commissie. Hierin zit helaas niet de zo succesvolle sociale-vertrouwensvraag. In de plaats daarvan is er een batterij vragen naar het vertrouwen in volken, met name van landen van de Europese Unie. Die vragen zijn met het oog op de Europese eenwording van bijzonder belang. Wederzijds vertrouwen van Europese volken is een voorwaarde voor Europese gemeenschapszin (Niedermayer 1995), verschaft legitimiteit aan politieke initiatieven op Europees niveau en het bevordert de ontwikkeling van coöperatieve relaties tussen de betrokken landen buiten de formele Europese politiek-bestuurlijke kaders.

Naast deze meting van sociaal vertrouwen is een omvangrijke batterij vragen beschikbaar om institutioneel vertrouwen te meten. De mate waarin specifieke kennis en ervaringen een rol spelen in de oordeelsvorming zal verschillen tussen organisaties en tussen respondenten ('de kerk' betekent voor wie bij de enige ware kerk hoort allicht iets anders dan voor de buitenkerkelijke die zich een beeld van de gemiddelde kerk probeert te vormen).

Tot slot zijn vragen beschikbaar over de geloofwaardigheid van bronnen van informatie

over het milieu en over biotechnologie. Voor beide politiek saillante onderwerpen kunnen burgers niet of nauwelijks op eigen ervaringen afgaan en zijn ze aangewezen op publieke informatiebronnen.

De volgende gegevens worden gebruikt: vertrouwen in volken in 1993 (EB 39.0) en in 1996 (EB 46.0), vertrouwen in bronnen van informatie over het milieu in 1995 (EB 43.1bis) en over biotechnologie in 1996 (EB 46.1) en vertrouwen in instituties in 1997 (EB 48.0). Omdat de indicatoren uit verschillende onderzoeken komen, is er over de samenhang op individueel niveau niets te zeggen, maar dat was ook niet de bedoeling.
Over de volgende landen wordt gerapporteerd: Nederland, Denemarken, Duitsland, België, Groot-Brittannië (Verenigd Koninkrijk zonder Noord-Ierland), Ierland, Frankrijk, Spanje, Portugal, Italië en Griekenland. Over deze landen zijn gegevens beschikbaar in alle gebruikte databestanden. Luxemburg blijft buiten beschouwing omdat er veel minder mensen ondervraagd worden dan de gebruikelijke ongeveer duizend per land (in Duitsland zowel in het oosten als in het westen 1000). In alle gevallen betreffen de gegevens de bevolking van 18 jaar en ouder; sociaal-demografische weegfactoren zijn gebruikt voor een grotere representativiteit.

3. VERTROUWEN IN VOLKEN

Allereerst bieden tabel 1 en 2 inzicht in het vertrouwen in vijftien volken. In de Eurobarometer is sinds 1970 een aantal keren de vraag gesteld hoeveel vertrouwen men heeft in diverse volken: “Wilt u mij voor elk volk dat ik u noem, zeggen of u er een groot vertrouwen, redelijk vertrouwen, niet zo’n vertrouwen of helemaal geen vertrouwen in heeft?”. De gegevens zijn zo bewerkt dat zoveel mogelijk een symmetrisch beeld ontstaat van vertrouwensrelaties. Vermeld worden percentages met een groot of redelijk vertrouwen van alle respondenten die over minstens elf van de vijftien volken een uitspraak deden (zie Dekker 1997 voor gemiddelde scores). Zolang het er niet meer zijn dan vier, worden ‘weet niet’-antwoorden beschouwd als indicaties voor het ontbreken van geen vertrouwen. Gearceerd zijn de cellen met vertrouwen in landgenoten.

Landgenoten genieten in het algemeen aanzienlijk meer vertrouwen dan andere volken. De Italianen vormen traditioneel de uitzondering op dit ‘etnocentrisme’. Italianen hebben net zo veel vertrouwen in Noord-Europeanen, Amerikanen en Japanners als in Italianen. Grieken en vreemd genoeg alleen in 1996 ook Portugezen hebben zeer weinig fiducie in andere volken. Kleine volken worden meer vertrouwd dan grotere volken. Denen en de bewoners van de Beneluxlanden – in eerder onderzoek samen met andere Scandinaviërs en Zwitsers (Inglehart 1991) – doen het daarom goed. Nederlanders hebben in beide jaren gemiddeld weinig vertrouwen in Italianen, Grieken en Russen. Het vertrouwen in Duitsers is niet lager dan vanwege het grootteverschil te verwachten was. Ook op basis van andere enquêtegegevens is er trouwens weinig aanleiding om in Nederland een sterk negatief Duitslandbeeld te veronderstellen (Dekker 1997). Opvallend voor Nederland is wel de daling van het vertrouwen in de Fransen tussen 1993 en 1996. Mogelijk speelden de Franse kernproeven en de Franse kritiek op het Nederlandse drugsbeleid hier een rol. Mogelijk onder invloed van schandalen (Dutroux vanaf augustus 1996) is het vertrouwen in Belgen in de enquête eind 1996 lager dan in 1993. Het vertrouwen in Russen daalt verhoudingsgewijs nog sterker.

Tabel 1. Vertrouwen in volken in 1993: 'groot'/'redelijk' vertrouwen in %, gemiddelden van landen en van volken.

	oordelend land											
beoordeeld volk	NL	DK	DE	BE	GB	IE	FR	ES	PT	IT	GR	gemiddeld
de Nederlanders	96	94	78	79	87	78	76	77	75	72	41	77
de Denen	89	98	79	79	84	77	73	68	70	63	42	75
de Duitsers	77	85	94	74	53	59	69	64	62	64	37	67
de Belgen	95	90	79	93	72	75	88	71	77	67	53	78
de Engelsen, de Britten	81	94	71	76	90	69	50	51	65	56	42	68
de Ieren	70	79	61	74	66	94	63	56	57	49	46	65
de Fransen	74	75	77	77	44	70	88	52	78	73	66	70
de Spanjaarden	63	61	62	63	57	60	69	83	62	66	66	65
de Portugezen	65	56	50	58	61	56	59	52	88	44	59	59
de Italianen	44	48	55	56	57	59	52	55	60	68	40	54
de Grieken	52	56	53	53	59	49	49	49	57	47	84	55
de Luxemburgers	95	88	82	95	68	71	86	68	79	61	49	77
de Amerikanen (uit de VS)	83	81	69	69	75	72	48	47	72	77	40	67
de Japanners	65	71	63	55	55	56	37	61	57	70	55	59
de Russen	44	46	35	40	51	43	28	43	48	39	36	41
gemiddeld	73	75	67	69	65	66	62	60	67	61	50	65
N	885	901	1680	852	780	679	915	755	700	746	804	

Bron: Eurobarometer 39.0 (maart-april 1993); respondenten met maximaal 4 'weet niet'-antwoorden (=geen vertrouwen).

Tabel 2. Vertrouwen in volken in 1996: 'groot'/'redelijk' vertrouwen in %, gemiddelden van landen en van volken.

	oordelend land											
beoordeeld volk	NL	DK	DE	BE	GB	IE	FR	ES	PT	IT	GR	gemiddeld
de Nederlanders	96	92	71	70	86	81	70	74	41	71	44	72
de Denen	93	99	75	76	86	78	73	69	38	66	46	72
de Duitsers	77	76	92	61	46	69	73	63	37	70	33	63
de Belgen	82	75	66	83	73	74	82	65	38	59	49	68
de Engelsen, de Britten	75	84	42	66	92	67	47	36	43	59	39	59
de Ieren	72	77	46	62	68	95	60	54	34	50	50	61
de Fransen	50	60	73	68	42	76	89	47	54	66	51	61
de Spanjaarden	70	60	56	61	59	71	67	91	42	61	58	63
de Portugezen	70	58	48	60	70	68	63	65	83	46	53	62
de Italianen	40	50	45	51	53	74	55	60	45	72	53	54
de Grieken	46	49	48	53	50	58	47	47	25	43	84	50
de Luxemburgers	94	81	77	90	72	77	80	64	46	62	46	72
de Amerikanen (uit de VS)	78	75	66	60	76	79	51	46	48	75	29	62
de Japanners	73	72	55	45	56	58	39	58	30	71	50	55
de Russen	33	31	14	25	42	39	22	34	16	30	44	30
gemiddeld	70	69	58	62	65	71	61	58	41	60	49	60
N	891	899	1655	824	782	658	846	694	738	800	844	

Bron: Eurobarometer 46.0 (oktober-november 1996); respondenten met maximaal 4 'weet niet'-antwoorden (=geen vertrouwen).

Met uitzondering van Ierland is overal het vertrouwen in andere volken in 1996 minder dan in 1993. De verhoudingen veranderen echter weinig. Op langere termijn is vertrouwen in volken een betrekkelijk stabiel gegeven (Niedermayer 1996) met als meest veranderlijke element de opvattingen over de niet-Europese volken (Inglehart 1991).

4. VERTROUWEN IN INFORMATIEBRONNEN

Tabel 3 biedt een beoordeling van de betrouwbaarheid van bronnen van informatie over het milieu in 1995. De gegevens zijn verkregen met een tweeledige vraag na overhandiging van een lijst met tien instituties: "Nu zou ik u willen vragen welke informatiebronnen volgens u de waarheid vertellen over hoe het met het milieu gesteld is. Wilt u uit die lijst de informatiebron kiezen waar u het meeste vertrouwen in heeft? ... Welke andere informatiebronnen vertellen volgens u de waarheid over hoe het met het milieu gesteld is?" In de tabel wordt geen onderscheid gemaakt tussen eerste en tweede keuzen. Vermeld wordt de frequentie van genoemde bronnen voor alle respondenten die ten minste één bron noemen of zeggen nergens vertrouwen in te hebben. Om de vergelijking te vergemakkelijken zijn in deze en de volgende tabel de informatiebronnen in volgorde van afnemend gemiddeld vertrouwen geplaatst.
De volgorde komt in alle landen sterk overeen: milieuorganisaties en de wetenschap scoren hoog en politieke partijen en de industrie bungelen onderaan.

Tabel 3. Vertrouwen in informatie over het milieu in 1995: bronnen in %, gemiddelden van landen en van bronnen.

	NL	DK	DE	BE	GB	IE	FR	ES	PT	IT	GR	gemiddeld
milieubeschermingsorganisaties	57	51	72	58	57	65	55	71	63	70	64	62
wetenschappers	53	56	51	50	43	49	53	58	43	57	50	51
consumentenorganisaties	45	50	51	33	35	29	52	42	40	39	27	40
docenten, op school of universiteit	22	21	22	33	35	42	25	32	47	30	41	32
de media	19	21	33	20	18	31	23	33	39	19	34	26
de overheid	12	23	17	10	13	13	10	9	9	9	9	12
organisaties voor toerisme	4	2	5	12	6	14	7	3	11	9	10	8
vakbonden	4	3	6	7	6	7	4	3	6	7	12	6
politieke partijen	4	4	6	5	3	4	3	4	4	2	2	4
de industrie	1	2	4	5	2	4	2	1	4	3	1	3
gemiddeld	22	23	27	23	22	26	23	26	27	25	25	24
N	930	949	1899	899	905	910	946	875	852	898	924	

Bron: Eurobarometer 43.1bis (april-mei 1995).

In 1996 is op overeenkomstige wijze met een tweeledige vraag geïnformeerd naar bronnen van informatie over moderne biotechnologie. De helft van de respondenten van de betreffende eurobarometer-module werd zo ondervraagd, de andere helft kreeg enkele minder bruikbare verzoeken voorgelegd om voor afzonderlijke biotechnische vraagstukken één bron van informatie aan te geven (EC 1997). In tabel 4 wordt over de

eerste variant gerapporteerd: het aantal mensen dat een informatiebron noemt als percentage van de respondenten die minstens één bron noemen of kenbaar maken nergens vertrouwen in te hebben.
De volgorde van te vertrouwen informatiebronnen komt opnieuw aardig overeen: consumenten- en milieuorganisaties genieten veel vertrouwen, vakbonden en politieke organisaties weinig.

Tabel 4. Vertrouwen in informatie over biotechnologie in 1996: bronnen in %, gemiddelden van landen en van bronnen.

	NL	DK	DE	BE	GB	IE	FR	ES	PT	IT	GR	gemiddeld
consumentenorganisaties	71	64	71	60	53	52	66	53	53	46	45	58
milieubeschermingsorganisaties	61	47	69	56	48	55	54	52	48	55	61	55
school en universiteit	48	47	29	41	42	33	25	31	32	36	42	37
dierenbeschermingsorganisaties	35	25	41	34	24	29	29	21	22	36	27	29
de overheid	27	24	15	13	17	19	17	26	17	13	29	20
religieuze organisaties	12	2	13	5	15	19	9	7	12	17	14	11
het bedrijfsleven	10	6	4	9	6	7	8	7	6	10	8	7
vakbonden	7	3	7	6	6	8	6	6	4	6	7	6
politieke organisaties	10	2	9	8	5	4	7	5	5	5	7	6
gemiddeld	31	24	29	26	24	25	25	23	22	25	27	26
N	462	494	948	477	487	478	489	478	474	489	483	

Bron: Eurobarometer 46.1 (oktober-november 1996).

Bij het milieu is de hiërarchie van informatiebronnen geprononceerder en vertoont zij meer gelijkenis tussen de landen dan bij de biotechnologie. Het laatste onderwerp is nieuwer en het leeft waarschijnlijk nog niet zo bij het grote publiek. Van informatie van de overheid hebben de burgers geen hoge dunk. Bij milieu haalt de overheid alleen in Denemarken het gemiddelde percentage van de tien organisaties, bij biotechnologie zit ze in vier landen in de buurt van het gemiddelde.

5. VERTROUWEN IN INSTITUTIES

Tabel 5 bevat een bewerking van de antwoorden op de vraag: “Ik wil u nu een vraag stellen over het vertrouwen dat u heeft in bepaalde instellingen. Zegt u mij voor elk van de volgende instellingen of u er eerder wel vertrouwen of eerder geen vertrouwen in heeft.” Dan worden zeventien instituties opgesomd. In tabel 5 zijn ze in volgorde gezet van het gemiddelde vertrouwen in de elf landen. Vermeld is het aantal mensen dat vertrouwen heeft in een institutie als percentage van alle respondenten die over minstens twaalf instituties een uitspraak deden. Maximaal vijf ‘weet niet’-antwoorden worden beschouwd als geen vertrouwen.
De overeenkomst tussen de landen is geringer dan bij de informatiebronnen en dat is niet een gevolg van het feit dat een langere lijst is voorgelegd. Als meest frequent vertrouwde instelling concurreren radio, leger, politie en charitatieve instellingen. De ver-

schillen zijn soms aanzienlijk, ook tussen verder verwante landen. Zo scoren kerk, leger, justitie en politie (Inglehart pleegt hier samenvattend van autoritaire instellingen te spreken) in Denemarken aanzienlijk hoger dan in Nederland, waar 'niet-overheidsinstellingen' zich in een grote populariteit mogen verheugen. Over de geringe betrouwbaarheid van politieke partijen bestaat wel grote overeenstemming. Met een landengemiddelde van 20% vertrouwen staan ze op afstand onderaan en alleen in Nederland en Groot-Brittannië sluiten ze niet de rij dankzij respectievelijk de Europese Unie en de geschreven pers (oftewel tabloids?).
In België springen regering, politie en justitie er internationaal uit door een laag vertrouwen, in overeenstemming met ander onderzoek en ongetwijfeld ten gevolge van de schandalen waar het land door geplaagd wordt (Elchardus en Smits 1998).

Tabel 5. Vertrouwen in instituties in 1997: 'eerder wel vertrouwen' in %, gemiddelden van landen en van instituties.

	NL	DK	DE	BE	GB	IE	FR	ES	PT	IT	GR	gemiddeld
de radio	79	79	64	72	70	73	62	70	64	51	48	67
het leger	54	75	65	35	80	85	54	60	65	58	85	65
de politie	70	90	70	30	72	79	50	64	52	59	70	64
liefdadigheidsinstellingen	69	60	50	34	69	67	43	70	70	62	70	60
de televisie	74	69	61	71	67	70	45	50	64	43	38	59
de kerk	44	79	51	30	56	59	36	49	69	56	77	55
de Verenigde Naties	60	71	42	33	53	71	44	59	57	56	52	54
niet-overheidsinstellingen	65	33	37	41	36	61	47	70	45	45	50	48
justitie, [nationale] rechtssysteem	54	73	54	15	50	60	36	41	45	32	64	48
overheidsinstellingen	59	59	40	29	51	66	47	41	41	26	42	46
het [nationale] parlement	66	62	39	22	51	40	38	49	50	31	52	45
de Europese Unie	37	38	32	30	26	62	45	57	56	53	60	45
de vakbonden	64	52	42	37	38	57	37	40	45	31	46	45
de geschreven pers	62	46	44	63	15	38	52	52	45	36	34	44
de [nationale] regering	68	54	31	16	49	40	38	44	50	29	49	43
grote ondernemingen	55	57	28	31	33	42	45	43	43	41	38	41
politieke partijen	41	32	14	10	19	21	12	23	16	14	21	20
gemiddeld	60	61	45	35	49	58	43	52	52	43	53	50
N	890	925	1578	885	801	837	914	843	718	814	930	

Bron: Eurobarometer 48.0 (oktober-november 1998); respondenten met maximaal 5 'weet niet'-antwoorden (=geen vertrouwen).

De verschillen in vertrouwensniveaus en diversiteit van patronen in tabel 5 verdienen nader onderzoek – de betrouwbaarheid van de Europese Unie zou wel eens aardig kunnen correleren met het aantal kilometers onlangs door de EU betaald asfalt – maar daarvoor ontbreekt hier de ruimte. De samenhang in individuele waarderingen moet geheel buiten beschouwing blijven. Wel wordt het tijd om naar de samenhang van nationaal scores op de vertrouwensindicatoren te gaan kijken..

6. DE SAMENHANG

Onderin de voorgaande tabellen zijn steeds gemiddelde percentages vermeld, maar dat zijn niet per se de beste samenvattingen. In het geval van vertrouwen in volken lijkt het beter om een onderscheid te blijven maken tussen vertrouwen in landgenoten en in andere volken. Bij de informatiebronnen lijkt het gemiddelde vertrouwen van minder belang dan dat er op zijn minst enkele bronnen worden vertrouwd. Er kunnen goede gronden zijn in de verschillende landen om bepaalde bronnen niet te vertrouwen. Dat geldt ook voor de instituties. Hier zal ook worden gekeken naar vertrouwen in twee soorten instituties die voor het functioneren van democratie van bijzonder belang zijn: de media en openbare instellingen. Deze dimensies van institutioneel vertrouwen laten zich ook goed onderscheiden op grond van factor- of schaalanalyses van individuele vertrouwensscores. Deze samenvattingen van de gegevens uit de voorafgaande paragrafen kunnen voor negen van de elf landen worden aangevuld met de bekende metingen van sociaal vertrouwen uit de *World values studies*.

Tabel 6. Nationale waarden op samengestelde vertrouwensindicatoren.

	NL	DK	DE	BE	GB	IE	FR	ES	PT	IT	GR
1. sociaal vertrouwen in 1981 (%)[a]	45	53	32	29	43	41	25	35	.	27	.
2. sociaal vertrouwen in 1990/'91 (%)[a]	54	58	44	34	44	47	23	34	.	35	.
3. vertrouwen in landgenoten in 1993 (%) (tabel 1)	96	98	94	93	90	94	88	83	88	68	84
4. vertrouwen in landgenoten in 1996 (%) (tabel 2)	96	99	92	83	92	95	89	91	83	72	84
5. gemiddeld vertrouwen in andere volken in 1993 (%) (tabel 1)	71	73	65	68	63	64	61	58	66	61	48
6. gemiddeld vertrouwen in andere volken in 1996 (%) (tabel 2)	68	67	56	61	63	69	59	56	38	59	46
7. vertrouwen in 2 of meer van 10 bronnen van infor matie over het milieu in 1995 (%) (tabel 3)	76	79	82	74	78	88	81	87	82	87	79
8. vertrouwen in 2 of meer van 9 bronnen van infor matie over biotechnologie in 1995 (%) (tabel 4)	89	82	87	79	79	80	79	70	68	76	86
9. vertrouwen in 7 of meer van 17 instituties in 1997 (%) (tabel 5)	82	84	59	38	70	78	56	66	69	55	67
10. vertrouwen in media[b]	72	65	56	69	51	60	53	57	58	43	40
11. vertrouwen in het openbaar bestuur[c]	62	69	50	24	59	62	44	50	51	39	60

[a] 'De meeste mensen zijn te vertrouwen'-meting uit World values (Duitsland= West-Duitsland) (Newton 1999).
[b] Gemiddelde % vertrouwen in de geschreven pers, de radio en de televisie (tabel 5).
[c] Gemiddelde % vertrouwen in justitie, politie, leger, overheidsinstellingen, regering en parlement (tabel 5).

De sociale-vertrouwensindicatoren geven enige grond om de Noordwest-Europese 'high trust societies' te onderscheiden van Zuid-Europese, katholieke of Romaanse 'low trust societies' (Fukuyama 1995). Bij de andere indicatoren ligt dat onderscheid minder voor de hand, zij het dat Denemarken en Nederland behalve bij de informatiebronnen wel steeds de hoogste vertrouwensniveaus tonen.

Tabel 7 brengt de samenhang tussen de nationale waarden van tabel 6 statistisch in kaart. Het gaat slechts om elf (negen bij de eerste twee indicatoren) waarnemingen. De significantie-aanduidingen hebben gezien de aard van de 'steekproef' statistisch weinig betekenis, maar halen wel een interessant patroon naar voren. De metingen van sociaal vertrouwen hangen sterk samen met vertrouwen in (andere) volken en met vertrouwen in instituties. De beide metingen van vertrouwen in informatiebronnen staan los van de andere vertrouwensindicatoren en opmerkelijk is toch ook dat er geen enkele samenhang tussen vertrouwen in het openbaar bestuur en vertrouwen in de media bestaat. Let wel, het gaat hier om relaties tussen nationale gemiddelden. Op individueel niveau zijn er in alle landen uitsluitend positieve verbanden in het vertrouwen dat respondenten toekennen aan de zeventien instituties. In een aantal gevallen, met name bij de media, zijn de verbanden wel zwak en niet-significant.

Tabel 7. Pearson correlaties x 100 tussen de nationale vertrouwensindicatoren van tabel 6.

	1	2	3	4	5	6	7	8	9	10
1. sociaal 1981										
2. sociaal 1990/'91	91**									
3. landgenoten 1993	58	55								
4. landgenoten 1996	75*	63	83**							
5. andere volken 1993	67*	78*	60	45						
6. andere volken 1996	72*	70*	35	53	51					
7. informatie milieu 1995	-18	-17	-47	-13	-27	-4				
8. informatie biotechnologie 1996	37	61	43	44	12	45	-44			
9. instituties 1997	86**	78**	37	65*	25	27	16	23		
10. media 1997	46	49	73*	56	83**	46	-36	13	x	
11. openbaar bestuur 1997	86**	76*	37	66*	8	20	11	36	x	0

x = niet vermeld want gebaseerd op dezelfde metingen; 'significantie': * p<0,05 en ** p<0,01.

Tabel 8 biedt als toegift op de landenvergelijking een vergelijking van binnenlandse regio's voor vertrouwen in volken en het institutionele vertrouwen. Lezing van Putnam (1993) doet verwachten dat het vertrouwen in het zuiden van Italië geringer is dan in het noorden. In tabel 8 wordt de grens gelegd tussen de noordelijke regio's met een boven-gemiddelde score op de 'civic community index' van Putnam (1993: 97) en de rest van Italië. Voor Duitsland ligt de verwachting voor de hand dat door de dictatoriale

Tabel 8. Nationale vertrouwensindicatoren uit tabel 6 voor enkele regio's[a]

	Nederland		Duitsland		België		Italië	
	west	rest	oost	west	Frans	Nl.	noord	zuid
4. landgenoten 1996	95	97	95	92	84	83	72	74
6. andere volken 1996	67	69	57	55	64	58	61	57
9. instituties 1997	80	84	54	60	40	36	55	55
10. media 1997	71	73	49	58	62	73	44	42
11. openbaar bestuur 1997	62	62	45	51	25	24	40	38

[a] In Nederland Noord- en Zuid-Holland en Utrecht tegenover de rest, in Duitsland de voormalige DDR tegenover de oude BRD, in België Franstalige tegenover Nederlandstalige respondenten, en in Italië de noordelijke regio's t/m Toscana, Umbria en Marche tegenover de rest.

ervaringen in het recente verleden het vertrouwen onder bewoners van de voormalige DDR nog altijd geringer is dan in het westen. België en Nederland zijn uit nieuwsgierigheid toegevoegd: is er verschil tussen Frans- en Nederlandstalige Belgen (vanwege taalverschillen in de enquête en/of verschillen tussen Vlaanderen en Wallonië) en wijken in Nederland de (cynischer of juist opener?) bewoners van de randstadprovincies misschien af van de rest van het land?
In Nederland zijn alle verschillen statistisch verwaarloosbaar. In Duitsland is er geen verschil in het vertrouwen in volken, maar bestaat in het oosten inderdaad wel significant minder vertrouwen in instituties (met name door verschillen bij politie, justitie, kerk en alle drie media). Franstalige Belgen hebben significant meer vertrouwen in andere volken en minder vertrouwen in de media (met name radio en tv) dan Nederlandstalige Belgen. In het noorden van Italië bestaat iets meer vertrouwen in andere volken dan in de rest van het land. Dat is het enige significante Italiaanse verschil in tabel 8. Van een groter wantrouwen tegenover landgenoten in het noorden van het land, in 1980 nog gesignaleerd door Inglehart (1991: 159), is geen sprake meer. Overigens bestaan er wel verschillen in de waardering van afzonderlijke instituties in tabel 5. Zo heeft men in het noorden meer vertrouwen in het justitieel apparaat en in politieke partijen en NGO's en minder in de kerk. Die verschillen liggen in de lijn van de door Putnam beschreven noord-zuidverschillen.

7. TOT BESLUIT

De gegevens in dit hoofdstuk laten zich redelijk samenvatten naar objecten van vertrouwen. Sommige volken (Denen, Luxemburgers, Nederlanders) worden meer vertrouwd dan andere (Russen, Grieken, Italianen), sommige bronnen van informatie (milieu- en consumentenorganisaties) worden eerder geloofd dan andere (vakbonden en bedrijfsleven) en sommige instellingen (de radio, het leger) worden betrouwbaarder geacht dan andere (politieke partijen en regeringen). Een samenvatting van de gegevens naar meer en minder vertrouwende landen is minder gemakkelijk en daarom ging het toch. Denen en Nederlanders schenken gemakkelijker vertrouwen dan andere volken, maar ze onderscheiden zich niet altijd, de verschillen zijn vaak niet groot en de overige landen variëren onderling nogal. Het materiaal biedt geen stevige basis voor een algemeen onderscheid van 'high trust' en 'low trust' landen.
Het vertrouwen in informatiebronnen vertoonde wel een duidelijke interne ordening, maar bood geen houvast voor de onderscheiding van landen. De internationale variantie van dit type vertrouwen was ook gering. Misschien deugde de meting niet, misschien is het idee van 'vertrouwen in publieke informatiebronnen' landenvergelijkend onzinnig. Het is dan gewenst om te zoeken naar andere typeringen en metingen van 'publiek vertrouwen' dat betekenis heeft voor de politieke meningsvorming en besluitvormingsprocessen. Het institutionele vertrouwen is daarvoor te algemeen; gangbare metingen van politiek wantrouwen en cynisme zijn meer op die processen toegespitst maar zijn vaak erg stereotiep (Schedler 1993).
Het is goed om te weten dat er een positieve correlatie bestaat tussen het niveau van sociaal vertrouwen van landen en regio's en hun politieke en economische *performance*, maar – verondersteld dat het geen schijnverband is – er moet wel een verhaal bij waarin mechanismen worden getoond tussen houdingen van burgers en prestaties van systemen. Met een algemeen beroep op cultuurverschillen en de theoretische

notie dat vertrouwen helpt collectieve-actieproblemen op te lossen, is niets verklaard, temeer daar niet kan worden aangenomen dat verschillende soorten vertrouwen van het intermenselijke tot het institutionele niveau in elkaars verlengde liggen. Empirisch blijkt dat zelden het geval te zijn en theoretisch ligt het ook niet voor de hand: "Democracy breeds the culture of trust by institutionalizing distrust" (Sztompka 1997). In aanvulling op onderzoek van houdingen in de bevolking is vergelijkend onderzoek nodig naar de 'organisatie' van maatschappelijk vertrouwen, met aandacht voor uiteenlopende zaken als anti-corruptiemaatregelen, processen van politieke en bestuurlijke schaalvergroting en de effecten van de berichtgeving in de media over schandalen.

LITERATUUR

Almond, G.A. en S. Verba, 1989. The civic culture; political attitudes and democracy in five nations. Newbury Park: Sage (oorspronkelijk 1963).

Banfield, E.C., 1958. The moral basis of a backward society. New York: Free Press.

Curtice, J. en R. Jowell, 1997. Trust in the political system. In: R. Jowell et al. (red.), British social attitudes; the 14th report. Aldershot: Ashgate, 89-109.

Dekker, P., 1997. Nationale beeldvorming van Duitsers en Nederlanders in Europees perspectief. Paper voor de voorjaarsconferentie van de NVPP. Rijswijk: SCP.

Dekker, P., en A. van den Broek, 1998. Civil society in comparative perspective. Voluntas 8/1.

EC, 1997. The Europeans and modern biotechnology; eurobarometer 46.1. Brussel: Europese Commissie (DG 12).

Edwards, B., en M.W. Foley, 1997 (red.). Social capital, civil society, and contemporary democracy. American Behavioral Scientist 40 (1997)/5.

Elchardus, M., en W. Smits, 1998. Vertrouwen. In: Vlaanderen gepeild. Brussel: Ministerie van de Vlaamse Gemeenschap, 45-91.

Fukuyama, F., 1995. Trust; the social virtues & the creation of prosperity. New York: The Free Press.

Gabriel, O.W., 1995. Political efficacy and trust. In: J.W. van deth en E. Scarbrough (red.), The impact of values. Oxford: Oxford University Press, 357-389.

Govier, T., 1997. Social trust and human communities. Montreal & Kingston: McGill-Queen's University Press.

Hollis, M, 1998. Trust within reason. Cambridge (UK): Cambridge University Press.

Inglehart, R., 1991. Trust between nations. In: K. Reif en R. Inglehart (red.), Eurobarometer; the dynamics of European public opinion. Basingstoke: MacMillan, 145-186.

Inglehart, R., 1997. Modernization and postmodernization. Princeton: Princeton University Press.

Misztal, B.A., 1996. Trust in modern societies. Cambridge (GB): Polity Press.

Newton, K., 1999. Social and political trust. In: P. Norris (red.). Critical citizens: Global support for democratic government. Oxford: Oxford University Press (te verschijnen).

Niedermayer, O., 1995. Trust and sense of community. In: O. Niedermayer en R. Sinnott (red.). Public opinion and internationalized governance. Oxford: Oxford University Press, 227-245.

Offe, C., 1997. How can we trust our fellow citizens? Paper voor de J.F. Kennedy School of Government Workshop on Confidence in democratic institutions, Washington DC.

Pew, 1998. Deconstructing distrust; how Americans view government. Washington: The Pew Research Center [*www.people-press.org/trustrpt.htm*].

Putnam, R.D. (m.m.v. R. Leonardi and R.Y. Nanetti), 1993. Making democracy work; civic traditions in modern Italy. Princeton: Princeton University Press.

Putnam, R.D., 1995. Tuning in, tuning out: the strange disappearance of social capital in America. PS: Political Science & Politics 28/4, 664-683 .

Schedler, A., 1993. Die demoskopische Konstruktion von 'Politikverdrossenheit'. Politische Vierteljahresschrift 34/3, 414-435.
Seligman, A.B., 1997. The problem of trust. Princeton: Princeton University Press.
Sztompka, P., 1997. Trust, distrust and the paradox of democracy. Paper voor het 17. IPSA wereldcongres, Seoul.
Uslaner, E.M., 1998. Vrijwilligerswerk en sociaal kapitaal. In: P. Dekker (red.). Vrijwilligerswerk vergeleken. Rijswijk/Den Haag: SCP/Vuga (te verschijnen).
Whiteley, P, 1998. Economic growth and social capital. ECPR News 9/3, 16-18.
Wuthnow, R., 1997. The role of trust in civic renewal. College Park (MD): University of Maryland (Working Paper 1 of the National Commission on Civic Renewal).
Yamigishi, T. en M. Yamigishi, 1994. Trust and commitment in the United States and Japan. Motivation and Emotion 18/2, 129-166.

14. De invloed van marktgerichtheid op het bedrijfsresultaat van industriële ondernemingen in Nederland

F. LANGERAK

SAMENVATTING

Onderzoeken naar het verband tussen marktgerichtheid en het bedrijfsresultaat zijn voornamelijk onder Amerikaanse ondernemingen uitgevoerd. Tegen deze achtergrond worden in dit artikel de resultaten uiteengezet van onderzoek naar het verband tussen marktgerichtheid en het bedrijfsresultaat onder 483 ondernemingen in uiteenlopende branches van de Nederlandse metaal- en elektrotechnische industrie. De resultaten tonen aan dat er een positief verband bestaat tussen de marktgerichtheid en de ROI, de groei van de omzet, de winst als percentage van de omzet en het succes van nieuwe producten. Bovendien blijkt dat het verband tussen marktgerichtheid en het bedrijfsresultaat wordt beïnvloed door verschillen tussen en binnen bedrijfstakken. Dit resultaat biedt managers van industriële ondernemingen een houvast bij de keuze of er wel of niet moet worden geïnvesteerd in marktgerichtheid.

1. INTRODUCTIE

Ondernemingen onderkennen het belang van marktgerichtheid, maar in de wetenschap en de praktijk worden uiteenlopende betekenissen aan het concept toegekend. Dit blijkt bijvoorbeeld uit de anekdote van een directievoorzitter die zijn topmanagers uitlegt dat de toegenomen concurrentie-intensiteit de onderneming dwingt om marktgerichter te gaan werken. De marketingmanager springt daar meteen op in: “Ik zeg al jaren dat wij meer marketinggericht moeten gaan werken”. De directievoorzitter snauwt daarop: “Ik zei marktgericht en niet marketinggericht”. Het is dan ook niet verwonderlijk dat marktgerichtheid veelvuldig tot controversen leidt (Jaworski en Kohli, 1996). Dit wordt versterkt doordat de empirische onderzoeken naar het verband tussen marktgerichtheid en het bedrijfsresultaat, alsmede naar de invloed van verschillen tussen en binnen bedrijfstakken op deze relatie, voornamelijk zijn uitgevoerd onder Amerikaanse ondernemingen (Jaworski en Kohli, 1993; Narver en Slater, 1990; Slater en Narver, 1994;1996). Tegen deze achtergrond komt de vraag op wat marktgerichtheid betekent en of er een positief verband kan worden aangetoond tussen de marktgerichtheid en het bedrijfsresultaat van Nederlandse ondernemingen in de metaal- en elektrotechnische industrie.

Het artikel is als volgt opgebouwd. In paragraaf 2 wordt de betekenis van marktgerichtheid uiteengezet. Vervolgens wordt in paragraaf 3 het verband tussen marktgerichtheid en bedrijfsresultaat besproken. In paragraaf 4 wordt de onderzoeksmethodologie uiteengezet. De resultaten van het onderzoek worden in paragraaf 5 besproken. In paragraaf 6 komen de implicaties voor managers van industriële ondernemingen aan de orde. Het artikel wordt in paragraaf 7 afgesloten met aanbevelingen voor vervolgonderzoek.

2. HET CONCEPT MARKTGERICHTHEID

In de marketingliteratuur bestaat er consensus over dat marktgerichte ondernemingen het marketingconcept hebben geïmplementeerd. Al meer dan 35 jaar worden managers gemaand om 'naar afnemers te luisteren', om de afnemer 'bovenin het organogram te plaatsen' en om het doel van de onderneming te definiëren als 'tevreden afnemers creëren en vasthouden' (Day, 1994). Toch is belangrijke kritiek op marktgerichtheid altijd geweest dat het meer een geloofsartikel was dan een praktische basis voor het managen van een onderneming. Dit is ook niet verwonderlijk want als de marketingliteratuur tot het eind van de jaren tachtig in ogenschouw wordt genomen, blijkt dat er weinig bekend was over de kenmerken van marktgerichte ondernemingen. Bovendien waren de bewijzen over de consequenties van marktgerichtheid voor het bedrijfsresultaat voornamelijk anekdotisch van aard. Daardoor hadden managers weinig richtlijnen voor verbetering van de marktgerichtheid van hun onderneming.

Gelukkig komt er nu verandering in deze situatie als gevolg van een herontdekking van marktgerichtheid in de Verenigde Staten begin jaren negentig (Kohli en Jaworski, 1990; Narver en Slater, 1990). Sindsdien is een aantal conceptuele onderzoeken verschenen waarin duidelijker wordt beschreven wat een marktgerichte onderneming is en waaruit deze bestaat. Volgens de meest recente literatuur staat marktgerichtheid voor een reeks opvattingen waarbij de markt als leidraad wordt genomen om verdedigbare concurrentievoordelen te realiseren (Deshpandé en Farley, 1998; Slater en Narver, 1998). Deze voordelen zijn gebaseerd op de lange termijn belangen van afnemers en worden gerealiseerd door afnemers superieure waarde te bieden. Van superieure waarde is sprake als afnemers het verschil tussen de (verwachte) benefits en de (verwachte) acquisitie- en gebruikskosten van een product hoger inschatten dan dat van concurrerende producten. Om superieure waarde te realiseren moeten ondernemingen relaties opbouwen met belangengroepen buiten en binnen de onderneming (Leeflang, 1995). Hiervoor moet de onderneming initiatieven nemen, onderhandelen en de tot stand gebrachte relaties onderhouden.

Deze reeks opvattingen vormt het leidende beginsel achter de gedragingen van een marktgerichte onderneming. Dit gedrag manifesteert zich in een drietal activiteiten, namelijk het verzamelen, verspreiden en gebruiken van marktinformatie. Bij de informatieverzameling gaat het om de activiteiten gericht op het verkrijgen van informatie over afnemers en andere relevante externe belangengroepen. Via verspreiding wordt informatie tussen interne belangengroepen uitgewisseld, waardoor het afbreukrisico met betrekking tot het gebruik kleiner wordt. De informatie wordt tenslotte door de interne belangengroepen gebruikt om door middel van een gecoördineerde toepassing van middelen superieure waarde voor afnemers te creëren.

Alhoewel er de laatste tijd dus meer inzicht is verkregen in wat een marktgerichte onderneming is, zijn er nog steeds leemten en tekortkomingen. Er bestaat bijvoorbeeld geen éénduidigheid over de vraag welke externe en interne belangengroepen zijn betrokken bij het creëren van superieure waarde voor afnemers. De bestaande marktgerichtheidsmodellen identificeren in dit verband afnemers en concurrenten als de externe belangengroepen, en de bedrijfsfuncties R&D, productie en marketing als de interne belangengroepen (Kohli en Jaworski, 1990; Narver en Slater, 1990). Bij toepassing van het marktgerichtheidsconcept op industriële ondernemingen moeten echter ook toeleveranciers tot de externe belangengroepen worden gerekend en maakt de bedrijfsfunctie inkoop zondermeer deel uit van de interne belangengroepen (Langerak, 1997). Dit betekent dat marktgerichtheid voor industriële ondernemingen verscherpt kan worden gedefinieerd aan de hand van de activiteiten van het verzamelen en verspreiden van informatie over afnemers (afnemersoriëntatie), concurrenten (concurrentenoriëntatie) en toeleveranciers (toeleveranciersoriëntatie) en het gebruik van deze informatie door de bedrijfsfuncties inkoop, R&D, productie en marketing (interfunctionele coördinatie) om door een gecoördineerde inzet van middelen superieure waarde voor afnemers te creëren. In het verlengde hiervan wordt de volgende hypothese geformuleerd:

H1 Marktgerichtheid voor industriële ondernemingen bestaat uit vier aan elkaar gerelateerde componenten, te weten de oriëntatie op afnemers, op concurrenten, op toeleveranciers en de interfunctionele coördinatie.

3. HET VERBAND TUSSEN MARKTGERICHTHEID EN HET BEDRIJFSRESULTAAT

Volgens de marketingliteratuur leidt het vermogen van marktgerichte ondernemingen om afnemers superieure waarde te bieden tot afnemerstevredenheid. Tevreden afnemers leiden tot afnemersloyaliteit en deze trouw komt tot uitdrukking in het vermogen van een onderneming om op lange termijn een bedrijfsresultaat boven het gemiddelde in de bedrijfstak te realiseren (Day, 1994). Het is daarom niet verwonderlijk dat een groeiende stroom van empirische onderzoeken onder Amerikaanse ondernemingen een positief verband heeft aangetoond tussen de marktgerichtheid en verschillende aspecten van het bedrijfsresultaat (Jaworski en Kohli, 1993; Narver en Slater, 1990; Slater en Narver, 1994;1996). De resultaten van de Amerikaanse onderzoeken worden in tabel 1 weergegeven.

Doordat de manier waarop ondernemingen worden ingericht voor een belangrijk deel cultureel wordt bepaald, komt de vraag op of een positief verband tussen marktgerichtheid en het bedrijfsresultaat ook kan worden aangetoond voor Nederlandse industriële ondernemingen. Om deze vraag te beantwoorden wordt de volgende hypothese geformuleerd:

H2 Er bestaat een positief verband tussen de mate van marktgerichtheid en het bedrijfsresultaat van Nederlandse industriële ondernemingen.

Om het verband tussen de marktgerichtheid en het bedrijfsresultaat éénduidig vast te kunnen stellen moet worden gecontroleerd voor verschillen tussen bedrijfstakken

Tabel 1. Het verband tussen marktgerichtheid en het bedrijfsresultaat.

Onderzoek:	De invloed van marktgerichtheid op het bedrijfsresultaat:	De invloed van toetredings- en mobiliteitsbarrières en isolerende mechanismen op het bedrijfsresultaat:	De invloed van toetredings-dings- en mobiliteitsbarrières en isolerende mechanismen op het verband tussen marktgerichtheid en bedrijfsresultaat:
Jaworski en Kohli (1993) sample 1:	Significant positief voor: Niet geïdentificeerd: - marktaandeel	Significant positief voor algehele prestatie: - productkwaliteit Significant positief voor marktaandeel: - macht toeleveranciers Significant negatief voor marktaandeel: - rivaliteit	Niet geïdentificeerd: - marktturbulentie - technologische turbulentie - rivaliteit
Jaworski en Kohli (1993) sample 2:	Significant positief voor: - algehele prestatie Niet geïdentificeerd: - marktaandeel	Significant negatief voor marktaandeel: - rivaliteit	Niet geïdentificeerd: - marktturbulentie - technologische turbulentie - rivaliteit
Narver en Slater (1990) sample 1:	Significant negatief voor: - ROA* Niet geïdentificeerd: - afnemerstrouw / omzetgroei - ROS**	Significant positief voor ROA: - macht afnemers	n.v.t.
Narver en Slater (1990) sample 2:	Significant positief voor: - ROA Niet geïdentificeerd: - afnemerstrouw / omzetgroei - ROS	n.v.t.	n.v.t.
Narver en Slater (1990) sample 1+2:	Niet geïdentificeerd: - ROA / afnemers trouw - omzetgroei / ROS	Significant positief voor ROA: - relatieve grootte Significant negatief voor ROA: - relatieve kosten / marktgroei - technologische turbulentie	n.v.t.
Slater en Narver (1994):	Significant positief voor: - ROA - succes nieuwe producten - omzetgroei	Significant positief voor ROA: - relatieve grootte Significant negatief voor ROA: - relatieve kosten - macht afnemers	Niet geïdentificeerd: - marktturbulentie op relatie marktgerichtheid en ROA - technologische turbulentie op relatie marktgerichtheid en succes nieuwe producten - marktgroei op relatie marktgerichtheid en omzetgroei

Vervolg tabel 1. Het verband tussen marktgerichtheid en het bedrijfsresultaat.

		Significant negatief voor omzetgroei: - toetredingsdrempels - rivaliteit Significant positief voor succes nieuwe producten: - relatieve grootte / kosten	- rivaliteit
Slater en Narver (1996):	Significant positief voor: - omzetgroei Niet geïdentificeerd: - ROA	Significant positief voor ROA: - relatieve grootte Significant positief voor omzetgroei: - relatieve grootte / marktgroei	n.v.t.

*ROA=return on assets; **ROS=return on sales

(interbedrijfstakverschillen) en binnen bedrijfstakken (intrabedrijfstakverschillen) die de hoogte van het bedrijfsresultaat kunnen beïnvloeden. Interbedrijfstakverschillen kunnen vanuit de "industriële organisatietheorie" worden verklaard door toetredings- en mobiliteitsbarrières die hun oorsprong vinden in verschillen in bedrijfstak- en marktstructuur, alsmede de marktmacht hierbinnen. Intrabedrijfstakverschillen kunnen vanuit de "resource-based theory of the firm" worden verklaard door isolerende mechanismen die hun oorsprong vinden in bedrijfsspecifieke investeringen in activa en capaciteiten. De toetredings- en mobiliteitsbarrières en isolerende mechanismen waarvoor in navolging van de Amerikaanse onderzoeken wordt gecontroleerd, worden in tabel 2 weergegeven.

Vanuit de "industriële organisatietheorie" en de "resource-based theory of the firm" bestaan verder goede gronden om te veronderstellen dat toetredings- en mobiliteitsbarrières en isolerende mechanismen ook de sterkte van het verband tussen marktgerichtheid en bedrijfsresultaat beïnvloeden (Jaworski en Kohli, 1993). Op markten waar toetredings- en mobiliteitsbarrières en isolerende mechanismen ontbreken, neemt de concurrentie-intensiteit namelijk toe. Hierdoor komt het vermogen van een onderneming om een bedrijfsresultaat boven het gemiddelde in de bedrijfstak te realiseren steeds meer af te hangen van de organisatorische vaardigheden om superieure waarde voor afnemers te creëren. Het is daarom opmerkelijk dat uit de resultaten van de Amerikaanse onderzoeken in tabel 1 blijkt dat "(…) there is little support for the proposition

Tabel 2. De toetredings- en mobiliteitsbarrières en isolerende mechanismen die het bedrijfsresultaat beïnvloeden.

Toetredings- en mobiliteitsbarrières:	Isolerende mechanismen:
– Macht afnemers – Macht toeleveranciers – Dreiging substituten – Dreiging van toetreding – Marktgroei – Concentratieratio – Rivaliteit tussen aanbieders	– Relatieve grootte van de onderneming – Relatieve integrale kostprijs – R&D-uitgaven

that the competitive environment has an effect on the strength of the market orientation-performance relationship" (Slater en Narver, 1994, p.53). Om te onderzoeken of verschillen tussen en binnen bedrijfstakken het verband tussen de marktgerichtheid en het bedrijfsresultaat van Nederlandse ondernemingen wel beïnvloeden, wordt de volgende hypothese geformuleerd:

H3a *De afwezigheid van toetredings- en mobiliteitsbarrières versterkt het verband tussen marktgerichtheid en het bedrijfsresultaat.*

H3b *De afwezigheid van isolerende mechanismen versterkt het verband tussen marktgerichtheid en het bedrijfsresultaat.*

4. METHODOLOGIE

Steekproef: De onderzoekspopulatie bestaat uit 2037 aselect getrokken industriële ondernemingen in uiteenlopende branches van de metaal- en elektrotechnische industrie (1) met afzonderlijke functies voor inkoop, R&D, productie en marketing en (2) met meer dan 20 werknemers. Er is gekozen voor een schriftelijke enquête om de benodigde gegevens te verzamelen. Hierbij is gebruik gemaakt van de single-informant techniek met leden van het managementteam als respondent. De respons bedroeg 483 volledig ingevulde vragenlijsten, ofwel een responspercentage van 23,7%. Van de respondenten bekleedt 69% de functie van algemeen directeur, terwijl 31% van de respondenten de positie van functioneel directeur voor de afdeling inkoop, R&D, productie of marketing vervult. Om de representativiteit van de steekproef te verifiëren is de verdeling van de respons op twee achtergrondkenmerken, te weten de branche en het aantal werknemers, vergeleken met de verdeling in de onderzoekspopulatie. Uit de analyse blijkt dat bij een overschrijdingskans kleiner dan 0,05 de responsverdeling de verhouding in de onderzoekspopulatie op beide achtergrondkenmerken volgt (χ^2=11,57; d.f.=6 en χ^2=9,25; d.f.=7).

Schalen: Om marktgerichtheid te meten is gebruik gemaakt van Langerak's (1997) schaal voor marktgerichtheid. De componenten van marktgerichtheid zijn alle vier gemeten met behulp van 5 items op een 6-punt Likertschaal. De meetschaal wordt in bijlage 1 weergegeven. Het bedrijfsresultaat is gemeten met behulp van objectieve en subjectieve criteria. De objectieve criteria zijn gemeten op ratioschalen en bestaan uit de ROI en de winst voor belastingen als percentage van de omzet. De subjectieve criteria zijn gemeten op 5-punt Likertschalen en behelzen het succes van nieuwe producten en de omzetgroei in vergelijking tot de belangrijkste concurrenten. Deze objectieve en subjectieve criteria vormen samen een goede indicatie voor het bedrijfsresultaat van industriële ondernemingen (Slater en Narver, 1996).

5. ANALYSE EN RESULTATEN

Op basis van de verkregen gegevens is de interne consistentie van de meetschalen voor de afzonderlijke componenten van marktgerichtheid geverifieerd door de (gecorrigeerde) item-to-total correlaties en de betrouwbaarheidscoëfficiënt uit te rekenen. Om de convergentie- en discriminantievaliditeit van en tussen de afzonderlijke componenten

van marktgerichtheid te verifiëren is met behulp van LISREL een confirmatieve factor analyse toegepast op een "multi-trait-multi-method" model (Bagozzi, Yi en Phillips, 1991). De traits bestaan hierbij uit de vier componenten van marktgerichtheid en de twee methods uit de respondentengroepen "algemeen directeuren" en "niet-algemeen directeuren". De validiteit van het "multi-trait-multi-method" model is getoetst aan de hand van de waarden van de rho (0,97), delta (0,98), en NCNF (0,91) indexen die boven de drempelwaarde van 0,9 moeten liggen. Uit de resultaten blijkt verder dat de t-waarden met betrekking tot de traits groter zijn dan 2 en dat een perfecte correlatie tussen de traits ontbreekt. Dit duidt erop dat convergentie- en discriminantievaliditeit van en tussen de afzonderlijke componenten van marktgerichtheid is bereikt. Om vast te stellen of de vier componenten convergeren naar één onderliggend construct van marktgerichtheid is een confirmatieve tweede orde factor model geschat. Opnieuw blijken de indexen boven de drempelwaarde van 0,9 (GFI=0,99; AGFI=0,97) te liggen. Dit betekent dat de vier afzonderlijke componenten van marktgerichtheid inderdaad convergeren naar één onderliggend construct van marktgerichtheid. Op basis hiervan wordt hypothese 1 bevestigd.

Om het verband tussen marktgerichtheid op het bedrijfsresultaat te analyseren is gebruik gemaakt van regressie-analyse. De resultaten worden in tabel 3 weergegeven.
Uit tabel 3 blijkt dat de betacoëfficiënten van marktgerichtheid significant en positief zijn in alle regressievergelijkingen. Op basis hiervan wordt hypothese 2 bevestigd. Het bedrijfsresultaat wordt daarnaast ook duidelijk beïnvloed door de toetredings- en mobiliteitsbarrières en de isolerende mechanismen. Opvallend hierbij is dat de aard van de toetredings- en mobiliteitsbarrières en isolerende mechanismen die van invloed zijn op het bedrijfsresultaat, verschillend zijn voor de afzonderlijke criteria voor het bedrijfsresultaat. Om te toetsen of de toetredings- en mobiliteitsbarrières en isolerende mechanismen ook de sterkte van het verband tussen marktgerichtheid en het bedrijfsresultaat beïnvloeden, is "moderated regressie analyse" toegepast (Sharma, Durand en Gur-Arie,

Tabel 3. Resultaten regressie analyse (beta coëfficiënten).

n=483	ROI	Omzetgroei	Winst/omzet	Succes nieuwe producten
Marktgerichtheid	0.21***	0.36***	0.10**	0.26***
R&D uitgaven	n.s.	n.s.	0.17***	0.18***
Relatieve kosten	n.s.	n.s.	–0.16***	n.s.
Relatieve grootte	n.s.	n.s.	0.11**	0.17***
Concentratiegraad	n.s.	n.s.	0.14**	0.10*
Marktgroei	n.s.	0.27***	0.15**	n.s.
Dreiging toetreding	n.s.	n.s.	n.s.	–0.10**
Macht afnemers	n.s.	n.s.	n.s.	n.s.
Macht toeleveranciers	n.s.	–0.03**	n.s.	n.s.
Rivaliteit	–0.10*	n.s.	n.s.	–0.13**
Dreiging substituten	–0.16***	–0.05**	n.s.	–0.10*
F-waarde	3.72***	10.44***	5.09***	5.94***
Adjusted R^2	0.09	0.24	0.14	0.16

*p<0.1; **p<0.05; ***p<0.01; n.s.= niet significant.

Tabel 4. Resultaten moderated regressie analyse (beta coëfficiënten).

n=483	ROI	ROI	ROI	ROI	Omzetgroei
Modererende variabele (z):	Relatieve kosten	Markt-groei	Macht afnemer	Dreiging substituten	Relatieve kosten
Marktgerichtheid (x)	0.11**	0,07**	0,10**	0,06*	0,07*
Interactie (xz)	0,71**	-0,66**	0,52**	0,59**	0,61**
R&D uitgaven	n.s.	n.s.	n.s.	n.s.	n.s.
Relatieve kosten	-0.49**	n.s.	n.s.	n.s.	-0.50**
Relatieve grootte	n.s.	n.s.	n.s.	n.s.	n.s.
Concentratiegraad	n.s.	n.s.	n.s.	n.s.	n.s.
Marktgroei	n.s.	0.55**	n.s.	n.s.	0.27***
Dreiging toetreding	n.s.	n.s.	n.s.	n.s.	n.s.
Macht afnemers	n.s.	n.s.	-0.44**	n.s.	-0.10**
Macht toeleveranciers	n.s.	n.s.	n.s.	n.s.	n.s.
Rivaliteit	-0.09*	-0.10*	-0.10*	-0.10*	-0.11**
Dreiging substituten	-0.16***	-0.16***	-0.16**	n.s.	n.s.
F-waarde	3.90***	3.87***	3.74***	3,74***	10.1***
Adjusted R^2	0.10	0.10	0.10	0.10	0.25

*p<0.1; **p<0.05; ***p<0.01; n.s.= niet significant

n=483	Omzetgroei	Omzetgroei	Winst/omzet	Succes nieuw product
Modererende variabele (z):	Toetreding	Macht afnemer	Dreiging substituut	Toetreding
Marktgerichtheid (x)	0.53***	0.10*	0.06	0.50***
Interactie (xz)	0.40**	0.76***	0.56*	0.59**
R&D uitgaven	n.s.	n.s.	0.17***	0.17***
Relatieve kosten	n.s.	n.s.	-0.15***	n.s.
Relatieve grootte	n.s.	n.s.	-0.10**	-0.18***
Concentratiegraad	n.s.	n.s.	0.14**	n.s.
Marktgroei	n.s.	n.s.	0.14**	n.s.
Dreiging toetreding	-0.36***	n.s.	n.s.	-0.66***
Macht afnemers	-0.12**	n.s.	n.s.	n.s.
Macht toeleveranciers	n.s.	n.s.	n.s.	n.s.
Rivaliteit	-0.10**	-0.10**	n.s.	-0.13**
Dreiging substituten	n.s.	0.05	n.s.	-0.10*
F-waarde	9.91***	10.65***	4.97***	6.01***
Adjusted R^2	0.24	0.26	0.14	0.20

*p<0.1; **p<0.05; ***p<0.01; n.s.= niet significant

1981). De resultaten worden in tabel 4 weergegeven. Uit tabel 4 blijkt dat voor de modererende variabelen (z) dreiging van toetreding, macht van afnemers en dreiging van substituten, de betacoëfficiënten van de interactie (xz) met marktgerichtheid (x) significant en positief zijn. Dit betekent dat het verband tussen marktgerichtheid en bedrijfsresultaat wordt versterkt naarmate de dreiging van toetreding, de macht van

afnemers en de dreiging van substituten toenemen. Uit de resultaten blijkt verder dat de betacoëfficiënt van de interactie tussen marktgerichtheid en de modererende variabele marktgroei significant en negatief is. Dit betekent dat het verband tussen marktgerichtheid en het bedrijfsresultaat wordt verzwakt naarmate de marktgroei hoger is. Op basis hiervan wordt hypothese 3a bevestigd. Verder blijkt dat de betacoëfficiënt van de interactie tussen marktgerichtheid en de modererende variabele relatieve kosten significant en positief is. Dit impliceert dat het verband tussen marktgerichtheid en bedrijfsresultaat wordt versterkt naarmate de relatieve kosten hoger zijn. Op basis hiervan wordt ook hypothese 3b bevestigd.

6. MANAGEMENTIMPLICATIES

De resultaten van het onderzoek tonen aan dat het aangescherpte concept van marktgerichtheid succesvol kan worden toegepast op industriële ondernemingen. Hierdoor is een duidelijker beeld ontstaan van de betekenis van marktgerichtheid voor industriële ondernemingen. In dit beeld wordt het accent gelegd op de reeks opvattingen die als leidraad dienen voor de organisatorische activiteiten van het verzamelen en verspreiden van informatie over afnemers, concurrenten en toeleveranciers en het gebruik van deze informatie door de bedrijfsfuncties inkoop, R&D, productie en marketing om door middel van een gecoördineerde inzet van middelen superieure waarde voor afnemers te creëren. In navolging van de Amerikaanse onderzoeken komt ook in dit onderzoek het vermogen van een onderneming om superieure waarde voor afnemers te creëren tot uitdrukking in het bedrijfsresultaat. De resultaten van de regressie-analyses tonen namelijk aan dat er een positief en robuust verband bestaat tussen de marktgerichtheid van ondernemingen in de Nederlandse metaal- en elektrotechnische industrie en de ROI, de groei van de omzet, de winst als percentage van de omzet en het succes van nieuwe producten. Deze resultaten zijn bemoedigend omdat nu vast staat dat industriële ondernemingen die beter zijn uitgerust om op marktomstandigheden te anticiperen, een langdurig concurrentievoordeel en een buitengewone rentabiliteit genieten. Bovendien bestaat nu de mogelijkheid om marktgerichte ondernemingen te identificeren, om te begrijpen wat zij doen en om programma's te ontwikkelen om marktgerichtheid op te bouwen.

In tegenstelling tot de Amerikaanse onderzoeken blijken verschillen tussen en binnen bedrijfstakken wel degelijk de sterkte van het verband tussen marktgerichtheid en het bedrijfsresultaat te beïnvloeden. Deze bevinding biedt managers van industriële ondernemingen een richtlijn bij hun keuze om te investeren in programma's om de marktgerichtheid te verbeteren. De resultaten van dit onderzoek geven namelijk aan dat bij een lage concurrentie-intensiteit, als gevolg van de aanwezigheid van toetredings- en mobiliteitsbarrières en isolerende mechanismen, het voor industriële ondernemingen niet per sé noodzakelijk is om marktgerichter te gaan werken. Er is echter een sterk argument om ook als de omstandigheden daartoe geen aanleiding geven in marktgerichtheid te investeren. Het zijn namelijk de schijnbaar statische omstandigheden van markten, waarin gevestigde of nieuwe concurrenten onverwachts de balans kunnen verstoren, afnemers hun latente macht kunnen gaan gebruiken en barrières in de mobiliteit verminderd kunnen worden, die het opbouwen van marktgerichtheid billijken. Ironisch genoeg is het juist dit onvermogen om als onderneming op veranderingen in de concurrentiekrachten te reageren, of op de behoeften van afnemers te anticiperen, het resultaat

van een tekortkoming in de marktgerichtheid van de onderneming. Het is daarom beter te investeren in marktgerichtheid als de zon schijnt, dan te wachten tot het regent.

7. AANBEVELINGEN VERVOLGONDERZOEK

Naar aanleiding van de resultaten van het onderzoek kan een aantal aanbevelingen voor vervolgonderzoek worden geïdentificeerd. Deze aanbevelingen hangen nauw samen met de beperkingen van dit onderzoek. Allereerst bestaat de onderzoekspopulatie uit industriële ondernemingen in de metaal- en elektrotechnische industrie. Alhoewel de metaal- en electrotechnische industrie een belangrijk deel uitmaakt van de industriële bedrijvigheid in Nederland moet voorzichtigheid worden betracht bij het generaliseren van de resultaten. Hiervoor is het noodzakelijk het onderzoek te herhalen in andere sectoren van de Nederlandse industrie. Ten tweede is er sprake van lage adjusted R-kwadraten in de regressievergelijkingen. Alhoewel dit inherent is aan de cross-sectionele opzet van het onderzoek suggereren de lage adjusted R-kwadraten een zwakke relatie tussen marktgerichtheid en het bedrijfsresultaat, terwijl dit in werkelijkheid niet zo is. Dit kan worden ondervangen door toekomstig onderzoek tot één sector te beperken. Ten derde zijn marktgerichtheid en het bedrijfsresultaat in dit onderzoek op hetzelfde tijdstip gemeten. Hierdoor worden vertragingen in de verbetering van het bedrijfsresultaat niet ondervangen. Longitudinaal onderzoek naar de relatie tussen marktgerichtheid en het bedrijfsresultaat behoort daarom tot de aanbevelingen voor vervolgonderzoek. Tenslotte is het huidige instrument om marktgerichtheid te meten ontwikkeld om managers zelf de marktgerichtheid van hun onderneming te laten beoordelen. In vervolgonderzoek verdient het aanbeveling om de marktgerichtheid ook door afnemers, concurrenten en toeleveranciers te laten beoordelen.

LITERATUUR

Bagozzi, R.P., Y. Yi and L.W. Phillips, 1991. Assessing construct validity in organizational research. Administrative Science Quarterly 36 (3), 421-458.
Day, G.S., 1994. The capabilities of market driven organizations. Journal of Marketing 58 (October), 37-52.
Deshpandé, R. and J.U. Farley, 1998. Measuring market orientation: Generalization and synthesis. Journal of Market Focused Management 2, 213-232.
Jaworski, B.J. and A.K. Kohli, 1993. Market orientation: Antecedents and consequences. Journal of Marketing 57 (July), 53-70.
Jaworski, B.J. and A.K. Kohli, 1996. Market orientation: Review, refinement, and roadmap. Journal of Market Focused Management 1, 119-135.
Kohli, A.K. and B.J. Jaworski, 1990. Market orientation: The construct, research propositions, and managerial implications. Journal of Marketing 54 (April), 1-18.
Kohli, A.K., B.J. Jaworski and A. Kumar, 1993. Markor: A measure of market orientation. Journal of Marketing Research 30 (November), 467-477.
Langerak, F., 1997. De invloed van marktgerichtheid op de bedrijfsprestatie van industriële ondernemingen. Amsterdam, Thesis Publishers.
Leeflang, P.S.H., 1995. Marketing aan het einde van het millenium. Holland Management Review 43, 20-31.
Narver, J.C. and S.F. Slater, 1990. The effect of a market orientation on business profitability. Journal of Marketing 54 (October), 20-35.

Pelham, A.M. and D. Wilson, 1996. A longitudinal study of the impact of market structure, firm structure, strategy and market orientation culture on dimensions of small firm performance. Journal of the Academy of Marketing Science 24 (1), 27-43.
Sharma, S., R.M. Durand and O. Gur-Arie, 1981. Identification and analysis of moderator variables. Journal of Marketing Research 18 (August), 291-300.
Slater, S.F. and J.C. Narver, 1994. Does competitive environment moderate the market orientation-performance relationship? Journal of Marketing 58 (January), 46-55.
Slater, S.F. and J.C. Narver, 1996. Competitive strategy in the market focused business. Journal of Market Focused Management 1, 159-174.
Slater, S.F. and J.C. Narver, 1998. Creating a market orientation. Journal of Market Focused Management 2, 241-255.

	Geheel mee oneens **1**	**2**	**3**	**4**	**5**	Geheel mee eens **6**
Uw onderneming…						
... heeft goed inzicht in het aankoopproces van afnemers	0	0	0	0	0	0
... voert regelmatig overleg met afnemers om de kwaliteit van de serviceverlening te vergroten	0	0	0	0	0	0
... speelt goed in op de fasen van het aankoopproces waarin afnemers zich bevinden	0	0	0	0	0	0
... heeft goed inzicht in het functioneren van de inkoopteams van afnemers	0	0	0	0	0	0
... handelt de klachten van afnemers adequaat af	0	0	0	0	0	0
Uw onderneming…						
... heeft voldoende inzicht in de manier waarop concurrenten hun marktdomein definiëren	0	0	0	0	0	0
... is goed op de hoogte van de specifieke vaardigheden van concurrenten	0	0	0	0	0	0
... houdt bij het opstellen van ons beleid nadrukkelijk rekening met de plannen van concurrenten	0	0	0	0	0	0
... analyseert de marketingactiviteiten van de concurrenten grondig	0	0	0	0	0	0
... houdt bij het nemen van beslissingen expliciet rekening met mogelijke tegenacties van concurrenten						
Uw onderneming…						
... verkent in overleg met de toeleveranciers de mogelijkheden om hun kosten te verlagen	0	0	0	0	0	0
... houdt de kwaliteitsnormen van alternatieve toeleveranciers scherp in de gaten	0	0	0	0	0	0
... voert periodiek onderzoek uit naar de financiële positie van toeleveranciers	0	0	0	0	0	0
... betrekt consequent toeleveranciers bij beslissingen die de onderlinge relatie ingrijpend beïnvloeden	0	0	0	0	0	0
... schakelt regelmatig toeleveranciers in bij het ontwikkelen van nieuwe producten	0	0	0	0	0	0
In uw onderneming…						
... stimuleert het management bewust de samenwerking tussen de verschillende functionele gebieden	0	0	0	0	0	0
... bezoeken managers van alle functionele afdelingen regelmatig afnemers	0	0	0	0	0	0
... worden de functionele activiteiten nauwgezet gecoördineerd	0	0	0	0	0	0
... komen vertegenwoordigers van verschillende functionele afdelingen periodiek samen om marktontwikkelingen te bespreken	0	0	0	0	0	0
... wordt de marktbenadering geformuleerd na uitvoerig overleg tusssen alle functionele afdelingen.	0	0	0	0	0	0

15. Meetmethoden voor corporate identity

J. VAN REKOM

SAMENVATTING

In de literatuur worden methodes aangedragen om de Corporate Identity van organisaties te meten, vanuit een drietal perspectieven: identiteit als door een onderzoeker vast te stellen eigenschappen, identiteit als eigenschappen die intern worden waargenomen, en de identificatie van werknemers met hun organisatie. Deze worden in dit artikel getoetst aan Albert en Whetten's criteria voor definities van Corporate Identity: de mate waarin de methode het centrale karakter van de organisatie weer geeft, de onderscheidende eigenschappen, en de continuïteit in de tijd. Geen enkele methode blijkt goed aan meer dan één criterium te voldoen. Daarom wordt op basis van de laddering-techniek een nieuwe methode ontworpen om Corporate Identity te meten.

1. INLEIDING

Het adequaat inspelen door de organisatie op de behoeften van de klanten vormt de kern van het marketing concept (Kotler, 1984:794). Traditioneel komt de toepassing van het marketing-concept neer op het analyseren van wat de klanten willen en het bedenken van een strategie om daarop in te spelen. Daarbij doet zich een fenomeen voor, dat even merkwaardig is als vanzelfsprekend: de aandacht ligt zo zwaar op wat de klant vindt en wat de klant wil, dat de rol van de organisatie zelf uit het spectrum verdwenen lijkt. Toch is wat de klant van de organisatie vindt een reflectie op wat de organisatie zelf is. De behoeften van de klant zijn ten dele geïnspireerd op wat de organisatie zelf mogelijk maakt. De rol van de organisatie zelf in de vorming van haar imago komt des te nadrukkelijker naar voren bij de aflevering van diensten. De klanten hebben dan persoonlijk contact met één of meer organisatieleden. Zij worden rechtstreeks geconfronteerd met de identiteit van de organisatie.

Als gevolg van het zware accent dat tot nu toe op de afnemers heeft gelegen is de identiteit van de organisatie in de literatuur en in de praktijk onderbelicht gebleven. Wanneer het management wil vaststellen wat de organisatie al doet en wat de eigen werknemers al waarmaken, dan vallen ze terug op hun intuïtie. Op precies dezelfde manier viel men voor de opkomst van marktonderzoektechnieken terug op de eigen intuïtie om in te schatten wat de klant wilde. Bij kleine bedrijven, waar iedereen elkaar kent is dat nog niet zo een groot probleem. Als de organisatie meer dan een paar honderd werknemers heeft is dit overzicht echter niet meer aanwezig. Het is dan moeilijk een helder inzicht te krijgen in waar de organisatie de facto voor staat (let wel: de organisatie als geheel, niet alleen het management) en waar het eventueel mis gaat in de interactie met

de klant. Hier zou het heel plezierig zijn als we zouden kunnen beschikken over eenzelfde instrumentarium om de identiteit van een organisatie te meten, zoals we beschikken over een instrumentarium om het imago ervan te meten. Dit artikel wil een stap in die richting zetten.

Allereerst zullen Albert en Whetten's (1986) criteria voor definities van Corporate Identity behandeld worden. Vervolgens komen meetmethoden aan bod, zoals die door auteurs vanuit verschillende invalshoeken ontwikkeld zijn. Deze worden daarna aan Albert en Whetten's criteria getoetst. Hoewel geen enkele tot nu toe ontwikkelde methode de toets van Albert en Whetten's criteria volledig doorstaat, biedt met name Weber's (1985) methode van 'verstehende Rekonstruktion' interessante perspectieven. Gebruik makend van Weber's basisprincipe en de in de marketing al langer bekende laddering-techniek wordt een nieuwe methode ontwikkeld. Deze wordt aan de hand van een voorbeeld geïllustreerd.

2. ALBERT EN WHETTEN'S CRITERIA VOOR DEFINITIES VAN CORPORATE IDENTITY

Veel inspanningen op het gebied van bedrijfscommunicatie zijn erop gericht een door het management gewenst imago te creëren, zowel buiten als binnen de organisatie. Dowling (1986) definieert imago als: "het geheel van betekenissen waardoor een object bekend is en aan de hand waarvan mensen het beschrijven, het zich kunnen herinneren en ermee omgaan". Het imago is het nettoresultaat van de interactie van iemands opvattingen, ideeën, indrukken en gevoelens met betrekking tot een object. Het zit in het hoofd van die persoon. Analoog aan Dowling's definitie van 'imago' valt 'identiteit' te omschrijven als "het geheel van betekenissen waardoor een object zich laat herkennen en waardoor het mensen in staat stelt het te omschrijven, het zich te herinneren en er mee om te gaan". Terwijl het 'corporate image' in het hoofd van de doelgroepen zit, zit de 'corporate identity' in de organisatie. Dit is een definitie, die nog steeds alle mogelijke invalshoeken op die identiteit toestaat. Als eerste stap om tot een eenduidige benadering te komen, formuleerden Albert en Whetten (1985) drie criteria waaraan beschrijvingen van de identiteit van een organisatie moet voldoen.

Hun eerste criterium is dat van 'centraal karakter'. Daarmee bedoelen Albert en Whetten, dat beschrijvingen de organisatie moeten kenschetsen op basis van iets dat belangrijk en essentieel is voor de organisatie zelf. Zij denken dan aan die eigenschappen die doorslaggevend zijn bij het nemen van belangrijke beslissingen. Managers proberen vaak de kerneigenschappen van een organisatie vast te stellen als een leidraad voor hoe ze moeten handelen en hoe ze tegenover andere organisaties moeten staan. Wat daarbij lastig is, is dat de meeste organisaties zich niet bewust zijn van hun centrale eigenschappen. Die liggen verankerd in de dagelijkse gang van zaken, en zijn zo vanzelfsprekend, dat men ze zelden of nooit onder woorden brengt.

Albert en Whetten's tweede criterium is dat van 'onderscheidend vermogen'. Het gaat erom, die eigenschappen te beschrijven waarin de organisatie herkenbaar van anderen verschilt. Dit criterium is lastig te operationaliseren. Hoe uniek een eigenschap is hangt helemaal af van met wie een organisatie vergeleken wordt. Als het om bijvoorbeeld de

vier naaste concurrenten gaat is dat heel anders, dan als het om alle organisaties in bijvoorbeeld de Benelux gaat. Dit betekent wel, dat meetmethodes open moeten staan voor wat een organisatie aan unieks te bieden heeft.

Albert en Whetten's derde criterium is dat van 'continuïteit in de tijd'. De te beschrijven eigenschappen moeten hetzelfde blijven in de loop van de tijd, ook onder veranderende omstandigheden. Uiteraard geldt dit binnen bepaalde grenzen. In de loop van de tijd verandert elke organisatie. Waar het om gaat is dat een meetinstrument voor corporate identity in elke geval in staat is, die elementen boven water te krijgen die relatief constant zijn in de tijd.

Beschrijvingen van organisaties die voldoen aan Albert en Whetten's criteria geven eigenschappen aan, die met redelijke zekerheid toe te schrijven zijn aan de organisatie: McArthur (1972) onderzocht onder welke omstandigheden toeschouwers geneigd waren gedrag aan mensen toe te schrijven. Zij constateerde, dat als mensen hetzelfde onder verschillende omstandigheden deden, als zij dat vaker herhaalden in de tijd en als alleen die mensen dat gedrag vertoonden, dat toeschouwers dat gedrag dan ook vaker aan hun toeschreven. De functie van Albert en Whetten's criteria is, dat ze die eigenschappen van de organisatie selecteren waarvan redelijkerwijs verwacht kan worden dat men ze zal herkennen als eigenschappen van de organisatie.

3. BESTAANDE EMPIRISCHE BENADERINGEN VAN CORPORATE IDENTITY

Albert en Whetten laten de invalshoek open van waaruit de eigenschappen van de organisatie beschreven moeten worden. Daar zijn meerdere mogelijkheden voor denkbaar. In de literatuur (van Riel, 1996, Dutton, Dukerich & Harquail, 1994) zijn drie verschillende benaderingen van het begrip "identiteit" te vinden: identiteit als bundel eigenschappen, zoals die door een onderzoeker worden vastgesteld, identiteit zoals het imago dat de eigen mensen van een organisatie hebben, en ten derde, de mate waarin men zich met een organisatie identificeert. Elk van deze benaderingen wordt hieronder kort besproken, met de daaruit voortvloeiende meetmethoden.

3.1. Identiteit benaderd als bundel door een onderzoeker vast te stellen eigenschappen

De literatuur biedt een drietal methoden die de identiteit van de organisatie als geheel meten, op een manier waarbij een onderzoeker zijn eigen meetresultaten verwerkt (van Riel, 1996). De methoden variëren in de mate waarin de onderzoeker zelf de resultaten ordent. Lux (1988) gebruikt daarvoor zeven vaste dimensies. Balmer (1996) richt zich specifiek op missie en strategie van de onderneming, om die als verder filter voor zijn meetmethode te gebruiken. Weber (1985) loopt als participerend observator door de organisatie en richt zich specifiek op hoe organisatieleden hun gedrag legitimeren.

3.1.1. Lux' stermethode

Lux' (1988) definieert Corporate Identity als "de optelsom van de eigenschappen van een organisatie". Als onderzoeker is Lux zelf degene die eigenschappen van een organisatie waarneemt. Hij deelt ze in in een zevental dimensies, die hij baseert op Guil-

ford's (1959) onderzoek naar basisdimensies van menselijke persoonlijkheid. Guilford had destijds over 4000 bijvoeglijke naamwoorden die mensen gebruikten om elkaar te beschrijven een factoranalyse uitgevoerd en was zo tot deze basisdimensies gekomen:

1. Behoeften: de interne en externe motieven en behoeften van mensen. Deze vormen volgens Lux de centrale aspecten van de persoonlijkheid van een organisatie en liggen aan de basis van het gedrag. Voorbeelden zijn: groei, zekerheid, en een gezond bedrijfsklimaat.
2. Competenties: de onderscheidende competenties ("distinctive competences") van een organisatie.
3. Attitude: de filosofie van de organisatie en de politieke achtergrond. Deze dimensie geeft weer, hoe de organisatie zichzelf en haar omgeving ziet.
4, Constitutie: de fysieke locatie, de structuur en de rechtsvorm van de organisatie.
5. Temperament: de manier waarop de organisatie dingen voor elkaar krijgt of juist laat zitten. Deze dimensie legt de sterkte, de intensiteit, de snelheid en de emotionaliteit van het gedrag van een organisatie vast.
6. Herkomst: het verband tussen de persoonlijkheid van de organisatie en haar verleden.
7. Belangen: het middellange- en langetermijn doelstellingen van de organisatie.

In wezen is Lux' methode een checklist bij het interviewen en observeren van werknemers en bij het onderzoek naar de documenten van en over de onderneming. Lux legt de manier waarop gegevens moeten worden verzameld niet vast. Hij schrijft alleen de wijze van samenvatten voor: alles moet tegen de zeven dimensies afgezet worden.

3.1.2. Balmer's Affinity Audit

Balmer (1996) ontwierp een methode om de krachten die de identiteit van de organisatie in stand houden te verklaren. Balmer gebruikt daarbij de principes van de 'grounded theory' (Strauss en Corbin, 1991) om de processen aan het licht te brengen die aan de identiteit van een organisatie ten grondslag liggen. Werknemers hebben een affiniteit met bepaalde waarden en aannames. Deze waarden en aannames leggen de identiteit vast. Balmer onderzoekt deze waarden en aannames door het dagelijkse taalgebruik, de ideologieën en de rituelen op te tekenen. Hij gebruikt een hele variëteit aan dataverzamelingsmethoden, voornamelijk semi-gestructureerde interviews, observatie en de door de organisatie geproduceerde documentatie. Balmer doet zijn onderzoek in vier fasen, waarvan de tweede de eigenlijke meting behelst:

1. stel de missie en de strategie van de organisatie vast
2. stel het dominante systeem van waarden en aannames in de organisatie vast
3. beoordeel het dominante systeem van waarden en aannames in het kader van de missie en de strategie
4. stimuleer die waarden en aannames die de missie en de strategie van de organisatie ondersteunen

3.1.3. Weber's methode van "verstehende Rekonstruktion"

Weber (1985) stelt een methode voor die hij de "verstehende Rekonstruktion" van de identiteit van een organisatie noemt. Werknemers voeren handelingen uit. Deze hande-

lingen zijn doelgericht. De taak van Weber's onderzoeker is, dit doelgerichte gedrag van organisatieleden te begrijpen. Daarvoor verplaatst de onderzoeker zich in het organisatielid. Hij maakt een 'rationele reconstructie' van de intuïtieve kennis van de organisatie. Weber gaat ervan uit, dat een handeling te begrijpen is vanuit het doel waar die handeling op gericht is. De onderzoeker gaat daarom een discussie met organisatieleden aan, waarin dezen hun gedrag als 'rationeel' moeten legitimeren. Weber draagt twee methoden aan om dit te bereiken: ten eerste stelt hij het 'narratieve interview' voor, waarin hij zijn respondent vraagt om alles te vertellen wat in hem opkomt. Zodra de respondent aan het eind van zijn verhaal komt, vraagt de interviewer verder. Zijn taak is om de respondent te laten praten. In aanvulling daarop stelt Weber antropologische methodes voor om de manieren aan het licht te brengen waarop werknemers "altijd al" hun sociale wereld gevormd en georganiseerd hebben. Hij baseert zich daarbij op Garfinkel's (1967) methode van doelbewuste vervreemding. Daarvoor stelt hij routine-handelingen uit het dagelijks leven bij wijze van 'onaangename verrassing' ter discussie. Weber stelt zijn respondenten provocerende vragen, doet alsof hij dingen niet begrijpt, en blijft koppig om uitleg vragen, waar organisatieleden vinden dat ze al lang alles uitgelegd hebben. Zo hoopt Weber datgene boven water te krijgen dat anders onbesproken blijft, omdat het zo vanzelfsprekend is.

Weber is vooral therapeutisch bezig. Hij vertelt organisatieleden niet, wat volgens hem hun identiteit is, maar hij helpt ze om zelf hun identiteit te reconstrueren. Hij gaat daarmee net zolang door, totdat een zekere consensus ontstaat onder de organisatieleden over wat de organisatie is. Weber geeft daarmee zijn organisatieleden inzicht in het waarom van wat ze doen, zonder dat hij overigens komt tot een empirische weergave van wat hij gemeten heeft.

3.2. Identiteit als waargenomen eigenschappen ("intern imago")

3.2.1. Lief's Working Impressions-vragenlijst

Lief (1995) heeft een vragenlijst ontwikkeld om de abstracte eigenschappen boven water te halen die organisaties kenmerken. In zijn visie komt betekenis altijd naar voren in een context de wordt gevormd door indrukken, houdingen en gevoelens. Lief heeft daarom een vragenlijst ontwikkeld, die bestaat uit 75 woordparen, zoals "warm-koud", "gekruid-mild". Hij vraagt werknemers van elk woordpaar eerst dat woord te omcirkelen dat weergeeft hoe ze zelf hun organisatie zien, en vervolgens van hetzelfde woordpaar dat woord waarmee een kwaadaardige concurrent de organisatie in kwestie zou omschrijven. Bij de analyses kijkt Lief alleen naar die woorden, die unaniem door alle respondenten omcirkeld zijn. Alleen die komen in de uiteindelijke verslaggeving terecht. Op die manier wordt consensus onder organisatieleden zichtbaar gemaakt. Zo ontwikkelt Lief een nieuwe vocabulaire voor de organisatie, die als een gemeenschappelijke taal kan dienen om over de organisatie te praten.

3.2.2. Atamer en Calori's "Diagnostic Général"

In Frankrijk ontwikkelden Atamer en Calori (1993) een methode voor exploratieve analyse van de identiteit van organisaties. Zij onderscheiden vier componenten van identiteit, die elk apart gemeten worden. Dit zijn de oriëntatie van de organisatie in zijn algemeenheid ("finalité"), de waarden en gedragsnormen ("culture"), het managementsysteem ("système de management") en de basisstrategie van de onderneming

("recette stratégique"). Bij het meten van de "finalité" vragen Atamer en Calori managers aan te geven, in welke mate zij prioriteit geven aan groei, rendement, en de maatschappelijke waarde van de organisatie, hoe ze tegenover het dragen van risico staan, hoe tegenover de onafhankelijkheid van hun onderneming en hoe tegenover prestige. Bij het meten van de waarden en normen van de organisatie vragen ze managers aan te geven, in welke mate de waarden uit het mission statement in de praktijk gebracht worden. Heeft de organisatie geen mission statement, dan geeft elk lid van het directieteam de vijf meest kenmerkende waarden van de organisatie weer. Voor elke waarde moeten ze dan twee manieren aangeven, waarop die waarde in de praktijk gebracht wordt. De derde component van de identiteit van een organisatie wordt gevormd door het management systeem. Hiervoor stellen Atamer en Calori de volgende vier vragen: welke concrete gebruiken het planningssysteem helder maken, welke gebruiken de houding tegenover risico weerspiegelen, welke de decentralisatie van beslissingen weergeven en welke juist de centralisatie. Atamer en Calori's vierde component van identiteit is de basisstrategie van de organisatie, zoals managers dat waarnemen: het domein van de activiteiten van de organisatie, de "distinctive competences" en de feitelijke strategie.

Atamer en Calori's meetmethode is geschikt om een consultant een indruk te geven van het strategische profiel van een organisatie. Maar daarvoor wordt alleen topmanagers naar hun mening gevraagd en geeft het eerder het image van de organisatie weer, zoals topmanagers dat erop nahouden. Atamer en Calori vragen managers te reflecteren op dimensies, die zij als auteurs van te voren als belangrijk hebben aangemerkt. Hun enige echt open vragen zijn die naar de dominante culturele kenmerken van de organisatie.

3.2.3. Bernstein's methode

Bernstein (1986) ontwikkelde een methode om identiteit te meten die heel eenvoudig is toe te passen bij een directieteam. Hij is vooral bruikbaar om snel overeenstemming te krijgen in een kleine groep over wat de identiteit van een organisatie inhoudt en waar eventueel actie moet worden ondernomen. Aan het begin van een sessie wordt elk van de directieleden gevraagd een lijst te maken van die organisatie-eigenschappen, die een belangrijke rol spelen en dat ook voor de toekomstige ontwikkeling van de organisatie zullen doen. Alle belangrijke waarden dienen meegenomen te worden. Daarop volgt een groepsdiscussie, waarin het totale aantal waarden wordt gereduceerd tot een achttal. De discussie is pas beëindigd, als alle aanwezigen het erover eens zijn welke acht waarden werkelijk belangrijk zijn. Elk van deze acht eigenschappen wordt vervolgens beoordeeld op een negenpunts-schaal, en wel op twee manieren. Ten eerste: de score die de organisatie volgens hun op deze eigenschappen zou moeten halen (de soll-scores). Ten tweede: de score die de organisatie op dit moment op deze eigenschappen haalt (de ist-scores). De belangrijkste meerwaarde van Bernstein's (1986) methode is managers bewust te maken van de dimensies waarin ze denken. Het is een goede ingang om discussie binnen de organisatie op gang te brengen. Het maakt interne conflicten binnen de directie zichtbaar en helpt bij het opstellen van een gewenste corporate identity. Een beperking van de methode is ook hier, dat alleen de mening van het topmanagement gevraagd wordt.

3.3. De mate van identificatie als identiteitsbenadering

Sommige auteurs concentreren zich op het effect van corporate identity op de mate waarin werknemers zich met hun organisatie identificeren (Keller, 1990, Elsbach en Glynn, 1996). Zij geven er de voorkeur aan, niet de identiteit van de organisatie sec te meten, maar de identificatie van de werknemer met de organisatie. Op dit terrein zijn twee methoden ontwikkeld en in de praktijk gevalideerd: de Mannheimer Corporate Identity test (Keller, 1990) en de Rotterdam Organization Identification Test (Smidts, van Riel en Pruyn, 1998).

3.3.1. Keller's Mannheimer CI-test

Keller (1987, 1990) richt zich op de gevoelens van identificatie die de identiteit van een onderneming bij doelgroepen oproept, zowel binnen als buiten de onderneming. Deze hebben hun weerslag op de cultuur van de organisatie, en daardoor op het gedrag (Keller, 1990). De door haar ontwikkelde Mannheimer CI-test onderscheidt vijf dimensies, die werknemersidentificatie meten. Op enkele dimensies vinden twee metingen plaats: één over de aanleg van werknemers om hogere scores op die dimensie te halen, en één met betrekking tot de score van de werknemers in de context van hun huidige baan bij hun huidige werkgever. De dimensies zijn de volgende:

1. Het vermogen van mensen om zich überhaupt met iets te identificeren, en met hun huidige werkgever in het bijzonder
2. De behoefte om te presteren: de bereidheid van werknemers om in het algemeen te presteren, en in hun huidige baan in het bijzonder.
3. Tevredenheid: de mate waarmee mensen tevreden zijn met hun werk. De tevredenheid met hun baan als zodanig wordt dan afgezet tegen hun tevredenheid met hun baan in de organisatie waar zij nu werken.
4. Organisatiestijl: de manier waarop informatie binnen de organisatie verstrekt wordt, de stijl van leiding geven en het bedrijfsklimaat.
5. De voorbeeldfunctie: de mate waarin mensen het de moeite waard vinden om zicht met de organisatie te identificeren.

In Duitsland heeft Keller een databank ontwikkeld, waar de scores van organisaties op deze dimensies vergeleken kunnen worden met het branchegemiddelde. De Mannheimer CI-test is vooral bruikbaar als een diagnose-instrument voor de gezondheid van de organisatie: als de scores van werknemers op de eerste drie dimensies bij de algemene variant van de vraag significant hoger ligt dan hun score als de huidige werkgever in het geding is, dan is dat een teken dat er iets mis is met de organisatie. Met name de dimensies van 'tevredenheid' en 'prestatiebehoefte' vervullen die functie.

3.3.2. Rotterdam Organization Identification Test (ROIT)

De Rotterdam Organization Identification Test (ROIT) is een korte vragenlijst van 15 items, die de identificatie van werknemers met hun organisatie vaststelt. Gebaseerd op de sociale identiteitstheorie (Ashforth & Mael, 1989) meet de vragenlijst, in welke mate organisatieleden:

- het gevoel hebben dat ze erbij horen
- vinden dat hun persoonlijke doelen overeenkomen met de doelen en waarden van de organisatie

– het fijn vinden om lid van de organisatie te zijn
– hun behoefte om zich bij een groep aan te sluiten
– voordelen ervaren van hun organisatielidmaatschap
– het gevoel hebben gesteund te worden
– een gevoel van erkenning hebben
– het gevoel hebben geaccepteerd te worden
– het gevoel van geborgenheid hebben

De beknoptheid van de vragenlijst maakt de ROIT erg geschikt voor herhalingsmetingen over een langere periode. Smidts, van Riel en Pruyn (1998) zijn bezig met het opbouwen van een databank waarin de gegevens van verschillende organisaties met elkaar vergeleken kunnen worden.

4. VERGELIJKING VAN DE METHODEN MET ALBERT EN WHETTEN'S CRITERIA

Elk van de acht besproken methoden is met een ander doel op de achtergrond ontwikkeld. Daardoor verschillen ze in de mate, waarin ze voldoen aan Albert en Whetten's (1985) criteria.

4.1. Albert en Whetten's criterium van 'centraal karakter'

Albert en Whetten's criterium van centraal karakter houdt in, dat de uitkomst van de meting de essentie van de organisatie moet weergeven. Lux (1988) zet alles wat hij waarneemt tegen zijn zeven dimensies af. Het is de vraag, in hoeverre de methode niet eerder weergeeft wat de onderzoeker zelf centraal vindt. Voor Balmer (1996) geldt dit in mindere mate. Weber (1985) heeft daar de meest doelgerichte systematiek voor ontwikkeld. Zijn continue provocaties dragen ervoor zorg, dat die aspecten boven water komen die mensen in hun gedrag aansturen. Atamer en Calori (1993) nemen genoegen met de inschattingen van topmanagers op door hen voorgelegde items. Bernstein (1986) ontwikkelt een mechanisme om de meest centrale eigenschappen eruit te filteren: hij laat zijn topmanagers net zo lang discussiëren tot die acht eigenschappen eruit komen, die iedereen het belangrijkst vindt voor de hele organisatie. De beide methoden om identificatie te meten en de methode van Lief (1995) zijn vaste vragenlijsten. Aan de hand van de scores is wel vast te stellen welke items meer aanslaan in de organisatie, maar niet of items uit de vragenlijst er echt centraal staan.

4.2. Albert en Whetten's criterium van 'onderscheidend vermogen'

Onderscheidend vermogen als zodanig is moeilijk aan de hand van de meting van één enkele organisatie vast te stellen. Wel verschilt de openheid van de meting sterk per methode. Open vragen stellen organisatieleden in staat om met items te komen waar de onderzoeker zelf niet aan gedacht heeft. Een tweede aspect van openheid is de manier waarop de onderzoeker de data samenvat. Hier kan hij zijn eigen dimensies aanbrengen, of de organisatieleden zelf de aanwijzingen laten geven hoe de data tot de kernpunten samen te vatten. Lux (1988) staat tijdens zijn dataverzameling open voor mogelijk organisatie-specifieke items. Alleen vat hij alles samen volgens zijn eigen zeven

dimensies – waardoor zijn meting aan organisatie-specificiteit inboet. De grounded theory die Balmer (1996) gebruikt staat hem toe zich te laten leiden door wat zijn onderzoeksobject hem vertelt. Voor Weber's (1985) methode van vragen, observeren en provoceren geldt hetzelfde. Bernstein (1986) laat de items volledig door zijn respondenten aandragen, en laat hen ook vrij in de keuze van de kernelementen daarin. Bernstein's methode is echter gevoelig voor wat Martin et al. (1983) de 'uniqueness paradox' noemen: als mensen hun eigen organisatie moeten omschrijven, komen ze bij voorkeur met die eigenschappen waarin ze juist niet van andere organisaties verschillen. De identificatie testen, de Mannheimer CI-test (Keller, 1987) en de ROIT (Smidts, van Riel en Pruyn, 1998) zijn vragenlijsten met vastliggende items. Net zoals de methode van Lief (1995) zijn zij zeer geschikt om organisaties onderling te vergelijken, maar niet om organisatie-specifieke kenmerken op te sporen.

4.3. Albert en Whetten's criterium van 'continuïteit in de tijd'

De mate waarin de gevonden eigenschappen van een organisatie duurzaam zijn is op twee manieren te benaderen. Het meest voor de hand ligt het herhalen van de meting. De vragenlijsten van Lief (1995) en de beide identificatietesten (Keller, 1987, Smidts, van Riel en Pruyn, 1998) zijn eenvoudig voor een herhalingsmeting in te zetten. Dat geldt met name voor de ROIT, die met 15 items heel snel ingevuld is (Smidts, van Riel en Pruyn, 1998). Weber (1985) zoekt structureel naar aanwijzingen voor continuïteit. Hij gaat ervan uit, dat die doelen en waarden die aan de basis liggen van het concrete gedrag, de duurzame kern van de identiteit van de organisatie vormen.

Tabel 1 geeft een overzicht van de methoden. Drie lijken er wat dichter bij de criteria van Albert en Whetten (1985) te komen: Balmer's Affinity Audit, Bernstein's methode en Weber's methode van 'verstehende Rekonstruktion'. Deze laatste methode biedt de meeste perspectieven. Weber's continue vraag tot legitimatie vormt het meest gerichte zoekalgorithme, en beperkt zich niet tot het topmanagement. Verder is Weber's methode niet gevoelig voor Martin's (1983) paradox van uniciteit, die inhoudt dat mensen graag dingen claimen waarin ze juist niet verschillen van anderen. Weber heeft een heel interessante aanzet voor een bruikbare meetmethode gegeven. Hij komt alleen niet tot een daadwerkelijke meting.

5. LADDERING ALS METHODE OM CORPORATE IDENTITY TE METEN

Het is de uitdaging, de kracht van Weber's methode te benutten voor een methode die de identiteit van de organisatie systematisch vastlegt. Weber's sterke punt is, dat hij zijn respondenten vrijuit laat praten, maar ondertussen steeds weer doorvraagt naar nieuwe legitimaties voor wat ze doen. Dit tweede aspect is kenmerkend voor de in de marketing al langer bekende laddering-techniek (Reynolds en Gutman, 1988). Om de gedachtenwereld van de respondent zo min mogelijk te beïnvloeden verdient het de aanbeveling de "zachte" versie van de laddering-techniek toe te passen (Grunert & Grunert, 1995). Deze variant laat de respondent zoveel mogelijk vrijuit kan praten. De openheid voor het denken van de respondent die dit met zich meebrengt, maakt deze techniek heel geschikt om corporate identity te meten. De combinatie van Weber's aan-

Tabel 1. Vergelijking van de bestaande methoden om corporate identity te meten.

Meetmethode	Doel waarvoor de meetmethode ontworpen is	Informanten	Soort vragen	Methodologie
Lux's ster-methode	Checklist om de identiteit van de organisatie vast te stellen	steekproef van alle werknemers	scores van de organisatie op zeven vastgeleg-de dimensies	observatie open interviews
Balmer's Affinity Audit	Overeenstemming tussen missie en het dominante waardesysteem vaststellen	mission statement and steekproef van werknemers	observatie, halfgestructu--reerde inter-views, docu-mentatie	grounded theory
Weber's 'verstehende Rekonstruktion'	Organisatieleden helpen zelf hun corporate identity vast te stellen	steekproef van alle werknemers	werknemers wordt gevraagd te legitimeren wat ze in hun werk doen	open gesprek-ken en con-fronterende discussies met werknemers
Lief's Working Impressions™ methode	Consensus over beschrijvingen van de organisatie vast stellen	steekproef van alle werknemers	vaste batterij van vragenlijst-items	enquête
Atamer and Calori's 'Diag-nostic Général'	Een snelle organisatiediagnose uitvoeren	managers	vaste vragenlijst	enquête
Bernstein's methode	Een discussie over de gewenste identiteit van de organisatie op gang brengen	managers	Dimensies die het meest rele-vant zijn voor de strategie van de organisatie	groepsdiscussie
Mannheimer CI. test	De identificatie van werknemers met de organisatie meten	steekproef van alle werknemers	vaste batterij van binaire schalen	enquête
Rotterdam Organization Identification Test (ROIT)	De identificatie van werknemers met de organisatie meten	steekproef van alle werknemers	vaste set van zevenpunts schalen	enquête

pak met de systematiek van de laddering-techniek lijkt een goede basis om een meet-methode voor corporate identity te ontwerpen en uit te testen.

Als uitgangspunt dient een zo divers mogelijke steekproef onder organisatieleden. Zo komt de maximale variëteit aan functies in de organisatie aan bod. De interviewer probeert eerst zo concreet mogelijk in beeld te krijgen wat de organisatieleden precies doen. Het antwoord op de vraag: "wat is uw functie?" is makkelijk te geven. De beschrijving van de functie, die de werknemer dan geeft, biedt niet altijd de beste aan-

	Wat wordt er gemeten	Uitkomst	Vermogen centraal karakter te meten	Vermogen onderscheidend vermogen te meten	Vermogen continuïteit te meten	Literatuur
	waarnemingen van werknemers en van de onderzoeker	scores van de organisatie op zeven vastliggende dimensies	matig	matig	gering	Lux (1988)
	overeenstemming van de waarden en aannames van werknemers en de missie	overeenstemming tussen missie, strategie en dominant systeem van waarden en aannames	matig	matig	matig	Balmer (1996)
	centrale waarden en legitimaties	het begrijpen van de identiteit van de organisatie	redelijk	matig	redelijk	Weber (1985)
	waarnemingen van werknemers	eigenschappen waarop volledige overeenstemming bestaat	matig	gering	gering	Lief (1995)
	waarnemingen van managers	score van de organiaatie op vastliggende schalen	gering	gering	gering	Atamer & Calori (1993)
	waarnemingen van managers	8 kerneigenschappen met de waargenomen scores van de organisatie daarop	redelijk	gering	matig	Bernstein (1986)
	identificatie van werknemers met de organisatie	diagnose van werknemersidentificatie	gering	gering	gering	Keller (1987)
	identificatie van werknemers met de organisatie	diagnose van werknemersidentificatie	gering	gering	matig	Smidts, van Riel & Pruyn (1998)

knopingspunten. Vaak geven werknemers dan eerder weer wat van hun verwacht wordt in plaats van wat ze concreet doen, of ze geven een zo algemene omschrijving, dat er heel moeilijk een beeld bij te vormen is, zoals bijvoorbeeld bij het antwoord: “ik zorg dat de zaken op deze afdeling soepel verlopen”. De interviewer blijft daarom eerst de vraag stellen: “wat moet ik me daar precies bij voorstellen?”. Zodra de interviewer als het ware kan naspelen wat het organisatielid doet, komt hij met de vraag: “Waarom doet u dat op die manier?”. Na het antwoord daarop stelt hij voor het eerst de typische laddering-vraag: “Waarom is dat belangrijk voor u?”. Deze vraag blijft de interviewer

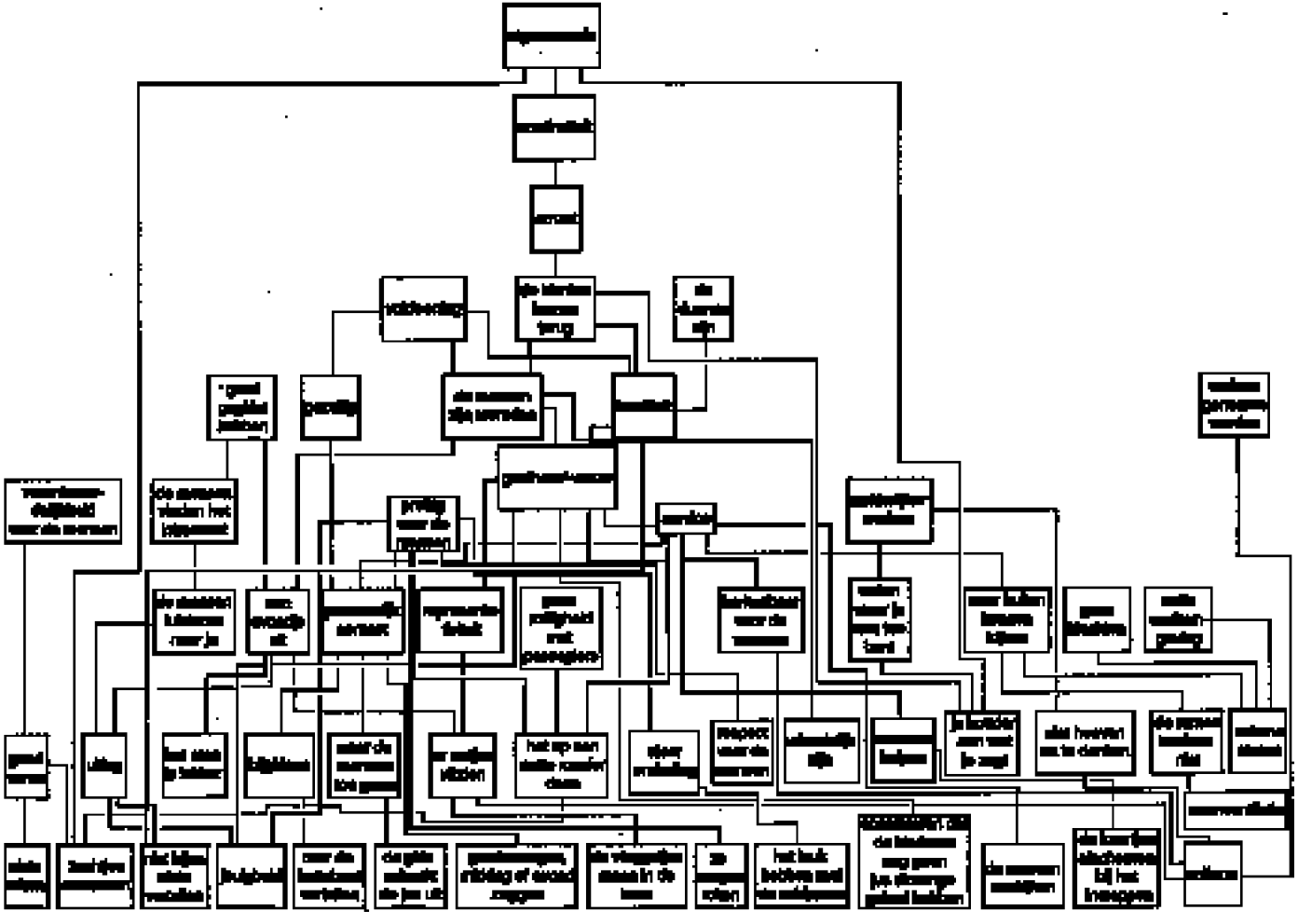
na elk antwoord opnieuw stellen, net zolang, tot het niet meer gaat. Als het interview zich dan verder weg beweegt van de concrete handeling, waar de keten mee begon, wordt het voor de respondent steeds moeilijker het antwoord onder woorden te brengen – juist omdat het vaak zo vanzelfsprekend is. Het vanzelfsprekende is precies het terrein, waar Weber (1985) de stabiele, duurzame kern verwacht van alles wat de organisatie doet.

Op het niveau van de individuele medewerkers worden zo middel-doelketens vastgelegd, die lopen van concrete handelingen naar de diepere achterliggende waarden. Op hoger niveau overlappen die ketens voor individuen afzonderlijk, en ook voor meerdere individuen als groep. De concepten, waarop deze ketens elkaar overlappen zijn het interessantst. Dit zijn de concepten, die meerdere organisatieleden gemeenschappelijk hebben, en die voor verschillende functies binnen de organisatie resulteren in verschillende uitwerkingen. Zij vormen de rode draad door de grote variëteit van handelingen van individuele organisatieleden.

Figuur 1 geeft de uitkomsten weer van zes laddering-interviews bij een rondvaartbedrijf in Amsterdam. Elk concept dat in de figuur is opgenomen is door minimaal twee respondenten genoemd, elke relatie door minimaal één respondent. Op organisatieniveau zijn die handelingen, doelen en waarden belangrijk, die door meer dan één persoon genoemd zijn. Anders zou er geen aanwijzing zijn voor een breder draagvlak binnen de organisatie. Figuur 1 geeft het resultaat weer van zes interviews bij de rederij. Alle aspecten die in figuur 1 vermeld zijn, zijn door minimaal twee werknemers genoemd. Alle relaties zijn door minimaal één medewerker genoemd. De middel-doel relaties vormen hele ketens, die door de figuur heen te volgen zijn: onderaan staat “de vlaggetjes staan in de kaas”. Een lijn verbindt dit met “er netjes uitzien”, en zo is het hele pad verder omhoog te volgen, via “representativiteit”, “gastheer/gastvrouw”, “de mensen zijn tevreden” tot “voldoening”.

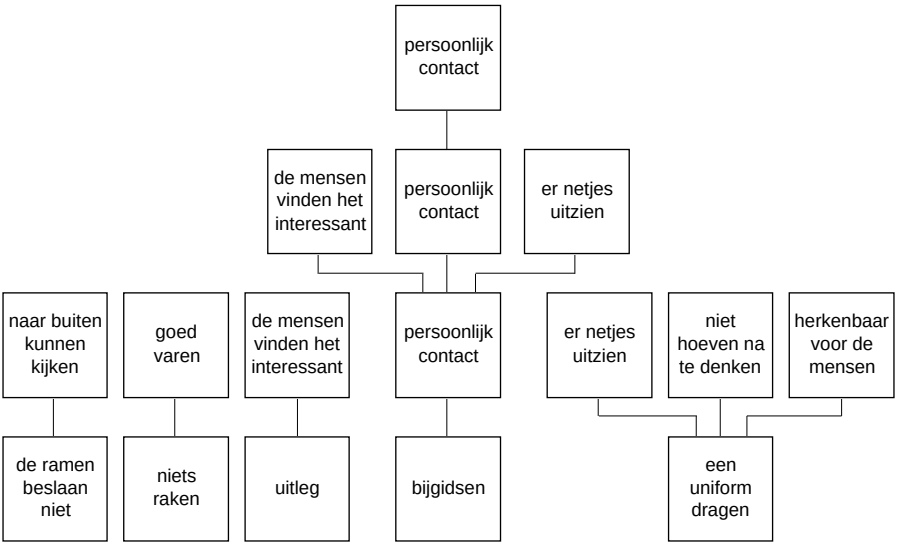
De concrete handelingen onderin de figuur zijn specifiek voor de functies van individuele werknemers. Helemaal links onderaan, bijvoorbeeld, zien we “zachtjes aanmeren”, dat een middel is om het doel “goed varen” te bereiken. Dat is typisch iets voor kapiteins. Hoe hoger de aspecten in de figuur staan, hoe algemener ze zijn. Op een gegeven moment wordt het zelfs zo algemeen, dat het niet meer alleen voor het specifieke bedrijf geldt. Zo zien we helemaal bovenaan “eigenwaarde”, en “omzet”. Dat zijn aspecten die bij veel meer organisaties in Nederland belangrijk zijn.

Voor een kernachtige samenvatting van de eigen identiteit is het middenstuk van figuur 1 het interessantst. Hier staan aspecten, die boven het functiespecifieke niveau uitstijgen, maar nog niet zo algemeen zijn dat ze ook voor elke andere organisatie gelden. Opvallend is hier de positie van “een goede gastheer/gastvrouw zijn”, centraal in de figuur. Alle concrete handelingen zijn namelijk hierop terug te voeren: van elk van de concrete handelingen onderaan loopt wel een pad naar dit “gastheerschap”. Bovendien het is een middel bij het bereiken van alle hogere doelen. Dit begrip vormt als het ware een “flessehals” waar het hele denken en doen van de werknemers van dit rondvaartbedrijf doorheen gaat.



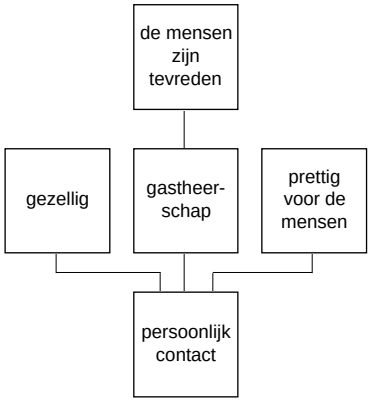
Figuur 1. Onderzoeksresultaten bij een rondvaartrederij in Amsterdam.

De figuur is vrij eenvoudig te versimpelen, door het aantal keren dat een aspect of een relatie minimaal genoemd moet worden te verhogen. Als het aantal mensen dat een relatie minimaal moet noemen ook op twee wordt gesteld, ontstaat de veel eenvoudigere figuur 2.



Figuur 2. De versimpelde figuur van Holland International rondvaart.

De structuur valt nu uiteen in een aantal verschillende verhalen, waarvan er een aantal duidelijk te herkennen zijn als typisch voor bepaalde functies. Het "niets raken" om "goed te varen" is bijvoorbeeld typisch iets voor kapiteins, net zoals "de ramen beslaan niet", zodat mensen "naar buiten kunnen kijken". Het geven van "uitleg", zodanig dat "mensen het interessant vinden" is weer typisch iets voor gidsen. Het enige deel dat voor alle functies geldt is het meest uitgewerkte gedeelte, waarin "persoonlijk contact" en "gastheerschap" centraal staan. Als we het minimum aantal mensen dat een aspect vermeld moet hebben nog verder opschroeven, naar minimaal 3, vallen de functiespecifieke verhalen weg. Dan blijft een heel eenvoudige harde kern over, die een breed draagvlak binnen de organisatie heeft (figuur 3).
Door het verhogen van het minimum aantal vermeldingen is heel makkelijk een bondige samenvatting van de onderzoeksgegevens te maken. Het nadeel van deze manier van werken is echter, dat aspecten die veel worden genoemd, maar steeds in wisselende contexten, net zo goed kenmerkend voor de organisatie kunnen zijn, maar zo ten onrechte uit de meetresultaten verdwijnen. De sociale netwerkliteratuur (Wasserman en Faust, 1994) biedt mogelijkheden om dit probleem te ondervangen door de centraliteit van de concepten te berekenen aan de hand van het relatiepatroon. Freeman (1979) stelt als maat voor het totale aantal relaties dat een concept met alle andere concepten heeft. Dit is de som van het aantal keren dat een concept als doel wordt genoemd en het aantal keren dat het als middel wordt genoemd (vgl. Pieters, Baumgartner en Stad, 1994). Op deze manier zijn vrij eenvoudig de centrale elementen in een structuur op te sporen. De centraliteit van een concept, op deze manier berekend, geeft aan, in hoeverre dat



Figuur 3. De allersimpleste figuur van Holland International Rondvaart.

concept in verband staat met alle andere handelingen die in een organisatie verricht worden. Hinkle (1965) vond, dat naarmate een concept vaker als doel of als middel werd genoemd, dat het dan ook stabieler was. De meest centrale concepten zouden daarmee wel eens de identiteit van de organisatie het beste kunnen samenvatten. Zo kwam bij de rondvaartrederij als meest centrale concept tevoorschijn het 'gastheer/gastvrouw zijn'. Dit concept staat ook grafisch ongeveer in het midden van de figuur. Alle concrete handelingen zijn erop terug te voeren (de lijnen van de concrete handelingen onderaan naar boven volgend), en het dient ook voor het bereiken van alle hogere doelen in de organisatie. Na 'gastheerschap' zijn 'persoonlijk contact', 'kwaliteit leveren' en 'er netjes uitzien' de meest centrale begrippen. Daarmee is in een viertal termen het geheel van wat de organisatie doet snel een eenvoudig weer te geven.

Deze methode lijkt aan de criteria van Albert en Whetten (1985) te voldoen. Ten eerste staan de doelen en waarden die in letterlijke zin het meest centraal staan in verband met alle heterogene handelingen die op de werkvloer plaatsvinden. Zij lijken de gemeenschappelijke motor achter het veel variabeler concretere gedrag van de organisatieleden. Uit onderzoek bij een grote automatiseerder bleek centraliteit berekend volgens de methode van Freeman (1979) ook significant te correleren met de mate waarin organisatieleden die doelen en waarden belangrijk vonden (een correlatie van 0.47, significant op 1 %; van Rekom, 1998). Centraliteit in letterlijke zin lijkt daarmee overeen te komen met Albert en Whetten's (1985) centraliteit in figuurlijke zin. Dit centraliteitsbegrip lijkt ook tegemoet te komen aan hun criterium van 'onderscheidend vermogen'. De hoogste waarden in de figuren, zoals 'omzet' en 'gevoel van eigenwaarde' zijn erg algemeen. De handelingen onderaan in de figuren, zoals 'goed varen' en 'bijgidsen', zijn heel functiespecifiek. Het middenniveau, waar we 'gastheerschap' en 'persoonlijk contact' vinden, lijkt op het niveau te zitten dat de organisatie als zodanig omvat, maar deze nog niet overstijgt. Centraliteit in deze structuur is ook een aanwijzing voor continuïteit van deze doelen en waarden in de tijd. Hinkle (1965) vond dat doelen en waarden die door veel middelen bereikt werden, en die veel andere doelen dienden, en dus centraal in een middel-doelstructuur stonden, veel minder onderhevig waren aan verandering. Van Rekom (1998) testte deze hypothese door het belang dat werknemers van een grote automatiseerder aan in laddering-onderzoek gevonden doelen en waarden

hechtten twee maal te meten, op een zevenpunts schaal, met een tijdsinterval van negen maanden. De verandering in het belang dat organisatieleden hechtten aan de gemeten doelen en waarden correleerde negatief met centraliteit (een correlatie van -0.47, 1 % significant). Dat wil zeggen: hoe meer een doel of waarde in de structuur centraal stond, hoe minder veranderlijk het belang dat eraan gehecht werd. In termen van Albert en Whetten's criteria lijkt de laddering-techniek een bruikbaar instrument om de identiteit van een organisatie te meten.[1]

6. CONCLUSIE

In de literatuur is een drietal invalshoeken op het meten van de identiteit van organisaties te onderscheiden. Ten eerste kan identiteit worden opgevat als het geheel van eigenschappen, zoals een onderzoeker die vaststelt. Een tweede invalshoek is, dat identiteit datgene is, dat de eigen organisatieleden vinden dat hun organisatie is. De derde stroming richt zich op het meten van de identificatie van werknemers met de organisatie waar ze werkzaam zijn. Methodes die de identiteit van een organisatie als geheel meten zijn in deze bijdrage kort weergegeven en vergeleken met de criteria die Albert en Whetten (1985) gebruiken voor definities van de identiteit van een organisatie: datgene dat tot uiting komt, moet centraal staan in het karakter van de organisatie, het moet specifiek voor de organisatie zijn, en het moet een zekere continuïteit in de tijd hebben.

De methodes die uitgaan van de identiteit, zoals die waargenomen wordt door de eigen organisatieleden worden daarbij gedwarsboomd door het verschijnsel, dat Martin et al. (1983) de 'paradox van de uniciteit' noemen: mensen claimen dat hun organisatie van andere organisaties verschilt in juist datgene, wat ze delen met andere organisaties. Dat wil zeggen: iedereen claimt dezelfde unieke eigenschappen. De organisatie-specificiteit is ook het zwakke punt van methodes, die met een gesloten vragenlijst de identiteit van een organisatie meten. Een gesloten vragenlijst is niet in staat om met items te komen die specifiek van toepassing zijn op die ene gemeten organisatie. Van de meer open methodes biedt de methode van 'verstehende Rekonstruktion' van Weber (1985) de meeste perspectieven. Hij zoekt heel systematisch naar de kern van onderliggende doelen en waarden, die organisatieleden aanstuurt in hun dagelijkse gedrag. Daarbij haalt hij, door zijn continue doorvragen, ook die doelen en waarden boven water, die zo vanzelfsprekend zijn, dat de organisatieleden er niet meer bij stilstaan. Weber zet zijn inspanningen echter niet om in een weergave van zijn meetresultaat.

De in de marketing al langer bekende laddering-techniek is bij uitstek geschikt om dit "koppige doorvragen naar de uitleg van betekenissen" systematisch aan te pakken en de structuur van die betekenissen in kaart te brengen. Deze structuur biedt een vruchtbare ingang om de identiteit van de organisatie te onderzoeken. De concepten die hier centraal in staan geven de identiteit van een organisatie het beste weer. De meetresultaten laten zien hoe centrale begrippen in verband staan met vrijwel alle concrete activiteiten van werknemers. Die doelen en waarden die een centrale positie in de structuur

[1] De laddering-techniek is als methode om de identiteit van organisaties te meten verder getest en gevalideerd in het proefschrift van Johan van Rekom (1998).

innemen lijken ook duurzamer in de tijd zijn. De toepassing van de laddering-techniek lijkt een geschikte methode op te leveren om de identiteit van de organisatie te meten. Deze biedt een bruikbaar uitgangspunt gemeenschappelijke vertrekpunten voor de corporate communicatie te formuleren (Van Riel, 1996) en de basis voor de communicatie naar binnen en naar buiten te leggen.

LITERATUUR

Albert, Stuart and David A. Whetten, 1985. Organizational Identity. In: L.L. Cummings and B.M. Staw (eds.): Research in Organizational Behaviour. Vol 7, 263-295.

Ashforth, B.E. and F. Mael, 1989. Social Identity Theory and the Organization. Academy of Management Review, Vol. 14, N° 1, 20-39.

Atamer, T. and R. Calori, 1993. Diagnostic et Décisions Stratégiques. Paris: Dunod.

Balmer, 1996. The nature of corporate identity: An explanatory study undertaken within BBC Scotland. Unpublished PhD Thesis. Glasgow: University of Strathclyde.

Bernstein, D., 1986. Company Image and Reality, a critique of Corporate Communications. Cassell Ltd., London.

Corporate Communication Centre, 1991. Applying the Mannheimer CI test in a Dutch company. Internal report.

Dowling, G.R., 1986. Managing Your Corporate Images. Industrial Marketing Management, 15, 109-115.

Dutton, Jane E., Janet M. Dukerich and Celia V. Harquail, 1994. Organizational Images and Member Identification. Administrative Science Quarterly, 39, 239-263.

Elsbach, K.D. and M.A. Glynn, 1996. Believing your own PR: Embedding identification in strategic reputation. Advances in Strategic Management, Vol. 13, 65-90.

Freeman, Linton C., 1979. Centrality in Social Networks. Conceptual Clarification. Social Networks, I, 1978/1979, 215-239.

Garfinkel, Harold, 1967. Studies in Ethnomethodology. Englewood Cliffs: Prentice Hall.

Grunert, Klaus G. and Suzanne C. Grunert, 1995. Measuring Subjective Meaning Structures by the Laddering Method: Theoretical Considerations and Methodological Problems. International Journal of Research in Marketing. Vol. 12, 209-225.

Guilford, J.P., 1959. Personality. New York: McGraw Hill.

Hinkle, Dennis N., 1965. The Change of Personal Constructs from the Viewpoint of a theory of Construct implications. Ph.D. Thesis, Ohio State University.

Keller, I., 1987. Corporate Identity, Elemente und Wirkung. Inauguraldissertation, Universität Mannheim.

Keller, I., 1990. Das CI-Dilemma, Abschied von falschen Illusionen. Wiesbaden: Gabler Verlag.

Kotler, Ph., 1984. Marketing Management: Analysis, Planning, and Control Englewood Cliffs, N.J.: Prentice Hall, 5th edition.

Lief, Todd, 1996. Common Ground, Common Cause and Common Sense. Design Management Journal, Summer 1995, 55-61.

Lux, Peter G.C., 1988. Zur Durchführung von Corporate Identity Programmen. In: Birkigt, K., M.M. Stadler and R. Funck (1988). Corporate Identity, Grundlagen, Funktionen, Fallbeispiele. Verlag Moderne Industrie, Landsberg/Lech, 4e Auflage.

Martin, Joanne, Martha S. Feldman, Mary Jo Hatch and Sim B. Sitkin, 1983. The Uniqueness Paradox in Organizational Stories. Administrative Science Quarterly, Vol. 28, 438-453.

McArthur, L.A., 1972. The how and what of why: Some determinants and consequences of causal attribution. Journal of Personality and Social Psychology, 22, 171-193.

Pieters, Rik, Hans Baumgartner and Huub Stad, 1994. Diagnosing means-end structures: The perception of wordprocessing software and the adaptive-innovative personality of managers. In:

Josée Bloemer, Jos Lemmink and Hans Kasper: Proceedings 23d Annual Conference of the European Marketing Academy, 18 t/m 20 Mei 1994 in Maastricht, 749-762.
Van Rekom, J., 1998. Corporate identity, development of the concept and a measurement instrument. Proefschrift. Rotterdam: Faculteit Bedrijfskunde, Erasmus Universiteit Rotterdam.
Reynolds, Th.J. and J. Gutman, 1988. Laddering Theory, method, analysis and Interpretation. Journal of Advertising Research. February/March 1988, 11-31.
Van Riel, C.B.M., 1996. Identiteit en imago, grondslagen van de corporate communication. Schoonhoven: Academic Press.
Smidts, A., C.B.M. van Riel and A.Th. Pruyn, 1998. The impact of employee communication on organizational Identification. In: Per Andersson (ed.): Proceedings 27th EMAC Conference, Track 3, pp. 279-290.
Strauss, Anselm and Juliet Corbin, 1991. Basics of Qualitative Research. Grounded Theory Procedures and Techniques. Newbury Park: Sage Publications.
Wasserman, Stanley and Katherine Faust, 1994. Social Network Analysis, Methods and Applications. Cambridge: Cambridge University Press.
Weber, Johannes, 1985. Unternehmensidentität und Unternehmenspolitische Rahmenplanung. Planungs- und Organisationswissenschaftliche Schriften. Herrsching: Prof.Dr. W. Kirsch.

16. Methoden voor winstmaximalisatie in direct marketing

H.R. VAN DER SCHEER[1]

SAMENVATTING

Het succes van een *direct-mailcampagne* wordt met name bepaald door (1) de selectie van de meest kansrijke adressen, de zogenaamde *targets*, en (2) het design van de mailing. In dit artikel worden voor deze twee aspecten methoden geïntroduceerd om de effectiviteit, en daarmee de winst, van een direct-mailcampagne te verhogen. Eerst wordt een overzicht gegeven van de selectiemethoden die vaak worden gebruikt om targets te selecteren. Deze methoden richten zich impliciet op mailings die bij respons een vaste opbrengst per respondent opleveren. Echter, veel direct-mailcampagnes genereren een respons waarbij de opbrengst, oftewel het *volume*, varieert over de respondenten. Er wordt een methode geïntroduceerd waarbij de responskans en het verwachte volume wordt bepaald. Op basis hiervan vindt dan de selectie plaats. De resultaten van de toepassing laten zien dat het inderdaad belangrijk is om ook het verwachte volume te modelleren. Daarna wordt een methode gepresenteerd waarmee op een effectieve manier het optimale *mailingdesign* kan worden bepaald. Deze methode wordt vervolgens uitgebreid door de interactie tussen de karakteristieken van de mailing en de karakteristieken van de huishoudens mee te nemen. Het resultaat is een methode die simultaan het mailingdesign optimaliseert en targets selecteert. De toepassing laat zien dat de winst hierdoor sterk toeneemt.

1. INTRODUCTIE

Direct marketing is een vorm van marketing die is gericht op het verkrijgen en onderhouden van een directe relatie tussen een aanbieder en haar afnemers. In de laatste decennia is direct marketing steeds belangrijker geworden. Dit komt naar voren door de hoogte van de bestedingen van organisaties aan direct marketing activiteiten en door het aantal (wetenschappelijke) artikelen dat over direct marketing zijn gepubliceerd. Redenen dat direct marketing belangrijk is geworden zijn onder andere: (1) een individualistische maatschappij waardoor traditionele marketing vaak niet meer toereikend is en (2) de ontwikkelingen in de informatietechnologie waardoor het mogelijk is om grote marketing databases aan te houden.

Het belangrijkste medium van direct marketing is (on)geadresseerde reclame, de zogenaamde direct-mail. Dit medium wordt op steeds grotere schaal gehanteerd door aan-

[1] Dit artikel is gebaseerd op het proefschrift "Quantitative Approaches for Profit Maximization in Direct Marketing" (1997) van Hiek van der Scheer, promotores: T.J. Wansbeek en P.S.H. Leeflang (Rijksuniversiteit Groningen).

bieders van fysieke producten (bijvoorbeeld Readers' Digest en Procter & Gamble), diensten (zoals Centraal Beheer) en ideeën (NOVIB en Greenpeace). Het voordeel van dit medium is dat huishoudens of klanten nauwkeurig kunnen worden geselecteerd. De kosten per contact zijn echter wel veel hoger dan bij de media van traditionele marketing (TV, radio, kranten). Hierdoor is het belangrijk dat een direct-mailcampagne voldoende respons genereert. Dit is te bereiken door de kritische succesfactoren van een campagne optimaal in te zetten.

De succesfactoren van een direct-mailcampagne zijn:
1. de selectie van de meest kansrijke adressen, de zogenaamde targets;
2. de karakteristieken van het aanbod;
3. de karakteristieken van de mailing, de zogenaamde *communicatie-elementen*;
4. het tijdstip waarop de mailing wordt verstuurd.

De selectie van de targets wordt zowel in theorie al in de praktijk als het belangrijkste element gezien. Een aanzienlijk gedeelte van direct-marketingonderzoek richt zich dan ook op technieken om targets te selecteren. Recentelijk is in de literatuur ook aandacht besteed aan het manipuleren van het design van de mailing (b.v. Smith en Berger 1996 en Van der Scheer, Vriens en Hoekstra 1996).

In dit artikel wordt een aantal methoden besproken om de effectiviteit, en daarmee de winst, van een direct-mailcampagne te verhogen. Deze methoden hebben zowel betrekking op de selectie van targets als het manipuleren van de karakteristieken van een mailing. Eerst wordt in paragraaf 2 een overzicht gegeven van de technieken die kunnen worden gebruikt om targets te selecteren. De meeste van deze selectietechnieken richten zich op mailings waarbij respons een vaste opbrengst oplevert. Echter, veel direct-mailcampagnes genereren niet zo'n simpele respons maar een respons waarbij de opbrengst, oftewel het volume, varieert over de huishoudens. In paragraaf 3 wordt een model geïntroduceerd dat zowel de responskans als het verwachte volume modelleert. In een toepassing wordt geïllustreerd hoe het model kan worden ingezet voor targetselectie. Paragraaf 4 gaat over het optimaliseren van het mailingdesign. Er wordt een methode gepresenteerd waarmee op een effectieve manier het optimale mailingdesign kan worden bepaald. Deze methode wordt uitgebreid door de interactie tussen de karakteristieken van de mailing en de karakteristieken van de huishoudens mee te nemen. Het resultaat is een methode die simultaan het mailingdesign optimaliseert en targets selecteert. De conclusies worden gegeven in paragraaf 5.

2. OVERZICHT SELECTIEMETHODEN

2.1. Introductie

Het selecteren van de targets, ook wel *list-segmentatie* genoemd, kan worden gezien als het proces om de markt, oftewel de *mailinglist*, te scheiden in twee groepen. Namelijk, een groep die een mailing moet ontvangen en een groep die geen mailing moet ontvangen. Aangezien een direct-mailcampagne alleen effectief is als de mailing de juiste targets bereikt, is veel aandacht besteed aan het ontwikkelen van selectiemethoden. In deze paragraaf wordt een overzicht gegeven van de methoden die in de praktijk veel-

vuldig worden toegepast en methoden die recentelijk in de literatuur zijn voorgesteld. Dit overzicht wordt afgesloten met een vergelijking van deze methoden.

Een cruciaal aspect voor de selectie van targets is de aanwezige informatie op de mailinglist, de zogenaamde selectievariabelen. Op basis van deze variabelen moet de selectie worden gemaakt. De volgende vier basis categorieën van selectievariabelen worden onderscheiden (b.v. Bult en Hoekstra, 1991; en Roberts en Berger, 1989, p. 106):
1. geografische variabelen (b.v. regio, urbanisatiegraad);
2. demografische variabelen (b.v. leeftijd, geslacht);
3. lifestyle variabelen (b.v. hobby's);
4. gedragsvariabelen (b.v. aankoophistorie).

In direct marketing wordt de laatste categorie als de belangrijkste beschouwd. Deze omvat de zogenaamde *RFM*-variabelen: *R*ecency, *F*requency en *M*onetary value. Met 'Recency' wordt o.a. de periode sinds de laatste aankoop bedoeld. 'Frequency' geeft het aantal aankopen in een bepaalde periode aan en 'Monetary value' het bedrag dat aan deze aankopen is besteed. Merk op dat deze variabelen alleen aanwezig zijn voor de huidige klanten. Dus, voor de acquisitie van nieuwe klanten moeten de andere type variabelen worden gebruikt.

2.2. Selectiemethoden

Dit overzicht van selectiemethoden begint met het RFM-model; dit is één van de eerste methoden die voor targetselectie is voorgesteld. Daarna worden twee eenvoudige methoden besproken, de kruistabel en CHAID, die geregeld worden gebruikt voor selectiedoeleinden maar die eigenlijk voornamelijk geschikt zijn voor een eerste dataanalyse. Vervolgens worden een aantal methoden besproken waarbij de binaire response (respons/nonrespons) wordt gemodelleerd: lineaire regressie, probitmodel en latente klasse analyse. Tenslotte wordt het neurale netwerk besproken dat als een alternatief voor de statistische methoden wordt beschouwd.

Bij de bespreking van de methoden wordt er impliciet vanuit gegaan dat de organisatie informatie over de respons heeft op basis van een *testmailing*. Deze informatie wordt vervolgens gebruikt om een methode te implementeren en op basis daarvan huishoudens te selecteren voor een *bulkmailing*.

RFM-model

Eén van de eerst modellen die werd ontwikkeld voor targetselectie is het zogenaamde RFM-model (zie b.v. Bauer 1988). Zoals de naam suggereert is dit model volledig gebaseerd op de RFM-variabelen. Elk van deze variabelen wordt in een aantal categorieën ingedeeld. Elk huishouden behoort per variabele tot één categorie. Per categorie wordt een score bepaald; deze is het verschil tussen het overall responspercentage en het responspercentage in die categorie (het responspercentage van de huishoudens behorende tot die categorie). De uiteindelijk score van een huishouden is de som van de scores van de drie categorieën waartoe het huishouden behoort. Het volgende fictieve voorbeeld van een postorderbedrijf illustreert dit.

Recency bestaat uit twee categorieën: afgelopen jaar een aankoop gedaan en de laatste aankoop was meer dan een jaar geleden; frequency refereert naar het aantal aankopen

in de laatste drie jaar: 1-4 aankopen en meer dan 4 aankopen; monetary value geeft het bedrag van de aankopen in de laatste drie jaar: t/m f 250,- en meer dan f 250,-. Bij de testmailing was de respons 8%. Onderstaande tabel geeft het responspercentage in elke categorie. De score is dit percentage minus 8; de laatste kolom geeft dit aan. Een klant die afgelopen jaar 1 aankoop van f 300,- heeft gedaan, heeft een score van -1 (2-4+1).

Variabele	Categorie	respons in testmailing	score
Recency	afgelopen jaar	10%	2
	meer dan jaar geleden	7%	-1
Frequency	1-4 aankopen	4%	-4
	meer dan 4 aankopen	12%	4
Monetary amount	t/m f 250,–	5%	-3
	meer dan f 250,–	9%	1

Deze scores kunnen vervolgens worden gebruikt voor selectie. Het grote voordeel van deze methode is de eenvoud. Twee nadelen van dit het model: (1) het houdt geen rekening met het feit dat variabelen met elkaar gecorreleerd zijn (b.v. veel aankopen impliceert hoog bedrag) en (2) eventuele andere informatie wordt genegeerd.

Kruistabel

De eenvoudigste methode om respons te modelleren is met een kruistabel. Twee of meer (gecategoriseerde) variabelen worden hierbij in een tabel tegen elkaar uitgezet en in de cellen worden responspercentages weergegeven. De cellen waarin het responspercentage hoog genoeg is, geven de huishoudens die moeten worden geselecteerd. Het nadeel is dat de indeling van de categorieën ad-hoc is en dat de methode onoverzichtelijk wordt als er veel variabelen worden meegenomen, met het risico dat er te weinig observaties per cel aanwezig zijn. Het is echter wel een zeer nuttige methode voor een eerste data-analyse om inzicht te krijgen in de data. Dit kan vervolgens als input dienen voor verdere analyses.

Chi-square automatic interaction detection (CHAID)

CHAID is een methode die de totale groep splitst in een aantal subgroepen op basis van de variabele die statistisch gezien de beste splitsing kan bewerkstelligen (Magidson, 1994). Elke subgroep kan weer verder worden opgesplitst. Zo ontstaat uiteindelijk een (CHAID-)boom die een aantal segmenten definieert. Deze segmenten zijn op basis van de gevonden splitsingen eenvoudig te interpreteren; dit wordt dan ook als het grootste voordeel gezien. Toch wordt het in de praktijk vaak gebruikt als een methode voor de eerste data-analyse, in het bijzonder om interactie-effecten te achterhalen, die als input dienen voor meer geavanceerde methoden (Shepard, 1995).

Linear probability model

Lineaire regressie met een binaire variabele als afhankelijke variabele, hier respons/nonrespons, wordt een linear probability model genoemd aangezien de voorspelde waarde als kans kan worden geïnterpreteerd. Hierbij wordt aangenomen dat de kans lineair afhangt van een aantal karakteristieken van het huishouden:

$$y_i = x_i \beta + u_i$$

waarbij y_i een random variabele is die de waarde één aanneemt als huishouden i respondeert en anders nul; x_i is de vector met karakteristieken van het huishouden (de geografische-, demografische-, lifestyle- en gedragsvariabelen), u_i is een random verstoring en β de onbekende parameter. De parameter kan worden geschat met de kleinste kwadraten methode. Op basis van de geschatte parameters kan voor ieder huishouden een responskans worden geschat. Deze responskans wordt vervolgens gebruikt om huishoudens voor een direct-mailcampagne te selecteren.

Dit model is eenvoudig toe te passen maar heeft een aantal nadelen (Judge et al. 1985). Bijvoorbeeld, de kleinste kwadraten methode geeft onnauwkeurige voorspellingen en is gevoelig voor de waarden van de onafhankelijke variabelen. Daarnaast kunnen de voorspellingen buiten het [0,1]-interval liggen.

Probit- en logitmodel

Een andere methode om discrete data te modelleren is met een probit- of logitmodel; deze komen tegemoet aan de nadelen van het linear probability model. Deze modellen nemen aan dat er een latente variabele is, y_i^*, die kan worden geïnterpreteerd als de '*bereidheid om te responderen*'. Aangenomen wordt dat deze latente variabele afhangt van de vector met karakteristieken van het huishouden, namelijk:

$$y_i^* = x_i \beta + e_i.$$

De daadwerkelijke respons wordt geobserveerd en niet de 'bereidheid om te responderen'. De geobserveerde respons is:

$$y_i = \begin{cases} 1 & y_i^* \geq 0 \\ 0 & y_i^* < 0. \end{cases}$$

Het probitmodel neemt aan dat de verstoringen, e_i, normaal zijn verdeeld en het logitmodel dat de verstoringen logistisch zijn verdeeld. Het model is te schatten met *maximum likelihood*; dit kan met vele statistische software pakketten, zoals SPSS. Evenals bij het lineaire regressiemodel kan op basis van de geschatte parameters voor ieder huishouden een responskans worden bepaald; deze wordt gebruikt om huishoudens voor een direct-mailcampagne te selecteren.

Bult en Wansbeek (1995) beargumenteren dat de verdeling van de verstoring een cruciale rol speelt bij de selectie van de huishoudens. Indien deze niet normaal zijn verdeeld, geeft het probitmodel inconsistente schattingen en daardoor een suboptimale selectie. Door niet à priori een verdeling aan de verstoring op te leggen maar het model te schatten met semiparametrische methoden kan de selectie, in ieder geval in theorie, worden verbeterd. Het nadeel van zulke schattingsmethoden is dat deze vaak lastig zijn te implementeren en dat hiervoor geen standaard software beschikbaar is.

Discriminant analyse

Bij een discriminant analyse wordt een functie bepaald die de respondenten het beste kan onderscheiden van de nonrespondenten (b.v. Maddala 1983, p. 16). Hierbij wordt de volgende verliesfunctie geminimaliseerd:

$$\ell = \sum_{i=1}^{n} \left(\hat{y}_i - y_i \right),$$

waarbij y_i één is als i respondeert en anders nul; $\hat{y}$ is één indien i als respondent wordt geclassificeerd en anders nul. Op basis van de geschatte discriminantfunctie worden huishoudens als respondent of nonrespondent geclassificeerd. Vaak wordt een lineaire discriminantfunctie gekozen die is gebaseerd op de normaliteitsaanname. Shepard (1995) laat zien dat bij een respons van minder dan 10%, de analyse gevoelig is voor deze normaliteitsaanname. Dit probleem kan in principe worden verholpen door een semiparametrische aanpak waarbij geen aannamen worden gedaan over de verstoring (zie Manski 1975, 1985).

Een ander aspect dat een belangrijke rol speelt bij de bovenstaande verliesfunctie is dat impliciet wordt aangenomen dat de kosten van misclassificatie symmetrisch zijn. Namelijk, een huishouden dat heeft gerespondeerd en als nonrespondent wordt geclassificeerd heeft dezelfde bijdrage aan de verliesfunctie als een huishouden dat niet heeft gerespondeerd maar wel als respondent wordt geclassificeerd. Voor de meeste praktische situaties is het echter zo dat de eerste vorm (de organisatie mist een respondent) van misclassificatie ernstiger is dan de tweede vorm (de organisatie stuurt een mailing te veel). Het is eenvoudig om deze asymmetrie in verlies mee te nemen in de analyse. Zie Bult (1993) voor een toepassing van een asymmetrische verliesfunctie die met een semiparametrische methode wordt geschat.

Latente klasse analyse

Een cruciale aanname in de methoden die hierboven zijn besproken is de homogeniteit van de huishoudens ten aanzien van hun responsgedrag. Namelijk, één model beschrijft de relatie tussen de onafhankelijke variabelen en de afhankelijke variabele. Dit kan een misleidend model opleveren indien er substantiële heterogeniteit aanwezig is. Het onderstaande voorbeeld verduidelijkt dit. Neem aan dat de populatie van huishoudens bestaat uit twee ongeveer even grote segmenten. Neem aan dat deze twee segmenten beschreven worden door deze twee regressie vergelijkingen:

$$y = x_1 \beta_1 - x_2 \beta_2 + \varepsilon$$

voor het eerste segment en

$$y = -x_1 \beta_1 + x_2 \beta_2 + \varepsilon$$

voor het tweede segment, waarbij x_1 en x_2 de twee verklarende variabelen zijn. Deze variabelen hebben dus in de twee segmenten een tegengesteld effect op de 'bereidheid om te responderen'. Indien één model wordt gebruikt, geldt dat $\beta_1 \approx \beta_2 \approx 0$; dit geeft aan dat één model voor de gehele populatie weinig zinvol is.

DeSarbo en Ramaswamy (1994) hebben een methode geformuleerd die rekening houdt met eventuele heterogeniteit. Hun methode specificeert simultaan de segmenten en per segment een responsfunctie. Dus in plaats van één model voor de gehele populatie, definieert deze aanpak een aantal segmenten en schat per segment een model. Verder wordt er een kans bepaald dat een huishouden tot een bepaald segment behoort, de zogenaamde *posterior* kans. Zo krijgt ieder huishouden een eigen responsmodel dat wordt bepaald door de responsmodellen per segment en de kans dat het huishouden tot dat segment behoort.

Hoewel het model theoretisch gezien is aan te bevelen, heeft het een aantal praktische problemen. Ten eerste is het vaak lastig om zo'n model te specificeren, met name als er veel observaties en veel variabelen zijn. Ten tweede moet er voor selectiedoeleinden een additionele vergelijking worden geschat om een posterior kans te genereren voor de huishoudens waaruit wordt geselecteerd. In het ideale geval wordt deze additionele vergelijking gebaseerd op variabelen die nog niet zijn gebruikt; dit is echter in de praktijk vaak niet haalbaar.

Neuraal netwerk
Vanaf de jaren tachtig zijn neurale netwerken in snel tempo ontwikkeld en op velerlei gebieden ingezet. Ook op het gebied van bedrijfseconomische vraagstukken is er in de laatste jaren een grote groei waar te nemen (zie b.v. Sharda, 1994). Sinds kort worden neurale netwerken toegepast op marketing vraagstukken (Kumar et al. 1995). Met betrekking tot targetselectie voor direct marketing zijn neurale netwerken zowel in theoretische als in praktische *journals* beschreven als een veelbelovend alternatief ten opzichte van traditionele statistische methoden (b.v. Zahavi en Levin, 1995, 1997). De reden hiervoor is dat neurale netwerken in staat zijn om complexe niet-lineaire relaties te achterhalen. Zie Van der Veer, Wierenga en Kluytmans (1994) voor een uitgebreide introductie van neurale netwerken.

Een neuraal netwerk kan worden gezien als een model waarbij een set van input variabelen (de onafhankelijke variabelen) via een hidden layer (een aantal niet observeerbare variabelen) worden verbonden met de output (de afhankelijke variabele). Het neurale netwerk bepaald een relatie tussen de input en output (via de hidden layer) zodanig dat een optimale fit wordt bepaald.

Nadat het netwerk is bepaald kan voor ieder huishouden een neurale netwerkscore worden bepaald. Deze score mag niet als kans worden geïnterpreteerd maar kan wel worden gebruikt om de huishoudens te ordenen. Degenen met de hoogste scores moeten worden geselecteerd.

Indien de gegevens veel 'ruis' bevatten (wat in direct marketing vaak het geval is) is de toegevoegde waarde van een neuraal netwerk ten opzichte van statistische methoden erg klein (zie volgende paragraaf). Dit komt onder andere doordat eventuele niet-lineaire verbanden slecht zijn te achterhalen (b.v. Den Uyl en Langendoen, 1997; Huizingh, De Boer en Wedel, 1995). Aangezien het niet eenvoudig is om een goed neuraal netwerk te specificeren, wordt het weinig gebruikt voor targetselectie.

2.3. Vergelijkingen beschreven in de literatuur

Hoewel de laatste tien jaar een groot aantal artikelen is gepubliceerd op het gebied van selecties voor direct-mailcampagnes, zijn er weinig artikelen die een vergelijking maken tussen de diverse methoden. De schaarse literatuur op dit gebied zal kort worden besproken.

Magidson (1988) vergelijkt CHAID met log-lineaire modellen. Zijn conclusie is dat CHAID goed werkt, maar hij vergelijkt het niet op basis van real-life data. Bult en Wansbeek (1995) laten zien dat de performance van CHAID beter is dan die van lineai-

re regressie, het logitmodel en een semiparametrische variant van het logitmodel voor de data waarop de modellen zijn geschat. Voor de validatieset is CHAID veruit de slechtste; het logitmodel laat de beste performance zien. Dit is verrassend, want er werd verwacht dat de semiparametrische variant het beter zou doen. Daarentegen, Bult (1993) laat zien dat de semiparametrische variant van de discriminant analyse het beter doet dan het logitmodel.

Met betrekking tot neurale netwerken tonen een aantal studies aan dat deze niet meer succesvol zijn dan de andere methoden. Bijvoorbeeld, Kumar et al. (1995) laten zien dat een logitmodel dezelfde resultaten geeft als een neuraal netwerk. In een onderzoek geïniteerd door de Stichting Research en DMSA, zijn neurale netwerken uitvoerig onderzocht (zie Den Uyl en Langendoen, 1997). Uit dit onderzoek blijkt dat de resultaten die met neurale netwerken worden bereikt weinig verschillen van de andere methoden. Lix en Berger (1995) hebben tegenstrijdige resultaten. Bij een toepassing is het neurale netwerk veel beter dan een logitmodel, terwijl bij een andere toepassing het logitmodel beter is. In een uitgebreide studie tonen Zahavi en Levin (1997) aan dat het logit-model ongeveer dezelfde performance heeft als een neuraal netwerk. Ze concluderen dan ook: “deze resultaten zijn niet bemoedigend voor de neurale netwerk aanpak aangezien het specificeren van een neuraal netwerk niet eenvoudig is”.

Naast het vergelijken van methoden is er ook onderzocht of het beter is om alle variabelen die op voorhand van belang lijken op te nemen of variabelen met een stapsgewijze procedure op te nemen. Twee studies presenteren tegenstrijdige resultaten: Magliozzi en Berger (1993) laten zien dat de laatste methode het best is terwijl Lix en Berger (1995) het tegengestelde concluderen.

Concluderend: zowel op basis van vergelijkende studies als op basis van theorie is het moeilijk een eenduidige uitspraak te doen over de beste techniek. Enerzijds komt dit doordat een techniek op vele, soms conflicterende, criteria worden beoordeeld. Anderzijds is het goed mogelijk dat de performance van een model afhangt van het specifieke gebruik en de beschikbare data.

3. MODELLEREN VAN RESPONSKANS EN VOLUME

3.1. Introductie

In de vorige paragraaf zijn een groot aantal methoden besproken om huishoudens voor een direct-mailcampagne te selecteren. De overeenkomst tussen deze methoden is het type afhankelijke variabele. Namelijk, er wordt gefocussed op respons versus nonrespons. Vervolgens wordt op basis van het geschatte model voor ieder huishouden de kans op respons bepaald. Deze responskans wordt daarna gebruikt om huishoudens te selecteren voor een direct-mailcampagne. Hierbij wordt vaak eerst een *afkapkans* bepaald; huishoudens met een responskans groter dan deze afkapkans worden geselecteerd voor de bulkmailing.

Bij veel direct-mailcampagnes is de respons niet binair maar moet een respondent een keuze maken over de hoogte van het volume. Hierbij valt te denken aan de totale beste-

ding bij een postorder bedrijf, het bedrag dat wordt gestort in een beleggingsfonds, of de donatie aan een charitatieve instelling. In veel gevallen wordt bij een direct-mail-campagne waar het volume bij respons ook een rol speelt het gemiddelde genomen van alle huishoudens die hebben gerespondeerd. Dit gemiddelde wordt dan gebruikt om een afkapkans te bepalen.

In de literatuur wordt er veelvuldig beargumenteerd dat het eigenlijk onvoldoende is om alleen de responskans te modelleren omdat ook het volume expliciet moet worden gemodelleerd. Zo liet bijvoorbeeld Courtheoux (1987) met een eenvoudig hypothetisch voorbeeld al zien dat het volume een belangrijk aspect is. Toch is er in de literatuur nog zeer weinig over gepubliceerd. Rao en Steckel (1995) merken op dat het volume kan worden geschat met een standaard lineair regressiemodel, maar hun empirische toepassing heeft betrekking op binaire respons.

In deze paragraaf wordt een model gegeven om zowel de responskans als het verwachte volume te modelleren; het is gebaseerd op Otter, Van der Scheer en Wansbeek (1997). De opbouw van deze paragraaf is als volgt. In paragraaf 3.2 wordt het model gegeven; dit wordt vervolgens geïllustreerd met een toepassing in paragraaf 3.3. In paragraaf 3.4 wordt besloten met enkele conclusies en een discussie.

3.2. Beschrijving van het model

Model

Beschouw een direct marketing organisatie die moet beslissen welke huishoudens een mailing moeten ontvangen. Indien een mailing wordt gestuurd is de winst, Π, simpelweg:

$$\Pi = AR - c,$$

waarbij $R=1$ als het huishouden respondeert en $R=0$ als het huishouden niet respondeert. A is de variabele die het volume geeft; deze varieert tussen de huishoudens; c zijn de kosten van een mailing. Dus de winst is $-c$ als een huishouden niet respondeert en is $A - c$ als een huishouden respondeert. De twee elementen van het model, R en A, geven de twee beslissingen die een huishouden moet nemen. De eerste beslissing is wel of niet responderen en de tweede beslissing is de hoogte van het volume. Deze twee beslissingen worden beide gemodelleerd.

Aangenomen wordt dat R kan worden gemodelleerd met een probitmodel (zie paragraaf 2). Dit betekent dat op het moment dat de parameters van dit model zijn geschat voor ieder huishouden op de mailinglist een responskans kan worden berekend. Dit model noemen we het *respons*model.

Als $R=1$, dus als een huishouden heeft gerespondeerd, wordt er vanuit gegaan dat het volume kan worden beschreven met een lineair model:

$$A = x\,y + v,$$

met de aanname dat

$$E\,(v \mid R = 1, x) = 0.$$

Dit definieert het zogenaamde *volume*model. De vector met verklarende variabelen, x, in het responsmodel en het volumemodel mogen hetzelfde zijn maar mogen ook andere variabelen bevatten. De verstoringen van het responsmodel en het volumemodel kunnen correleren maar dat speelt verder geen rol meer. Deze wijze van modelleren van de kans en het volume wordt een *two-part model* genoemd (b.v. Duan et al. 1983).

Optimale selectie

Op basis van een testmailing kan het bovenstaande model eenvoudig worden geschat. Het responsmodel wordt geschat met maximum likelihood en het volumemodel met lineaire regressie. Vervolgens kan, op basis van het geschatte model, voor alle huishoudens een responskans p en het verwachte volume a worden bepaald. Voor elk huishouden dat de testmailing niet heeft ontvangen moet nu op basis van hun schatting van p en a worden besloten of een mailing wordt gestuurd. De eenvoudige selectieregel die hierbij wordt gebruikt is: selecteer een huishouden als $ap > c$. De combinaties van p en a waarbij de organisatie indifferent is met betrekking tot deze keuze, $ap = c$, definiëren de optimale *afkapcurve* ($a = c/p$).

Indien alleen de kans, p, wordt gemodelleerd, dan wordt voor a vaak het gemiddelde, $\bar{a}$, van de respondenten (van de testmailing) genomen. In dat geval is de schatting voor a voor alle huishoudens gelijk. De selectieregel is nu: selecteer een huishouden als $p > c / \bar{a}$. Nu is er geen optimale afkapcurve te definiëren maar een afkap*kans*, $p_c \equiv c / \bar{a}$. Dus alle huishoudens met een p groter dan deze afkapkans moeten worden geselecteerd. Dit is de methode die veelvuldig wordt toegepast. Het is dus een bijzonder geval van het hierboven beschreven model.

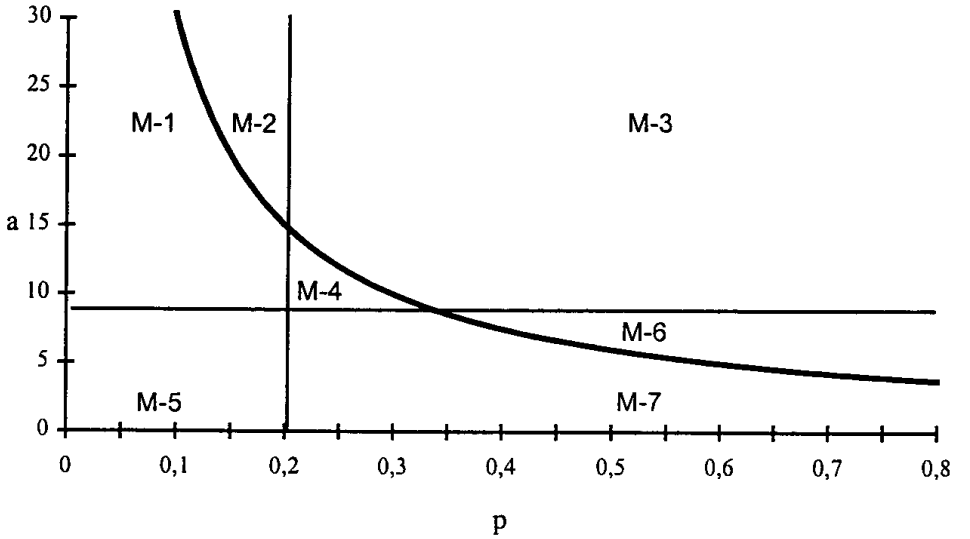
Een ander bijzonder geval wordt verkregen door alleen het volume te modelleren en de responskans voor ieder huishouden gelijk te nemen. Deze responskans, $\bar{p}$, is dan het percentage respondenten in de testmailing. In plaats van een afkapcurve of een afkapkans kan nu het afkap*volume* worden gedefinieerd: $a_c \equiv c / \bar{p}$. Alle huishoudens met een verwacht volume groter dan dit afkapvolume moeten worden geselecteerd voor de direct-mailcampagne.

Samenvattend betekent het bovenstaande dat de volgende drie methoden worden geanalyseerd:

- Alleen de responskans modelleren;
- Alleen het volume modelleren;
- De responskans en het volume modelleren.

Figuur 1 illustreert de verschillen in de selectie. De x-as geeft de responskans en de y-as het verwachte volume. Ieder huishouden kan op basis van hun combinatie van p en a in de figuur (als punt) worden weergegeven. De huishoudens ten noordoosten van de (afkap-)curve zouden een mailing moeten ontvangen. De verwachte winst is daar namelijk positief.

Om deze (optimale) selectie te vergelijken met de twee andere selectieregels, zijn in de figuur zeven regio's, M-1,…, M-7, weergegeven (met gefingeerde getallen). Deze regio's worden gescheiden door de rechte lijnen en de curve. Als alleen de responskans wordt gemodelleerd dan moeten alle huishoudens met een kans rechts van de verticale



Figuur 1. Selectie van huishoudens op basis van a *en* p.

lijn (p = 0,2) een mailing ontvangen, dus de regio's M-3, M-4, M-6 en M-7. Dit betekent dat de huishoudens in regio M-4 en M-7, die negatieve winst generen, worden geselecteerd. Daarnaast ontvangen huishoudens in M-2 ten onrechte geen mailing. Deze regio bevat de huishoudens die een lage responskans hebben maar wel een groot volume genereren bij respons: dit is natuurlijk een zeer interessante groep.

Indien alleen het volume wordt gemodelleerd kan een soortgelijke analyse van de regio's worden gegeven. In dit geval moeten de huishouden met een a boven de horizontale lijn $a = a_c = 9$ worden geselecteerd. Dit zijn de regio's M-1, M-2, M-3 en M-4. Nu worden de huishoudens uit de regio M-1 en M-4 geselecteerd terwijl die een negatieve winst genereren.

Het daadwerkelijke verschil tussen de drie methoden voor een bepaalde campagne hangt af van het aantal huishoudens in elk van de regio's. In theorie moet het modelleren van zowel de responskans als het volume de beste selectie opleveren en dientengevolge de hoogste winst. Van de andere twee methoden is op voorhand niet te zeggen welke de beste selectie zal opleveren. Dit moet empirisch worden bepaald.

3.3. Toepassing

Het model wordt geïllustreerd met een voorbeeld van een charitatieve instelling die door middel van direct-mailcampagnes fondsen werft. De dataset bestaat uit 40 000 huishoudens. Al deze huishoudens hebben al eens een keer gedoneerd aan de organisatie. De afhankelijke variabele in het responsmodel is de respons/nonrepons en in het volumemodel het gedoneerde bedrag. Van de 58 potentiële verklarende variabelen zijn na een grondige data-analyse o.a. de volgende variabelen geselecteerd: het gedoneerde bedrag in de afgelopen twee jaren, de interactie tussen deze donaties, aantal jaren op de mailinglist, grote van de familie en de urbanisatiegraad.

De respons was 33,9% ($\bar{p}$) en de gemiddelde donatie (bij respons) was Fl. 17,04 ($\bar{a}$).

Om de performance van het model te bepalen is het model geschat op basis van 1 000 observaties; de rest van de observaties vormen de validatieset. Dit betekent dat op basis van het geschatte model wordt bepaald welke huishoudens in de validatieset zouden moeten worden geselecteerd. De winst wordt berekend door bij de geselecteerde huishoudens de werkelijke bedragen te sommeren en de kosten van de mailing er af te trekken. Bij de drie methoden verschillen de huishoudens die worden geselecteerd en dus de behaalde winst.

Het resultaat, de winst in de validatieset, kan sterk afhangen van de splitsing van de dataset in een analyseset en validatieset. Dit kan betekenen dat bij een bepaalde splitsing de eerste methode het beste resultaat geeft terwijl bij een andere splitsing de derde methode de grootste winst genereert. Dit probleem is op te lossen door een groot aantal keren de volgende analyse volledig uit te voeren:

1. analyseset en validatieset construeren;
2. model schatten met behulp van de analyseset;
3. de huishoudens bepalen die op basis van het geschatte model zouden zijn geselecteerd (in de validatieset); dit moet voor elk van de drie methoden worden uitgevoerd;
4. de winst bepalen op basis van geselecteerde huishoudens.

Op deze manier wordt een inzicht verkregen in de gemiddelde performance van de drie methoden. Daarnaast kan het aantal keren worden bepaald dat de ene methode meer winst genereert dan een andere.

Tabel 1 geeft het gemiddeld aantal geselecteerde huishouden en gemiddelde winst van de drie methoden op basis van 500 analyse- en validatiesets. De resultaten maken duidelijk dat het modelleren van de responskans en het volume de hoogste winst genereert. Ten opzichte van het modelleren van alleen de reponskans neemt de winst met 11,5% toe! Wat verder opvalt is dat zelfs als alleen het volume wordt gemodelleerd de winst sterk toeneemt. Dit impliceert dat de huidige praktijk, alleen de responskans modelleren, een essentieel onderdeel van de respons niet meeneemt voor de selectie van huishoudens.

Tabel 1. Performance van de drie methoden.

Methode	**Aantal geselecteerde huishoudens**	**Winst (in guldens)**
Alleen responskans modelleren	38 068	89 163
Alleen volume modelleren	31 225	96 672
Responskans en volume modelleren:	27 016	99 433

In plaats van naar de gemiddelde performance van de drie methoden te kijken, kan ook de performance per splitsing in analyse- en validatieset worden vergeleken.
Deze vergelijking laat zien dat ‘responskans en volume modelleren’ in respectievelijk 100% en 85% van de gevallen meer winst genereert dan ‘alleen responskans modelleren’ en ‘alleen volume modelleren’. Deze percentages maken duidelijk dat het belangrijk is om naast de responskans ook het volume te modelleren. De extra kosten van het

specificeren van het model voor het volume zijn relatief klein als al een responsmodel wordt bepaald. Dus ook als de kosten van het specificeren van het uitgebreide model worden meegenomen, betekenen de bovenstaande resultaten dat een organisatie meer winst kan maken door beide elementen van respons te modelleren.

3.4. Conclusie en discussie

In deze paragraaf is een methode geïntroduceerd om zowel de responskans als het volume te modelleren. De empirisch bevindingen impliceren dat veel extra winst te behalen valt door niet alleen de responskans te modelleren. Deze resultaten moeten natuurlijk wel voorzichtig worden geïnterpreteerd want het is niet uitgesloten dat bij andere toepassingen de verschillen tussen de drie methoden minder groot zijn.

De selectieregels die hierboven zijn geïntroduceerd kunnen worden verfijnd door in de specificatie van de verwachte winst een, nog te bepalen selectieregel, op te nemen. De technische aspecten van deze methode worden hier achterwegen gelaten, deze zijn te vinden in Otter, Van der Scheer en Wansbeek, 1997, en Van der Scheer, 1997. Het resultaat is een complexe selectieregel waarbij nog keuzes moeten worden gemaakt ten aanzien van de operationalisatie. De eenvoudigste manier van operationaliseren geeft de hierboven beschreven selectieregel. Bij twee andere (complexere) manieren van operationaliseren steeg de (gemiddelde) winst naar respectievelijk Fl. 99 924 en FL. 100 123. Conclusie: de geavanceerde methoden genereren inderdaad meer winst maar de toename is marginaal.

Een voordeel van het two-part model is dat de twee delen, het responsmodel en het volumemodel, afzonderlijk kunnen worden geschat. De verstoringen van deze delen mogen gecorreleerd zijn maar dat beïnvloed het schatten niet. Een model dat wel expliciet rekening houdt met (mogelijke) afhankelijkheid van de verstoringen is het zogenaamde *sample-selection model* (b.v. Amemiya, 1985). Hierbij wordt uit gegaan van bivariaat normaal verdeelde verstoringen. Indien dat het geval is en de correlatie niet gelijk is aan nul, geeft het two-part model geen zuivere schattingen. Dit zou dus pleiten voor het gebruik van het sample-selection model. Toch zijn er redenen om het two-part model te prefereren (Van der Scheer, 1997):

1. de schattingen van de parameters zijn van ondergeschikt belang omdat het model in principe alleen wordt gebruikt voor voorspellen en selecteren;
2. het sample-selection model is gevoelig voor misspecificatie van de verstoringen, in tegenstelling tot het two-part model (b.v. Manning et al. 1987);
3. zelfs als de verstoringen bivariaat normaal zijn verdeeld, is de voorspellende performance van het two-part model bijna gelijkwaardig met de voorspellende performance van het sample-selection model (b.v. Duan et al. 1983 en Manning et al. 1987);
4. het two-part model is zeer eenvoudig te schatten in tegenstelling tot het sample-selection model waarbij een complexe likelihood functie moet worden gemaximaliseerd.

Het is echter interessant om de performance van het sample-selection model te vergelijken met het two-part model.

4. MAILINGDESIGN

4.1. Introductie

Een belangrijk aspect van het succes van een direct-mailcampagne zijn de karakteristieken van de mailing, oftewel de *communicatie-elementen* (b.v. Fraser-Robinson, 1989; Throckmorton, 1992; Van der Scheer, Hoekstra en Vriens, 1998, en Vögele, 1992). Deze elementen hebben betrekking op alle onderdelen van de mailing: de enveloppe, de brief, de brochure en de responskaart. Het mailingdesign heeft betrekking op de inhoud en lay-out van de mailing, gebruik van illustraties, etc. Om een optimaal design te bepalen moet de direct marketing marketeer eigenlijk weten welke karakteristieken in welke mate bijdragen aan de respons.

In de praktijk wordt veel aandacht besteed aan het ontwikkelen van een mailingdesign. De nadruk ligt hierbij op de creatieve aspecten; daarnaast wordt gebruik gemaakt van verscheidene vuistregels zoals te vinden in Fraser-Robinson (1989) en Vögele (1992). Ervaring speelt daarbij natuurlijk ook een belangrijke rol.

Onderzoek op dit gebied komt voornamelijk uit de hoek van de *behavioral decision theory* met betrekking tot charitatief gedrag. Hierbij is bijvoorbeeld het effect onderzocht (DeJong en Oopik, 1992; Schibrowsky en Peltier, 1995; Smith en Berger, 1996) van:
- de hoogte van de voorgestelde donatie(s);
- feitelijke versus anekdotisch informatie;
- positieve versus negatieve informatie.

Conclusies van dergelijke studies kunnen worden gebruikt voor het ontwerpen van een mailingdesign.

De combinatie van creativiteit, vuistregels, ervaring en bepaalde studies leiden tot een bepaald mailingdesign dat wordt ingezet voor de direct-mailcampagne. Vaak worden enkele mailings ontworpen die door een groep experts en/of potentiële ontvangers worden beoordeeld om tot een definitieve keuze te komen. Een andere manier om tot een definitieve keuze te komen is door meerdere type testmailings te versturen. De mailing die de hoogste respons oplevert wordt vervolgens voor de bulkmailing gebruikt.

Nadat een bepaald mailingdesign is gekozen, rijst de vraag hoe dit design kan worden aangescherpt om de effectiviteit te vergroten. Met andere woorden, hoe kunnen de karakteristieken van die bepaalde mailing zodanig worden gemanipuleerd opdat een hogere respons wordt gegenereerd? In deze paragraaf wordt een methode besproken om een optimaal mailingdesign te bepalen. In 4.2 wordt de methode besproken; deze wordt in 4.3 geïllustreerd met een toepassing. Door rekening te houden met de interactie tussen de karakteristieken van de mailing en de karakteristieken van de huishoudens kan het mailingdesign worden verfijnd. Deze uitbreiding wordt besproken in 4.4 en geïllustreerd in 4.5.

4.2. Methode om design te bepalen

Een veel toegepaste methode om een mailingdesign te verbeteren is door elk karakteristiek apart te onderzoeken. Uitgaande van een bepaald design wordt de mailing op

één karakteristiek gemanipuleerd door twee of meer type mailings te ontwikkelen die alleen op dat karakteristiek van elkaar verschillen. Bijvoorbeeld, het karakteristiek dat een organisatie wil onderzoeken is het gebruik van illustraties, waarbij de keuze gaat tussen geen, één of twee illustraties. De drie type mailings (een zonder illustraties, een met één illustratie en een met twee illustraties) worden vervolgens aselect verstuurd naar een groep huishoudens. Op basis van de respons op de drie type mailings wordt bepaald wat de optimale keuze is met betrekking tot het gebruik van illustraties. Nadat deze keuze voor dit karakteristiek is gemaakt, wordt een andere karakteristiek gekozen waarvoor op soortgelijke wijze wordt bepaald hoe dat karakteristiek moet worden ingevuld. Na verloop van tijd ontstaat een effectievere mailing.

Bovenstaande methode is een inefficiënte manier om tot een effectievere mailingdesign te komen omdat:
- een groot aantal mailings nodig zijn om tot het optimale design te komen;
- het niet mogelijk is om de interactie-effecten tussen de mailingkarakteristieken mee te nemen; hierdoor bestaat het risico dat niet het optimale design wordt bereikt.

Met behulp van een *statistisch design*, ook wel een *conjoint field-experiment* genoemd, kan in één test voor een groot aantal karakteristieken de optimale keuze worden bepaald. Eerst worden de karakteristieken (b.v. illustratie) en de relevante niveau's hiervan gekozen (b.v. geen, één of twee illustraties). Daarna wordt een statistisch design bepaald dat met zo min mogelijk verschillende type mailings het mogelijk maakt om voor elk niveau van een karakteristiek het effect op de respons te bepalen. Het statistisch design geeft aan hoe de mailingdesigns voor de test er uit moeten zien. Elk van deze verschillende type mailingdesigns worden naar een groep huishoudens gestuurd. Op basis van de respons bij de verschillende type mailings die in de test worden gebruikt kunnen voor *alle* combinaties van de niveau's van de diverse karakteristieken de (verwachte) respons worden bepaald. Dus ook voor de type mailings die niet expliciet in de test zijn gebruikt. Dit in tegenstelling tot een methode waarbij per karakteristiek het optimale niveau wordt bepaald.

4.3. Toepassing

De methode om met behulp van een statistisch design het optimale mailingdesign te specificeren wordt geïllustreerd aan de hand van een toepassing bij een charitatieve instelling. Deze instelling wil, mede door de enorme concurrentie op de charitatieve markt, het mailingdesign optimaliseren. Uitgangspunt is de mailing die ze tot dan toe hebben ingezet voor hun direct-mailcampagnes. In overleg met deze organisatie zijn de belangrijkste karakteristieken onderscheiden. Tabel 2 geeft de karakteristieken en niveau's die zijn gebruikt in de test.
De *acceptgiro* zit of aan de brief vast of is los bijgevoegd. De *brochure* geeft algemene informatie over de charitatieve instelling. De brief kan worden voorzien van een *illustratie* die op verschillende plaatsen kan worden afgedrukt. In de brief kan gebruik worden gemaakt van *vetgedrukte woorden* om bepaalde informatie te benadrukken. Deze kunnen veel, weinig of niet voorkomen. De *post scriptum* kan een samenvatting van de brief bevatten of nieuwe informatie. De *handtekening* onderaan de brief is of van een professor of van de directeur van de charitatieve instelling. Het *adres* van het huishouden kan op de brief worden gedrukt of op de acceptgiro.

Tabel 2. Gebruikte karakteristieken in de test.

Karakteristiek		**Niveau**	
Acceptgiro	Vast	Los	
Brochure	Wel	Geen	
Illustratie	Links	Rechts	Geen
Vetgedrukte woorden	Veel	Weinig	Geen
Post scriptum	Samenvatting	Nieuwe informatie	
Handtekening	Professor	Directeur	
Adres	Brief	Acceptgiro	

Er zijn dus zeven karakteristieken waarvan vijf met twee niveau's en twee met drie niveau's. Op basis hiervan zijn er 288 ($2^5 3^2$) mogelijke verschillende type mailings te construeren. Elk van deze mailings zou naar een groep huishoudens kunnen worden gestuurd en de mailing met de grootste respons kan dan worden ingezet voor een landelijke mailing. Echter, met behulp van het statistische design hoeven maar 16 verschillende type testmailings te worden verstuurd om toch over alle 288 mailings de verwachte respons te kunnen bepalen en dus het optimale mailingdesign voor de bulkmailing te kiezen.

In de test is elk van de 16 verschillende type mailings naar zo'n 3 000 huishoudens gestuurd; in totaal zijn bijna 48 000 mailings verstuurd. De overall respons was 54,4%. Een probitmodel is gebruikt om het effect van de karakteristieken gelijktijdig te bepalen. Hieruit blijkt dat alleen *adres* geen significant effect heeft op de respons (details zijn te vinden in Van der Scheer, 1997 en Vriens, Van der Scheer, Bult en Hoekstra, 1998). Bij het bepalen van het optimale design hoeft met dit karakteristiek dus geen rekening te worden gehouden. Tabel 3 geeft het optimale niveau van elk van de karakteristieken; dit definieert het optimale mailingdesign.

Tabel 3. Optimale mailingdesign.

Karakteristiek	**Niveau**
Acceptgiro	Los
Brochure	Geen
Illustratie	Geen
Vetgedrukte woorden	Veel
Post scriptum	Samenvatting
Handtekening	Professor

De verwachte respons bij het optimale mailingdesign is 59,7% (dit is te bepalen met behulp van het probitmodel), dit is een relatieve toename van de respons met 9,8%! Hieruit blijkt het belang aan van een goed ontworpen mailingdesign en het geeft aan dat een verfijning van het design op basis van een test, zoals hier beschreven, lonend is.

4.4. Interactie

Bij de bovenstaande methode wordt gekeken naar de wijze waarop het optimale mailingdesign kan worden bepaald. Nadat dit is gebeurd, worden de huishoudens voor de

campagne geselecteerd. De twee stappen, optimalisatie van mailingdesign en selectie van huishoudens, vinden dus na elkaar plaats. Het is echter beter om beide stappen te integreren. In dat geval wordt per huishoudens (bij selectie) het optimale mailingdesign bepaald. Met andere woorden, de bovenstaande methode wordt uitgebreid door rekening te houden met het interactie tussen de karakteristieken van de huishoudens en de karakteristieken van mailing. Het is namelijk aannemelijk dat deze interactie aanwezig is. Bijvoorbeeld, huishoudens die ieder jaar aan de charitatieve instelling doneren zullen geen behoefte hebben aan een algemene brochure van de instelling terwijl huishoudens die nog niet eerder zijn benaderd deze informatie juist wel op prijs stellen. Dit impliceert dat een brochure alleen bij huishoudens die voor het eerst een mailing ontvangen wordt bijgevoegd. Door rekening te houden met dit soort interacties worden verschillende typen mailings naar verschillende huishoudens gestuurd.

De interactie kan worden gemodelleerd binnen een probitmodel. Details hierover zijn te vinden in Bult, Van der Scheer en Wansbeek, 1997, en Van der Scheer, 1997. Hier wordt alleen ingegaan op de toepassing.

4.5. Toepassing (vervolg)

De toepassing van 4.3 wordt nu uitgebreid door de interactie mee te nemen. Van de huishoudens zijn vier karakteristieken bekend: het gedoneerde bedrag in de afgelopen twee jaren, de periode sinds de laatste donatie en de periode dat het huishouden op de mailinglist staat. Voor de mailingkarakteristieken zijn negen dummy-variabelen aangemaakt: voor elk van de karakteristieken met twee niveau's is één dummy-variabele geconstrueerd en voor de karakteristieken met drie niveau's zijn twee dummy-variabelen geconstrueerd. In totaal zijn in het model 50 parameters geschat: een intercept (1 parameter), het effect van een de karakteristieken van de mailing (9 parameters), het effect van de karakteristieken van de huishoudens (4 parameters) en interactie-effecten (36 parameters). Meer dan de helft van de interactie-effecten zijn statistisch significant. Dit geeft aan dat er inderdaad sprake is van interactie.

Voor ieder huishouden kan voor elk van de 288 mogelijke type mailings de responskans en de daaruit volgende verwachte winst worden bepaald. Hierbij moet expliciet rekening worden gehouden met de kosten van de mailing aangezien deze kunnen verschillen. Voor ieder huishouden kan dus het optimale mailingdesign worden bepaald; dit is de mailing die de hoogste verwachte winst voor dat huishouden genereert. Uit de toepassing bleek dat 52 mailings voor minimaal één huishouden als optimaal werd aangeduid. Zoals verwacht, het aantal huishoudens behorend bij een bepaalde optimale mailing varieerde sterk. De grootste aantallen waren 10585, 6157 en 5365, terwijl 13 mailings (van de 52) voor minder dan 10 huishoudens het optimale design was.

Het spreekt voor zich dat zo'n fine-tuned, gepersonaliseerde mailing methode niet optimaal is als ook alle vaste kosten worden meegenomen. De charitatieve instelling besloot daarom dat een mailing mocht worden meegenomen als deze voor minimaal één procent van huishoudens de optimale mailing is. Dit betekent dat er 13 type mailings van de 52 overblijven. Merk op: deze 13 mailings hoeven niet de mailings te zijn die in de test zijn gebruikt. Tabel 4 geeft de verwachte winst voor een aantal verschillende scenario's.

Tabel 4. Verwachte winst.

Scenario	**Verwachte winst (in guldens)**
Slechtste mailing	247 879
Traditionele mailing	388 180
Beste mailing	443 647
Top 13 mailings	484 515

Als benchmark is de slechtste mailing van de 288 mogelijke mailings gegeven; de verwachte winst is hierbij Fl. 247 879. De mailing die traditioneel door de charitatieve instelling wordt gebruikt geeft een winst van Fl. 388 180. Het verschil in winst bij deze twee scenario's geeft het belang van de karakteristieken van een mailingdesign aan. Indien de beste mailing wordt gebruikt, stijgt de winst met 14% tot Fl. 443 647. De winst stijgt zelfs tot Fl. 484 515 als de beste 13 mailing worden gebruikt. Dit is een toename van de winst met 25% ten opzichte van de mailing die traditioneel wordt ingezet. Deze cijfers geven aan dat een test om tot een optimaal design(s) te komen lucratief is.

5. CONCLUSIES

In dit artikel zijn een tweetal methoden gegeven die de effectiviteit en daarmee de verwachte winst van een direct-mailcampagne vergroten.

De eerste is een selectiemethode waarbij niet alleen de responskans van (potentiële) klanten wordt geschat maar ook het verwachte volume. De methoden die in de literatuur zijn beschreven en in de praktijk worden gebruikt, focussen echter alleen op de responskans. Dit leidt tot een suboptimale selectieregel. De resultaten van de toepassing laten zien dat het inderdaad belangrijk is om ook het verwachte volume te modelleren.

De andere methode om de effectiviteit van direct-mailcampagne te vergroten is door een optimaal mailingdesign te construeren. De traditionele aanpak om het effect van meerdere karakteristieken te analyseren is het apart beschouwen van elk karakteristiek. Meestal is zo een methode inefficiënt. In dit artikel is een methode geïntroduceerd om een optimaal design op een efficiëntere manier te bepalen. Deze methode is uitgebreid door de interactie tussen de karakteristieken van de mailing en de karakteristieken van de huishoudens mee te nemen. Het resultaat is een methode die simultaan het mailingdesign optimaliseert en targets selecteert. De toepassing laat zien dat er inderdaad sprake is van interactie; de optimale strategie resulteert in dertien verschillende soorten mailings. Elke target ontvangt één van deze mailings. Door deze strategie neemt de verwachte winst toe met 25% ten opzichte van de traditionele mailing en met 9% als één optimale mailing voor alle targets wordt gebruikt.

LITERATUUR

Amemiya, T. (1985), *Advanced Econometrics*, Harvard University Press, Cambridge, Massachusetts.
Bult, J.R. (1993), "Semiparametric versus parametric classification models: An application to direct marketing", *Journal of Marketing Research*, 30, 380-390.
Bult, J.R. en J.C. Hoekstra (1991), *Marktonderzoek en direct marketing*, in Jaarboek van de Nederlandse Vereniging van Marktonderzoekers 1991-92, De Vrieseborch, Haarlem.
Bult, J.R., H.R. van der Scheer en T.J. Wansbeek (1997), "Interaction between target and mailing characteristics in direct marketing, with an application to health care fund raising", *International Journal of Research in Marketing*, 14, 301-308.
Bult, J.R. en T.J. Wansbeek (1995), "Optimal selection for direct mail", *Marketing Science*, 14, 378-394.
Bauer, C.L. (1988), "A direct mail customer purchase model", *Journal of Direct Marketing*, 2 (3), 16-24.
Courtheoux, R.J. (1987), "Database modeling: Maximizing the benefits'', *Direct Marketing*, 3, 44-51.
DeJong, W. en A.J. Oopik (1992), "Effects of legitimizing small contributions and labeling potential donors as 'helpers' on responses to a direct mail solicitation for charity", *Psychological Reports*, 71, 923-928.
Den Uyl, M.J. en E. Langendoen (1997), *De inzet van adaptieve analysetechnieken in direct marketing*, in Jaarboek 1997 van de Nederlandse Vereniging voor Marktonderzoek en Informatiemanagement, De Vrieseborch, Haarlem.
DeSarbo, W.S. en V. Ramaswamy (1994), "CRISP: Customer response based iterative segmentation procedures for response modeling in direct marketing", *Journal of Direct Marketing*, 8 (3), 7-20.
Duan, N., W.G. Manning, C.N. Morris en J.P. Newhouse (1983), "A comparison of alternative models for the demand for the medical care", *Journal of Business & Economic Statistics*, 1, 115-126.
Fraser-Robinson, J. (1989), *The Secrets of Effective Direct Mail*, McGraw-Hill, London.
Huizingh, K.R.E., Th.W. de Boer en M. Wedel (1995), *Neurale netwerken versus statistische methoden: een vergelijking*, in Jaarboek van de Nederlandse Vereniging van Marktonderzoekers 1995-96, De Vrieseborch, Haarlem.
Judge, G.G., W.E. Griffiths, R.C. Hill, H. Lütkepol en T.C. Lee (1985), *The Theory and Practice of Econometrics*, second edition, Wiley, New York.
Kumar, A., V.R. Rao en H. Soni (1995), "An empirical comparison of neural network and logistic regression models", *Marketing Letters*, 6, 251-263.
Lix, T.S. en P.D. Berger (1995), "Analytic methodologies for database marketing in the US", *Journal of Targeting, Measurement and Analysis for Marketing*, 4, 237-248.
Maddala, G.S. (1983), *Limited Dependent Variables in Econometrics*, Cambridge University Press, Econometric Society Monographs 3, Cambridge.
Magidson, J. (1988), "Improved statistical techniques for response modeling", *Journal of Direct Marketing*, 2 (4), 6-18.
Magidson, J. (1994), "The CHAID approach to segmentation modeling: CHi-squared Automatic Interaction Detection", in R.P Bagozzi, editor, *Advanced Methods of Marketing Research*, Basil Blackwell, Cambridge, MA, 118-159.
Magliozzi, T.L. en P.D. Berger (1993), "List segmentation strategies in direct marketing", *OMEGA International Journal of Managment Science*, 21, 61-72.
Manning, W.G., N. Duan en W.H. Rogers (1987), "Monte Carlo evidence on the choice between sample selection and two-part model", *Journal of Econometrics*, 35, 59-82.
Manski, C.F. (1975), "Maximum score estimation of the stochastic utility model of choice", *Journal of Econometrics*, 3, 205-228.

Manski, C.F. (1985), “Semiparametric analysis of discrete response, asymptotic properties of the maximum score estimator”, 27, 313-333.
Otter, P.W., H.R. van der Scheer en T.J. Wansbeek (1997), “Direct mail selection by joint modeling of the probability and quantity of response”, *Som Research Report, 97B25, University of Groningen*.
Rao, V.R. en J.H. Steckel (1995), “Selecting, evaluating, and updating prospects in direct mail marketing”, *Journal of Direct Marketing*, 9 (2), 20-31.
Roberts, M.L. en P.M. Berger (1989), *Direct Marketing Management*, Prentice Hall, Englewood Cliffs, NJ.
Schibrowsky, J.A. en J.W. Peltier (1995), “Decision frames and direct marketing offers: A field study in a fundraising context”, *Journal of Direct Marketing*, 9 (2), 8-16.
Sharda, R. (1994), “Neural networks for the MS/OR analyst: An application bibliography”, *Interfaces*, 24 (2), 116-122.
Shepard, D. (1995)*, The new direct marketing: How to implement a profit-driven database marketing strategy*, second edition, Business One Irwin, Homewood, IL.
Smith, G.E. en P.D. Berger (1996), “The impact of direct marketing appeals on charitable marketing effectiveness”, *Journal of the Academy of Marketing Science*, 24, 219-231.
Throckmorton, J. (1992), “Creative methods: An overview”*,* in E.L. Nash (ed.), *The Direct Marketing Handbook*, McGraw-Hill, New York.
Van der Scheer, H.R. (1997*), Quantitative Approaches for Profit Maximization in Direct Marketing*, Ph.D. thesis, University of Groningen.
Van der Scheer, H.R., J.C. Hoekstra en M. Vriens (1996), “Using ‘opt-out’ reply cards in direct mail: Optimal design, target selection and profit implications”, *Journal of Direct Marketing*, 10 (3), 18-27.
Van der Scheer, H.R., J.C. Hoekstra en M. Vriens (1998), *De ‘opt-out’ antwoordkaart in direct marketing: privacy als selectie instrument*, in Jaarboek 1998 van de Nederlandse Vereniging voor Marktonderzoek en Informatiemanagement, De Vrieseborch, Haarlem.
Van der Veer, R.C.P., B. Wierenga en J.C.H.W. Kluytmans (1994), *Neurale netwerken in marketing*, in Jaarboek van de Nederlandse Vereniging van Marktonderzoekers 1994-95, De Vrieseborch, Haarlem.
Vögele, S. (1992), *Handbook of Direct Mail; Dialogue method of direct written sales communication*, Prentice Hall Int. (UK) Ltd, Hemel Hempstead.
Vriens, M., H.R. van der Scheer, J.R. Bult en J.C. Hoekstra (1998), “Conjoint experiments for direct mail response optimization”, *European Journal of Marketing*, 32, 323-339.
Zahavi, J. en N. Levin (1995), “Issues en problems in applying neural computing to target marketing”, *Journal of Direct Marketing*, 9 (3), 33-45.
Zahavi, J. en N. Levin (1997), “Applying neural computing to target marketing”, *Journal of Direct Marketing*, 11 (1), 5-22.

17. Vaststelling van de effecten van marketing-communicatie op het Internet

L.P. HOEK en W.F. VAN RAAIJ

SAMENVATTING

In dit artikel wordt een overzicht gegeven van de factoren die van belang zijn om de effecten van marketingcommunicatie op het Internet te meten. De effecten betreffen de kennis, houding en het (koop/bestel)gedrag van de doelgroep. De bestaande methoden om reclame-effecten te meten, vaak teruggaand tot de jaren '60 (Lucas en Britt, 1966), moeten worden aangepast om de effecten van marketing-communicatie op het Internet te kunnen vaststellen. Op basis van een verricht practijk-onderzoek wordt een aanzet aangegeven tot een meetmethode van deze effecten.

1. INLEIDING

Het artikel ziet er als volgt uit. In paragraaf 2 geven we een overzicht van de instrumenten en media voor marketingcommunicatie (MC), met enkele relevante eigenschappen en hun geschiktheid om thematische danwel actiematige communicatiedoelstellingen te bereiken. Vaak is synergie te bereiken door een combinatie van het Internet met bestaande MC-instrumenten en media. In paragraaf 3 volgt een overzicht van onderzoek naar de effecten van de vermelding van Internet-adressen, de waarde van banners, en de effecten van website-ontwerpen. In paragraaf 4 geven we de randvoorwaarden aan waaraan een meetinstrument voor de effecten van het Internet moet voldoen. De wijze waarop surfers door het Internet navigeren is te volgen met 'click-streams'. Naast dit surfproces zijn effectmetingen mogelijk wat betreft de veranderingen in kennis, attitude en gedrag als gevolg van het Internetgebruik. Dit staat in paragraaf 5. In paragraaf 6 wordt de opzet van een meetinstrument geschetst en in paragraaf 7 volgen enkele conclusies en aanbevelingen..

2. MARKETINGCOMMUNICATIE

De laatste jaren zijn verscheidene instrumenten en media toegevoegd aan de mogelijkheden die ter beschikking staan voor het marketingcommunicatiebeleid. Sportsponsoring, organisatie en sponsoring van evenementen, non-spot advertising (product placement en merkvermelding in TV-programma's) en het Internet zijn voorbeelden van de nieuwe communicatie-mogelijkheden. Een MC-instrument is een geheel van samenhangende mogelijkheden om in contact te komen met doelgroepen en informatie te verschaffen over merken van producten en diensten. In Tabel 1 zijn de mogelijke

MC-instrumenten vermeld met hun relatieve kosten per contact, hun mate van individualisatie en hun mogelijkheden voor thematische en actiematige marketingcommunicatie. Individualisatie is de tegenpool van massacommunicatie en betreft de aanpassing van de boodschap aan de individuele behoeften en wensen van consumenten of andere afnemers. Thematische communicatie betreft de beïnvloeding van kennis en attitude van consumenten. Actiematige communicatie betreft de beïnvloeding van het gedrag. Zie ook Floor en Van Raaij (1998).

2.1. Marketingcommunicatie-instrumenten en -media

In Tabel 1 staan de MC-instrumenten aangegeven met enkele eigenschappen. Hieronder bespreken we ze in meer detail.

Tabel 1. Marketingcommunicatie-instrumenten.

MC-instrumenten	Kosten	Individualisatie	Thema	Actie
Reclame	laag	–	++	+
Marketing-PR	laag	–	++	–
Sponsoring	laag	+	++	–
Sales promotie	gem.	+	+	++
Direct marketing	gem.	++	++	++
Winkelcommunicatie	laag	–	+	++
Persoonlijke verkoop	hoog	++	+	++
Beurzen en tentoonstellingen	hoog	++	++	+
Internet	gem.	++	++	++

Reclame in massamedia betreft advertenties in dag- en vakbladen en tijdschriften, en commercials op televisie en radio. De kosten per contact (CPM (cost per Mille) of GRP (Gross Rating Points)) voor reclame zijn relatief laag, maar de thematische en actiematige effecten zijn doorgaans niet zo groot. Door herhaling van de boodschap kunnen geleidelijk thematische en gedragseffecten ontstaan. De mate van individualisatie is laag; alle consumenten zien in principe dezelfde boodschap. Hooguit kunnen voor verschillende doelgroepen verschillende boodschappen in verschillende media worden geplaatst. Door toevoeging van response-mogelijkheden (postadres, Internet-adres of telefoon/faxnummer) kan reclame een basis worden voor de opbouw van een database en voor direct marketing. Bij veel campagnes is reclame het hoofdinstrument.

Bij *marketing-PR* worden tussenpersonen, bijvoorbeeld journalisten, ingeschakeld om de doelgroep te bereiken. Bij de introductie van een nieuw product of bij een product-incident zijn goede relaties met de pers van levensbelang. De kosten en de mate van individualisatie 'per tussenpersoon' zijn hoog, maar de kosten per bereikte doelgroeppersoon kunnen relatief laag, soms zelfs gratis (free publicity) zijn. De mate van individualisatie per doelgroeppersoon is laag. De thematische effecten kunnen hoog zijn, aangezien de redactionele informatie in een medium als betrouwbaarder en objectiever wordt gezien dan reclame.

Sponsoring heeft vooral een thematische doelstelling: de vergroting van merkbekendheid en het creëren van gunstige associaties en goodwill. De kosten per doelgroepper-

soon zijn in het algemeen laag. Er is enige individualisatie mogelijk door specifieke gebeurtenissen of evenementen te sponsoren, gericht op een bepaalde doelgroep. Sponsoring is doorgaans een aanvullend MC-instrument.

Sales promotie is het tijdelijk gunstiger maken van de prijs/waarde-verhouding van een product of dienst, bijvoorbeeld door een prijsverlaging of een gratis toegift (premium). SP heeft vooral een gedragsdoelstelling; het bevorderen van probeeraankopen (trial) bij nieuwe klanten en volumevergroting en trouw bij bestaande klanten. De mate van individualisatie is laag, tenzij de SP wordt gekoppeld aan een loyalty-programma zoals bij Albert Heijn en Edah. In dat geval profiteren vooral de trouwe klanten van de aanbiedingen.

Direct marketing kan een zelfstandig marketingsysteem zijn met eigen MC-communicatie en distributie. De DM-communicatie maakt gebruik van media als post (catalogus) en interactieve media zoals telefoon (voice response systemen, telemarketing) en het Internet. De individualisatie kan bij DM hoog zijn, doordat de boodschap aangepast kan worden aan de individuele omstandigheden en behoeften. Vooral bij interactieve media is er sprake van een snelle terugkoppeling bij de verwerking van informatie en beantwoording van vragen. Communicatie en transactie gaan naadloos in elkaar over. De kosten per post- of telefoon-contact zijn gemiddeld; de thematische en actiematige effecten zijn relatief hoog.

Winkelcommunicatie omvat de merchandising van producten in de winkel: productpresentatie, schapindeling en verpakking. De mate van individualisatie is gering, maar de effecten op het koopgedrag kunnen groot zijn, aangezien veel consumenten pas in de winkel beslissen welk product of merk ze kopen.

Persoonlijke verkoop is relatief duur per contact. Vertegenwoordigersbezoek kost honderden guldens en ook persoonlijke bediening in de winkel is relatief duur. Anderzijds is dit persoonlijke contact vaak effectief om klanten te helpen, hun keuzeproces te beïnvloeden en relaties te onderhouden. Persoonlijk contact is in feite het oudste interactieve medium.

Beurzen en tentoonstellingen (B&T) vormen een complex instrument, waarin veel andere MC-instrumenten zoals persoonlijke verkoop, merchandising, direct marketing en marketing-PR een rol spelen. Ook komen verscheidene media in het kader van B&T voor: gedrukte media zoals catalogi, folders en brochures, voorts video- en 'life'-productdemonstratie en telefoon bij de follow-up van beurscontacten. De kosten per contact zijn door de individualisatie relatief hoog, maar de thematische en gedragseffecten kunnen ook hoog zijn, vooral in de business-to-business context.

<u>Media</u>

In het kader van deze MC-instrumenten komen *media* voor om met de doelgroep te communiceren. Het betreft de volgende media (Floor en Van Raaij, 1998):

- *gedrukte media*, zoals dagbladen, vakbladen, tijdschriften, brochures, catalogi en telefax.
- *ethermedia*, zoals televisie en radio.
- *rechtstreekse media*, zoals persoonlijk contact en direct mail.

- *interactieve media*, zoals het Internet en telefoon (voice response systemen).
- *winkelmedia*, zoals verpakkingen, schapkaarten, posters en winkelwagentjes.
- *buitenmedia*, zoals billboards, vrachtauto's en luchtballonnen.
- *bioscopen.*
- *reclameartikelen*, zoals gratis balpennen en kalenders, en kleding met reclame-opdruk.

Internet: een MC-instrument of een medium?
Het belang van het *Internet* als communicatiemedium neemt toe. Het Internet is in feite meer dan een medium; het is een MC-instrument omdat het niet alleen communicatie- maar ook transactie- en distributiemogelijkheden biedt. 'Digitale' producten zoals software, tekst, muziek, beelden, financiële en toeristische diensten, kunnen via het Internet worden besteld, betaald en geleverd. Het Internet is een interactief medium/instrument, dus met een grote mate van individualisatie. Gebruikers kunnen zelf het initiatief nemen op de tijd en plaats die ze zelf wensen. Het medium/instrument is actueel aangezien er een voortdurende updating van de informatie mogelijk is. Het Internet is thematisch vanwege de informatie die verstrekt wordt en actiematig vanwege de rechtstreekse bestel- en betaalmogelijkheden via het Internet. De kosten per contact liggen zelfs lager dan bij direct mail of telemarketing.

2.2 Synergie uit combinaties met het Internet

Het Internet kan in veel combinaties met andere MC-instrumenten en -media een rol spelen. Vaak levert een combinatie van een 'traditioneel' medium met Internet een bepaalde meerwaarde of synergie op. Voorbeelden van deze combinaties zijn vermelding van het Internet-adres in advertenties en commercials, en het gebruik van het Internet om actuele informatie te verkrijgen over het assortiment, de prijzen, de leverbaarheid van goederen en de after-sales service. Een Internet-site kan een vervanging zijn voor een catalogus of kan informatie op CD-ROM completeren met actuele prijs- en leverbaarheids-informatie.

Het Internet kan een zelfstandig marketingsysteem zijn, als informatie via het Internet verspreid wordt, transactie via het Internet gedaan kunnen worden en ook de goederen via het Internet geleverd kunnen worden. Dit betreft 'digitale' goederen zoals software, teksten, afbeeldingen, informatie-, reserverings-, boekings- en financiële diensten, muziek en video.

Stadia van een transactie
In een transactiecyclus kunnen de volgende stadia worden onderscheiden: (1) attendering, (2) uitwerking, (3) herinnering (reminder), (4) feitelijke transactie (aankoop) en (5) after-sales.

1. In de *attenderingsfase* spelen instrumenten als reclame, marketing-PR, sales promotie, merchandising en eventueel direct marketing een grote rol. Het zijn instrumenten die tegen relatief geringe kosten grote groepen kunnen bereiken. Goede attenderingsmedia zijn televisie, radio, advertenties in bladen, winkelmedia en direct mail. Door direct response-mogelijkheden in te bouwen, kunnen consumenten reageren. Daarna kunnen deze respondenten met meer toegespitste en geïndividualiseerde informatie

worden benaderd. Attendering kan betrekking hebben op een nieuw merk, een verbeterd product, een probleemoplossing of een speciaal aanbod. Attendering is dus deels een zelfselectieproces van consumenten die interesse hebben voor een bepaald product of een bepaalde dienst.

Banner ads bieden ook een attenderingsfunctie. Surfers op het Internet komen op hun doelgerichte of recreatieve zoektocht deze banners ongevraagd tegen als een soort virtuele buitenreclame. Door de banner aan te klikken, is onmiddellijk meer informatie te verkrijgen.

2 In de *uitwerkingsfase* vervullen gedrukte media (brochures, catalogi) en het Internet een belangrijke rol. Deze media verschaffen uitgebreide informatie aan consumenten die in een bepaald onderwerp zijn geïnteresseerd en zich willen oriënteren over het aanbod. Deze media bieden ook transactiemogelijkheden, als een consument een keuze heeft gemaakt. Door middel van 'intelligent search agents' kan een consument zonder al te veel moeite informatie vergaren op het Internet. Ook traditionele mogelijkheden zoals winkelbezoek of bezoek van een vertegenwoordiger bieden de mogelijkheid om informatie te vergaren.

3 Een *reminder* is een herhaalde attendering of het is een reminder om de informatie uit de brochure, catalogus of Internet-site te gebruiken. Het is een follow-up om in een tweede ronde 'meer' uit de markt te halen bij personen die informatie hebben opgevraagd, maar (nog) geen transactie hebben gedaan.

4 In de *transactiefase* besluit de consument tot een aankoop (gesloten propositie) of tot het opvragen van meer informatie. Direct mail en telefax hebben hierbij een trage terugkoppeling, terwijl snelle interactieve media als persoonlijke verkoop/bediening, telefoon en het Internet een snelle terugkoppeling mogelijk maken om de transactie te bevestigen en de levering en betaling af te spreken.

5 De *after-sales fase* omvat de elementen na de transactie die de klant helpen optimaal nut en tevredenheid te halen uit het product of de dienst. Denk hierbij aan garanties, gebruiksaanwijzingen, onderhoudstips, lijst van dealers en reparateurs en catalogi van complementaire producten en diensten. Het Internet kan hier een nuttige functie vervullen. Gebruikers van een product of dienst kunnen via het Internet hun ervaringen uitwisselen in een 'virtual community'. Via het Internet kunnen klantentevredenheid, klantentrouw, herhalingsaankopen en cross-selling van complementaire producten bevorderd worden.

Het Internet kan dus vooral in de uitwerkings-, transactie- en after-sales fase benut worden, meestal in combinatie met andere instrumenten en media. Aangezien het Internet een instrument is waarbij de consument het initiatief tot communicatie neemt, zijn attendering en reminding vooral voorbehouden aan traditionele media, behalve bij banner ads voor wie op het Internet surft. In dit geval kunnen we spreken van sequentiële integratie van instrumenten en media (Van Raaij, 1994).

In het bovenstaande is uitgegaan van een informatieve benadering van het transactie- en koopproces. Hierbij speelt informatie de hoofdrol om een beslissing te kunnen nemen. Hiernaast bestaat het Internet ook als een medium voor amusement. Ook commerciële web-sites kunnen entertainment bieden. Denk aan de sites van Heineken en Grolsch met wedstrijden, puzzels en informatie over het uitgaansleven en films. Voor

deze merken is het Internet een middel om positieve associaties en loyaliteit aan het merk te koppelen.

3. BESTAAND ONDERZOEK

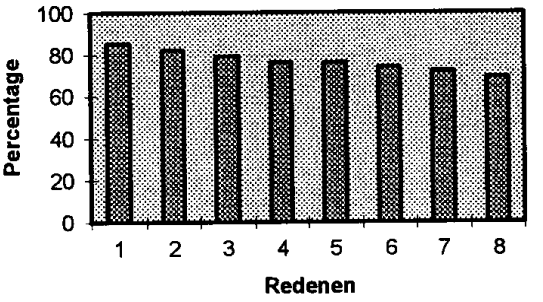
Er is reeds onderzoek gedaan naar de effectiviteit van het Internet in het algemeen en ook naar de effectiviteit van commerciële boodschappen op het Internet. Meestal zijn de uitgangspunten van deze onderzoeken anders dan dit onderzoek. Er zijn echter wel overeenkomsten en de resultaten geven een indicatie van de te definiëren meetmethode. Ook de methodologie van de hieronder beschreven onderzoeken geven inzicht in de verschillende methodieken die toegepast kunnen worden. Zo is er onderzoek gedaan naar de effectiviteit van het vermelden van het Internetadres in traditionele media (Maddox en Mehta, 1997, blz. 47-59). Dit onderzoek geeft aan dat met name de combinatie van verscheidene media een effectieve communicatie-strategie kan zijn. Ook is er onderzoek gedaan naar de waarde die consumenten hechten aan commerciële boodschappen op het Internet en waar deze *waarde* door wordt bepaald (Ducoffe, 1996, blz. 21-35). In de laatste studie wordt nagegaan welke fysieke ontwerpen van een webpagina of website het meest effectief zijn (Dreze en Zufryden, 1997, blz. 77-91). Een webpagina is een pagina met informatie op het World Wide Web (WWW), een onderdeel van het Internet.

3.1 Het vermelden van het Internetadres in andere media

Om consumenten te laten weten dat een organisatie een website op het Internet heeft, wordt de laatste tijd steeds vaker de URL (Uniform Resource Locator of web-adres: http://www.etc) genoemd in bijvoorbeeld televisiereclame, in een advertentie in een krant of tijdschrift of op de radio. Het noemen van een web-adres op de radio geeft overigens problemen omdat een dergelijk adres vaak moeilijk te onthouden is.

In Amerika is het effect van het vermelden van een URL in de traditionele media onderzocht (Maddox en Mehta, 1997). Dit onderzoek geeft met name aanknopingspunten hoe de combinatie van traditionele media en het Internet kan worden geanalyseerd. Uit het onderzoek komt naar voren dat de meeste respondenten een vermelding van de URL in traditionele media opmerken (83%), waarbij vooral de computer-, telecommunicatie- en de financiële branche hoog scoren. Verder worden URL's met name gezien op de televisie (60%), op afstand gevolgd door tijdschriften (20%) en kranten (17%). Op zich is dit niet verbazingwekkend, daar het televisiegebruik ook vele malen hoger ligt dan het gebruik van andere media. 13% van de respondenten geeft aan dat zij een web-site daadwerkelijk bezoeken na het zien van een vermelding van het web-adres in een bepaald medium. Het mag duidelijk zijn, dat hoewel veel mensen URL's in media opmerken, slechts weinigen daadwerkelijk de vermelde Internetsite bezoeken. Het is dus de vraag of de combinatie van traditionele media en het Internet effectief is.

Een belangrijke vraag is waarom respondenten na het zien van een advertentie met een verwijzing naar een bepaalde site, deze site daadwerkelijk bezoeken. In figuur 1 staat een overzicht van de belangrijkste redenen.



Bron: 1997, blz. 47-59

1. Zoeken van gedetailleerde informatie; 2. Interesse in productcategorie; 3. Interesse in merk; 4. Wil meer leren / informatie; 5. Koopintentie; 6. After-sales support; 7. Gemakkelijkste weg om aan informatie te komen; 8. Wenst geen persoonlijk contact
Bron : Maddox en Mehta, Journal of Advertising Research, 1997, blz. 47-59

Figuur 1. Redenen om een Internetsite te bezoeken na het zien van een URL in traditionele media.

Zoals blijkt uit de figuur is met name de interesse voor meer informatie van doorslaggevend belang om een Internetsite te bezoeken. Of het nu gaat om interesse in meer details, interesse in het product of merk, of gewoon de wens om meer te weten te komen over een bepaald product, dienst of merk, informatie is de belangrijkste beweegreden een Internetsite te bezoeken. Ook scoren de intentie een aankoop te doen, of het krijgen van na-aankoop service hoog bij deze respondenten. Het mag duidelijk zijn dat de conclusie van dit onderzoek is dat een commerciële uiting op het Internet vooral zal moeten bestaan uit relevante informatie.

De allernieuwste ontwikkeling op het gebied van de combinatie van traditionele media en het Internet is het gebruik van barcodes (Automatiseringsgids, 1997). Een bepaald artikel in een tijdschrift, maar ook bijvoorbeeld een plaatje van een product, zijn voorzien van een barcode. Door gebruik te maken van een barcode-lezer die in verbinding staat met de PC, kan direct toegang worden verkregen tot de website van de betreffende organisatie. Deze ontwikkeling is met name handig bij Internetpagina's die vrij 'diep' gesitueerd zijn. Informatie over bijvoorbeeld een deel van een productassortiment vereist een aantal handelingen (mouse-clicks) om te kunnen bereiken vanaf de homepages. Een barcode brengt de gebruiker direct bij de betreffende pagina, zonder dat allerlei ingewikkelde URL's moeten worden ingetypt en de route naar de site behoeft te worden afgelegd.

3.2 De waarde van banners op het Internet

Een ander onderzoek heeft zich met name geconcentreerd op de vraag of 'banner ads' of 'banners', dat zijn kleine commerciële uitingen die een link bevatten naar een Internetsite van een onderneming, effectief zijn om een boodschap over te brengen (Briggs en Hollis, 1997). Dit onderzoek heeft zich voornamelijk geconcentreerd op de technologie- en mannenkleding-branches. Er is onderzocht of het aanklikken van de banners

ook daadwerkelijk effect sorteert. Het onderzoek gaat verder dan het analyseren van de zogenaamde *click-through rates*, dat wil zeggen de statistieken die aangeven hoeveel mensen een banner aanklikken. Briggs en Hollis (1997) proberen te meten of het aanklikken van een banner de attitude en het gedrag van consumenten beïnvloedt. Het onderzoek bestond uit het ondervragen van een experimentele groep, die een Internet-adres bezocht, en een controlegroep, die geen reclame-uiting onder ogen kregen. Op deze wijze was men in staat een goede vergelijking te maken tussen respondenten die een reclame-uiting hebben gezien en respondenten die deze uiting niet hebben gezien. Verder werd er onderscheid gemaakt tussen effecten die onmiddellijk waarneembaar zijn (na een exposure) en vertraagde effecten, die het resultaat aangeven van meerdere exposures in de loop van de tijd. In dit laatste geval worden campagnes in de tijd gevolgd (tracking) en worden effecten over een bepaalde periode vastgesteld (Hoogerbrugge, 1998).

In het onderzoek was sprake van een laboratoriumsituatie. Hiermee wordt bedoeld dat de effectiviteit van de banners werd onderzocht in een omgeving die anders was dan de 'normale' omgeving waarin Internet-surfers met banners wordt geconfronteerd. De externe validiteit van het onderzoek is twijfelachtig. Gedragen respondenten zich in een 'normale' situatie net zo als in het laboratorium? Kunnen we de resultaten van een laboratorium-onderzoek generaliseren naar de praktijk? Het bovenstaande onderzoek is goed 'online' uit te voeren zonder dat de respondenten zich anders gaan gedragen dan in een 'normale' situatie. In de toekomst zullen marktonderzoekers zich steeds meer gaan bezighouden met online dataverzameling, al dan niet aangevuld met surveys en interviews. Maar ook deze surveys en interviews zullen in de toekomst via het interactieve medium zelf worden afgenomen.

3.3 Effectiviteit van een website ontwerp

Naast bovenstaande onderzoeken is er onderzoek gedaan naar welke fysieke elementen de effectiviteit van een website beïnvloeden (Dreze en Zufryden, 1997, blz. 77-91). Op basis van een conjuncte analyse zijn een aantal kenmerken van een website geanalyseerd. Deze analyse is onderverdeeld in twee niveaus: die van de individuele webpagina en die van een gehele *site*. Het onderzoek heeft de elementen 'achtergrond' (geen achtergrond, groene of geblokte achtergrond), 'omvang plaatjes' (klein of groot), 'geluid' (generiek of computer-afhankelijk) en het gebruik van bekende personen (presentatoren) met betrekking tot communicatie via het Internet geanalyseerd. Er werden verschillende webpagina's ontwikkeld, elk met een andere lay-out. Door combinaties van pagina's aan te bieden kon worden achterhaald welke fysieke kenmerken het meest geschikt zijn voor commerciële uitingen. Respondenten kregen een eigen URL tot hun beschikking en moesten aangeven hoe ze de betreffende pagina's waardeerden. Op deze wijze ontstond er een beeld welke fysieke kenmerken het meest effectief zijn. De effectiviteit werd bepaald aan de hand van de tijd die op een pagina werd doorgebracht en het aantal pagina's dat werd opgevraagd.

Aan dit onderzoek zijn een aantal voorwaarden verbonden (Dreze en Zufryden, 1997, blz. 78-79) :

- Het onderzoek moest *unobtrusive* (niet-verstorend) zijn. Dit houdt in dat de meetmethode geen invloed moet hebben op de resultaten van de meting. Veel van de be-

staande meetmethoden zijn duidelijk herkenbaar voor de respondenten en dit kan de resultaten van de meting beïnvloeden. Deze voorwaarde kon worden bereikt door het observeren van *click-streams*, het feitelijke surfgedrag van de Internetgebruiker zonder dat de gebruikers wisten dat het surfgedrag werd vastgelegd. Uit ethische overwegingen moest dan aan de respondenten achteraf toestemming worden gevraagd om hun gegevens in het onderzoek te gebruiken.
- Externe validiteit. Dit houdt in dat de meetmethode de inhoud van een website analyseert door het feitelijke gedrag op zo'n website te meten, en dus niet in een laboratoriumsituatie.
- Gebaseerd op grote steekproeven. De traditionele conjuncte analyse wordt door de complexiteit en hoge kosten meestal toegepast op hooguit honderd respondenten. Door de methode te implementeren op het Internet is het mogelijk het gedrag van alle bezoekers te analyseren en hieruit een grotere steekproef te trekken. Dit verhoogt de betrouwbaarheid van de resultaten.

Deze methode van onderzoek heeft voorts de volgende voordelen:
- Het kan onmiddellijk 'online' worden verricht. Door te meten op het moment van exposure zijn de resultaten snel beschikbaar.
- Door het feitelijke surfgedrag op het moment van exposure te meten is het niet nodig kunstmatige situaties te creëren. Hierdoor ontstaat dus een hoge externe validiteit.
- De registratie van de gegevens is online en dus relatief goedkoop.

4. RANDVOORWAARDEN

Het te ontwikkelen meetinstrument zal aan een aantal voorwaarden moeten voldoen, om praktisch hanteerbaar, efficiënt én effectief te zijn. De voorwaarden zijn onder te verdelen in algemene randvoorwaarden, randvoorwaarden die voortvloeien uit de specifieke kenmerken van het Internet en praktische randvoorwaarden, die uit het praktijkonderzoek naar voren zijn gekomen.

4.1. Algemene randvoorwaarden

Het meetinstrument moet trachten zoveel mogelijk *unobtrusive* (niet-verstorend) zijn. Dit houdt in dat de meetmethode geen invloed mag hebben op de resultaten van de meting. Veel van de bestaande meetmethoden zijn duidelijk herkenbaar voor de respondent en dit kan de resultaten van de meting beïnvloeden. Deze voorwaarde kan worden bereikt door het observeren van *click-streams*, het feitelijke surfgedrag van de Internetgebruiker.

Hieruit vloeit voort dat het instrument een zekere externe validiteit moet bevatten. Dit houdt in dat de meetmethode de inhoud van een website analyseert door het feitelijke gedrag op een website te meten., en niet op basis van kunstmatig gecreëerde meetsituaties.

Verder moet het meetinstrument het gedrag liefst onmiddellijk (on-line of real-time) meten. Hierdoor zijn de kosten lager en kunnen de resultaten sneller bekend zijn.

4.2. Randvoorwaarden Internet

De belangrijkste randvoorwaarde die kan worden afgeleid uit de kenmerken van het Internet, is dat Internet de mogelijkheid biedt voor twee-weg communicatie. Gebruikers (consumenten) kunnen reageren op aangeboden informatie, waarna de aanbieder van informatie (producent) op basis van deze reactie het informatie-aanbod kan aanpassen. Voor dit onderzoek betekent dit, dat het mogelijk is om op het moment dat een Internetgebruiker een commerciële boodschap onder ogen krijgt, de reactie op deze boodschap te meten. Het moet derhalve mogelijk zijn om via het medium Internet direct te achterhalen welk effect een commerciële boodschap heeft op een consument. Een dergelijke methode zorgt ervoor, dat de informatie wordt achterhaald in een *real-life situation*, in plaats van een kunstmatige situatie, zoals in een laboratorium-onderzoek. Hierdoor zullen de omstandigheden waaronder een onderzoek wordt uitgevoerd, nagenoeg identiek zijn aan de 'normale' situatie waarin een respondent een commerciële boodschap onder ogen krijgt. Bijkomende voordelen om direct via het Internet de effectiviteit van een reclame-uiting te meten is dat het goedkoper en sneller is.

Uit de analyse van het Internet is verder naar voren gekomen dat consumenten die commerciële informatie opvragen op het Internet, al een bepaalde belangstelling hebben voor deze informatie. Daar de Internetgebruiker zelf beslist welke informatie hij/zij wil zien, in welke volgorde en voor hoe lang (internal pacing; zie Pieters en Van Raaij, 1992, hoofdstuk 18), is het uitgangspunt anders dan bij traditionele media. Traditionele media zoals televisie worden gekenmerkt door externe pacing, d.w.z. dat de zender bepaalt hoeveel en in welke volgorde en snelheid de informatie wordt uitgezonden. De consument heeft deze informatie maar te accepteren zoals het uitgezonden wordt.

Bovendien is het bij televisie en radio onduidelijk of een consument geïnteresseerd is in de boodschap, en welke houding (attitude) hij heeft ten opzichte van het aangeboden product. Een surfer op het Internet die zelf informatie opvraagt, heeft waarschijnlijk al bij voorbaat meer interesse voor de aangeboden informatie. Het zal dus noodzakelijk zijn om inzicht te krijgen hoe een bezoeker van een website op het moment van 'binnenkomst' (vóórmeting) denkt over het product of de dienst (attitude) alsmede het kennisniveau ten aanzien van de productklasse en het individuele merk. Daarnaast kan inzicht in het gedrag uit het verleden met betrekking tot de aangeboden producten, alsmede inzicht in de gedragsintentie, een ijkpunt opleveren voor de effectiviteitsmeting. Als nu deze beginwaarden voor de te meten variabelen bekend zijn, kunnen aan het einde van het bezoek aan de *site* deze variabelen opnieuw worden gemeten, waardoor de effectiviteit van de reclameboodschap het verschil tussen begin- en eindwaarde van de te meten variabelen vormt.

Een laatste voorwaarde voor het te ontwikkelen meetinstrument heeft consequenties voor de manier waarop de meting wordt verricht. Daar gebruikers van het Internet verwachten dat een website informatie, liefst zelfs gecombineerd met entertainment, biedt, zal het meetinstrument aan deze voorwaarden moeten voldoen. Mogelijkheden om aan deze voorwaarden te voldoen is het meten in spelvorm, of het aanbieden van een *incentive*, in de vorm van bijvoorbeeld een te winnen prijs of een beloning voor de medewerking.

Naast de bovenstaande randvoorwaarden is nog één aspect onbesproken gebleven: de huidige technische mogelijkheden van het Internet. Uiteraard zal ieder meetinstrument rekening moeten houden met de technologische sterkten en zwakten van het Internet. Voor dit onderzoek betekent dit, dat de methode waarop commerciële uitingen kunnen worden beoordeeld, technisch haalbaar moet zijn.

Op dit moment bestaan methoden om de *click-streams* van Internet-gebruikers te volgen. Zo zijn er log-files beschikbaar, die inzicht geven in wie de site bezoekt en hoe een surfer 'binnenkomt'. Deze log-files zijn (Allen, 1998):
1 – *Access files* (file names, IP adressen, datum en tijd)
2 – *Referer files* (URL of linking website)
3 – *Browser files* (type browser software)
4 – *Error files* (incomplete file-requests en error-codes).

Het probleem met deze log-files is bijvoorbeeld dat bij bezoekjes van één netwerk (server) door verschillende mensen de log-files slechts één URL-code vermelden, terwijl waarschijnlijk meerdere mensen de betreffende site hebben bezocht. Een aantal providers heeft een beperkt aantal gateways (toegangswegen) naar het Internet, en dit levert dus hetzelfde probleem op. De huidige log-software is nog niet geheel in staat deze problemen op te lossen (Allen, 1998).

Naast deze log-files is het verder mogelijk voor een onderneming om *cookies* aan te brengen op de harde schijf van een website-bezoeker. Dit zijn kleine bestanden die worden aangebracht tijdens het bezoek van een surfer aan een bepaalde site. De 'cookies' kunnen bijvoorbeeld informatie bevatten over de geraadpleegde bestanden of de ingevoerde informatie. Het koppelen van 'cookies' van eerdere bezoekjes met een hernieuwd bezoek kan verschillen in surfgedrag blootleggen. Het probleem met 'cookies' is dat veel gebruikers van het Internet hun computer afschermen tegen het plaatsen en veranderen van deze bestanden.

4.3 Praktijkonderzoek

Uit het uitgevoerde praktijkonderzoek kunnen de volgende conclusies worden getrokken.

De meeste ondernemingen gebruiken het Internet op dit moment als communicatie-instrument, waarbij het productaanbod een belangrijke rol speelt. Daarnaast wordt het Internet vooral van belang geacht in combinatie met andere media. Ook het vergroten van de merkbekendheid en merkkennis zijn doelstellingen die in de praktijk worden gehanteerd. Het Internet wordt voornamelijk gebruikt om de kennis-component van consumenten te beïnvloeden. De te ontwikkelen meetmethode zal dus deze kennis-component moeten meten, eventueel in relatie met andere media.

Daarnaast is een belangrijke doelstelling van de aanwezigheid op het Internet het beïnvloeden van gedragsgerelateerde eigenschappen. Dit heet *gedrags- of koopfacilitatie*. Gedragsfacilitatie (Floor en Van Raaij, 1998) is het verschaffen van informatie waar en hoe het product of de dienst verkregen kan worden: verkrijgbaarheid, betaalmogelijkheden, leveringsvoorwaarden. Daar het nog niet echt mogelijk is het daadwerkelijke

gedrag onmiddellijk te beïnvloeden, kan worden geconcludeerd dat bedrijven op deze wat indirecte wijze de consumenten proberen aan te sporen actie te ondernemen. Naast de kenniscomponent probeert men dus ook de gedragscomponent te beïnvloeden. Dit element zal dus tevens terug moeten komen in de meetmethode.

Veel ondernemingen hebben niet alleen een eigen website, maar hebben tevens op meerdere plaatsen verwijzingen aangebracht naar de eigen site. De belangrijkste boodschap, en dus het belangrijkste object van onderzoek is de inhoud van de eigen website, gekoppeld aan de specifieke doelstellingen. Tevens zal echter moeten worden gekeken wat de betekenis is van een consument die via een banner of een zoeksysteem bij de betreffende website terecht is gekomen. Dit kan een indicatie zijn van een bepaalde interesse, een bepaalde kennis of een bepaalde attitude. Daarnaast kan het natuurlijk zo zijn dat er reeds een intentie tot aankoop aanwezig is, en dat men op zoek is naar specifieke gedragsgerelateerde informatie.

Het meetinstrument moet dus aan de volgende voorwaarden voldoen:
1. Kennis (merkbekendheid en merkkennis) meten.
2. Attitude (attitude en evaluatie van de website en het merk) meten.
3. Gedragsfacilitatie en gedragsintentie meten.
4. De inhoud van de eigen website, alsmede het verkeer van en naar deze site vastleggen.
5. Gebaseerd zijn op het feitelijk gedrag ten aanzien van website, waardoor geen kunstmatige situatie hoeft te worden gecreëerd (kosten) én de metingen een hoge externe validiteit bezitten.
6. Rekening houden met andere beginwaarden van interesse en kennis van consumenten bij het Internet in vergelijking met massamedia, doordat de consument bij het Internet zelf bepaalt wat hij/zij wil zien (internal pacing). Indien de bewuste medewerking van consumenten noodzakelijk is, zal de informatie- en amusementswaarde zo groot mogelijk moeten zijn om consumenten te overreden mee te werken. Ook kan een *incentive* consumenten overreden mee te doen.
7. Rekening houden met het feit dat het Internet als een complementair marketingcommunicatie-instrument wordt gebruikt en dus onderdeel is van de totale marketingcommunicatie-mix in bepaalde stadia van het oriëntatie- en transactieproces.

5. EFFECTMETING INTERNET

Bij de effectmeting worden cognitieve (kennis), evaluatieve (attitude) en gedragscomponenten onderscheiden. Kennis en attitude moeten door ondervraging van respondenten gemeten worden; gedrag kan niet-verstorend worden geregistreerd. In de volgende secties komen deze drie componenten aan de orde.

5.1. Kenniscomponenten

Er bestaan een aantal meetmethoden om kenniscomponenten als merkbekendheid en merkkennis te meten. Met name de posttests van communicatie-effecten zijn in dit geval relevant. Posttest-methoden meten na het zien van een commerciële uiting de impact die zo'n uiting heeft gehad op de kennis en houding van de consument. Herken-

nings- en herinneringstests zijn de meest relevante van deze meetmethoden (Pieters en Van Raaij, 1992, hoofdstuk 10).

Herkenningstests (recognition) zijn geschikt om verschillende advertenties met elkaar te vergelijken om zodoende inzicht te krijgen welke uitingen het meest opvallen en herkend worden. Deze methoden meten in hoeverre een uiting wordt herkend en welke elementen opvallen. Het feit dat een commerciële uiting wordt herkend, geeft aan dat de consument bekend is met het merk, en wellicht zelfs enige kennis van de eigenschappen van dit merk heeft (dit hoeft uiteraard niet). Deze methode kan worden gebruikt bij *banners* en vermeldingen in zoeksystemen door middel van het grafisch weergeven van de essentie van een bepaalde reclame (bv. logo of beeldmerk). Het aanklikken van zo'n logo betekent dan dat de consument hoogstwaarschijnlijk het beeldmerk herkent én dat hij of zij meer wil weten over het betreffende merk. Het meten van de *click-through* rates van deze beeldmerken / logo's geeft dus een indicatie van de reeds aanwezige merkbekendheid. Door deze *click-through rates* te vergelijken met het bezoekersaantal van een bepaalde *site* waarop deze banner is geplaatst, wordt een eerste indruk verkregen omtrent de merkbekendheid én de merkkennis van het betreffende merk.

Bij het analyseren van de resultaten van deze test moet echter rekening worden gehouden met een aantal zaken. Ten eerste zal de relevante omgeving waarin het beeldmerk wordt vertoond, van invloed zijn op de herkenning van het betreffende beeldmerk. Veel logo's op één plaats maken het moeilijker om op te vallen, zeker als er aanmerkelijke verschillen zijn in omvang, kleur en achtergrond. Ten tweede geeft het aanklikken van een *banner* een indicatie van de effectiviteit van de meer traditionele media. Het zien van een televisie-commercial of een gedrukte advertentie kan consumenten ertoe aanzetten meer informatie te zoeken en dus Internet te gebruiken. Bovenstaande meetmethode zorgt ervoor, dat het gedrag van surfers niet-verstorend (*unobtrusive*) wordt geanalyseerd.

Herinneringstests (recall) zijn vergelijkbaar met herkenningstests, met het verschil dat er in tegenstelling tot herkenningstests geen hulp wordt geboden. Er wordt gevraagd zelf bepaalde reclameboodschappen uit het geheugen op te halen (retrieval). Deze methode is tevens geschikt om merkbekendheid te testen (de respondenten die in staat zijn reclame-uitingen te herinneren kennen de reclame en dus het merk.). Daarnaast geeft herinnering een goede indicatie van de kennis van een merk, zeker als er details van een bepaalde reclameboodschap worden gevraagd. Er zijn drie niveaus van effecten te onderscheiden:

1. Bewezen merkregistratie: Het percentage consumenten dat de advertentie kan herinneren en beschrijven.
2. Idee-penetratie: Het percentage consumenten dat de inhoud van de advertentie kan reproduceren.
3. Koopintentie: Het percentage consumenten dat van plan is het product te kopen.

Deze methode kan als volgt worden gebruikt. Na het bezoek aan een *site* kan door middel van een aantal korte vragen worden achterhaald welke uitingen zijn blijven hangen en welke kennis is opgeslagen. De kennis die wordt overgedragen kan bestaan uit ken-

nis die reeds aanwezig was én nieuwe, op de *site* aanwezige en opgedane, kennis. Door nu onderscheid te maken tussen de op de *site* geleverde informatie en elders verkregen informatie, wordt geanalyseerd hoe effectief de boodschap is. Door middel van verschillende antwoordcategorieën kan enerzijds worden gekeken welke informatie van de *site* is overgenomen en welke reeds aanwezig was. Hierdoor ontstaan ten eerste inzicht in de 'beginkennis' van de betreffende consument en ten tweede de verandering in kennis veroorzaakt door de (commerciële) boodschap op de *site*. Het nadeel van deze methode is dat de consument zich bewust wordt van het feit dat zijn gedrag wordt geanalyseerd en daardoor wellicht zijn antwoorden aanpast. Daarnaast zijn consumenten wellicht niet bereid mee te werken en zullen de *site* mijden. Een bepaalde beloning, in de vorm van bijvoorbeeld een *incentive* zoals een gratis proefmonster kan de kans op medewerking van consumenten vergroten. Daarnaast is het mogelijk de vragen op een speelse wijze, bijvoorbeeld in de vorm van een quiz (met eventueel prijzen), te stellen waardoor de consument bereid is mee te werken.

Herkenningstests kunnen worden gebruikt door een specifiek kenmerk van een reclameboodschap als *banner* te gebruiken op een *traffic site*, waardoor inzicht wordt verkregen in de merkkennis van consumenten. Ook geeft dit een indicatie van de interesse in een bepaald product, dienst of merk. Herinneringstests kunnen worden gebruikt door na het bezoek aan een *site* te vragen welke elementen van een boodschap zijn blijven hangen alsmede welke elementen reeds aanwezig waren (door elementen uit andere media op te nemen in de antwoordcategorieën). Incentives en een speelse manier van ondervragen vergroten de kans op medewerking.

5.2. Evaluatieve componenten

Naast kenniscomponenten zijn ook evaluatieve of attitude-componenten belangrijk. Een website beïnvloedt niet alleen de kennis maar ook de houding van consumenten. Met name de aantrekkelijkheid van de website en de amusementswaarde die een site biedt, zullen een effect hebben op de attitude van de consument ten opzichte van het merk en de website. Net als bij reclame-onderzoek is het onderscheid tussen attitude t.o.v. de site en t.o.v. het merk relevant.

De attitude t.o.v. de website is de appreciatie, waardering of likability (Pieters en Van Raaij, 1992) van de site. Een gunstige waardering van de site creëert een positieve bias en heeft in het algemeen een positieve invloed op de informatieverwerking. Net als bij reclame, verblijven consumenten langer in een gunstig gewaardeerde site. Bovendien hebben ze de neiging de gegeven informatie in die site positiever te waarderen en meer op positieve dan op negatieve informatie te letten. "Biased processing" van informatie verhoogt de kans op een positieve uitkomst, d.w.z. een gunstiger attitude, koopintentie of zelfs koopgedrag.

De attitude t.o.v. de website is, net als de kennis-elementen, vast te stellen na het bezoek aan de site door te vragen naar een overall-oordeel over de site en een beoordeling van de aspecten / attributen van de site. Een voormeting is niet nodig, aangezien het niet is te verwachten dat personen al een oordeel hebben voordat ze de site hebben bezocht. Ervaring met het Internet kan hier een modererende factor zijn. Dit wil zeggen dat personen die veel ervaring hebben met websites, in het algemeen kritischer zullen

zijn dan personen met weinig Internet-ervaring. Het is te verwachten dat een website die gunstig scoort op herkenning en herinnering, dus op kenniscomponenten, ook gunstig gewaardeerd wordt. De attitude t.o.v. de website is niet alleen een esthetische waardering, maar ook een waardering van het nut en de informatiewaarde van de website.

De attitude t.o.v. het merk of de organisatie achter het merk moet wel met een voor- en een nameting moeten worden vastgesteld. Uit het verschil kan geconcludeerd worden in hoeverre de website bijdraagt aan een attitudeverbetering t.o.v. het merk of de organisatie.

5.3. Gedragscomponenten

Bij het analyseren van verandering in gedragscomponenten is het belangrijk op te merken dat het meten van daadwerkelijk gedrag relatief gemakkelijk is. Door de mogelijkheid aan te bieden een product te bestellen, een offerte aan te vragen, of een proefaankoop te doen (open en gesloten proposities) kan direct worden gemeten welke consumenten daadwerkelijk gedragsmatig hebben gereageerd op de reclame-boodschap. Er is echter een valkuil. Het zou kunnen zijn dat consumenten reeds een bepaald gedrag wilden vertonen (gedragsintentie), en het Internet gebruiken om dit gedrag te realiseren. In deze situatie hebben andere media (of bijvoorbeeld verkopers) reeds het beoogde effect bereikt en wordt het Internet dus gebruikt als een alternatief aankoopkanaal. Deze valkuil kan worden vermeden door bij het 'betreden' van de *site* meteen de mogelijkheid te bieden om iets te bestellen, een offerte aan te vragen en dergelijke. Consumenten die meteen voor deze optie kiezen zijn dus door andere communicatie-instrumenten overtuigd en gebruiken Internet dus als aankooppunt in plaats van bijvoorbeeld een retailer. Als daarentegen consumenten eerst een reclameboodschap zien, en daarna besluiten iets te bestellen, kan deze actie voor een groot deel worden toegeschreven aan de betreffende boodschap op de *site*.

Het simpelweg analyseren van het feitelijke gedrag kan dus reeds een aanwijzing zijn over de effectiviteit van een commerciële boodschap op het Internet. Omdat de respondenten in het praktijkonderzoek hebben aangegeven, dat met name de gedragsfacilitatie en gedragsintentie belangrijk zijn, moet deze methode enigszins worden uitgebreid. Onder gedragsfacilitatie verstaan we het voor de consument creëren van de gelegenheid om een bepaald gedrag (bv. aankoop) te vertonen. Dit houdt in, dat de Internetsite informatie bevat over bijvoorbeeld het dichtstbijzijnde verkooppunt, telefoonnummers van vertegenwoordigers, etc.. Om er nu achter te komen of deze informatie is overgedragen en begrepen kan een onderneming ervoor zorgen, dat er automatisch een bericht wordt verstuurd naar een verkooppunt als er een consument om deze informatie vraagt. De detaillist of vertegenwoordiger krijgt dan een sein dat er een potentiële klant heeft geïnformeerd naar zijn adres. Door nu deze detaillist te vragen door te geven als er contact met hem wordt gezocht door de betreffende consument, kan worden geanalyseerd welk percentage opvragers daadwerkelijk een winkel bezoekt of telefonische informatie opvraagt. Gedragsintentie is het voornemen een product aan te schaffen of informatie op te vragen. Meestal zal deze fase voorafgaan aan de gedragsfacilitatie. Het feit dus dat een consument informatie opvraagt omtrent verkooppunten of vertegenwoordigers, geeft aan dat er een gedragsintentie bestaat. Door nu te analyseren wanneer een consu-

ment om deze informatie vraagt, direct bij het 'betreden' van de site of pas na het zien van informatie, kan worden geconcludeerd welk medium voor het beoogde effect heeft gezorgd : de meer traditionele media of Internet. Het is tevens niet ondenkbaar, dat een consument de *Internetsite* betreedt via een link met een bepaalde detaillist of ander distributiekanaal. Wellicht wil deze persoon meer informatie, of is zelfs bezig met de aanschaf van een bepaald product / dienst. Het feit dat de betreffende consument zich op de *site* van een verkooppunt bevindt, kan een indicatie zijn dat hij of zij van plan is een product of dienst aan te schaffen.

Gedragsintentie is af te leiden uit enerzijds het aantal consumenten dat om gedragsfaciliterende informatie vraagt. Hierbij kan een onderscheid worden gemaakt tussen de directe en indirecte vraag naar deze informatie. De directe vraag houdt in dat een consument op een andere manier tot de betreffende intentie is gekomen. Dit geeft dus de effectiviteit van de rest van de marketingcommunicatie-mix weer. De indirecte vraag impliceert dat consumenten eerst informatie tot zich nemen (op de *site*) en daarna om gedragsfaciliterende informatie vragen. Zij zullen voor een (groot) deel zijn overtuigd door de boodschap op de *site*. Gedragsfacilitatie kan worden geanalyseerd door te bekijken of de informatie die wordt opgevraagd leidt tot gedrag. Door aan detaillisten door te geven welke consumenten er informatie hebben opgevraagd en deze informatie terug te koppelen in geval van een winkelbezoek, kan worden nagegaan hoeveel consumenten daadwerkelijk handelen naar aanleiding van de aangeboden gedragsfacilitatie. Als dit scenario om privacy-redenen niet mogelijk is, kunnen verkopers aan potentiële klanten vragen of zij de Internetsite hebben bezocht. Zo ja, dan kan er vanuit worden gegaan dat de boodschap op het Internet redelijk effectief is geweest.

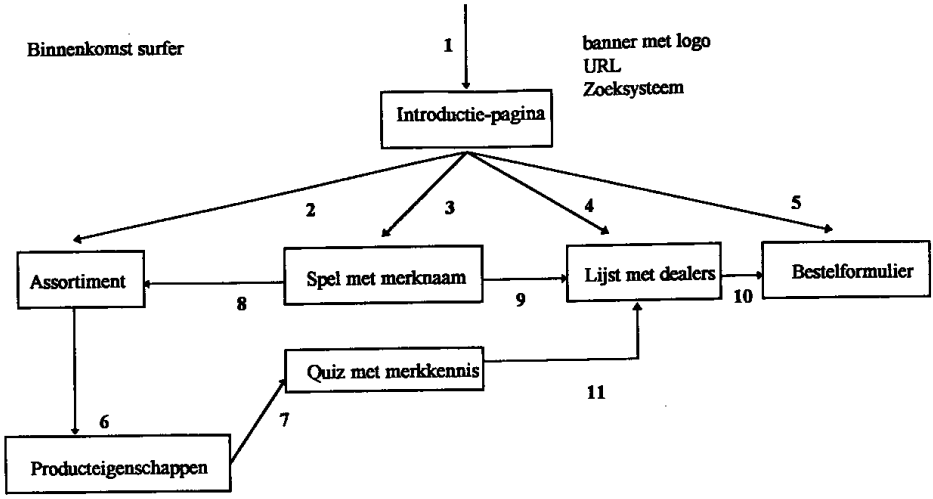
6. OPZET VAN EEN MEETINSTRUMENT

Uit de voorgaande paragrafen kan worden afgeleid dat de website (het informatie-aanbod) zodanig kan worden ingericht dat door het simpelweg volgen van de click-streams veel effecten van een commerciële uiting kunnen worden bepaald. Ter illustratie volgt hier een voorbeeld hoe zo'n website er uit zou kunnen zien.

Meubelen zijn een product waarbij consumenten zich in het algemeen hoog betrokken voelen. Een meubelproducent wil in de Internetsite de volgende mogelijkheden bieden.

- het assortiment te tonen
- de merkbekendheid te vergroten
- de merkkennis te vergroten
- de attitude positiever te maken
- koopfacilitatie te bieden, d.w.z. informatie geven over waar en hoe een product of dienst kan worden aangeschaft.
- de mogelijkheid te bieden om een aankoop te doen.

De website kan dan als volgt worden ingericht :



Figuur 2. Opzet van een website.

Primaire keuze-mogelijkheden :

1 – Binnenkomst van de surfer (consument) en identificatie via *access files*. De plaats van herkomst achterhalen via *referer files*. Dit geeft een indicatie van de effectiviteit van traditionele media (in geval van binnenkomst met URL) of van de verwijsmogelijkheden als banners en zoeksystemen, in geval van binnenkomst via een andere website.

2 – Keuze van een assortiment geeft aan dat de consument het merk kent, verder wil leren, en geïnteresseerd is in de producten onder deze merknaam. Wellicht is dit een eerste stap tot een aankoop. Verdere verdieping in producteigenschappen (**6**) geeft aan dat consument merkkennis tot zich wil nemen. Uitgang via quiz (**7**) om herinnering te testen (vrijblijvend). Ten behoeve van de gedragsfacilitatie volgt een lijst met dealers (**11**).

3 – Spel met merknaam om de merkbekendheid te vergroten en tevens een positieve attitude op te bouwen. Deze keuze geeft aan dat de consument geïnteresseerd is in het merk. Een andere mogelijkheid is dat de consument vermaakt wil worden. Dit is te ondervangen door een link naar het productassortiment te leggen (**8**). Dit duidt op verdere interesse. Deze mogelijkheid niet wederzijds maken, ervan uitgaande dat bij de keuze voor assortiment reeds voldoende merkbekendheid aanwezig is. Wel link naar lijst met dealers (**9**) om de consument in staat te stellen een showroom te bezoeken.

4 – Direct naar lijst met dealers (koopfacilitatie). Geeft aan dat producten bekend zijn en ook het merk bekend is, wellicht door eerdere bezoeken of traditionele media. Gedragsintentie aanwezig (bezoek aan dealer). Alleen een link naar het bestelformulier (**10**), ervan uitgaande dat assortiment bekend is en merk tevens voldoende bekendheid geniet.

5 – Direct naar het bestelformulier. Het gehele beslissingsproces is doorlopen. Consument is gereed om een aankoop te doen.

Door nu het click-gedrag van individuele surfers te analyseren, kan de meubel-onderneming veel te weten komen over de effectiviteit van de website. Ook de manier waarop de consument bij een website terechtkomt, geeft aan hoe effectief de verwijzingen in traditionele media of op andere Internetsites zijn.

De structuur van deze website zal waarschijnlijk niet compleet zijn. Er zijn verscheidene links te leggen en verscheidene componenten toe te voegen. De figuur laat zien dat door de inrichting van de website een onderneming veel te weten kan komen over de effecten van de boodschap, zowel op het Internet als in de traditionele media, ook al is het op basis van bovenstaande analyse moeilijk deze effecten te kwantificeren. Door verder onderzoek zal wellicht in de toekomst deze kwantificering gerealiseerd kunnen worden.

7. CONCLUSIES EN AANBEVELINGEN

Het informatiezoekgedrag op het Internet is met een analyse van click-streams goed te volgen en te analyseren. Het proces en de gevolgde routes bij het surfen op het Internet wordt hiermee vastgelegd. Het bezoek aan een Internet-site en ook de duur ervan kan hiermee worden bepaald. Daarnaast is nuttig en nodig bezoekers te vragen naar hun oordeel over het ontwerp en de inhoud van de website. Met voor- en nametingen kunnen de effecten van de website in termen van kennis, attitude en (koop)gedrag worden vastgesteld.

Internet-adressen in traditionele media zijn vaak de aanleiding om een website te bezoeken. Op het Internet zelf zijn banners een dergelijke aanleiding om door te klikken naar de website.

Het Internet wordt bijna altijd gebruikt in een communicatiemix met andere MC-instrumenten en media. Vooral in de uitwerkings- , transactie- en after-sales fase van een transactiecyclus kan het Internet een nuttige aanvullende rol vervullen in de marketing-communicatiemix. Het Internet is een informatie- en transactiemedium, maar kan ook entertainment verschaffen.

In het artikel is een voorbeeld geschetst van de opzet van een website met informatie, entertainment en de mogelijkheid om ‘online’ de effecten van de site vast te stellen. Marktonderzoekers zullen in de toekomst steeds meer door online registratie gegevens verzamelen over de kennis, houding en het gedrag van consumenten.

LITERATUUR

Allen Interactive, http://www.allen.com, 1998.
Automatiseringsgids, Reclame op het Web is niet succesvol. *Automatiseringsgids*, 17 oktober 1997
R. Briggs, en N. Hollis, Advertising on the Web: Is there really response before click-through? *Journal of Advertising Research*, 37, maart-april 1997, 33-45.
X. Dreze en F. Zufryden, Testing web site design and promotional content. *Journal of Advertising Research*, 37, maart-april, 1997, 77-91.
R.H. Ducoffe, Advertising value and advertising on the web. *Journal of Advertising Research*, 36, september-oktober 1996, 21-35.
J.M.G. Floor en W.F. van Raaij, *Marketing-communicatie Strategie*. Houten: Educatieve Partners Nederland, 1998, derde editie.
L. Hoek, *Het Nut van Internet Gemeten.* Doctoraal-scriptie, Rotterdam: Erasmus Universiteit, 1998.
M. Hoogerbrugge, 'Tracking-onderzoek'. Hoofdstuk 7 in: G. Franzen (red.), *Merken & Reclame*. Deventer: Kluwer BedrijfsInformatie, blz. 255-296.
D.B. Lucas en S.H. Britt, *Het Effect van Reclame*. Utrecht: Het Spectrum, 1966 (Marka 44).
L.M. Maddox en D. Mehta, The role and effect of web addresses in advertising. *Journal of Advertising Research*, 37, maart-april 1997, 47-59.
R.G.M. Pieters en W.F. van Raaij, *Reclamewerking*. Houten: Educatieve Partners Nederland, 1992.
W.F. van Raaij, *Geïntegreerde Communicatie. Een Ontvangersperspectief.* Utrecht: Lemma, 1994
J.R. Rossiter en L. Percy, *Advertising Communications & Promotion Management*. Boston: McGraw-Hill, 1998, tweede editie.
Sterne, J., *World Wide Web Marketing. Integrating the Internet into your Marketing Strategy* New York: John Wiley, 1995.

Medewerkers

H. Anker was verantwoordelijk voor onderzoek en strategie in de PvdA-campagne van 1998. Hij studeerde politicologie aan de Universiteit van Amsterdam en promoveerde in 1992 aan diezelfde universiteit op het proefschrift *Normal Vote Analysis*. Daarna werkte hij als universitair docent empirische politicologie aan de Universiteit van Nijmegen en op part-time basis voor de Partij van de Arbeid. In 1996 werkte hij als opiniepeiler bij Greenberg Research in Washington, D.C., alwaar hij onder meer betrokken was bij de Clinton-campagne en de Britse verkiezingen van 1997.

T. Bijmolt is universitair hoofddocent Marketing aan de Katholieke Universiteit Brabant. Hij promoveerde in 1996 aan de RuG op een proefschrift over Meerdimensionale Schaaltechnieken. Zijn onderzoek richt zich op het ontwikkelen en toepassen van statistische methoden binnen marketing-onderzoek. Tenslotte is het vermeldenswaardig dat hij lid is van de WTG.

M. van Birgelen studeerde bedrijfseconomie aan de Universiteit Maastricht met als specialisatie Marketing en Marktonderzoek. Sinds 1997 is hij als assistent in opleiding verbonden aan de sectie Marketing en Marktonderzoek van de Univerisiteit Maastricht. Zijn promotieonderzoek richt zich met name op mogelijkheden en problemen bij internationaal marktonderzoek en het gebruik van marktonderzoek informatie door managers.

A. van den Broek is verbonden aan het Sociaal en Cultureel Planbureau. Hij houdt zich bezig met vraagstukken op het terrein van (vrije)tijdsbesteding en cultuur.

P. Dekker is politicoloog en werkt bij het Sociaal en Cultureel Planbureau. Hij publiceert over sociale en politieke houdingen en participatie in Nederland, vaak in internationaal perspectief. Onderwerpen van onderzoek zijn thans de publieke opinie over het milieu, vrijwilligerswerk, civil society en non-profit sector.

W. Dijkstra is universitair hoofddocent bij de vakgroep Methoden en Technieken van Sociaal Wetenschappelijk Onderzoek van de Vrije Universiteit. Al ruim 25 jaar houdt hij zich bezig met vragenlijst onderzoek.

H. Gruijters studeerde aan de Katholieke Universiteit Brabant, afstudeerrichting Marktonderzoek. Zijn afstudeerscriptie betrof een meta-analyse over de relatie tussen reclamegeheugen en merkattitude. Momenteel werkt hij als marketing analist bij OTTO B.V. te Tilburg.

L. Halman is als Senior Onderzoeker verbonden aan *WORC*, het Work & Organization Research Centre van de Sociale Faculteit van de Katholieke Universiteit Brabant. Hij is onder meer betrokken bij een grootschalig, internationaal en longitudinaal vergelijkend onderzoek naar waarden in Europa: de European Values Study (*EVS*), en bij een internationaal vergelijkende studie naar religieus en moreel pluralisme (*RAMP*). Hij is Program Director van de 1999 *EVS* waardenstudie.

L.P. Hoek studeert binnenkort af bij de faculteit Bedrijfskunde van de Erasmus Univrsiteit Rotterdam op het onderwerp: 'Het effect van commerciele commmunicatie op Internet'. Interesses liggen voornamelijk op marketing-gebied, met name de kansen en bedreigingen die nieuwe media bieden voor een onderneming. Ook de gevolgen van nieuwe media voor de marketing behoren tot het interesseveld. Op dit moment werkzaam bij de branche-vereniging van de meubelfabrikanten (CBM) met als taken het begeleiden en uitwerken van Electronic Commerce projecten in de meubelindustrie alsmede het analyseren en interpreteren van marktinformatie ten behoeve van de aangesloten leden.

J.C. Hoekstra is als bijzonder hoogleraar direct marketing verbonden aan de economische faculteit van de Erasmus Universiteit Rotterdam en als universitair hoofddocent marktkunde aan de economische faculteit van de Rijksuniversiteit Groningen

K.R.E. Huizingh is universitair hoofddocent marketing en informatie bij de Economische Faculteit van de Rijksuniversiteit Groningen. In 1992 promoveerde hij op het onderzoek 'Marketing informatiesystemen: Perspectief voor een ontwikkelingspad'. Zijn onderzoek richt zich op het raakvlak tussen marketing en informatietechnologie, met onderwerpen als Internet-marketing, database marketing en marketing decision support systemen.

K. Koelemeijer is research fellow bij het Centre for Supply Chain Management van de Universiteit Nijenrode. Haar onderzoek richt zich op het modelleren van het gedrag van consumenten en ondernemingen in distributiekanalen, in het bijzonder met betrekking tot assortimenten en dienstverlening.

F. Langerak studeerde bedrijfseconomie aan de Erasmus Universiteit Rotterdam. Sinds 1993 is hij verbonden aan de Erasmus Universiteit Rotterdam, aanvankelijk als assistent-in-opleiding en universitair docent bij de capaciteitsgroep Commerciële Beleidsvorming van de faculteit der Economische Wetenschappen en op dit moment als universitair docent bij de vakgroep Marketing Management van de faculteit Bedrijfskunde. Daarnaast is hij als freelance wetenschappelijk onderzoeker verbonden aan de Vlerick School voor Management. Verder is hij vanaf 1994 hoofdredacteur van Marketing Wise.

A.J. Olivier is Algemeen Directeur van Research International Nederland. Werkte als Product Manager en Research Manager bij Unilever na zijn studie bedrijfseconomie in Rotterdam. Promoveerde in Groningen bij Peter Leeflang met een proefschrift over Marketing Decision Support Monitoren. Was in de afgelopen zes jaar voorzitter van de VMO en is tevens voorzitter van de Examencommissie van de NVMI en eindverantwoordelijk voor de marktonderzoek-adviseurs-examens van de NVMI.

H. Oppewal is senior lecturer bij de vakgroep Marketing van de Faculty of Economics van de Universiteit van Sydney in Australië en universitair docent bij de vakgroep Stedebouw van de Faculteit Bouwkunde van de Technische Universiteit Eindhoven. Zijn onderzoek is gericht op het modelleren van individueel keuzegedrag in toepassingen op het gebied van marketing en ruimtelijke ordening. Hij publiceerde onder andere in de Journal of Marketing Research, Journal of Economic Psychology, en Journal of Retailing and Consumer Services.

R. Pieters is hoogleraar Marketing aan de Katholieke Universiteit Brabant. Zijn onderzoek richt zich op consumentengedrag, effectiviteit van marketing-communicatie, en kwaliteit van dienstverlening. Hij bereidt een onderzoek voor naar het effect van gedrukte reclame op de waargenomen kwaliteit van de bediening aan de toog van Kempische bierhuizen.

L.M. Pondman is methodoloog en verrichtte haar promotieonderzoek bij de vakgroep Methoden en Technieken van sociaal-wetenschappelijk onderzoek aan de Vrije Universiteit te Amsterdam. Daarnaast was zij als onderzoeker werkzaam bij KPN Research. Momenteel werkt zij bij het Business Intelligence Centre van KPN Telecom.

P.H. Punter is onderzoeksdirecteur van OP&P Product Research. Hij heeft waarnemings-psychologie gestudeerd in Utrecht met als specialisatie de reukzin. Na 10 jaar werkzaam te zijn geweest bij de Universiteit van Utrecht heeft hij in 1987 OP&P opgericht. Is sinds 1985 bestuurslid van de Sensory and Consumer Science Group van de Society of Chemical Industry in London en mede-oprichter van de NVMI vakgroep Sensorisch Onderzoek.

W.F. van Raaij studeerde psychologie en statistiek aan de Rijksuniversiteit Leiden. Daarna was hij verbonden aan de Technische Hogeschool Twente, de Katholieke Hogeschool Tilburg en de University of Illinois in Urbana -Champaign (U.S.A.). Van 1979 tot 1993 was hij hoogleraar Economische Psychologie aa de Erasmus Universiteit Rotterdam. Van 1988 tot 1992 was hij tevens bijzonder hoogleraar Reclamewetenschap aan deze universiteit. Sinds 1993 is hij hoogleraar Marketing aan de Faculteit Bedrijfskunde van de Erasmus Universiteit.
Recente boeken van zijn hand zijn: Consumentengedrag (met G. Antonides, 1997, Utrecht: Lemma) en Consumer Behaviour (met Gerrit Antonides, 1998, Chichester: Wiley).

J. van Rekom is als Universitair Docent verbonden aan de vakgroep Marketing Management aan de Faculteit Bedrijfskunde van de Erasmus Universiteit Rotterdam. Hij studeerde Bedrijfseconomie aan de Vrije Universiteit in Amsterdam. Hij houdt zich bezig met onderzoek op het gebied van organisatie-identiteit en imago, en op het gebied van toerisme. In januari 1998 promoveerde hij aan de Erasmus Universiteit op zijn proefschrift "Corporate Identity, development of the concept and a measurement instrument". Zijn methodologische specialisatie betreft de betekenisstructuuranalyse in brede zin.

C. van Rooy is adjunct directeur van Veldkamp/Marktonderzoek.
Onderzoeker vanaf begin jaren zeventig. Specialisatie: media- en reclame-onderzoek. Publiceert geregeld over nieuwe ontwikkelingen, bijvoorbeeld op het terrein van direct marketing. Stond samen met Fred Bronner aan de basis van MediaMerkbeleving, het onderzoeksmodel dat in 1998 de Hans du Chatinierprijs kreeg toegekend. Vanuit zijn specifieke interesse regelmatig betrokken bij (non-)respons analyses. In 1998 verscheen van zijn hand bij Kluwer in de serie Marketing Wijzer de bundel over marktonderzoek.

K. de Ruyter is als universitair hoofddocent verbonden aan de sectie Marketing en Marktonderzoek van de Universiteit Maastricht. Zijn onderzoeksinteresses liggen voornamelijk op het gebied van de marketing van diensten en daarbinnen de relaties tussen aanbieders en afnemers.

H.R. van der Scheer studeerde econometrie aan de Rijksuniversiteit Groningen. Daarna werkte hij als Assistent in Opleiding bij de vakgroep Econometrie en de sectie Marktkunde en Marktonderzoek aan de R*u*G. In 1997 promoveerde hij op het proefschrift *Quantitative Approaches for Profit Mazimization in Direct Marketing.* Sinds 1 januari 1998 werkt hij als senior marktonderzoeker bij Research International USA.

J.H. Smit is logistiek manager en methodoloog bij de Longitudial Aging Study Amsterdam (LASA) Vrije Universiteit Amsterdam. In die functie is hij verantwoordelijk voor de gegevens verzameling van zeer omvangrijke mondelinge en schriftelijke surveys.

J.C. van Steenis is projectmanager Market Research bij HJ Heinz BV. Beschikt als bedrijfsonderzoeker over een ruim twintigjarige ervaring. De laatste acht jaar hebben haar werkzaamheden een sterk internationaal karakter gekregen. Lid van de Examencommissie van de NVMI. Sinds 1997 bestuurslid van het Voorlichtingscentrum voor Markt- en Opinieonderzoek Uw Mening Telt.

H.C.J. Vrolijk is als universitair docent verbonden aan de vakgroep marktkunde en marktonderzoek aan de economische faculteit (Rijksuniversiteit Groningen). In 1996 promoveerde hij op het onderzoek 'Methoden en regels voor de projectplanning: een toepassing van beslissingsondersteunende systemen'.

M. Wetzels is sinds 1998 als universitair docent verbonden aan de sectie Marketing en Marktonderzoek van de Universiteit Maastricht. Zijn onderzoek richt zich met name op kwaliteit van dienstverlening en kwaliteitsmanagement binnen dienstverlenende organisaties.

Summary

M. van Birgelen, K. de Ruyter and M. Wetzels: What determines market research information effectiveness: a user-based study
It is beyond question that accurate information, obtained by doing market research, is of crucial importance to the success or failure of modern companies. Despite this importance of market research and the growing amount of information that is available nowadays, it has remained unclear to what extent the management of organizations actually use it for decision-making purposes. This article focuses on effective use of market research information by organizations and factors that might influence this use. When doing so, market research done by a research agency can be considered a service that is provided to a principal. An empirical study among NVMI members who have somebody conduct market research on a regular basis was done to test several hypotheses on effective use of market research information.

L.M. Pondman: Reduction of the refusal rate in telephone survey interviews
The levels of nonresponse that are the consequence of subject refusals in telephone survey interviews are still increasing. This jeopardizes the representativeness of survey outcomes and thereby the degree to which surveys provide accurate portraits of the population is reduced.
This study focuses on some aspects of interviewer behavior and on the influence of these aspects on the subject's decision whether or not to participate. The interviewer-subject interactions that take place before the actual interview starts are closely examined. This examination has resulted in the development of a persuasion strategy. An experiment was carried out in order to test whether this strategy did indeed induce lower refusal rates.
It is shown that the likelihood of participation increases if the persuasion strategy is applied. Furthermore, it appears that it is very important to motivate interviewers. Interviewer who are motivated to do the best they can in trying to persuade subjects to participate obtain significant lower refusal rates than interviewers who are not motivated.

H.C.J. Vrolijk, K.R.E. Huizingh and J.C. Hoekstra: Relevant Characteristics of Web users
The goal of the empirical research described in this paper is to explore the user charachteristics that affect the perception of a Web site. For this purpose existing theories with respect to advertising effectiveness, information search behavior and consumer navigation in on-line computer environments are integrated into one research model. The results show that attention is mainly explained by product knowledge, involvement, challenge and flow. Involvement and flow play a major role in the explanation of cognition, affection and conation.

J.H. Smit and W. Dijkstra: Collecting survey data among elderly respondents
In the past decennia the interest for the elderly part of the population (55+ years) has grown. This -among other things- stimulated scientific research among the elderly. It is quite remarkable that methodological research into the conditions under which datacollections among the elderly can be optimised is near to non-existant in the methodological literature. In this study a number of basic principles concerning nonrespons and instrument development are briefly discussed. Also an indication is given about the data quality obtained in a large scale survey among the elderly. The results indicate dat data quality could be a serious problem in surveys among the elderly and indications were found that this quality tends to decrease as respondents grow older.

P.H. Punter: In the end it's the sensory evaluation
Increasing competition, the disappearance of trade barriers and an increasing internationalization force the Food Industry to faster product development and innovation. However, the success rate

of new product formulations and innovations in the Food Industry is very low. This is partly caused by the fact that many new products fail to meet consumer expectations, specifically with respect to taste. Success is basically determined by fundamental insight in consumer expectations with regard to taste and the capability to translate these insights into product formulations. Although traditional market research partly provides this type of information, it is no guarantee for success. Sensory evaluation is originally a scientific discipline within R&D and Quality Control. Using experts or trained assessors, the sensory characteristics of products can be quantified and effects of for instance changes raw material or processing can be measured. In the past 50 years, sensory professionals have developed and tested a large number of methods for the quantification of sensory reactions to stimuli. Because of the association of sensory evaluation with experts or trained assessors, the possible contribution to market research (specifically to product- and concept testing) has never fully materialized. In the past decade, more emphasis has been placed on consumer oriented sensory evaluation and on providing information which would facilitate the communication between marketing and R&D. This article presents an introduction to sensory evaluation, an overview of the methodology and of the applications and an example of 'consumer guided product development'.

K. Koelemeijer and H. Oppewal: Modelling the effects of assortment composition on consumer behavior: An application to florist stores
In this paper we present an approach to measuring the interdependencies within and among product categories or assortments. Using the extended logit model we explicitly model the utility of an item in a hypothetical assortment as a function of the availability and characteristics of the other items in the assortment, thus capturing substitution, complementarity and asymmetric dominance effects. The model is estimated from choice experiments to allow control over the composition of the assortments and the decision context in terms of store ambience and purchase goal. The approach is demonstrated in an application to modelling consumer choices from cut flower assortments in florist stores. For this product category we find that in-store decisions of customers about whether and where to purchase an item are affected by the size and composition of the assortment in the present store, and by the presence of a competing store. Their decisions are not much affected however by the ambience of the present store, nor by the ambience or choice range of a competing store, nor by the travel distance to that store. Substitution occurs between flowers of the same variety, but not to the same extent for all varieties.

R. Pieters, T. Bijmolt and H. Gruijters: Advertising memory and brand attitude: a meta-analysis
Within the field of advertising research, there is much debate whether or not advertising memory is a valid measure of advertising effectiveness. In this article, we discuss the background of this debate. We present a meta-analysis, that is a quantitative analysis of the literature, on the relation between advertising memory and brand attitudes. The results show that there is a small, positive, significant relation between these two concepts. Moreover, the size of the relationship depends on a number of moderator variables. We discuss the implications of our results and directions for future research.

A. van den Broek: The place of shopping on the agenda
The amount of time the Dutch spend shopping in an average week has been remarkably stable over the 1975-1995 period. While much more time was spent on certain (leisure)activities (e.g. sports and watching tv) and while other (leisure)activities were on the decline (e.g. social contacts and reading), the time-consumption of shopping hardly changed.
Beneath the surface, however, things were changing. More people, in particular more men, joined the ranks of the regular shoppers. Shopping follows suit in a general trend in the use of (leisure)time: growing numbers of people get involved in growing numbers of activities. Shopping has become part of the grown repertoire of activities of more people.

C. van Rooy and J.C. van Steenis: "Ring, ring, who's calling?"
More and more people are indicating irritation about the telephone harassment from the telebusiness sector. Not only the telebusiness industry is bothered by the irritation and fatigue of the consumer, but also within the market research world problems are encountered. The difficulties grow further because of the problematics of non-response.
Notably is that most discussions about consumer harassment are primarily held on principles of vague notions and prejudices. The lack of adequate facts and figures has been the main obstacle.
In commission from the Dutch Association of Direct Marketing, Distance Selling and Sales Promotion and the Information Center 'Your Opinion Counts', Veldkamp/Market Research has conducted a study about this telephone pressure.
Each Dutch person of 18 years or older, is being called approximately 2.2 times per year for market research and 3.1 times per year for outbound telebusiness. For inbound telebusiness, the average is 6.7 times per year.

A.J. Olivier: Research methods for Trademark Infringements
This article deals with the ill defined and troublesome relation between market research and trademark legislation disputes.
Lack of knowledge and rules places market research in a false light and most judges are inclined to ignore the outcome of market research.
We distinguish a number of research methodologies and research protocols and for each we define their strength and weaknesses and indicate in what situation they can be used in a court of law.

L. Halman: What is happening with the family? An exploration of European attitudes towards family, marriage, parenthood, parent-child-relations and women's roles
Recently family is put back on the political agenda and has become a popular topic in public discourse. The recent demographic developments seem to indicate that family-life has diappeared as the cornerstone of society. Individualization and traditional family solidarity are considered to be phenomena that can not coexist.
Such ideas are examined in this article and we explore the relevance of family-life in contemporary, modern, secular, individualized and above all emancipated society. Our empirical analyses seem to reveal that family-life has not disappeared and is still highly valued. The idea that traditional family values are on the wane can not be confirmed. Although modernisation, individualization, secularization and emancipation have gradually transformed human society, people's attitudes and values towards family-life, marriage parenting and gender roles have remained rather stable. As such the persistence of traditional attitudes are an important exception to the ideas of modernization and postmodernization.

H. Anker: Strong and social. Message development for Labour in the 1998 election campaign.
This paper is about political message development and political message testing in the Netherlands. This process is illustrated by means of the Dutch Labour Party's 'strong and social' message in the 1998 parliamentary elections. A new type of message test that so far had not been applied within Dutch political research played a pivotal role in developing this message. This test is described in detail, along with the text and substantive backgrounds of four campaign messages that were submitted to this new test. As expected, a message that combined issues that were conceivably 'owned' and 'not owned' by the Labour Party produced the best test results. This message was then further refined in many different ways, culminating in the message that Labour used for its campaign.

P. Dekker: Levels of trust in countries of the European Union
This chapter focuses on national levels of various kinds of trust from a social capital perspective. Eurobarometer data from eleven countries are used to measure trust toward nationalities, credibility of sources of information, and institutional trust. The basic idea is that to get large scale economic and political consequences of basic interpersonal trust this trust should be generalized to

the real actors in public discussions and democratic politics. European countries show similar patterns of trust (e.g. Danes are more trusted than Russians, consumer organisations and media more than unions and government). However, it appears that countries with higher levels of social trust do not necessarily have more confidence in the information public actors give about both the environment and biotechnology. More research is needed to validate differences in social capital between countries and to link social trust to political and economic performance.

F. Langerak: The effect of market orientation on the performance of industrial businesses in the Netherlands

There is equivocality in the results of studies that have addressed industry and market structures and investigated them in relation to the market orientation performance relationship. Against this background this article reports the results from an empirical investigation into the market orientation performance relationship for a sample of 483 businesses competing in the Dutch metal and electromechanical industry. The results indicate that there is a positive relationship between market orientation and the ROI, sales growth, profits as a percentage of sales and new product success. Furthermore the results indicate that differences between and within industry and market structures impact the market orientation performance relationship. This result provides managers with valuable insights with regard to the choice whether or not to invest in building a market orientation.

J. van Rekom: Methods for measuring corporate identity

Literature offers methods to measure Corporate Identity from three basically different perspectives: identity as a set of characteristics to be established by a researcher, identity as set of characteristics perceived by organization members, and the perspective of employee identification with their organization. The measurement methods flowing forth from those perspectives are tested on Albert and Whetten's criteria for definitions of corporate identity: expression of central character, distinctiveness and continuity over time. So far, no method seems to live up to more than one of these criteria simultaneously. Therefor, a new method is proposed to measure corporate identity, based on the laddering technique.

H.R. van der Scheer: Profit maximization in direct marketing

An effective direct marketing campaign aims at selecting those targets, offer and communication elements that maximize the net profits. The list of households to be mailed, i.e. the targets, is considered to be the most important component. Therefore, a large amount of direct marketing research focuses on target selection techniques. In recent literature some papers have been published on improving the communication elements.

This article starts with an overview of the various selection techniques, which have either been proposed in the literature or are commonly employed by direct marketeers. Nearly all of these selection techniques focus on the cases of fixed revenues to a positive reply and hence concentrate on binary choice modeling (response versus nonresponse). However, many direct mailing campaigns do not generate simple binary response, but rather a response where the quantity differs between households. This article specifies a model that incorporates the quantity of response as well as the probability of response. The results of the empirical application suggest that adding quantity modeling to probability modeling can contribute significantly to profitability.

The second part of the article focuses on the communication elements. A traditional approach to evaluate the effect of many direct mailing characteristics is by analyzing them separately, which is obviously inefficient. An approach is presented to determine the optimal mailingdesign in a more efficient manner. An extention of this approach is to combine target selection with design optimalization. A model is proposed which simultaneously optimizes the mailing design and selects the targets. The application shows that this method increases the net profits.

L.P. Hoek and W.F. van Raaij: Assessing the effects of marketing communication on the Internet

This is a review of the factors that are important to assess and measure the effects of marketing communication on the Internet. These effects concern the awareness, knowledge, attitude and behavior of the target group of consumers or business buyers. The existing methods of assessing communication effects have to be adapted to assess the effects of marketing communication on the Internet.

Often, a synergy can be obtained in combining the Internet with traditional communication instruments, such as advertising, marketing PR, sponsorship, sales promotion, direct marketing, in-store communication, personal selling, trade fairs & exhibitions, and media, such as printed media, television, radio, direct media, outdoor media, and interactive media.

The synergy is described in terms of het stages of a transaction: awareness, elaboration, reminder, transaction, and after sales. The present research is on the inclusion of Internet addresses in traditional media, the functions of banner ads, and the website design.

The Internet offers the possibility of unobtrusive and externally valid research on knowledge and attitude change, recognition and recall tests, and information request and buying behavior. An example is included to design a website with a measurement instrument for a furniture producer.