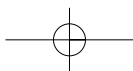
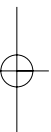


Ontwikkelingen in het marktonderzoek

Jaarboek 2006





ONTWIKKELINGEN IN HET MARKTONDERZOEK

Jaarboek 2006
MarktOnderzoekAssociatie

Redactie

A.E. Bronner

P. Dekker

E. de Leeuw

L.J. Paas

K. de Ruyter

A. Smidts

J.W. Wieringa



SPAAR EN HOUT – HAARLEM

Copyright © 2006 De auteurs
p/a MarktOnderzoekAssociatie
Arlandaweg 92 – 1043 EX Amsterdam
email: info@moaweb.nl
website: www.moaweb.nl

Eindredactie: Prof. dr. A.E. Bronner

Alle rechten voorbehouden. Niets uit deze uitgave mag worden
verveelvoudigd, opgeslagen in een geautomatiseerd gegevensbestand,
of openbaar gemaakt, in enige vorm of op enige wijze, hetzij elektronisch,
mechanisch, door fotokopieën, opnamen of op enig andere manier,
zonder voorafgaande schriftelijke toestemming van de uitgever.

Uitgeverij Spaar^{en}Hout
Postbus 5229 – 2000 CE Haarlem
email: info@vrieseborch.nl
website: www.vrieseborch.nl

*All right reserved. No part of this publication may be reproduced, stored
in a retrieval system, or transmitted, in any form or by any means,
electronic, mechanical, photocopying, recording, or otherwise,
without the prior written permission of the publisher.*

NUR 686
ISBN 90 8683 001 3

Inhoud

Ten geleide	6
I Marketinginstrumenten	8
1. Reacties van Consumenten op Product Placements in Televisieprogramma's J. Bleijerveld en J. Lemmink	9
2. Virtuele communities voorbij de hype: Nieuwe inzichten voor managers, marketeers en marktonderzoekers K. de Valck, G.H. van Bruggen en B. Wierenga	27
3. Loyaliteitsprogramma's in de Nederlandse detailhandel: Adoptie en effectiviteit J. Leenheer	43
II Dataverzameling	62
4. De interviewer, hulp of hindernis? J. van der Zouwen	63
5. De levende of digitale interviewer: Een vergelijking tussen CASI, CAPI en CATI A.E. Bronner en T. Kuijlen	77
6. Landenvergelijkende surveys: Vergelijking tussen landen en tussen surveys I.A.L. Stoop	95
III Marketing en Marktonderzoek	108
7. Meten op markten P.S.H. Leeftang	109
8. Marketing van producten die passen bij de genetische structuur van het individu A. Rontelap, S. Wendel, H.C.M. van Trijp en B.G.C. Dellaert	127
9. Publicatiedrift van Nederlandse Marketingwetenschappers P.C. Verhoef en S. Stremersch	143
IV Analysetechnieken	160
10. Multi-niveau latente klasse analyse: Met een toepassing bij het simultaan clusteren van landen en consumenten J.K. Vermunt, T.H.A. Bijmolt en L.J. Paas	161
V Fusie en imputatie	174
11. Imputeren in plaats van amputeren: De voordelen van meervoudige imputatie J.A. Voerman	175
12. Integrated Campaign Evaluation: Vragen naar mediagebruik op basis van fusie als alternatief voor recall A. van Meurs	191
VI Segmentatie	206
13. Burgersegmenten en vertrouwen in de regering M.C.M. Arnold	207
Medewerkers	221
Summary	227

Ten geleide

Dit jaar 2006 is een bijzonder jaar voor het Jaarboek. Niet omdat het boek uitkomt, want dat bent u van ons gewend bij de start van het nieuwe jaar. Bijzonder, omdat het kenniscentrum van de MOA ten doop wordt gehouden en daarin speelt het Jaarboek een belangrijke rol. Op de website zijn alle eerder in het Jaarboek verschenen artikelen te doorzoeken. Gewoonlijk kijken wij in dit Ten Geleide vooruit naar de artikelen in het nieuwe boek. Maar omdat nu alle artikelen weer toegankelijk zijn eerst even een terugblik. Zeker ook voor de jongeren die niet over de oude boeken beschikken, maar nu opeens wel alle informatie uit voorbije jaren kunnen benutten.

Het eerste Jaarboek verscheen in 1975 en is daarna elk jaar in boekvorm uitgekomen. Het boek was de opvolger van het Marktonderzoek Kwartaalschrift en had tot doel nieuwe ontwikkelingen op het gebied van marktonderzoek en consumentengedrag te signaleren en te analyseren. Het boek had en heeft een brugfunctie tussen de academische en commerciële marktonderzoekwereld. Vooral Arie van der Zwan, Mathieu Meulenberg en Gery van Veldhoven hebben zich ingespannen bij het tot ontwikkeling brengen van de nieuwe formule, waarmee een professioneel Nederlandstalig publicatie-medium tot stand kwam. Nog enkele getallen:

- 31 boeken verschenen
- omvattend 432 artikelen
- in totaal in 30 jaar 25 verschillende redacteurs
- 3 verschillende hoofdredacteurs.

Er is altijd veel aandacht besteed aan (a) methoden en technieken (conjunct meten, segmentatieanalyse, LISREL, HOMALS, etc.), (b) methoden van dataverzameling, (c) marketinginstrumenten en (d) de maatschappelijke omgeving van de consument. Proefschriften die op het terrein van het marktonderzoek verschijnen hebben altijd een belangrijke plaats ingenomen en de promovendus wordt uitgenodigd in een artikel de essentiële punten uit de dissertatie op een toegankelijke wijze aan te geven en vooral de relevantie voor de praktijk te schetsen.

Nu naar het Jaarboek 2006. De artikelen zijn verdeeld over zes rubrieken. Wij openen met drie artikelen over marketingcommunicatie. Steeds meer lopen 'content' en reclame door elkaar. Een voorbeeld is het plaatsen van producten in televisieprogramma's, de zogeheten product placement. Product placements variëren van de plaatsing van het product op een plek in de achtergrond van een scene tot het gebruik van het product door de hoofdpersonen à la de door 007 gereden BMW's. Alom is de verwachting dat dit type reclame in de toekomst alleen maar zal toenemen, zeker als consumenten zich met digitale harddiskrecorders aan commercials kunnen onttrekken. Daarom aandacht voor dit onderwerp in het Jaarboek. Het tweede artikel wil aangeven hoe managers en marktonderzoekers strategisch gebruik kunnen maken van virtuele internetcommunities. Virtuele communities die zijn georganiseerd rondom consumptie-gerelateerde interesses zijn zeer populair. De community invloed blijkt aanzienlijk te zijn, vooral in de fasen van het zoeken naar informatie en het afwegen van alternatieven. Anderzijds kunnen virtuele communities een waardevol marketinginstrument zijn en een locatie voor marktonderzoek. Nokia bijvoorbeeld heeft een speciaal team dat internet afstruint op zoek naar conversaties tussen consumenten over mobiele telefonie in het algemeen en hun producten in het bijzonder. De auteurs geven zelfs het advies aan marktonderzoekers dat zij er goed aan zouden doen om zelf virtuele communities op te starten. Een ander marketinginstrument waarmee relaties met klanten versterkt kunnen worden

is een loyaliteitsprogramma. Leenheer deed hiernaar in het kader van haar proefschrift een zeer uitgebreide studie. Beloningen in een spaarprogramma kunnen stimulerend zijn en haar studie is een pleidooi voor langlopende programma's.

Dan over naar de dataverzameling. De rol van de interviewer, hulp of hindernis en de levende versus de digitale interviewer. Hindernis vooral als de interviewer inadequate antwoorden accepteert of incorrecte herstellpogingen doet. En uit de vergelijking tussen levende en digitale interviewer blijkt dat sociale wenselijkheid een veel grotere rol speelt bij de levende interviewer. In veel onderzoeken blijft tegenwoordig de dataverzameling niet tot één land beperkt. Internationaal vergelijkend onderzoek neemt steeds meer toe. Ineke Stoop gaat in op problemen bij landen-vergelijkend onderzoek. Landen-vergelijkend onderzoek is lastig, maar zij geeft in het artikel een aantal te volgen aanbevelingen.

De derde rubriek betreft de relatie tussen marketing en marktonderzoek. In de bijna 200-jarige geschiedenis van de KNAW is er in 2004 voor het eerst een lezing gegeven op het terrein van marketing en marktonderzoek en wel door onze oud-redacteur Peter Leeflang. In zijn bijdrage aan dit Jaarboek is een verkorte en aangepaste versie van deze lezing gegeven waarin hij op basis van de literatuur tot generalisaties over marktwerking tracht te komen. Een ander actueel onderwerp waarover wij graag een bijdrage wilden is de marketing van producten die passen bij de genetische structuur van het individu. Nutrigenomics is één van de gebieden waar op genetische samenstelling gebaseerde producten nu reeds in de markt beschikbaar beginnen te komen. Een consument stuurt wanglijm en een ingevulde vragenlijst over eetgewoonten en levensstijl naar de fabrikant en ontvangt dan adviezen over voedingssupplementen. Dienstverlening op basis van genetica is echte één op één marketing, waar consumenten individueel worden geïdentificeerd en aangesproken. Nederlandse marketingwetenschappers behoren volgens de analyse van Verhoef en Stremersch tot de wereldtop. Bedrijven kunnen de in dit artikel gepresenteerde informatie gebruiken bij het zoeken naar geschikte universiteiten voor contractonderzoek en de werving van nieuw personeel onder marketingstudenten. Deze bijdrage eindigt met een filosofie over hoe de kloof tussen praktijk en theorie te dichten.

In de methodensectie proberen wij altijd aandacht aan nieuwe bruikbare technieken te geven. Een veel belovende nieuwe segmentatietechniek is latente klasse-analyse. En daarbinnen kan multi-niveau latente klasse-analyse worden benut om op twee niveaus te segmenteren, bijvoorbeeld eerst consumenten en vervolgens landen. Missing data geven de onderzoeker altijd buikpijn, zeker als die op cruciale variabelen voorkomen. Imputatietechnieken vormen een oplossing. Lex van Meurs benadrukt dat wij voor fusie geen fobie moeten hebben en dat het een zeer bruikbaar alternatief is. Beter een fusie voor bepaalde schattingen dan onduidelijke ad hoc vraagstellingen. Het boek eindigt met een artikel over de maatschappelijke context van de consument. Vertrouwen in de Nederlandse economie en regering, politieke betrokkenheid, politiek cynisme en andere variabelen uit de belevingsmonitor van de overheid worden met elkaar in verband gebracht.

Kortom een nieuw Jaarboek en alle verschenen artikelen digitaal toegankelijk. Instrumenten om de kwaliteit van het marktonderzoek te optimaliseren.

Namens de redactie,

A.E. Bronner

1. Reacties van Consumenten op Product Placements in Televisieprogramma's

J. BLEIJERVELD en J. LEMMINK*

SAMENVATTING

Het plaatsen van producten in televisieprogramma's, de zogenaamde product placement, is een relatief nieuw en in toenemende mate gebruikt marketinginstrument. In dit artikel wordt eerst een typologie voor deze product placements gepresenteerd. Daarnaast worden de invloeden van drie types placement in drie soorten televisieprogramma's betreffende drie productklassen op het gedrag van consumenten getoetst met behulp van een steekproef onder consumenten en met gebruikmaking van conjunct meten en beschrijvingen van product placements. Tevens wordt gecontroleerd voor invloeden van sekse, kijkfrequentie en initiële programma- en productvoorkeuren van consumenten.

De resultaten van het onderzoek wijzen uit dat een product placement van het "programma bepalend" type in een soap en betreffende een FMCG het succesvolst is in het genereren van een koopimpuls bij de consument. Deze combinatie van product placement type en programma leidt ook tot de meeste aandacht en de meest positieve attitude, al kan niet uit het onderzoek worden afgeleid welke productklasse hier het best bij is gebaat. De resultaten laten verder zien dat als de consument een groter aantal uren tv kijkt, dat een positief effect heeft op de attitude van de consument ten aanzien van de product placement. Sekse is daarentegen alleen van invloed wanneer het geplaatste product zogenaamde geslachtsindicatoren vertoont van hetzelfde geslacht als de consument. Deze zal dan meer aandacht, een betere attitude en een hogere koopintentie ontwikkelen voor dat product. Tenslotte blijkt dat consumenten met een initiële voorkeur voor het product of het programma waarin het geplaatst is, meer aandacht, een betere attitude en een waarschijnlijker koopintentie voor het geplaatste product te ontwikkelen.

Op basis van deze resultaten kunnen managers nu beter voorspellen welke product placement noodzakelijk is voor welk effect. Toekomstig onderzoek zou deze uitkomsten verder moeten valideren vanwege het experimentele karakter van dit onderzoek.

1. INTRODUCTIE

Sponsoring en reclame bestaan al langere tijd in een grote verscheidenheid aan vormen. Een van de oudste voorbeelden van een vorm van sponsoring is de relatie tussen de

* Jeroen Bleijerveld is promovendus aan de Universiteit Maastricht, Faculteit Economische Wetenschappen en Bedrijfskunde, Dep. Marketing. Het artikel is gebaseerd op zijn afstudeerscriptie: jfj.bleijerveld@mw.unimaas.nl

Prof. dr. Jos Lemmink is hoogleraar marketing aan de Universiteit Maastricht, Faculteit der Economische Wetenschappen en Bedrijfskunde, Departement Marketing: j.lemmink@mw.unimaas.nl

artiest en diens patroon, zoals die tussen Publius Vergilius Maro en Gaius Cilnius Maecenas ten tijde van het Romeinse rijk. Een voorbeeld van reclame uit dezelfde tijd is het organiseren van de fameuze Romeinse spelen, hetgeen naast het “kalm houden” van het volk ook het vergroten van de keizerlijke populariteit tot doel had.

In de huidige maatschappij zijn reclame en sponsoringen niet meer uit het dagelijks leven weg te denken. Van een bezoek aan een willekeurige voetbalwedstrijd tot een bezoek aan de opera en van de ochtendkrant tot het RTL nieuws, overal worden consumenten geconfronteerd met uitingen van reclame of sponsoring. Vaak is er zelfs sprake van een combinatie van reclame en sponsoring, zoals wanneer de reclameborden langs het voetbalveld worden gecombineerd met een reclame op het grote scherm tijdens de rust.

Waar de zojuist genoemde reclame uitingen nog relatief duidelijk als zodanig te onderscheiden zijn, zal deze paper zich concentreren op een vorm van reclame die wat minder bekend en opvallend is, namelijk *product placement*. Product placement is simpelweg gezegd de plaatsing van een product of merklogo in een film, televisie programma of computerspel in ruil voor enige vorm van compensatie met als uiteindelijke doel om de kijker te motiveren het product te kopen. Voorbeelden van deze product placements in films zijn onder andere het drinken van Coca Cola door de ruimtereizigers in de vijftigerjaren film ‘Destination moon’ (Vollmers & Mizerski, 1994) en de door 007 gereden BMWs en Aston Martins in recente James Bond films. Voorbeelden in Nederlandse televisieprogramma’s zijn dagelijks te vinden in bijvoorbeeld ‘Goede Tijden Slechte Tijden’, waarin regelmatig busjes Pringles chips, Philips Senseo koffiezetapparaten en DE koffiekopjes worden gebruikt door de personages.

Waar deze placements vroeger werden gebruikt voor het verhogen van de realiteit van de filmscene, worden ze nu openlijk gebruikt voor commerciële doeleinden. Tegenwoordig worden ze dan ook frequent gecombineerd met reguliere reclame voor een nog groter effect of bijvoorbeeld om een nieuw product te lanceren. Zo combineerde Philips de plaatsing van haar nieuw te lanceren Philishave Sensotec in de Bondfilm ‘Die Another Day’ met reguliere reclames, waarin Bond het apparaat gebruikte om zich met zichtbare tevredenheid van zijn baardharen te ontdoen.

Het succes van deze placements kan worden geïllustreerd door het feit dat de verkopen van het pindaas snoepje ‘Reese’s pieces’ van Hershey’s met 65 % stegen, nadat het snoepje te zien was geweest in de film ‘ET, the Extra Terrestrial’ (Higgins, 1995). Het succes van product placements kan worden verklaard aan de hand van de volgende drie eigenschappen van het instrument. Ten eerste wordt er bij product placements niet weggezappt omdat de reclame deel uitmaakt van het programma en de kijker bij wegzappen dus het risico loopt iets van het programma te missen (Meenaghan, 1991). Ten tweede is het product in een product placement automatisch verbonden aan de lifestyle van het personage dat het gebruikt. Kijkers verzamelen actief deze informatie over de lifestyle van hun favoriete televisiepersonages en gebruiken deze om koopbeslissingen te nemen (DeLorme, Reid & Zimmer, 1994). Tenslotte is aangetoond dat televisiekijkers tijdens een programma *intentioneel leren*, terwijl zij bij de start van het reclameblok overschakelen op een *incidentele leermodus*. Het gevolg hiervan is dat kijkers zich product placements beter kunnen herinneren dan reguliere reclames (Gupta en Lord, 1998).

Het komt dus niet als verassing dat de praktijk van product placement zich in de afgelopen jaren heeft uitgebreid naar de Nederlandse televisieprogramma’s, zoals ‘Goede Tijden Slechte Tijden’ (GTST), ‘Eigen Huis en Tuin’, ‘Baantjer’, ‘Born2cook’, ‘De Grote

Beurt', etc. De mate waarin product placements worden gebruikt in Nederlandse televisieprogramma's is voor velen misschien wel een verrassing, maar voor IP bedroegen de inkomsten uit product placements in 1997 al honderd miljoen gulden (€ 45,5 miljoen) ofwel 15 % van hun totale reclame omzet (van Gelder, 1997).

Hoewel product placement in televisieprogramma's in de praktijk dus reeds een substantiële populariteit geniet, heeft de meerderheid van de academische literatuur zich vooralsnog geconcentreerd op de plaatsing van producten in (bioscoop)films. Wij besteden daarom juist aandacht aan product placements in televisieprogramma's. Het onderzoek richt zich op de invloed van het type placement, het soort televisieprogramma en de productklasse op het gedrag van consumenten.

2. TYPOLOGIE VAN PRODUCT PLACEMENTS

Product placements variëren van een simpele plaatsing van het product op een plank in de achtergrond van een scène, zoals de 'DE' koffiekopjes op de plank achter de counter in café 1753 in 'GTST', tot het gebruik van het product door de hoofdpersonages, zoals het eten van chips uit het 'Pringles' busje door vrijwel alle hoofdpersonen in hetzelfde 'GTST'. Er zijn zelfs voorbeelden van televisieprogramma's die compleet lijken te bestaan uit een product placement, zoals de serie 'Airline' waarin werknemers van 'EasyJet' gevolgd worden tijdens hun dagelijkse werkzaamheden en waarbij alle facetten van het product 'EasyJet' duidelijk in beeld komen.

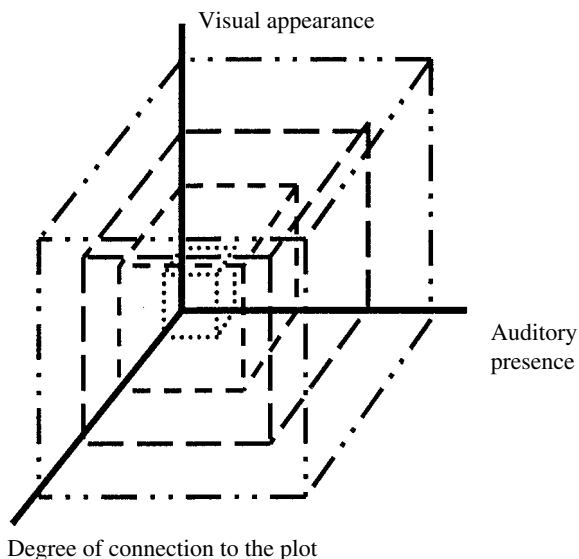
Gezien deze grote verscheidenheid aan vormen is het nodig om een typologie voor product placements te ontwikkelen. In de bestaande literatuur zijn reeds een aantal verschillende typologieën te vinden. Deze kunnen, hoewel er ook andere vormen bestaan (Shapiro, 1993), grofweg worden ingedeeld in twee groepen.

De eerste groep typologieën maakt onderscheid tussen product placements op basis van de zintuigen die worden geactiveerd door de product placement, het zicht en het gehoor. Auteurs die gebruik maken van deze typologie onderscheiden meestal drie vormen van product placement, namelijk 'zichtbare', 'hoorbare' en 'zicht- en hoorbare' product placements (Gupta en Lord, 1998). De eerste categorie bestaat uit producten die zichtbaar zijn in het programma, maar die niet worden benoemd of beschreven, zoals wanneer een hoofdpersoon een blikje 'Coca Cola' zou drinken met het logo duidelijk zichtbaar. Een ander voorbeeld is een pak 'Wasa knäckebröd' dat op de achtergrond op een plank staat in de keuken in 'GTST'. In de tweede categorie wordt het geplaatste product alleen genoemd zonder dat het in beeld te zien is. Een voorbeeld hiervan is te vinden in de film 'Wallstreet', waarin Martin Sheen's personage "Get this kid a Molson Light" schreeuwt naar een serveerster (Gupta en Lord, 1998). In de derde categorie tenslotte worden de geplaatste producten zowel genoemd of beschreven als zichtbaar in beeld gebracht. Een voorbeeld hiervan komt uit de film 'Wayne's World', waarin de hoofdpersoon de slogan "The choice of a new generation" roept alvorens een slok te nemen van zijn blikje 'Pepsi Cola' (Gupta en Lord, 1998).

De tweede groep typologieën onderscheidt product placements op basis van de mate van integratie in het programma. Auteurs die gebruik maken van deze typologie maken veelal onderscheid tussen 'impliciete', 'geïntegreerde expliciete' en 'niet-geïntegreerde expliciete' product placements (d'Astous en Séguin, 1999). Een product placement is 'impliciet' wanneer het product onnadrukkelijk aanwezig is in het programma en het min of meer contextueel wordt gebruikt (d'Astous en Séguin, 1999). Voorbeelden hier-

van zijn vrachtwagens met merklogo's die voorbijrijden in de achtergrond of reclameposters die in een bushokje op de achtergrond te zien zijn. De tweede categorie bestaat uit producten die nadrukkelijk aanwezig zijn en gebruikt worden door een of meerdere hoofdpersonen (d'Astous en Séguin, 1999). Een voorbeeld hiervan is het eten van Pringles chips tijdens een dialoog door verschillende hoofdpersonen uit 'GTST'. Tenslotte, een product placement is 'niet-geïntegreerd expliciet' wanneer het product of het merklogo nadrukkelijk wordt gepresenteerd in het programma terwijl het niet is geïntegreerd in het script (d'Astous en Séguin, 1999). Voorbeelden hiervan zijn de zogenaamde billboards en reclamebumpers, die verschijnen aan het begin van het programma of tussen de reclameblokken en het programma, zoals wanneer de voeten van vier vrouwen voor een televisietoestel met daarop het logo van 'Crystal Clear' in beeld komen voordat de serie 'Desperate Housewives' begint.

Hoewel het mogelijk is om de bestaande typologieën in deze twee groepen onder te brengen, betekent dit niet dat deze groepen *mutually exclusive* zijn. Een 'impliciete' product placement kan bijvoorbeeld met hetzelfde gemak een 'zichtbare' placement genoemd worden wanneer het gaat om het pak 'Wasa knäckebröd' op de plank in de achtergrond in 'GTST'. Dit betekent echter niet dat een 'zichtbare' product placement ook altijd 'impliciet' genoemd kan worden. De billboards bijvoorbeeld kunnen alleen van het 'zichtbare' type placement zijn, maar in de typologie van d'Astous en Seguin (1999) behoren ze tot de 'niet-geïntegreerd expliciete' en niet de 'impliciete' categorie. Gebaseerd op deze inzichten ontwikkelde Russell (1998) een bruikbaar raamwerk waarin zij de mate van integratie van de product placement in het script combineert met de visuele en auditieve aanwezigheid ervan. Op basis van haar raamwerk hebben wij onderscheid gemaakt tussen vier types product placement (fig. 1), die, wanneer men van de kleinste kubus naar de grootste gaat, telkens hoger scoren op de mate van belangrijkheid *en integratie in het script* en de *visuele en auditieve aanwezigheid*.

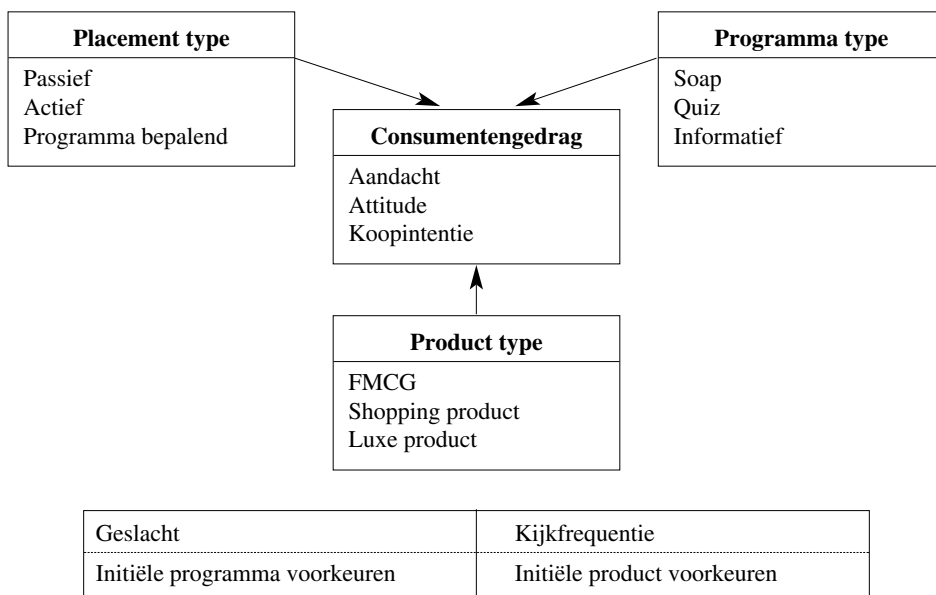


Figuur 1. Russell's (1998) raamwerk met de placement types.

De kleinste kubus vertegenwoordigt het eerste placement type genaamd 'Billboards' en behelst de billboards en reclamebumpers, die voor het programma en als overgang naar de reclameblokken worden getoond. De op één na kleinste kubus behelst product placements die slechts passief in het programma voorkomen. Producten in dit 'Passief' type placement zijn zichtbaar op voor- of achtergrond en kunnen worden besproken, maar ze worden niet actief gebruikt door de personages in het programma. Wanneer dat laatste wel het geval is, kan de placement worden geplaatst in de volgende kubus of categorie, die "Actief" wordt genoemd. Placements van dit 'Actief' type zijn dan, vanwege het feit dat ze gebruikt worden door de hoofdpersonages, ook belangrijker voor het script. De laatste categorie is het 'Programma bepalend' type product placement. In dit type zijn de geplaatste producten eigenlijk meer de kern van het televisieprogramma. Het product wordt met al zijn kenmerken getoond en besproken in een normale consumptie situatie. Het voorbeeld van 'EasyJet' in het programma 'Airline' uit het begin van dit hoofdstuk is een goed voorbeeld van dit type placement.

3. HET ONDERZOEKSMODEL: PRODUCT PLACEMENT EN HET GEDRAG VAN CONSUMENTEN

In dit onderzoek zullen, zoals eerder gemeld, de invloeden van drie product placement variabelen op het effect van een product placement op het consumentengedrag worden getest. Daarnaast wordt getest op vier demografische factoren die ook van invloed kunnen zijn op dit effect. Al deze variabelen en factoren zijn samengevat in het onderzoeksmodel in figuur 2.



Figuur 2. Het onderzoeksmodel.

3.1. Gedragsmodel

Als beschrijving van het consumentengedrag maken wij gebruik van de variabelen 'Aandacht' (Attention), 'Attitude' en 'Koopintentie' (Buying Intention) (McGuire, 1973). In de eerste stap wordt de kijker zich bewust van de product placement. In de tweede stap heeft de kijker zich een mening gevormd over de product placement en het geplaatste product. In de derde stap tenslotte heeft de kijker een zekere intentie het geplaatste product te kopen. Hoewel er veel kritiek is geweest op de volgorde van stappen in deze 'hierarchy of effects' modellen (bijv. Politz, 1975 en Krugman, 1977 en 2000), zijn de stappen op zich desalniettemin valide. De validiteit van deze stappen voor onderzoek naar product placements is aangetoond in eerder onderzoek door Babin en Carder (1996), Gould, Gupta en Grabner-Kräuter (2000), Morton en Friedman (2002) en Russell (2002) en wordt tevens bevestigd door de 'IP Sponsoring Effect Meter' (Akkerman, 2003). Dit laatste onderzoek toont aan dat negentig procent van de ondervraagden zich bewust is van de plaatsing van producten in televisieseries, terwijl de meerderheid een positieve mening vormt ten opzichte van het geplaatste product en zelfs zevenendertig procent aangeeft een aankoop te overwegen.

3.2. Type Product Placement

Eerder zijn vier types placement geïdentificeerd, maar het 'billboard' type zal niet worden getest in het onderzoek vanwege het feit dat in tegenstelling tot de andere vormen de producten in deze vorm niet echt in de serie worden geplaatst. Getest worden dus de 'passief', 'actief' en 'programma bepalend' placement types.

Het is waarschijnlijk dat de drie placement types verschillen in hun effectiviteit om de aandacht van de consument te trekken. In Steertz's (1987) onderzoek bijvoorbeeld herkende (recall) 57% van de respondenten de product placements van het 'zicht- en hoorbare' type, terwijl 51% zich die van het 'hoorbare' type herinnerden en slechts 33% die van het 'zichtbare' type. Wanneer dit laatste type in de achtergrond werd gebruikt, herkende zelfs slechts 8% de product placement. Deze bevindingen worden ondersteund door Gupta en Lord (1998), die vonden dat prominente placements leiden tot betere herkenning bij de kijker dan meer subtiele placements. Zij gaven verder aan dat de prominentie van placements groeit naarmate het product groter wordt, het midden in het beeld geplaatst wordt, het zowel visueel als auditief gepresenteerd wordt en het tenslotte beter in de actie van de scène geïntegreerd wordt. Vertaald naar de typologie, die is ontwikkeld in deze paper, duiden deze resultaten erop dat de aandacht van kijkers voor een product placement toeneemt van de het 'passief' type placement, waar het product niet zo goed is geïntegreerd in de actie van de scène, tot het 'programma bepalend' type, waar dit veel meer het geval is. Dit leidt tot de eerste hypothese:

H₁ De aandacht van kijkers voor een product placement neemt toe over het continuüm van het 'passief' type placement tot het 'programma bepalend' type.

Het tweede doel van een product placement is dat de kijker zich een positieve attitude vormt ten aanzien van het geplaatste product. Terwijl kijkers een positieve attitude blijken te hebben ten aanzien van het plaatsen van producten in films in het algemeen (Zimmer & DeLorme, 1997), is het waarschijnlijk dat de verschillende types placement leiden tot verschillende reacties van de kijker. Russell (2002) suggereert dat kijkers

zich alleen een positieve attitude vormen over het geplaatste product wanneer het product op een voor het verhaal zinvolle manier is geplaatst. Wanneer dit niet zo is, zien zij de placement als een hinderlijke en verborgen reguliere reclame. Zij suggereert verder dat een 'zichtbare' type placement dus niet te duidelijk in beeld moet worden gebracht omdat dit niet veel kan toevoegen aan het verhaal en het dus het risico loopt aangezien te worden voor hinderlijke reclame. Deze suggestie is echter niet noodzakelijk correct. DeLorme, Reid and Zimmer (1994) rapporteren bijvoorbeeld dat kijkers juist actief op zoek zijn naar elke vorm van informatie over de levensstijl van de personages uit films om die dan te kunnen vergelijken met die van zichzelf. Deze zoektocht wordt gefaciliteerd door product placements van elk type. Een placement van het 'zichtbare' type kan dus belangrijker zijn voor de kijker dan Russell (2002) suggereert, omdat deze de kijker kan helpen bij het vormen van een beeld over de personages. Dit zou betekenen dat de kijker het type placement prefereert dat het meeste informatie biedt, maar dat kent ook zijn grenzen. Zoals Boulet en Piquet (d'Astous and Bitz, 1995) vonden voor sponsoring is het waarschijnlijk dat een product placement zijn geloofwaardigheid verliest op het moment dat het lijkt dat de product placement de "raison d'être" is van het televisieprogramma. Concluderend is het dus waarschijnlijk dat de relatie tussen het type placement en de attitude van de kijker niet lineair verloopt. De attitude zal positiever worden wanneer men van een 'passief' naar een 'actief' type gaat, omdat de hoeveelheid informatie over de levensstijl van de personages toeneemt. De attitude wordt echter minder positief wanneer het gaat om een 'programma bepalend' type placement, omdat deze lijdt aan een gebrek aan geloofwaardigheid en dus wordt aangezien voor reguliere commercial.

H₂ De attitude van kijkers ten aanzien van een product placement is het meest positief wanneer het een placement van het 'actief' type betreft.

Het uiteindelijke doel van een product placement is om de kijker te verleiden het geplaatste product ook daadwerkelijk te gaan kopen. Het feit dat het inderdaad mogelijk is om dit te bereiken, blijkt wel uit het onderzoek van DeLorme en Reid (1999), waarin verschillende bezoekers van een film aangeven door de plaatsing in de film geïnspireerd te zijn geraakt tot het kopen van het product. Uit de commentaren van de film bezoekers, die zij verzamelden, blijkt verder dat de kijkers de producten identificeren met het personage uit de film en diens levensstijl in de film. Hoe meer informatie de product placement dus verschaft over de levensstijl van het personage, hoe groter de kans dat de kijker wordt verleid tot het kopen van het product. Maar ook hier geldt dezelfde grens als bij de attitude stap. Wanneer de placement de "raison d'être" lijkt voor het programma, verliest de product placement zijn geloofwaardigheid in de ogen van de kijker en wordt deze gezien als hinderlijke reclame. Dit leidt tot de derde placement type hypothese:

H₃ Kijkers zijn het meest geneigd een koopintentie te ontwikkelen voor een geplaatst product wanneer de placement van het 'actief' type is.

3.3. Type televisieprogramma

Uit de literatuur blijkt dat de geloofwaardigheid van de bron een belangrijke rol speelt in de attitude stap van het consumenten gedragsmodel (McGuire, 1973; d'Astous &

Bitz, 1995). Het doel van een product placement is om een positieve attitude te genereren en daarom is het belangrijk dat het programma een geloofwaardig medium is voor de placement. d'Astous en Séguin (1999) hebben dezelfde drie types programma getest, die ook in dit onderzoek zijn getest, namelijk quiz/variëteiten programma's, zoals 'Cash en Carlo', miniseries/drama's, zoals de gemiddelde soap, en informatieve programma's, zoals 'Barend en van Dorp'. Zij wijzen erop dat de kijker verschillende motieven heeft voor het kijken van deze programma's. Naar het eerste type kijkt men over het algemeen voor vermaak, terwijl kijkers van het tweede type ook kijken vanuit de wens zich te identificeren met de personages uit de serie. Kijkers van het derde type zijn vooral uit op informatie en opinies. Uit het onderzoek (d'Astous & Séguin, 1999) bleek verassend genoeg dat de respondenten het meest negatief waren over product placement in miniseries. De verklaring van de auteurs hiervoor was dat de kijkers van miniseries product placements in hun serie waarschijnlijk alleen accepteren als deze filantropisch en niet commercieel georiënteerd zijn. Het lijkt echter waarschijnlijker dat kijkers zich juist de meest positieve attitude vormen jegens de placement wanneer deze plaatsvindt in een soap, omdat de placement de kijker dan waardevolle informatie kan verschaffen over de levensstijlen van de personages met wie hij of zij zich wil identificeren. Kijkers zullen zich dan al minder positieve attitudes vormen wanneer het een placement in een quiz betreft. De kijker zal de minst positieve attitude vormen voor placements in informatieve programma's, omdat hij van deze programma's objectieve informatie verwacht. Hoewel er tot op heden weinig onderzoek is gedaan naar de invloed van het soort programma op de andere twee stappen van het gedragsmodel is het desalniettemin interessant ook hiernaar te kijken. Dit leidt tot de volgende hypothesen:

H_{4A} Het gebruik van verschillende televisieprogramma's voor een product placement zal leiden tot verschillende aandachtsniveaus van de kijker.

H_{4B} De attitude van kijkers ten aanzien van de product placement is het meest positief wanneer deze plaatsvindt in een soap, gevolgd door die in een quiz en tenslotte die in een informatief televisieprogramma.

H_{4C} Het gebruik van verschillende televisieprogramma's voor een product placement zal leiden tot verschillende niveaus voor de koopintenties van de kijker.

3.4. Type product

Hoewel er meerdere product karakteristieken, zoals ethische lading (Gupta & Gould, 1997) en het imago van het bedrijf achter het product (Mackenzie and Lutz, 1989; Mitchell and Olson, 1981), van invloed kunnen zijn op het effect van een placement op het consumentengedrag, zal de focus in dit onderzoek liggen op de volgende drie productklassen (Dibb, Simkin, Pride & Ferrell, 1997). De eerste klasse bestaat uit de zogenoemde Fast Moving Consumer Goods (FMCG). Dit zijn relatief goedkope producten, die regelmatig worden gekocht en geconsumeerd en die niet veel inspanning vereisen van de consument om ze te kopen. De tweede categorie behelst de 'shopping' goederen. Dit zijn de relatief al wat duurdere producten, die nog wel redelijk regelmatig worden gekocht, maar waarvan de aankoop wel wat meer wordt gepland en overwogen door de consument. De derde klasse behelst de 'luxe' producten. Dit zijn de relatief dure en slechts af en toe aangeschafte producten, die consumenten pas kopen na veel

vergelijken, nadenken en planning. Het lijkt waarschijnlijk dat een product placement het meest succesvol is, wanneer het een product betreft waarvoor de kijker sowieso al niet veel inspanning zou verrichten om het te kopen. De kijker heeft dan genoeg aan enkele relatief simpele product karakteristieken om zijn of haar keuze te maken. Wanneer het een product betreft waarvan de kijker meer informatie vereist om zijn of haar keuze te kunnen maken, lijkt het waarschijnlijk dat een product placement minder effectief verleidt. Dit leidt tot de volgende hypothese:

H₅ Verschillende types producten zullen, wanneer geplaatst in een televisieprogramma, leiden tot verschillende niveaus van aandacht, attitude en koopintentie bij de kijker.

3.5. Demografische factoren

In dit onderzoek zal tevens worden gecontroleerd voor een aantal demografische factoren, namelijk sekse, de frequentie van televisie kijken en eventuele initiële programma en productvoorkeuren.

Het is vrij waarschijnlijk dat sekse van invloed is op de perceptie van een product placement, vooral wanneer het geplaatste product geslachtsindicatoren heeft. Hiermee wordt bedoeld dat het product bepaalde attributen heeft, die het op consumenten doet overkomen als typisch iets voor of mannen of vrouwen. Een voorbeeld van zo'n product is de auto, die over het algemeen meer wordt gezien als iets voor mannen dan voor vrouwen (Kanungo & Pang, 1973). Het blijkt verder uit onderzoek dat mensen relatief positiever reageren op producten met indicatoren van het eigen geslacht (Iyer & Debevec, 1986). Gould, Gupta & Grabner-Kräuter (2000) vonden vervolgens dat mannen over het algemeen vaker geneigd zijn koopintenties te ontwikkelen voor producten, die zij eerder in de film hebben gezien. Dit alles leidt tot hypothese zes:

H_{6A} Mannen en vrouwen verschillen in hun reacties op een product placement, wanneer die een product met geslachtsindicatoren behelst.

H_{6B} Mannen zullen vaker aangeven een koopintentie te hebben voor een product dat zij eerder hebben gezien in een televisieprogramma.

Hoe meer een persoon televisie kijkt, hoe bekender deze is met product placements en dus hoe sterker diens attitude ten opzichte van het concept. Eagly & Chaiken (1995) concludeerden reeds dat een attitude sterker wordt naarmate deze gebaseerd is op meer informatie en reflectie daarop. Gupta & Gould (1997) toonden vervolgens aan dat er een positieve relatie bestaat tussen de attitude jegens het concept product placement en de attitude ten opzichte van een specifieke placement. Er is echter nog onduidelijkheid over of meer ervaring met product placements leidt tot een positievere of negatievere attitude jegens het concept van product placement. Gupta en Gould (1997) stellen dat een verhoogde frequentie van televisiekijken inderdaad leidt tot een positievere attitude ten opzichte van product placements in het algemeen en dus ook jegens specifieke product placements en de geplaatste producten zelf. Het is echter ook voor te stellen dat een hogere kijkfrequentie kan leiden tot een negatievere houding ten opzichte van het plaatsen van producten in televisieprogramma's. Dit idee is gebaseerd op het 'Persuasion Knowledge Model' (PKM) (Friestad & Wright, 1994), hetgeen inhoudt dat de

kijker zich meer en meer bewust wordt van het achterliggende commerciële doel van product placements, naarmate hij of zij ze meer tegenkomt door bijvoorbeeld vaker televisie te kijken. Hoe vaker een persoon dus televisie kijkt, hoe vaker hij of zij product placements zal zien en hoe beter hij of zij op de hoogte zal zijn dat deze zijn bedoeld om hem of haar te verleiden tot aankoop van het geplaatste product. Een persoon die vaak televisie kijkt zal daarom kritischer of negatiever staan tegenover product placement dan iemand die minder vaak kijkt. Samenvattend is er dus nog onduidelijkheid over of een hogere kijkfrequentie leidt tot positievere of negatievere attitudes jegens product placement. Dit leidt tot hypothese zeven:

H₇ De attitude van de kijker jegens product placement wordt beïnvloed door de frequentie van televisiekijken.

4. METHODE VAN ONDERZOEK

Er worden in dit onderzoek dus drie variabelen getest op elk drie niveaus en daarnaast wordt er gekeken naar vier demografische factoren (fig. 2). Om het aantal van 27 mogelijke combinaties van deze drie variabelen terug te brengen tot een beter testbaar aantal is gebruik gemaakt van conjunct meten (Luce & Tuckey, 1964). Deze multivariate techniek is ontwikkeld voor marketing onderzoek vanuit het idee dat elk product eigenlijk een bundel kenmerken is (Hair, Anderson, Tatham & Black, 1995; Lancaster, 1966) en maakt het in de 'full-profile' 'fractional factorial design' vorm mogelijk alle variabelen op alle niveaus te testen met behulp van slechts 9 combinaties (Malhotra, 1999). In dit onderzoek wordt verder niet getest op mogelijke interacties tussen de drie variabelen, hetgeen betekent dat de basale compositie regel kan worden toegepast. Dit betekent dat het effect van een product placement op het consumentengedrag gelijk is aan de som van de effecten op het consumentengedrag van het placement type, het programma type en het product type. Dit alles houdt in dat de respondenten 9 verschillende combinaties van alle drie variabelen moeten evalueren om een goed beeld te krijgen van de invloeden van alle drie de variabelen op al hun niveaus (Malhotra, 1999). Op basis van Addelman's (1962) tweede basisplan is bepaald welke negen combinaties nodig en voldoende zijn om de effecten te schatten. Een voorbeeld van zo'n combinatie is het 'passief' plaatsen van een FMCG in een soap. Deze negen combinaties zijn vervolgens verwerkt in negen korte en realistische beschrijvingen van scènes, zoals je ze op televisie zou kunnen zien. Als programma's voor deze beschrijvingen zijn de goed bekende programma's 'Goede Tijden Slechte Tijden' (soap), 'Airline' (soap), 'Cash en Carlo' (quiz), 'Barend en van Dorp' (informatief) en 'Wannahaves' (informatief) gebruikt. Ook is er een fictieve quiz bedacht, omdat de beschrijving van één zo'n combinatie onnatuurlijk zou overkomen wanneer gebruik was gemaakt van een bestaande quiz. Deze programma's zijn gekozen, omdat zij naast bekend ook realistische programma's zijn om producten in te plaatsen. Als herkenbare producten zijn gebruikt een bus 'Pringles' chips, een blikje 'Coca Cola', een 'Iglo' diepvries maaltijd (alle drie FMCG), een 'Helly Hansen' jas, 'Hilfiger' kleding, 'Benetton' kleding (alle drie 'shopping' producten), een 'Fiat' auto, een 'Rolex' horloge en een reis naar de Caraïben met 'EasyJet' (alle drie luxe producten). De 'Fiat' auto en 'Benetton' kleding zijn specifiek gekozen vanwege hun geslachtsindicatoren en zullen dus gebruikt worden voor het testen van hypothese 6A.

Respondenten werden benaderd met een vragenlijst waarin deze beschrijvingen waren opgenomen. Deze vragenlijst bestond verder per beschrijving uit twee vragen gericht op aandacht (Bucholz & Smith, 1991; Duncan & Nelson, 1985), twee vragen gericht op de attitude (Bagozzi, Baumgartner & Yi, 1992; Henthorne, LaTour and Nataraajan, 1993) en één gericht op de koopintentie (Petroshius and Monroe, 1987). Tevens zijn vragen toegevoegd om te peilen hoe de attitude van de respondenten van tevoren was ten opzichte van de programma's en producten die gebruikt zijn. De vragenlijst werd tenslotte afgesloten met vragen gericht op de demografische factoren van de respondent. Als antwoordschaal is de zeven-punt Likert (1932) schaal gebruikt. De vragenlijst is vervolgens op een website gezet en een link hiernaar is via e-mail gestuurd aan een poule van Nederlandse studenten, die zich vrijwillig voor deze poule hebben opgegeven.

5. ANALYSE EN CONCLUSIE

In totaal hebben 121 respondenten de vragenlijst ingevuld. Deze gegevens zijn vervolgens geanalyseerd met behulp van een (dummy) regressie analyse, waarvan de resultaten zijn gepresenteerd in tabel 1. Als basis categorie voor deze analyse is de combinatie van het 'passief' type placement, het soap type programma en het FMCG type product genomen. Verder is het belangrijk om te vermelden dat de regressieanalyses zijn gedaan voor elke vraag afzonderlijk. Verder is voor de test van hypothese 6A gekozen voor het berekenen van de correlaties tussen het gegeven antwoord op de vraag en het geslacht van de respondent.

De resultaten van de regressieanalyse ondersteunen hypothese 1, hetgeen betekent dat de kijker de meeste aandacht heeft voor placements van het 'programma bepalend' type (β 's van 0.147 en 0.221), gevolgd door respectievelijk die van het 'actief' (β 's van 0.080 en 0.137) en 'passief' type (tabel 1). Hypothese 2 wordt daarentegen verworpen, maar alleen op basis van de resultaten voor de tweede attitudevraag. Hoewel de bij $p < 0.05$ niet-significante resultaten van de eerste attitudevraag er al op wijzen, laten alleen de resultaten van de tweede attitudevraag significant ($p < 0.05$) zien dat de kijker de meest positieve attitude vormt voor een product placement van het 'programma bepalend' type (β van 0.495), gevolgd door die van het 'actief' (β van 0.057) en 'passief' type (tabel 1). Uit deze analyse is dus gebleken dat het 'programma bepalend' type product placement niet, zoals was verwacht, lijdt aan een gebrek aan geloofwaardigheid (d'Astous & Bitz, 1995). Integendeel, de analyse ondersteunt de gedachte dat juist het type dat de meeste informatie biedt (DeLorme, Reid & Zimmer, 1994) en het best is geïntegreerd in het programma (Russell, 2002), leidt tot de meest positieve attitude bij de kijker. Ook hypothese drie wordt verworpen, want ook voor koopintentie is geen bewijs gevonden voor het feit dat een 'programma bepalend' type product placement zou leiden aan een gebrek aan geloofwaardigheid (d'Astous & Bitz, 1995). De resultaten van de regressieanalyse tonen aan dat de kans het grootst is dat een kijker een koopintentie ontwikkelt wanneer het een product placement van het 'programma bepalend' type (β van 0.310) betreft, gevolgd door die van het 'actief' (β van 0.142) en tenslotte 'passief' type (tabel 1).

Voor hypothese 4 laten de resultaten een wat meer gemengd beeld zien. Hypothese 4A wordt ondersteund, maar alleen door de tweede aandachtsvraag. Waar bij de eerste aandachtsvraag alleen de β (- 0.080) voor informatieve programma's significant is

Tabel 1. Resultaten regressieanalyse (alle dikgedrukte β s met * zijn significant bij een betrouwbaarheidsniveau van 95%).

		Aandacht vraag 1	Aandacht vraag 2	Attitude vraag 1	Attitude vraag 2	Attitude Vr. 1&2 combi	Koop Intentie
Placement type	Actief	0.080*	0.137*	0.008	0.057*	0.041	0.142*
	Programma bepalend	0.147*	0.221*	0.135*	0.495*	0.390*	0.310*
Programma type	Quiz	0.002	-0.071*	-0.050	-0.090*	-0.084*	-0.097*
	Information	-0.080*	-0.095*	-0.144*	-0.078*	-0.125*	-0.192*
Product type	Shopping	-0.077*	-0.088*	0.014	-0.067*	-0.036	-0.120*
	Luxe	0.002	0.004	0.047	0.004	0.027	-0.205*
Demografie	Man	-0.008	0.027	-0.019	-0.069*	-0.054*	0.003
	Kijk-frequentie	0.038	0.019	0.019	0.138*	0.099*	-0.01
Initiële voorkeuren	Programma	0.138*	0.123*	0.215*	0.130*	0.195*	0.075*
	Product	0.158*	0.192*	0.227*	0.118*	0.193*	0.430*
Model Samenvatting	Adjusted R ²	0.078	0.121	0.135	0.303	0.262	0.380
	F-waarde	10.21	15.87	17.89	48.15	39.52	67.29
	p	< 0.001	< 0.001	< 0.001	< 0.001	< 0.001	< 0.001
	n.	121	121	121	121	121	121

($p < 0.05$), laten de resultaten voor de tweede vraag zien dat de kijker de meeste aandacht heeft voor een product placement in een soap, gevolgd door die in een quiz (β , van -0.071) en tenslotte die in een informatief programma (β van -0.095) (tabel 1). De resultaten voor hypothese 4B zijn echter wat onduidelijker. Wederom is bij de eerste vraag alleen de β (-0.144) van het informatief programma significant ($p < 0.05$), terwijl de β (-0.050) voor het quiz programma toch nog een betrouwbaarheid van $p < 0.15$ haalt. Dit is belangrijk om te vermelden omdat de resultaten voor de tweede attitudevraag bij $p < 0.05$ aangeven dat de kijker de positiefste attitude vormt voor een product placement in een soap, gevolgd door die in een informatief programma (β van -0.078) en tenslotte die in een quiz (β van -0.090) (tabel 1). Waar de resultaten voor de eerste attitudevraag de hypothese dus lijken te ondersteunen, is dit niet het geval voor de significantere resultaten van de tweede vraag. Daarom is een extra regressie gedaan, waarbij de beide attitude vragen samen zijn genomen. De resultaten hiervoor

tonen bij $p < 0.01$ aan dat de consument de positiefste attitude vormt voor een product placement in een soap, gevolgd door die in een quiz (β van -0.084) en tenslotte die in een informatief programma (β van -0.125) (tabel 1). Op basis van al deze uitkomsten lijkt het toch dat hypothese 4B kan worden geaccepteerd. De resultaten voor hypothese 4C zijn weer duidelijker en tonen duidelijk aan dat de kans op een koopintentie bij de consument het hoogst is, wanneer deze placement is gedaan in een soap gevolgd door die in een quiz (β van -0.097) en tenslotte die in een informatief programma (β van -0.192) (tabel 1). Hiermee wordt hypothese 4c geaccepteerd.

Voor hypothese 5 laat de regressieanalyse alleen significante resultaten zien voor koopintentie. Uit die resultaten blijkt dat de kans dat een kijker daadwerkelijk een koopintentie ontwikkelt voor een product uit een product placement, het grootst is voor FMCG, gevolgd door 'shopping' producten (β van -0.120) en tenslotte luxe goederen (β van -0.205) (tabel 1). De resultaten voor aandacht en attitude lijken een indicatie te geven dat het plaatsen van een luxe product het best is om aandacht en een positieve attitude bij de kijker te bewerkstelligen, maar omdat de resultaten niet het vereiste betrouwbaarheidsniveau halen dient dit met voorzichtigheid te worden bekeken. Hypothese 5 kan daarom alleen worden geaccepteerd voor koopintentie.

De resultaten voor hypothese 6A (tabel 2) laten zien dat er inderdaad onderscheid gemaakt kan worden tussen de aandachtsniveaus van mannen en vrouwen wanneer het gaat om producten met geslachtsindicatoren. Vrouwen hebben in vergelijking met mannen minder aandacht voor in programma's geplaatste auto's (r van -0.212 en -0.249 bij $p < 0.05$). Voor in televisieprogramma's geplaatste kleding lijken zij meer aandacht te hebben dan mannen (r van 0.103 en 0.104), maar deze resultaten zijn slechts significant bij $p < 0.30$. Voor attitude is dit effect minder duidelijk in de resultaten te vinden, daar alleen de correlatie voor tweede attitudevraag en alleen voor de placement van kleding het $p < 0.05$ significantieniveau haalt. Daarom zijn ook de correlaties berekend voor de combinatie van beide vragen (tabel 2). Dit levert significante resultaten ($p < 0.05$) op voor de placement van kleding, hetgeen aantoont dat vrouwen een meer positieve attitude vormen voor de plaatsing van kleding in televisieprogramma's dan mannen. De resultaten voor de koopintentie tenslotte tonen aan dat bij $p < 0.10$ vrouwen in vergelijking met mannen minder snel geneigd zijn koopintenties te ontwikkelen voor geplaatste auto's (r van -0.151) en sneller geneigd zijn dit te doen voor kleding

Tabel 2. Invloed van gedrag op geslacht (alle dikgedrukte correlaties met * voor significantie van $p < 0.05$).

	Aandacht vraag 1	Aandacht vraag 2	Attitude vraag 1	Attitude vraag 2	Attitude Vr. 1&2 combi	Koop Intentie
Beschrijving 2 met auto	-0.212*	-0,249*	-0.175	0.022	-0.065	-0.151
p-waarde	0.020	0.006	0.055	0.813	0.315	0.098
Beschrijving 8 met kleding	0.103	0.104	0.026	0.234*	0.139*	0.384*
p-waarde	0.261	0.254	0.777	0.010	0.031	0.000

(r van 0.384) (tabel 2). Concluderend kan worden gesteld, dat bij $p < 0.10$ kijkers zoals verwacht meer aandacht, een positievere attitude en een grotere kans op een koopintentie hebben voor geplaatste producten met geslachtsindicatoren van hun eigen geslacht.

De regressieanalyse voor hypothese 6B (tabel 1) laat geen significante resultaten zien die de hypothese ondersteunen. Mannen geven dus niet vaker dan vrouwen aan de intentie te hebben een product te kopen dat zij gezien hebben in een televisieprogramma. Zij verschillen overigens ook verder niet significant in hun aandacht of attitude reacties op product placements (tabel 1).

Hypothese 7 wordt ondersteund door de resultaten van de regressieanalyse voor de tweede attitudevraag (β van 0.138) en voor de combinatie van beide attitudevragen (0.099) (tabel 1). Hoe vaker iemand televisie kijkt, hoe positiever diens attitude ten opzichte van het concept product placement. Voor de andere stappen van het gedragsmodel is geen invloed gevonden van de kijkfrequentie.

Uit de resultaten van de regressieanalyses komt tenslotte duidelijk naar voren dat zowel initiële programmavoorkeuren als initiële productvoorkeuren een positieve invloed hebben op aandacht, attitude en koopintentie (tabel 1). De β 's zijn zelfs relatief groot in vergelijking met die voor de geteste product placement variabelen.

6. DISCUSSIE

Een eerste beperking van het onderzoek behelst de presentatie van de stimuli. In dit onderzoek zijn de respondenten namelijk niet geconfronteerd met scènes met product placements in een daadwerkelijk getoond televisieprogramma, maar met beschrijvingen van dergelijke scènes op papier. Dit is duidelijk niet de gebruikelijke manier waarop mensen product placements ervaren. Tevens is uit eerder onderzoek gebleken dat mensen voor het kijken van televisie de andere hersenhelft gebruiken dan voor het lezen van tekst en dat de cognitieve prestaties van deze helften verschillen (Krugman, 2000). Het gebruik van product placement beschrijvingen in plaats van echte voorbeelden op televisie kan de resultaten van dit onderzoek dus hebben beïnvloed (Russell & Puto, 1999b). Een andere limitatie van dit onderzoek komt voort uit de veronderstelling dat mensen altijd een bepaalde drempelwaarde van betrokkenheid (involvement) halen bij het programma dat ze kijken. Hoewel, dit in onze opinie zeker geldig is, wil dat niet zeggen dat vervolgens de mate van betrokkenheid voor de programma's altijd hetzelfde is (Russell & Puto, 1999a). Men kan bijvoorbeeld naar een programma kijken omdat "er niets beters op is" of omdat men perse "dat interview met Jack Spijkerman" wil zien. In deze gevallen is er toch zeker een verschil in de mate van betrokkenheid (involvement) en dit zal waarschijnlijk ook leiden tot een verschil in de perceptie van product placements (Russell & Puto, 1999a).

Toekomstig onderzoek naar het effect van product placements zou dan ook kunnen controleren voor de mate van betrokkenheid bij het programma in kwestie. Een andere suggestie is de respondenten in toekomstig onderzoek de product placements te presenteren in werkelijke of nagespeelde scènes op de televisie. Ten derde is het voor toekomstig onderzoek ook interessant om de leeftijdsgroep waarvoor wordt getest zowel naar boven als beneden uit te breiden. Dit onderzoek heeft namelijk alleen getest op 121 mensen in de leeftijd van 18 tot 35 jaar, waarvan 56% man en 44% vrouw was. Als laatste suggestie is het tevens interessant om te kijken hoe kijkers omgaan met product

placements naarmate zij meer ervaring met deze vorm van reclame krijgen. Dit zou kunnen worden onderzocht met behulp van een longitudinaal onderzoek.

Ter afsluiting van deze paper is het interessant te kijken wat de uitkomsten van dit onderzoek kunnen betekenen voor de praktijk. Het eerste wat dit onderzoek duidelijk maakt is dat de combinatie van een ‘programma bepalend’ type placement in een soap en met een FMCG de grootste kans zal hebben om de kijker tot aankoop van het product te verleiden. Voor het creëren van aandacht voor, of een positieve attitude jegens het product, blijkt de combinatie van een ‘programma bepalend’ type placement en een soap programma het beste te werken. Het blijft echter onduidelijk welk type product hier het beste voor is. Een tweede uitkomst van dit onderzoek die interessant is voor de praktijk, is dat men rekening dient te houden met de programmavoorkeuren van de doelgroep. Het is duidelijk aangetoond dat kijkers, die een voorkeur hebben voor een bepaald programma, ook meer aandacht hebben, een positievere attitude vormen en waarschijnlijker een koopintentie ontwikkelen voor een product placement in dat programma. Dit betekent dus dat een bedrijf haar producten het beste kan plaatsen in het favoriete programma van haar doelgroep en dan niet alleen om meer consumenten te bereiken, maar tevens omdat de product placement dan effectiever is. Ten derde blijkt uit dit onderzoek dat een toegenomen kijkfrequentie leidt tot positievere attitudes voor product placement bij de kijker. Het is dus voor de effectiviteit van de product placement belangrijk een programma te kiezen dat vooral bekeken wordt door mensen, die al veel televisie kijken. Tot slot van deze paper is de laatste praktische implicatie van dit onderzoek dat marketing managers er goed aan doen hun product te analyseren op geslachtsindicatoren alvorens het te plaatsen in een televisieprogramma. Een placement is namelijk alleen effectief wanneer het product die indicatoren vertoont die passen bij de demografische kenmerken van het publiek. Producten met mannelijke indicaties, zoals auto’s, dienen dus te worden geplaatst in programma’s met een overwegend mannelijk publiek, terwijl die met vrouwelijke indicaties beter geplaatst kunnen worden in de door overwegend vrouwen bekeken programma’s.

LITERATUUR

- Adelman, S. (1962, February). Orthogonal main-effect plans for asymmetrical factorial experiments. *Technometrics*, 4, February, 21 – 46.
- Akkerman, I. (2003). Account Manager Sponsoring for Endemol Netherlands. Lecture at the “Behind the scenes: Let the imagination guide you” congress at the University of Maastricht, organized by 3MA.
- Babin, L. A. & Carder, S. T. (1996). Advertising Via the Box Office: Is Product Placement Effective? *Journal of Promotion Management*, 3 (1-2), 31 – 51.
- Bagozzi, R P., Baumgartner, H. & Yi, Y. (1992, March). State versus action orientation and the theory of reasoned action: An application to coupon usage. *Journal of consumer research*, 18 (4), March, 505 – 518.
- Bucholz, L. M. & Smith, R. E. (1991). The role of consumer involvement in determining cognitive response to broadcast advertising. *Journal of advertising*, 20 (1), 4 – 17.
- d’Astous, A. & Bitz, P. (1995). Consumer evaluations of sponsorship programmes. *European Journal of Marketing*, 29 (12), 6 – 22.
- d’Astous, A. & Séguin, N. (1999). Consumer Reactions to Product Placement Strategies in Television Sponsorship. *European Journal of Marketing*, 33 (9/10), 896 – 910.
- DeLorme, D. E. & Reid, L. N. (1999, summer). Moviegoers’ Experiences and Interpretations of Brands in Films Revisited. *Journal of Advertising*, 28 (2), summer, 71 – 95.

- DeLorme, D. E., Reid, L. N. & Zimmer, M. R. (1994). Brands in Films: Young Moviegoers' Experiences and Interpretations. In K. W. King (Ed.), *Proceedings of the 1994 Conference of the American Academy of Advertising* (p. 60). Athens, USA: American Academy of Advertising.
- Dibb, S., Simkin, L., Pride, W. M. & Ferrell, O. C. (1997). *Marketing: Concepts and Strategies* (3rd ed.). Boston, USA: Houghton Mifflin.
- Duncan, C. P. & Nelson, J. E. (1985). Effects of humor in a radio advertising experiment. *Journal of advertising*, 14 (2), 33 – 40.
- Eagly, A. H. & Chaiken, S. (1995). Attitude Strength, Attitude Structure and Resistance to Change. In R. E. Petty & J. A. Krosnick (Eds.). *Attitude Strength: Antecedents and Consequences* (pp. 247 – 282). Hillsdale, USA: Erlbaum.
- Friestad, M. & Wright, P. (1994, June). The Persuasion Knowledge Model: How People Cope with Persuasion Attempts. *Journal of Consumer Research*, 21 (1), 1 – 31.
- Gould, S. J., Gupta, P. B. & Grabner-Kräuter, S. (2000, winter). Product Placements in Movies: A Cross-Cultural Analysis of Austrian, French and American Consumers' Attitudes Toward This Emerging International Promotional Medium. *Journal of Advertising*, 29 (4), winter, 41 – 58.
- Gupta, P. B. & Gould, S. J. (1997, spring). Consumers' Perceptions of the Ethics and Acceptability of Product Placements in Movies: Product Category and Individual Differences. *Journal of Current Issues and Research in Advertising*, 19 (1), spring, 37 – 50.
- Gupta, P. B. & Lord, K. R. (1998, spring). Product Placement in Movies: The Effect of Prominence and Mode on Audience Recall. *Journal of Current Issues and Research in Advertising*, 20 (1), spring, 47 – 59.
- Hair Jr., J. F., Anderson, R. E., Tatham, R. L. & Black, W. C. (1995), *Multivariate Data Analysis with Readings* (4th ed.). New Jersey, USA: Prentice-Hall Inc.
- Henthorne, T. L., LaTour, M. S. & Natarajan, R. (1993). Fear appeals in print advertising: An analysis of arousal and ad respons. *Journal of advertising*, 22 (2), 59 – 69.
- Higgins, K. T. (1985). There's Gold in Silver Screen Product Plugs. *Advertising Age*, 19 (21), p. 6.
- Iyer, E. and Debevec, K. (1986). Gender Stereotyping of Products: Are Products Like People? *Developments in Marketing Science*, 9, 40 – 45.
- Kanungo, R. N. & Pang, S. (1973, April). Effects of Human Models on Perceived Product Quality. *Journal of Applied Psychology*, 57 (2), April, 172 – 178.
- Krugman, H.E. (1977, August). Memory without Recall, Exposure without Perception. *Journal of Advertising Research*, 17 (4), August, 7 – 12.
- Krugman, H.E. (2000, November/December). Memory without Recall, Exposure without Perception. *Journal of Advertising Research*, 40 (6), November/December, 49 – 54.
- Lancaster, K. J. (1966). A new approach to consumer theory. *Journal of Political Economy*, 74 (2), 132 – 157.
- Likert, R. (1932). A technique for the measurement of attitudes. *Archives of Psychology*, 140.
- Luce, R. D. & Tuckey, J. W. (1964). Simultaneous conjoint measurement: A new type of fundamental measurement. *Journal of Mathematical Psychology*, 8 (1), 1 – 27.
- MacKenzie, S. B. & Lutz, R. J. (1989, April). An Empirical Examination of the Structural Antecedents of Attitude Toward the Ad in an Advertising Pretesting Context. *Journal of Marketing*, 53, April, 48 – 65.
- Malhotra N. K. (1999), *Marketing Research: An applied orientation* (3rd ed.). New Jersey, USA: Prentice-Hall Inc.
- McGuire, W. J. (1973). Persuasion, Resistance and Attitude Change. In I. Pool, W. Schramm, F. Frey, N. Maccoby & E. B. Parker (Eds.). *Handbook of Communication* (pp. 216 – 252). Chicago, USA: Rand McNally.
- Meenaghan, T. (1991). Sponsorship – legitimizing the medium. *European Journal of Marketing*, 25 (11), 5 – 10.

- Mitchell, A. A. & Olson, J. C. (1981). Are Product Attribute Beliefs the Only Mediator of Advertising Effects on Brand Attitude? *Journal of Marketing Research*, 18, August, 318 – 332.
- Morton, C. R. & Friedman, M. (2002). I Saw It In The Movies: Exploring the Link Between Product Placement Beliefs and Reported Usage Behavior. *Journal of Current Issues and Research in Advertising*, 24 (2), 33 – 40.
- Petroshius, S. M. & Monroe, K. B. (1987). Effect of product-line pricing characteristics on product evaluations. *Journal of consumer research*, 13 (4), March, 511 – 519.
- Politz, A (1975). Creativeness and Imagination. *Journal of Advertising*, 4 (3), 11 – 14.
- Russell, C. A. (1998). Toward a Framework of Product Placement: Theoretical Propositions. *Advances in Consumer Research*, 25 (1), 357 – 362.
- Russell, C. A. (2002). Investigating the Effectiveness of Product Placements in Television Shows: The Role of Modality and Plot Connection Congruence on Brand Memory and Attitude. *Journal of Consumer Research*, 29 (3), December, 306 – 318.
- Russell, C. A. & Puto, C. P. (1999a). Rethinking Television Audience Measures: An Exploration into the Construct of Audience Connectedness. *Marketing Letters*, 10 (4), 387 – 401.
- Russell, C. A. & Puto, C. P. (1999b). Novel Experimental Methods: Opportunities and Challenges. *Advances in Consumer Research*, 26 (1), 590 – 600.
- Shapiro, M. (1993). *Product Placement in motion pictures* (working paper). USA: Northwestern University.
- Steertz, E. M. (1987). *The Cost Efficiency and Communication Effects Associated With Brand Name Exposure Within Motion Pictures* (unpublished master's thesis). USA: University of West Virginia.
- Van Gelder, H. (1997). Reclame buiten de spot. *NRC Handelsblad*, October 1st.
- Vollmers, S. & Mizerski, R. (1994). A Review and Investigation Into the Effectiveness of Product Placements in Films. In K.W. King (Ed.), *Proceedings of the 1994 American Academy of Advertising Conference* (pp. 97 – 102). Athens, USA: American Academy of Advertising.
- Zimmer, M. R. and DeLorme, D. E. (1997). *The Effects of Brand Placement Type and a Disclaimer on Memory for Brand Placements in Movies*. Paper presented at the Association for Education in Journalism and Mass Communication Conference, Chicago, USA.

2. *Virtuele communities voorbij de hype: Nieuwe inzichten voor managers, marketeers en marktonderzoekers*

K. DE VALCK, G.H. VAN BRUGGEN en B. WIERENGA¹

SAMENVATTING

Consumenten leggen steeds vaker contact met elkaar via het Internet. Elektronische discussieforums, mededelingenborden, nieuwsgroepen, e-maillijsten, chatrooms en weblogs bieden consumenten, waar dan ook ter wereld, de mogelijkheid om hun kennis, ervaringen en meningen te delen. De populariteit van deze elektronische contacten tussen consumenten blijkt uit de grote hoeveelheid virtuele communities die georganiseerd zijn rondom consumptiegerelateerde interesses. Dit soort communities vertegenwoordigen uitgebreide netwerken van consumptiekennis en kameraadschap die invloed hebben op consumentengedrag. In dit artikel gaan we in op een aantal nieuwe inzichten ten aanzien van virtuele communities die managers, marketeers en marktonderzoekers beter in staat stellen om strategisch gebruik te maken van de vele mogelijkheden die nu nog vaak onbenut worden gelaten.

1. INTRODUCTIE

“Ik ben lid geworden in de zomer van 2001, omdat ik graag gebruik wilde maken van de receptendatabase. Op dat moment had ik geen idee van alle andere dingen die je kan doen op SmulWeb. Ik wist helemaal niet dat het een virtuele community was. Ik ben begonnen met het invoeren van één recept en ik was helemaal enthousiast toen het op mijn persoonlijke pagina verscheen. Dus toen ben ik meer recepten gaan toevoegen en ik ben gaan rondsurfen om te kijken wat ik nog meer kon doen. Ik bezocht de persoonlijke pagina’s van andere leden die mooi gemaakt waren met plaatjes en andere dingen, en het leek me leuk om ook mijn eigen pagina aan te kleden. Dat is hoe ik betrokken ben geraakt bij SmulWeb. Toen ik net begon, zei ik tegen mezelf dat ik niet teveel tijd online mocht zijn en ik heb me dan ook echt beperkt. Maar tegenwoordig ben ik de hele dag online. Ik heb twee computers op mijn werkplek, zodat ik op de ene kan werken en de andere kan gebruiken om in een verloren half uurtje deel te nemen in de SmulWeb-discussieforums. Ik log ook in vanaf mijn computer thuis, elke dag een half

¹ Dit artikel is gebaseerd op het proefschrift *Virtual Communities of Consumption; Networks of Consumer Knowledge and Companionship* (De Valck 2005). De auteurs danken redacteur Ale Smidts voor zijn commentaar op eerdere versies van dit artikel.

uur of zo, om artikelen en recensies toe te voegen". (Interview met Julia [1], SmulWeb-lid, op 5 april 2002 te Amsterdam).

SmulWeb is Nederlands grootste culinaire virtuele community. Opgericht in 1998 door een onafhankelijk bedrijf dat online community-marketingconcepten ontwikkelde, telt SmulWeb op dit moment ongeveer 160.000 leden, 275.000 recepten, 14.000 restaurant-recensies, 15.000 culinaire artikelen, 200 subcommunities en zes discussieforums. In 2004 is SmulWeb verkocht aan portal-exploitant Internet Pleinen B.V. Het succes van SmulWeb zit 'm in de toewijding van leden zoals Julia, die veel tijd en energie steken in het leveren van bijdragen. Samen met andere leden die haar passie voor koken en lekker eten delen, vormen zij een hechte gemeenschap. Men kan bij elkaar terecht voor culinair advies, maar ook voor een gezellig kletspraatje of voor emotionele steun in moeilijke tijden. SmulWeb is één van de vele virtuele communities die zijn georganiseerd rondom consumptiegerelateerde interesses. In dit soort communities draait het bovenal om het delen van kennis en ervaringen, het uitwisselen van tips en meningen, het bediscussiëren van gewoonten en gebruiken. Soms met betrekking tot een consumptieactiviteit in het algemeen, maar dikwijls ook met betrekking tot specifieke producten, diensten of merken. In SmulWeb gaat het dus over kooktechnieken, eetrituelen en biologische productie, maar ook over restaurant X, Senseo Crema en Coca Cola.

Voor marketeers vormen dit soort virtuele communities een potentiële bron van waardevolle informatie. Immers, de bijdragen van de leden geven vaak zeer gedetailleerd inzicht in consumptiepatronen, wensen en behoeften, en in de symbolische betekenis die wordt toegekend aan de consumptieactiviteit en haar producten. Een groot voordeel van virtuele communities is dat ze ongemerkt geraadpleegd kunnen worden. Anders dan met traditionele methoden zoals focusgroepen, interviews en enquêtes stelt het de marketeer in staat om de conversaties tussen consumenten in een natuurlijke, real-time setting te observeren. Natuurlijk kunnen marketeers ook actief deelnemen in communities. Zij verlenen hen immers toegang tot een groep loyale en enthousiaste consumenten. Door deze groep te voorzien van laatste nieuwtjes, speciale aanbiedingen en achtergrondinformatie over bedrijf en product kan hun loyaliteit (en hun mond-tot-mond reclamepotentieel!) verhoogd worden. Ook kan deze groep geconsulteerd worden waar het gaat om productontwikkeling en gebruikstips. Op die manier kunnen via virtuele communities relaties ontstaan tussen producent en consument die breder, dieper en sterker zijn dan relaties die enkel transactiegericht zijn.

Virtuele communities lijken dus niets dan goeds te brengen. Op het hoogtepunt van de Internethype zijn zij zelfs door velen gepropageerd als de ultieme online marketingapplicatie (o.a., Hagel III & Armstrong 1997; Jolink 2000; McWilliam 2000; Oostveen 2001; Williams & Cothrel 2000). Veel bedrijven hebben dus indertijd een community-functionaliteit aan hun websites toegevoegd in de verwachting hiermee een groot aantal bezoekers te trekken en te binden. Echter, in de praktijk bleven vele discussieforums en chatrooms leeg. Vaak was er geen geld om personeel aan te stellen dat de bijdragen op gang kon brengen en, als het wél liep, kon monitoren en analyseren. Als gevolg hiervan is menig community-initiatief een stille dood gestorven. Het laatste jaar is de belangstelling voor het fenomeen virtuele community echter weer groeiende. Dit is onder andere in de hand gewerkt door het succes van *social networking* software, dat bijvoorbeeld wordt gebruikt door dating-sites zoals OneHello en Meetup, en door de populariteit van *blogging*, het bijhouden van een virtueel dagboek. Het commerciële succes van communities als eBay en iVillage (resp. 1,1 miljard en 21,1 miljoen dollar winst in Q2-

2005), de bewezen impact van online consumentenrecensies op sites zoals Amazon en Filmcritic (Chevalier & Mayzlin 2003; Dellarocas, Farag Awad & Zhang 2004), alsmede de impact die uitgaat van georganiseerde bloggers die online hun beklag doen over een bedrijf of product en die hun krachten kunnen bundelen om gezamenlijk actie te ondernemen (Businessweek, 2 mei 2005; Financial Times, 7 februari 2005), hebben bedrijven wederom de ogen geopend voor de exploitatiemogelijkheden die een virtuele community kan bieden.

In tegenstelling tot een aantal jaar geleden is er veel meer bekend over de werking van virtuele communities. Verschillende onderzoekers hebben systematisch in kaart gebracht waarom mensen deelnemen aan virtuele communities, hoe ze eraan deelnemen en met welk effect (o.a., Bagozzi & Dholakia 2002; De Valck 2005; Dholakia, Bagozzi & Klein Pearo 2004; Hennig-Thurau, Gwinner, Walsh & Gremler 2004). Daarnaast zijn er twee belangwekkende artikelen gepubliceerd over de manier waarop marketinginformatie onttrokken kan worden aan online conversaties tussen consumenten (Kozinets 2002; Godes & Mayzlin 2004). Marketeers kunnen hun voordeel doen met deze kennis, zodat (hernieuwde) community-initiatieven de gewenste vruchten afwerpen.

Het doel van dit artikel is om een aantal nieuwe inzichten op een rij te zetten en aan te geven hoe managers, marketeers en marktonderzoekers er strategisch gebruik van kunnen maken. Ten eerste bespreken we verschillende typen community-platformen, daarbij aandacht schenkend aan specifieke organisatiefactoren die bijdragen aan de optimalisatie van de community als communicatiemedium tussen consumenten en tussen consumenten en producenten. Ten tweede geven we een classificatie van community-leden op basis van hun participatieprofielen die marketeers in staat stelt om interessante doelgroepen te lokaliseren en hen inzichten biedt in hoe deze doelgroepen aangesproken kunnen worden. Tenslotte laten we zien hoe marktonderzoekers door middel van netnografisch onderzoek niet alleen inzicht kunnen krijgen in de werking van de virtuele community als referentiegroep, maar ook in de afwegingen die ten grondslag liggen aan de aankoopbeslissingen van de community-leden.

2. COMMUNITY-PLATFORMEN

2.1. *Community-platformen: doelen en functionaliteiten*

Er zijn virtuele communities in vele soorten en maten; variërend in onderwerp van triviaal tot zeer serieus, variërend in toegankelijkheid van open tot zeer gesloten, variërend in redactie van streng tot helemaal niet gecensureerd, en variërend van opgericht met winst oogmerk tot het onbaatzuchtig delen van kennis. Virtuele communities worden bovendien geïnitieerd door bedrijven en individuen, professionals en amateurs, kenners en kwanselaars. Kortom, het veld is breed en omvangrijk. De vraag die rijst, is met welk type community moet je je als marketeer inlaten? En hoe moet je je opstellen om door de community-leden geaccepteerd te worden?

Gezien de grote hoeveelheid communities die al voorhanden is, is het allereerst belangrijk een afweging te maken tussen aansluiting vinden bij één of meerdere bestaande communities of zelf een community creëren. Het grote voordeel van de eerste optie is dat er geen of slechts beperkte energie gestoken hoeft te worden in opbouw en onderhoud. Dit is vooral aantrekkelijk voor bedrijven die niet over voldoende financiële mid-

delen beschikken om een community-platform in te richten en een campagne te voeren om bezoekers naar de site te trekken. Het creëren van een eigen community kost veel tijd en energie, maar het grote voordeel van deze optie is dat men de zeggenschap over inhoud en organisatie in eigen handen heeft. Op die manier kan de community optimaal benut worden als marketinginstrument, CRM-platform, inspiratiebron voor productontwikkeling, et cetera.

Tabel 1. Community-doelen en ondersteunende functionaliteiten.

Informatie Mededelingenborden, discussieforums, gespecialiseerde databases, zoekfunctie	Recreatie Chatrooms, spelletjes, wedstrijden, nominaties, fotogalerie
Relatie Ledenpagina's, ledenprofielen, e-mail, gastenboeken, verjaardagskalender, cyclische evenementen, <i>RL</i> ontmoetingen	Transactie Veilingen, reputatiesysteem, bestel- en betalingsfunctionaliteiten

De meest bekende classificatie van virtuele communities is afkomstig van Hagel en Armstrong (1997). Zij onderscheiden communities naar het doel dat ze dienen, namelijk; *communities of interest* (informatie), *communities of fantasy* (recreatie), *communities of relationships* (relatie) en *communities of transaction* (transactie). Om deze doelen te realiseren zijn er verschillende community-software-systemen nodig (Kozinets 1999). Tabel 1 geeft een overzicht van de vier community-doelen en de functionaliteiten die deze doelen ondersteunen.

Het is van belang de gewenste doelen en de daarvoor benodigde functionaliteiten met elkaar in overeenstemming te brengen. Als een bedrijf bijvoorbeeld kiest voor aansluiting bij een bestaande community, zoals een Yahoo-groep of een club van Het Net, dan moet men zich realiseren dat de community-functionaliteit van deze portals alleen uit discussieforums bestaat. Dit maakt deze communities dus meer geschikt voor informatie-uitwisseling dan voor het opbouwen van relaties tussen community-leden die elkaar nog niet kennen. De van de communities losstaande chatrooms op deze portals trekken bezoekers op zoek naar vluchtige, dikwijls anonieme recreatie, die zich dus niet gauw tot toegewijde community-leden zullen ontwikkelen.

Bedrijven kunnen natuurlijk ook kiezen voor aansluiting bij of creatie van een geïntegreerd platform met meerdere functionaliteiten, zodat verschillende doelen gecombineerd kunnen worden. Een goed voorbeeld hiervan is eBay. De leden kunnen transacties met elkaar verrichten met behulp van een veiling- en betalingssysteem en het door eBay ontwikkelde reputatiemechanisme waarbij kopers en verkopers elkaars gedrag kunnen beoordelen. Daarnaast kunnen zij bij elkaar terecht met vragen over een gedeelde interesse via vele gespecialiseerde discussieforums. Tenslotte ondersteunt eBay de ontwikkeling van relaties tussen leden door middel van ledenprofielen en de mogelijkheid elkaar persoonlijke berichten te sturen. Ook SmulWeb is opgebouwd volgens het principe van een geïntegreerd community-platform dat een combinatie van informatie-, recreatie- en relatiedoelen dient.

Of er nu gekozen wordt voor aansluiting bij een bestaande community of het creëren van een eigen community, in beide gevallen moeten bedrijven er zich bewust van zijn dat community-platformen niet gelijk zijn aan reclamemedia. Virtuele communities

zijn ontstaan uit de behoefte te communiceren. Al gedurende het eerste computernetwerk-experiment in de jaren zestig, werd de mogelijkheid om onafhankelijk van tijd en plaats met elkaar te communiceren volop benut door de betrokken onderzoekers die verspreid waren over verschillende onderzoeksinstituten in de V.S.. De allereerste community was een e-maillijst gevormd door een aantal van die onderzoekers, allen geïnteresseerd in science fiction, om informatie uit te wisselen over boeken, films, et cetera (Hauben & Hauben 1999; Rheingold 1993). Je hebt het als bedrijf bij het verkeerde eind als je virtuele communities alleen gebruikt als middel om reclameboodschappen te verspreiden. Alhoewel community-leden best openstaan voor banners die toegespitst zijn op hun behoeften, zijn zij meer geïnteresseerd in achtergrondinformatie die iets toevoegt aan hun expertise (Oostveen 2001). Bedrijven die willen participeren in de online discussies tussen de community-leden zullen moeten openstaan voor een echte dialoog. Shell reageert inhoudelijk op alle berichten die achtergelaten worden op haar community-platform. De discussianten voelen zich daardoor serieus genomen en zij zijn bereid een negatieve houding bij te stellen. Op die manier bereikt Shell wellicht meer dan met een ophemelende imagocampagne.

2.2. *Strategisch management van community-platformen*

Howard Rheingold, geestelijk vader van de term virtuele community, heeft één van de eerste belangwekkende boeken over communities geschreven (*The Virtual Community*, 1993). In dit boek zet hij uiteen dat virtuele communities als het ware vanzelf ontstaan daar waar mensen de beschikking hebben over interactieve communicatienetwerken. Deze technologisch-deterministische visie is een utopie gebleken in de huidige netwerksamenleving waarin er een overdaad is aan communicatiemogelijkheden. Het toevoegen van een chatroom of een discussieforum aan een corporate website heeft niet automatisch tot gevolg dat er gebruik van wordt gemaakt, laat staan dat er een community-gevoel tussen de deelnemers ontstaat. De vraag is dan ook, hoe kunnen community-managers de functie van de community als communicatiemedium tussen consumenten en tussen consumenten en producenten optimaliseren?

Hennig-Thurau et al. (2004) hebben onderzocht wat consumenten motiveert om bijdragen te leveren aan online opinieplatforms. De belangrijkste reden die voor alle *posters* geldt, is de wens om anderen met hun kennis en ervaring te helpen. Vervolgens onderscheiden zij vier groepen *posters* op basis van hun secundaire motivaties om bijdragen te leveren. De eerste groep zijn de echte altruïsten die alleen gedreven worden door de wens andere consumenten helpen. De tweede groep wil zowel consumenten als producenten laten profiteren van hun kennis en expertise. De derde groep wordt gedreven door economische prikkels; bijdragen worden bijvoorbeeld beloond met punten die opgespaard kunnen worden voor een online aanschaf. De vierde groep heeft verscheidene secundaire motivaties, waaronder positieve eigenwaarde, sociaal profijt, reciprociteit en het helpen van de producent en de community-manager. Deze laatste groep levert de meeste bijdragen, terwijl de eerste en tweede groep het minst aanleveren. Hennig-Thurau et al. concluderen dat, om het aantal bijdragen aan de community te verhogen, community-managers vooral moeten wijzen op het nut van deze bijdragen voor anderen, zijnde zowel consumenten als het bedrijf. Daarnaast gaat het erom het eigen voordeel voor de consument te benadrukken, zoals een verhoogde sociale status, erkenning van expertise en wederkerigheid wanneer zij zélf op zoek zijn naar informatie of advies. In de laatste plaats kunnen er dan ook financiële incentives aan het leveren van bijdragen gekoppeld worden.

Het aanleveren van bijdragen is de ene kant van de medaille, het gebruik maken van deze bijdragen de andere. Over het algemeen neemt men aan dat slechts 20% van de community-leden bijdragen levert, terwijl 80% alleen maar leest wat anderen hebben toegevoegd. Een eerste vereiste om het gebruik van de community als informatiebron te stimuleren is ervoor te zorgen dat er een effectieve en efficiënte zoekfunctie is. Maar wat kan de marketeer die de impact van de community op de aankoopbeslissingen van haar leden wil verhogen nog meer doen? Om hierin inzicht te krijgen, hebben De Valck, Van Bruggen en Wierenga (2003) onderzoek gedaan naar de determinanten en effecten van de invloed die uitgaat van virtuele communities op het consumptiebeslissingsproces van consumenten. Zij zijn de eersten die empirisch aantonen dat community-invloed aanzienlijk is, vooral in de fases van het zoeken naar informatie en het afwegen van alternatieven (De Valck 2005). Door actief deel te nemen in de virtuele community en haar te voorzien van achtergrondinformatie, uitleg en commentaar kunnen marketeers de communicatie tussen de community-leden ondersteunen, bijsturen en, waar nodig, weerleggen. Op die manier kunnen zij de waarde van de community als informatiebron verhogen en haar impactpotentieel optimaal benutten.

Het onderzoek van De Valck, Van Bruggen en Wierenga brengt verder aan het licht dat de belangrijkste determinanten van community-invloed gelijk zijn aan die van interpersoonlijke beïnvloeding in de traditionele context. Dat wil zeggen dat naarmate leden een sterkere sociale band hebben met de community en naarmate ze de community vaker bezoeken (ze zich vaker in haar invloedssfeer bevinden) community-invloed op het beslissingsproces toeneemt. Om de impact van de community als informatiebron te optimaliseren zouden managers van virtuele communities dus aandacht moeten besteden aan het verhogen van de sociale betrokkenheid van de leden bij de community en het verhogen van hun bezoekfrequentie. Sociale betrokkenheid kan gestimuleerd worden door het toevoegen van functionaliteiten waarmee leden relaties met elkaar kunnen aangaan (zie Tabel 1). De bezoekfrequentie kan verhoogd worden door ervoor te zorgen dat er regelmatig nieuwe en interessante bijdragen worden toegevoegd. Naast de door Hennig-Thurau et al. (2004) voorgestelde methoden kan het management dit stimuleren door goede bijdragen eruit te lichten en te belonen, zodat leden worden aangespoord deze voorbeelden te volgen.

3. COMMUNITY-LEDEN

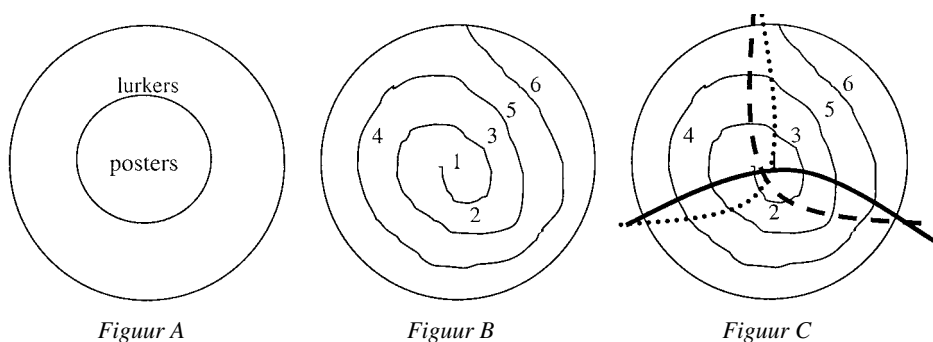
3.1. *Community-leden: een classificatie*

Community-managers, marketeers en marktonderzoekers moeten zich realiseren dat virtuele communities op verschillende manieren gebruikt worden door hun leden. In de quote aan het begin van dit artikel omschrijft Julia hoe haar lidmaatschap van SmulWeb zich langzaam ontwikkeld heeft. Ze herinnert zich de aantrekkingskracht van de uitgebreide persoonlijke pagina's van andere leden en de verbazing toen haar bijdrage op haar eigen pagina verscheen. Ze besluit meer bijdragen te gaan leveren en ook haar eigen pagina uit te breiden met persoonlijke informatie en illustraties. Al snel heeft ze contact gelegd met andere actieve leden door middel van gastenboekberichten, forumdiscussies en privé e-mailuitwisselingen. De tijd en energie die ze in haar lidmaatschap steekt, zijn aanzienlijk toegenomen sinds haar aanmelding. Van een incidenteel be-

zoekje volgt ze nu dagelijks urenlang de discussies en nieuwe bijdragen, waar ze in de avonduren ook haar steentje aan bijdraagt.

Natuurlijk zijn niet alle community-leden zo actief als Julia. Een groot deel van de leden laat zich zelfs helemaal niet horen in de community, maar maakt er wel stilletjes gebruik van. De meest gangbare classificatie van typen leden is dan ook de dichotomie *poster* versus *lurker*; dat wil zeggen, leden die actief bijdragen aan de inhoud van de community versus leden die alleen maar lezen wat anderen bijgedragen hebben (Rheingold 1993; Hagel & Armstrong 1997). Niet alleen wordt algemeen aangenomen dat de eerste groep veel kleiner is dan de tweede, maar ook dat de kleinere groep posters de meest waardevolle groep leden is. Onderzoek heeft onder andere uitgewezen dat posters de community vaker bezoeken, er meer bij betrokken zijn en er ook meer online kopen (Brown, Tilton & Woodside 2002; Okleshen & Grossbart 1998). Marketeers zouden zich dus alleen op de posters kunnen richten, maar daarmee gaan ze, wellicht ten onrechte, voorbij aan het grootste deel van de community-leden. Verder kunnen we ons afvragen of alle posters even interessant zijn. De simpele tweedeling van posters en lurkers geeft geen inzicht in deze zaken.

De Valck, Van Bruggen en Wierenga (2005) hebben een classificatie ontwikkeld die recht doet aan de verschillen tussen hoe vaak, hoe veel en in welke vorm leden informatie bijdragen en onttrekken aan de community (zie ook De Valck 2005). Daartoe werd een grootschalige online enquête uitgevoerd onder leden van de SmulWeb-community. In totaal hebben 1007 leden vragen beantwoord over hun SmulWeb-lidmaatschap, hun deelnamepatroon en hun persoonlijke karakteristieken. De classificatie is gebaseerd op vijf factoren. Twee factoren geven inzicht in de mate van participatie, namelijk bezoekfrequentie en bezoekduur. De andere drie factoren geven inzicht in de vorm van participatie; i.e., de mate waarin leden informatie aan de community onttrekken en toevoegen via de databases met recepten, recensies en artikelen en de mate waarin ze actief discussiëren met andere leden via de discussieforums en chatroom. De classificatie is uitgevoerd door middel van een clusteranalyse en zij is gevalideerd via drie verschillende procedures. Het resultaat is een classificatie van zes verschillende typen community-leden. Deze zes typen kunnen we in een denkbeeldige cirkel plaatsen die de community representeert.



Posters: (1) kernleden; (2) conversatieliefhebbers; (3) informatieliefhebbers; (4) hobbyisten;
Lurkers: (5) functionalisten; (6) opportunisten

Figuur 1. Classificatie van community-leden.

Figuur A laat de traditionele tweedeling van posters en lurkers zien, d.w.z. een kleine kern leden die bijdragen levert, omringt door een grote groep leden die niets toevoegt, maar alleen informatie onttrekt aan de community. De poster-lurker dichotomie komt overeen met de meest extreme typen van de community-ledenclassificatie van De Valck, Van Bruggen en Wierenga. Aan de ene kant zijn daar de kernleden die een behoorlijke portie tijd en energie steken in verschillende aspecten van het virtuele gemeenschapsleven, zoals het toevoegen van informatie en het aangaan en onderhouden van relaties met andere leden. Zij zijn de ultieme posters. Daartegenover staan de opportunisten die af en toe langskomen voor wat informatie, maar die nooit een ander zullen helpen door eigen ervaringen of adviezen toe te voegen. Dit zijn de ultieme lurkers. Echter, tussen deze extremen onderscheiden De Valck, Van Bruggen en Wierenga nog vier andere typen community-leden. Figuur B representeert hun classificatie.

De kernleden (1) bevinden zich in het midden. Deze leden bezoeken de SmulWeb-community dagelijks gedurende anderhalf uur en zij zijn zowel gericht op het ontrekken, het bijdragen, als het bediscussiëren van informatie. Door hun grote aantal bijdragen (en hun kennis!) over het centrale thema zijn dit de opinieliders van de community. Naast de kernleden zijn er nog drie typen leden die informatie toevoegen, namelijk de conversatieliefhebbers (2), de informatieliefhebbers (3) en de hobbyisten (4). Conversatieliefhebbers en informatieliefhebbers hebben dezelfde mate van participatie, namelijk drie à vier keer per week gedurende een half uur, maar de vorm van hun participatie is verschillend. Conversatieliefhebbers participeren vooral in de discussieforums en chatroom, terwijl informatieliefhebbers zich vooral richten op het toevoegen en ontrekken van informatie via de databases. De hobbyisten vormen weer een andere groep posters. Zij zijn, na de kernleden, het meest actief in de community met een bezoekfrequentie van vijf à zes keer per week en een gemiddelde bezoekduur van één uur. Echter, deze leden houden zich niet zo bezig met het uitwisselen van culinaire informatie. In plaats daarvan besteden ze veel tijd aan het mooi maken van hun persoonlijke ledenpagina's met illustraties en achtergrondmuziekjes en aan het rondsturen van gastenboekberichten met een hoog poëziealbumgehalte.

De functionalisten (5) en opportunisten (6) bevinden zich in de periferie van de community. Deze leden zijn meestal nog niet zo lang lid (net zoals de hobbyisten), zij voelen zich niet sociaal betrokken bij de community en zij voegen nauwelijks informatie toe. Alhoewel de functionalisten en opportunisten dus als lurkers bestempeld kunnen worden, is er wel een groot verschil tussen de beide typen community-leden. Functionaristen zijn gemotiveerd om hun culinaire expertise op te bouwen. Daartoe bezoeken zij de community wekelijks gedurende een half uur voor het opzoeken van zowel recepten, recensies, als artikelen. De opportunisten zijn slechts geïnteresseerd in het nu en dan opzoeken van een recept. Zij besteden niet langer dan vijftien minuten aan zo'n incidenteel bezoek.

De cirkelende lijn in Figuur B geeft de dynamische natuur weer van het lidmaatschap van een virtuele community dat wordt gekarakteriseerd door verschuivingen in focus en een zich uitbreidend patroon van participatie. De positie van de typen community-leden representeert min of meer de volgorde waarin de leden zich van de ene naar de andere rol ontwikkelen. Hoe dichter bij de kern, hoe groter de mate van participatie, hoe breder de variatie van ondernomen online activiteiten, hoe hoger de sociale betrokkenheid bij de community en hoe langer het lidmaatschap al geduurd heeft. Figuur C laat zien dat de zes typen community-leden gerelateerd kunnen worden aan de drie community-platformdoelen van SmulWeb: informatie, recreatie en relatie (zie para-

graaf 2.1). De opportunisten (6), functionalisten (5), en informatieliefhebbers (3) richten zich op informatie in de vorm van recepten, recensies en/of artikelen. De hobbyisten (4) richten zich op recreatie; zij maken technische kunstwerkjes van hun persoonlijke pagina's en sturen poëziealbumachtige berichten rond. De conversatieliefhebbers (2) richten zich op relaties door actief in gesprek te gaan met andere leden via de discussieforums en chatroom. Hierdoor ontstaan hechte banden tussen de leden, die in dit soort gesprekken vaak veel meer met elkaar delen dan alleen hun laatste geslaagde recept. De kernleden bevinden zich in de overlap van de drie oriëntatiesferen, wat betekent dat zij zich op alle drie community-doelen richten [2].

3.2. Strategisch management op basis van participatieprofielen

Net zoals de poster-lurker dichotomie is de classificatie van de zes typen community-leden gebaseerd op variabelen die geanalyseerd kunnen worden met behulp van een softwaresysteem dat automatisch het deelnamepatroon van de leden bijhoudt [3]. Op die manier kunnen community-managers dus regelmatig checken hoe de samenstelling van haar ledenbestand zich ontwikkelt. Ook kunnen ze adequaat op deze ontwikkelingen inspelen; als er bijvoorbeeld veel functionalisten en opportunisten zijn, dan moet er wellicht geïnvesteerd worden in een snellere zoekfunctie; en als het aantal hobbyisten toeneemt, dan moet de servercapaciteit berekend zijn op het opladen van grote muzieken beeldbestanden. Alleen maar bijhouden wie bijdragen levert en wie niet geeft onvoldoende inzicht in dergelijke community-dynamiek. Maar marketeers en marktonderzoekers kunnen ook nog op een andere manier profiteren van de classificatie van de zes typen community-leden ten opzichte van de poster-lurker dichotomie. Het stelt ze namelijk in staat om onderscheid te maken tussen opinieleiders en andere posters en tussen lurkers die echt geïnteresseerd zijn het uitbreiden van hun kennis en lurkers die hiertoe niet gemotiveerd zijn. Dit inzicht is belangrijk voor het bepalen van de juiste doelgroep voor marketingacties of marktonderzoek. Door bovendien rekening te houden met de verschillende oriëntatiesferen van de typen leden, kunnen marketeers communicatiestrategieën bedenken die inspelen op specifieke behoeften.

De kernleden zijn de experts van de community. Marketeers kunnen hun aandacht trekken met geavanceerde producten en innovatieve productontwikkelingen die iets toevoegen aan hun kennis en ervaring. Wanneer de kernleden enthousiast zijn over een product, dienst of bedrijf, dan zullen zij er graag over praten. Marketingacties gericht op deze kleine groep kunnen via hun mond-tot-mondreclame dus een veel groter bereik hebben. De conversatieliefhebbers, informatieliefhebbers en hobbyisten beschikken over het algemeen over minder expertise dan de kernleden. Dit betekent dat zij onvankekelijk zullen zijn voor informatie en aanbiedingen met betrekking tot minder geavanceerde producten en ervaringen. De conversatieliefhebbers zullen geneigd zijn interesse te tonen wanneer ze mee kunnen praten over een product, dienst of bedrijf. Om hen te bereiken, zouden marketeers dus een buzz-marketingactie op touw kunnen zetten door een nieuwtje bekend te maken in de discussieforums. Informatieliefhebbers kunnen voorzien worden van achtergrondinformatie over producten en diensten in de vorm van artikelen of een themapagina. De aandacht van de hobbyisten, tenslotte, kan getrokken worden door een online wedstrijd of opiniepeiling te organiseren. Dit sluit aan bij hun recreatieoriëntatie, terwijl het hen tegelijkertijd informeert over een product, dienst of bedrijf. Aangezien deze laatste groep leden zich niet zo bezig houdt met het toevoegen van recepten, recensies en artikelen en ook niet actief deelneemt in de discussieforums,

is hun mond-tot-mond reclamepotentieel kleiner vergeleken bij dat van de conversatieliefhebbers en informatieliefhebbers.

De functionalisten en opportunisten maken een niet te onderschatten deel uit van de community (tezamen vormen zij 50% van de SmulWeb-steekproef). Alhoewel zij nauwelijks bijdragen leveren aan de community, kan hun mond-tot-mond reclamepotentieel buiten de community aanzienlijk zijn. Immers, zij zijn het publiek van de communicatie tussen de andere leden. Door informatie van de community te onttrekken en haar buiten de community verder te verspreiden, vormen deze twee groepen een brug tussen de kleine klik van community-experts en het grotere publiek van alledaagse koks. De functionalisten zijn een interessantere doelgroep voor marketeers dan de opportunisten vanwege hun oprechte interesse in het onderwerp van de community. Als gevolg daarvan staat de groep functionalisten waarschijnlijk meer open voor informatie afkomstig van marketeers dan de groep opportunisten. Net zoals bij de informatieliefhebbers kan de aandacht van de functionalisten getrokken worden met feitelijke informatie. Vanwege hun korte bezoeken aan de community moet deze informatie echter wel op een opvallende, compacte manier gebracht worden.

4. NETNOGRAFIE

4.1. *Community-intelligence: netnografisch onderzoek*

Eén van de grote voordelen die virtuele communities marktonderzoekers te bieden hebben, is de mogelijkheid om onopvallend bij te houden wat de community-leden met elkaar bespreken. Op die manier kunnen marktonderzoekers veel te weten komen over hun consumptiegedrag en de processen die daaraan ten grondslag liggen. Julia en haar kompanen praten bijvoorbeeld dikwijls over hun liefde voor kokkerellen en de daaruit voortvloeiende afkeer van kant-en-klare maaltijden, die in hun mening weinig fantasierijk en smaakvol zijn. Een onderzoek van een voedingsmiddelenfabrikant onder een select gezelschap SmulWeb-leden leverde dan ook als resultaat op dat één van hun productontwikkelingen beter op de markt gebracht kon worden als maaltijdstarter dan als kant-en-klare maaltijd, zodat er ruimte blijft voor de creativiteit van de kok (dhr. Oostveen, oprichter SmulWeb, 4 juli 2003). Voor marktonderzoekers die overwegen om virtuele communities te gebruiken als onderzoekslocatie om inzicht te krijgen in consumentengedrag zijn twee zaken van belang. Wiens discours moet je analyseren? En hoe voer je zo'n onderzoek systematisch uit?

Het is duidelijk dat met het monitoren van community-bijdragen alleen de posters onderzocht kunnen worden. De classificatie van zes typen community-leden heeft aan het licht gebracht dat de posters zich op verschillende manieren manifesteren in de community. De meest inzichtelijke bijdragen zijn te vinden in de discussieforums, omdat de leden hierin aan elkaar uitleggen hoe ze zich gedragen en waarom. Dit betekent dus dat met een monitoringsstrategie gericht op deze discussieforums alleen de kernleden en conversatieliefhebbers in het onderzoek betrokken zijn. Zeker waar het gaat om het onderzoeken van nieuwe trends en opiniërende meningen zijn deze typen leden, en dan vooral de kernleden, zeer geschikt. Als de marktonderzoeker er echter opuit is meer te weten te komen over het gedrag van de grote massa of de achterlopers, dan moet hij zich meer richten op de functionalisten en opportunisten. Omdat deze

typen leden geen bijdragen leveren, moeten zij onderzocht worden met traditionele methoden zoals enquêtes of interviews.

De manier waarop de discussieforums van virtuele communities systematisch onderzocht kunnen worden, is uitgebreid beschreven door Kozinets (2002). Zijn methode is gebaseerd op kwalitatieve onderzoekstechnieken afkomstig uit de culturele antropologie en aangepast aan de eisen van een virtuele onderzoekslocatie. De onderzoeksmethode heet daarom *netnografie*. Kozinets' richtlijnen omvatten vier hoofdpunten: (1) entree, (2) datacollectie, (3) analyse en interpretatie en (4) onderzoeksethiek. We zullen elk van deze hoofdpunten kort bespreken.

De eerste stap in een netnografisch onderzoek is uiteraard het kiezen van één of meerdere discussieforums die geanalyseerd zullen worden. Deze selectie hangt in de eerste plaats af van het feit of het onderwerp waarnaar men onderzoek wil doen besproken wordt. Daarnaast is het van belang na te gaan hoe uitgebreid het onderwerp besproken wordt en door hoeveel leden. Interessante inzichten kunnen al verkregen worden door een inhoudsanalyse van een discussie tussen slechts een handvol community-leden. Echter, in het algemeen geldt dat een forum met veel bijdragen, afkomstig van een grote groep discussianten, informatie oplevert die beter generaliseerbaar is. Voordat overgegaan wordt tot datacollectie moet de onderzoeker zich bekend maken met de community, dat wil zeggen dat hij zoveel mogelijk informatie moet verzamelen over de community, de forums en de individuele deelnemers. Kozinets spreekt in dit verband over *prolonged engagement*. Het is noodzakelijk langere tijd bij een community betrokken te zijn om de toon en aard van de discussies beter te kunnen duiden.

De datacollectie, analyse en interpretatie vinden vaak gelijktijdig plaats. Met in het achterhoofd de onderzoeksvraag maakt de onderzoeker een eerste selectie van de te analyseren discussies op basis van de titel die de discussiestarter gekozen heeft. Deze titel moet raken aan het onderwerp van onderzoek. Vaak zal een geselecteerde discussie uiteindelijk over iets heel anders gaan en ook komt het onderwerp van onderzoek aan bod in discussies die er op het eerste gezicht niets mee te maken lijken te hebben. Daarom is het noodzakelijk langere tijd bij de community betrokken te zijn, zodat men een goed gevoel ontwikkeld heeft over waar de meest interessante bijdragen te vinden zijn. Deze eerste selectie is het uitgangspunt van de data-analyse, maar zij zal steeds aangevuld worden met data verkregen uit nieuw geselecteerde discussies om te kijken of de getrokken conclusies standhouden of verworpen moeten worden.

Het analyseren van de data is een hermeneutisch proces waarbij de overeenkomsten en verschillen tussen complete discussies, delen van discussies en individuele bijdragen in kaart worden gebracht. Op die manier kunnen er thema's en categorieën in de data worden ontdekt. De data-analyse wordt afgewisseld met nieuwe ronden datacollectie om te kijken of de thema's hout snijden. De interpretatie van de data kan bevorderd worden door met meerdere onderzoekers samen te werken, zodat verschillende visies aan een nader onderzoek onderworpen kunnen worden. Voor een juiste interpretatie is het verder van cruciaal belang dat de community-leden zélf om een reactie op de gevonden resultaten gevraagd wordt. Daartoe is het noodzakelijk om de netnografie in z'n geheel aan de community-leden voor te leggen en hen om commentaar, aanvullingen en verbeteringen te vragen.

Kozinets' richtlijnen gaan tenslotte in op de ethische aspecten van netnografisch onderzoek. De grote vraag die academici zich op dit moment stellen is: kunnen publieke bijdragen aan het Internet beschouwd worden als vrij te exploreren onderzoeksmateriaal of moet rekening gehouden worden met het feit dat de schrijver nooit bedoeld heeft dat

zijn bijdragen gelezen en zelfs geanalyseerd worden door anderen dan zijn mede-community-leden? Academics zijn het niet eens over het antwoord op deze vraag. Kozinets roept daarom op een conservatieve houding aan te nemen en leden ALTIJD om toestemming te vragen voordat men bijdragen gebruikt voor onderzoek en ervoor te zorgen dat leden anoniem blijven. De internationale Association of Internet Researchers (AoIR) neemt een andere positie in. Zij stelt dat wanneer er duidelijk sprake is van een publiek karakter van de forums (wanneer dat bijvoorbeeld uiteengezet is in de statuten), het recht van de leden op privacy, vertrouwelijkheid en toestemming voor gebruik van hun bijdragen afneemt (Ess & AoIR 2002). Onderzoekers zullen zelf een afweging moeten maken of het onderzoek er baat bij heeft als ze de community-leden over een onderzoek informeren voorafgaand of na afloop van de datacollectie. In beide gevallen moeten zij volledig inzicht geven in hun affiliaties, het onderzoeksdoel, de onderzoeksprocedure en de te volgen, of de gevolgde, onderzoeksethiek.

4.2. *Netnografie nader belicht*

De richtlijnen van Kozinets zijn het uitgangspunt geweest voor een grootschalig netnografisch onderzoek in de discussieforums van SmulWeb in het kader van het promotieonderzoek van De Valck (2005). Het onderzoek richtte zich op de groep kernleden die een centrale rol spelen in de informatie-uitwisseling tussen de leden. Om het proces van interpersoonlijke beïnvloeding in virtuele communities inhoudelijk beter te begrijpen heeft zij onderzocht waarover en hoe de kernleden met elkaar communiceren in de discussieforums. De resultaten van dit onderzoek zijn niet alleen interessant voor het SmulWeb-management en voor marketeers, beleidsmakers en anderen werkzaam in de culinaire sector. De inzichten in de dynamiek tussen de discussianten zijn relevant voor community-managers, marketeers en onderzoekers in het algemeen. Daarom besluiten we dit artikel met een bespreking van het onderzoek en de inzichten die het heeft opgeleverd.

Voor aanvang van de datacollectie heeft De Valck zich gedurende een periode van drie jaar bekend gemaakt met de gebruiken en gewoonten van de community en haar deelnemers door rond te surfen op de site, bijdragen te lezen, interviews te houden met leden en management en door deel te nemen aan zogenaamde SmulWeb-dagen waarop de leden elkaar in het echt ontmoeten. Na deze inwerkperiode is begonnen met de datacollectie. Als eerste zijn alle forumonderwerpen van 2003 geïnventariseerd. Hieruit is een eerste selectie van onderwerpen gemaakt. Criteria hierbij waren dat het onderwerp regelmatig moest terugkeren, dat het door meerdere personen te berde werd gebracht en dat het voldoende reacties uitlokte. Gedurende de collectie en analyse van data spitste het onderzoek zich meer en meer toe op het discours van de leden omtrent koken en eten. Het uiteindelijke volume van de verzamelde data is 3163 forumbijdragen die afkomstig zijn van 82 discussianten.

Om de interpretatie van de data te bevorderen is van al deze 82 discussianten een e-profiel gemaakt, dat wil zeggen dat hun persoonlijke pagina's binnen SmulWeb zijn geanalyseerd om meer inzicht te krijgen in hun culinaire expertiseniveau, hun relaties met andere leden, hun zelfconcept en hun persoonlijke gegevens. De analyse en interpretatie van de forumdiscussies zijn als artikel aan de database van SmulWeb toegevoegd en alle leden zijn uitgenodigd commentaar te geven. Zestien leden hebben aan deze oproep gehoor gegeven; alle reacties waren positief en bevestigden de interpreta-

tie. Gedurende de hele onderzoeksperiode heeft De Valck een persoonlijke pagina binnen SmulWeb bijgehouden waarop achtergrondinformatie over het onderzoek en haarzelf gevonden kon worden. De privacy van de leden is gewaarborgd door pseudonimen toe te passen voor alle leden van wie een quote is gebruikt.

De netnografie heeft in de eerste plaats duidelijk gemaakt dat alhoewel de communityleden een interesse in hetzelfde onderwerp delen hun meningen en consumptiegedrag uiteenlopen. Dit wordt veroorzaakt door het feit dat er aanzienlijke verschillen zijn tussen de leden wat betreft leeftijd, nationaliteit, opleiding, inkomen, woonlocatie, gezinssituatie, beroepssituatie en levensstijl. Anders dan bij traditionele referentiegroepen, waar geldt 'soort zoekt soort', is de virtuele community bij uitstek de plaats waar mensen samenkomen die elkaar in het echte leven niet zo gauw zouden kennen vanwege geografische, demografische, economische en sociale barrières. Deze verschillen bevorderen het leereffect tussen de leden, maar ze zorgen ook voor onderlinge spanningen. In de tweede plaats heeft de analyse vier interactiedynamieken binnen de discussieforums in kaart gebracht die het meest relevant zijn voor marktonderzoek; leden nemen deel aan discussies om (1) kennis te delen, (2) te onderhandelen over normen, (3) uitdrukking te geven aan tegenovergestelde waarden en (4) overeenkomsten te vieren.

De verschillen tussen de discussianten geven aanleiding tot het delen van kennis, bijvoorbeeld tussen leden van verschillende leeftijdsgroepen of nationaliteiten. Om elkaar van hun expertise te overtuigen beroepen de leden zich op allerlei autoriteiten zoals de media en wetenschappelijk onderzoek. Deze bronnen verliezen echter snel in validiteit zodra een lid over een 'natuurlijke' autoriteit beschikt; als hij bijvoorbeeld in een bepaalde regio opgegroeid is en dus precies weet welke ingrediënten er in een streekgerecht horen. Dit onderstreept het voordeel dat marketeers kunnen hebben wanneer zij natuurlijke experts in de community faciliteren mond-tot-mondreclame te maken.

De verschillen tussen de discussianten leiden ook tot het onderhandelen over normen en waarden met betrekking tot de consumptieactiviteiten die centraal staan in de community. Deze discussies zijn verfijnd en uitvoerig. Ze bevatten gedetailleerde beschrijvingen van gedrag en de onderliggende motivaties. Een tactiek die de leden hanteren om uiting te geven aan een door hen gehanteerde norm is het vertellen van verhalen over anderzins gedrag dat afwijkt van deze norm. Deze verhalen, bijvoorbeeld over het prototype afgrijselijke schoonmoeder, zouden goed kunnen dienen als uitgangspunt voor een reclamecampagne. Als men over z'n eigen afwijkende gedrag spreekt dan wordt dat gedrag altijd in een context geplaatst. Opnieuw betekent dit een bron van rijke informatie voor de marketeer die op zoek is naar een juiste positionering of *USP*. De netnografie heeft bijvoorbeeld precies aan het licht gebracht welk soort diepvriesvoedsel onder welke omstandigheden door de discussianten, die vers de norm vinden, geoorloofd wordt.

Gedurende deze discussies bereiken de leden soms een gedeelde standaard, maar in andere gevallen blijken de tegenstellingen te groot waardoor het debat polariseert. Om ervoor te zorgen dat het gemeenschapsgevoel hierdoor niet ondermijnd wordt, communiceren de forumleden daarnaast veelvuldig over dat wat hen bindt. Ze bespreken herkenbare tradities en rituelen en ze delen geheime passies met betrekking tot de centrale consumptieactiviteiten. Ook het anonieme karakter van de virtuele community werkt dit soort 'identificatieconversaties' in de hand. De leden geven zich makkelijker bloot aan een onzichtbaar publiek dat dikwijls zeer begripvol is omdat zij de slechte gewoonten en geheime passies delen. Opnieuw leveren deze discussies interessante inzichten op in consumentengedrag en de afwegingen die daarbij gemaakt worden. De forumle-

den bespreken bijvoorbeeld met elkaar wat te doen tegen nachtelijke koelkastplunderingen van familieleden. ‘Verstoppen’ is het devies van sommigen. Anderen geven juist aan dat voedsel dat verstopt is een nog grotere aantrekkingskracht op lekkerbekken heeft. Welke marketeer is de eerste die met een koelkast-schutkleurverpakking op de markt komt?

5. CONCLUSIE

De hype rond virtuele communities is voorbij. Dat is waarschijnlijk terecht. Virtuele communities zijn niet zonder meer succesvol en nuttig. Het vereist inspanning om ze te creëren, te onderhouden, te exploiteren en om er informatie aan te onttrekken. Dit artikel heeft in vogelvlucht een aantal nieuwe inzichten besproken op het gebied van platformmanagement, marketing op basis van participatiegedrag en netnografisch onderzoek. Door de juiste functionaliteiten aan de community toe te voegen en door het gebruik ervan te stimuleren kunnen community-managers het aantal bijdragen en de invloed die ervan uitgaat aanzienlijk vergroten. Door rekening te houden met de manier waarop leden bijdragen leveren en/of onttrekken aan de community (via databases, discussieforums, chatrooms, etc.), en door te kijken naar hoe vaak en hoe lang ze dat doen, kunnen marketeers op een gerichte en toegespitste manier marketingacties in de community uitvoeren. Door het toepassen van de netnografische onderzoeksmethode, tenslotte, kunnen marktonderzoekers de rijkdom aan getuigenissen, beschrijvingen, verhandelingen en emoties die in communities gedeeld worden systematisch onderzoeken en duiden.

Als de juiste inspanning geleverd wordt dan vormen virtuele communities waardevolle marketinginstrumenten en locaties voor marktonderzoek. Zo is Adobe erin geslaagd enthousiaste gebruikers van hun producten aan zich te binden in een forum op hun website. Adobes productontwikkelaars participeren actief in dit forum om innovatieve ideeën op te doen en uit te testen. Een ander voorbeeld is Nokia, dat haar traditionele marktonderzoek deels vervangen heeft met een speciaal team dat stelselmatig het Internet afstruint op zoek naar conversaties tussen consumenten over mobiele telefonie in het algemeen en hun producten in het bijzonder. Op die manier kunnen ze doorlopend een vinger aan de pols houden wat betreft marktontwikkelingen. Bij al deze inspanningen moet uiteraard wel rekening gehouden worden met het feit dat niet alle consumenten zich met virtuele communities inlaten. Andere marketingkanalen en onderzoeksmethoden blijven daarom belangrijk naast community-marketing en netnografie.

Gezien de hernieuwde belangstelling voor virtuele communities in het bedrijfsleven is het voor marktonderzoekers van cruciaal belang zich te bekwamen in de netnografische onderzoeksmethode. In een goedlopende community kan het aantal bijdragen enorm oplopen. Zonder een gedegen monitoringsstrategie en kennis van de interactiedynamieken tussen de leden verdrinkt men al gauw in grote hoeveelheden data of ziet men zich geconfronteerd met slechts nietszeggende eenregelige bijdragen van het type “ik vind dat product helemaal niets”. In plaats van een opdracht tot het monitoren van een producentensite af te wachten, zouden marktonderzoekers er goed aan doen ook zelf virtuele communities op te starten. Op deze wijze kunnen zij waardevolle panels creëren waarin ervaringen en reacties langdurig in de tijd gevolgd kunnen worden.

Natuurlijk valt er nog veel meer aan de weet te komen over virtuele communities. We verwijzen graag naar de complete artikelen en onderzoeken die we hier aangehaald

hebben [4]. Maar virtuele communities zijn ook een fenomeen dat gekend kan worden door het aan den lijve te ondervinden. Neem dus eens een kijkje bij SmulWeb of een andere community. We verzekeren u: er gaat een wereld voor u open! Een virtuele wereld, maar wel met verstrekkende mogelijkheden met een zeer reële impact.

NOTEN

1. Julia is een pseudoniem.
2. De classificatie van zes typen community-leden is gebaseerd op een enquête uitgevoerd in één virtuele community. De classificatie hangt voor een deel af van de functionaliteiten waarover een community-platform beschikt. Dit betekent dat in communities met een ander soort community-platform, meer, minder, of andere typen leden gevonden kunnen worden. De classificatie van De Valck, Van Bruggen en Wierenga is dan ook niet normatief bedoeld, maar wil er op wijzen dat als community-managers en marketeers verder kijken dan alleen het onderscheid tussen posters en lurkers, zij veel meer gedifferentieerde participatiepatronen zullen ontdekken. Met dit inzicht kunnen zij hun voordeel doen wat betreft de ontwikkeling van management- en marketingstrategieën.
3. Bij gebrek aan toegang tot dergelijke logfile-data hebben De Valck, Van Bruggen en Wierenga hun classificatie gebaseerd op een enquête waarin de leden naar hun participatiegedrag gevraagd is. Het voordeel van deze methode was dat de leden ook gevraagd kon worden naar andere aspecten van hun lidmaatschap en hun persoonlijke karakteristieken. Het voordeel van een automatisch *tracking* systeem is dat het de deelnamepatronen meer accuraat en gedetailleerder in kaart kan brengen.
4. Het proefschrift *Virtual Communities of Consumption; Networks of Consumer Knowledge and Companionship* (De Valck 2005) is online beschikbaar op de ERIM website via de volgende link: <http://hdl.handle.net/1765/6663>. Ook kunt u een paperbackversie bestellen door 20 euro over te maken op girorekening 1198965 t.n.v. K. de Valck te Parijs onder vermelding van uw naam en adresgegevens. Het proefschrift zal u dan zo spoedig mogelijk toegestuurd worden.

LITERATUUR

- Bagozzi, R.P. and U.M. Dholakia, 2002. Intentional Social Action in Virtual Communities. *Journal of Interactive Marketing* 16 (2), 2-21.
- Brown, J.J., A. Tilton and D.M. Woodside, 2002. The Case of Online Communities. *McKinsey Quarterly* (1).
- Chevalier, J. and D. Mayzlin, 2003. The Effect of Word-of-Mouth on Sales: Online Book Reviews. *Working Paper Yale School of Management*, New Haven, CT.
- Dellarocas, C., N. Farag Award and X. Zhang, 2004. Exploring the Value of Online Reviews to Organizations: Implications for Revenue Forecasting and Planning. Paper Presentation *Marketing Science Conference*, Rotterdam, Erasmus Universiteit Rotterdam.
- De Valck, K., 2005. *Virtual Communities; Networks of Consumer Knowledge and Companionship*. ERIM PhD Series in Management 50, RSM Erasmus University. Online beschikbaar via <http://hdl.handle.net/17651/>
- De Valck, K., G.H. van Bruggen and B. Wierenga, 2003. Using Virtual Communities as Reference Groups: Antecedents and Consequences in the Context of Consumer Decision-Making. *Proceedings of the 32nd European Marketing Conference*, Department of Marketing, University of Strathclyde, Glasgow, Scotland.
- De Valck, K., G. H. van Bruggen and B. Wierenga, 2005. Patterns of Participation in Virtual Communities of Consumption; a Member Typology. *Proceedings of the 34th European Mar-*

- keting Conference, Universita Bocconi, Milan, Italy. And in: *Proceedings of the 32nd International Research Seminar in Marketing*, La Londe les Maures, France.
- Dholakia, U.M., R.P. Bagozzi and L. Klein Pearo, 2004. A Social Influence Model of Consumer Participation in Network- and Small-Group-Based Virtual Communities. *International Journal of Research in Marketing* 21, 241-263.
- Ess, Ch. and AoIR Ethics Working Committee, 2002. *Ethical Decision-Making and Internet Research: Recommendations from the Association of Internet Research Ethics Working Committee*. Online beschikbaar via <http://www.aoir.org/reports/ethics.pdf>
- Godes, D. and D. Mayzlin, 2004. Using Online Conversations to Study Word-of-Mouth Communication. *Marketing Science* 23 (4), 545-560.
- Hagel III, J., and A.G. Armstrong, 1997. *Net Gain: Expanding Markets Through Virtual Communities*. Boston, MA, Harvard Business School Press.
- Hauben, M. and R. Hauben, 1999. *On the History and Impact of Usenet and the Internet*. Los Alamitos, CA, IEEE Computer Society Press.
- Hennig-Thurau, T., K.P. Gwinner, G. Walsh, and D.D. Gremler, 2004. Electronic Word-of-Mouth Via Consumer-Opinion Platforms: What Motivates Consumers to Articulate Themselves on the Internet? *Journal of Interactive Marketing* 18 (1), 38-52.
- Jolink, D., 2000. *Virtual Communities: Digitale Doelmarkten voor E-Commerce*. Groningen, Gopher Publishers.
- Kozinets, R.V., 2002. The Field Behind The Screen: Using Netnography For Marketing Research in Online Communities. *Journal of Marketing Research* 39 (February), 61-72.
- McWilliam, G., 2000. Building Stronger Brands through Online Communities. *Sloan Management Review* (Spring).
- Okleshen, C. and S. Grossbart, 1998. Usenet Groups, Virtual Community and Consumer Behavior. In: J.W. Alba and J.W. Hutchinson (eds.), *Advances in Consumer Research*. Provo, UT: Association for Consumer Research: 275-282.
- Oostveen, R., 2001. *Virtuele Gemeenschappen: de Killer App voor Online Marketing*. Amsterdam, Pearson Education.
- Rheingold, H., 1993. *The Virtual Community: Homesteading on the Electronic Frontier*. New York, NJ, HarperCollins.
- Williams, R. L. and J. Cothrel, 2000. Four Smart Ways to Run Online Communities. *Sloan Management Review* (Summer).

3. Loyaliteitsprogramma's in de Nederlandse detailhandel: Adoptie en effectiviteit

J. LEENHEER*

SAMENVATTING

Loyaliteitsprogramma's kunnen een goed instrument zijn om klantloyaliteit en klantkennis te verbeteren in de detailhandel. Loyaliteitsprogramma's worden vooral ingezet door klantgeoriënteerde ondernemingen in branches waarin consumenten met enige regelmaat aankopen maken en het voor bedrijven moeilijk is om onderscheidend te zijn met het productassortiment. Bedrijven lijken technologische uitdagingen onvoldoende mee te wegen bij hun adoptiebeslissing, waardoor de analyse van klantenkaartdata vaak tegenvalt. Uit analyses in de supermarktsector blijkt dat met loyaliteitsprogramma's de gedragsloyaliteit van deelnemers (share-of-wallet) is te verhogen. Wel wordt de effectiviteit gemakkelijk overschat als geen rekening wordt gehouden met zelfselectie van deelnemers of tweezijdige causaliteit. Toch blijkt per saldo de netto-omzet toe te nemen als gevolg van deelname aan een loyaliteitsprogramma. De effectiviteit van loyaliteitsprogramma's kan worden verbeterd door een optimale inzet van het design van het programma. Zo blijkt dat beloningen die klanten ertoe prikkelen nieuwe aankopen te doen (bijv. winkelvouchers) behoorlijk kostendrukkend kunnen zijn.

1. INLEIDING

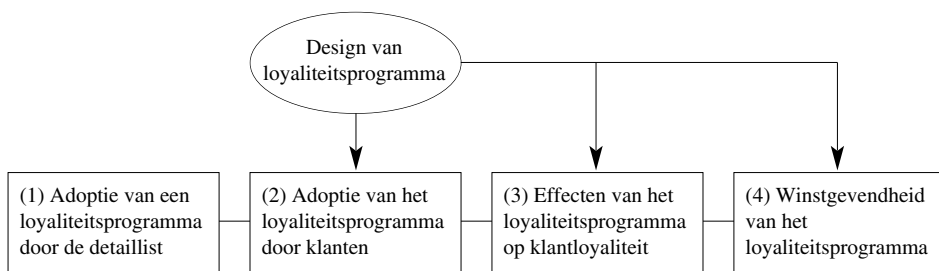
In de detailhandel wordt veelvuldig gebruik gemaakt van loyaliteitsprogramma's in de vorm van klantenkaarten. Een loyaliteitsprogramma geeft vorm aan een strategie van klantrelatie-management (customer relationship management). Klantrelatiemanagement stelt zich tot doel om doorlopend in contact te staan met klanten en met een persoonlijke behandeling van de meest waardevolle klanten, de effectiviteit van marketinginitiatieven en klantretentie te vergroten. Het creëren van klantloyaliteit is een belangrijk element van klantrelatiemanagement. Loyale klanten voelen zich verbonden met het bedrijf, hetgeen zich gedragsmatig uit in herhalingsaankopen en een hoge share-of-wallet [1]. Voor de meeste bedrijven geldt dat hun klanten onderling sterk verschillen in loyaliteit en winstgevendheid. Daarom is klantdifferentiatie een belangrijk element van klantrelatiemanagement.

In de detailhandel is het vormgeven van klantrelatiemanagement niet gemakkelijk

* Dit artikel is gebaseerd op het proefschrift van de auteur: "The Adoption and Effectiveness of Loyalty Programs in Retailing." CentER Dissertation Series no.135. Het proefschrift werd op 22 september 2004 verdedigd aan de Universiteit van Tilburg. Promotor is Prof. Dr. Tammo H.A. Bijmolt.

omdat klanten meestal anoniem blijven, er slechts beperkt persoonlijk contact bestaat met klanten en klantrelaties niet-contractueel zijn. Om klantspecifieke informatie te verwerven is het opbouwen van databases onontbeerlijk. Daarnaast is er behoefte aan relationele marketinginstrumenten om te komen tot een gedifferentieerde en persoonlijke klantbenadering. Relationele marketinginstrumenten kunnen relaties versterken door het verschaffen van economische of sociale waarde.

Een loyaliteitsprogramma (LP) is een relationeel marketinginstrument met een vrij generiek karakter. Het kan worden omschreven als een geïntegreerd systeem van marketingacties die tot doel hebben om de loyaliteit te versterken van de deelnemers aan het programma. Deze acties kunnen bestaan uit bijvoorbeeld prijskortingen, spaarprogramma's, loterijen, persoonlijke post, etc. De specifieke acties die onder het programma plaatsvinden, bepalen de exacte vormgeving (design) van het loyaliteitsprogramma. Meestal wordt gebruik gemaakt van klantenkaarten. Met het scannen van deze klantenkaarten wordt klantspecifieke informatie verkregen.



Figuur 1. De adoptie en effectiviteit van loyaliteitsprogramma's in de detailhandel.

De centrale vraag is welke rol loyaliteitsprogramma's kunnen spelen in de marketingmix van detaillisten. Daarbij dienen een aantal deelvragen te worden beantwoord (zie Figuur 1). Wat drijft detaillisten om een loyaliteitsprogramma te adopteren? Welke factoren zetten vervolgens klanten ertoe aan om het programma te adopteren, dat wil zeggen deelnemer te worden? De daaropvolgende vraag luidt hoe loyaliteitsprogramma's invloed hebben op de loyaliteit en de winstgevendheid van klantenkaarthouders. Belangrijk is ook hoe het design van het loyaliteitsprogramma van invloed is op de adoptiebeslissing van klanten en op de effectiviteit en winstgevendheid van het programma. Aan de hand van de bespreking van een drietal empirische studies zullen de bovenstaande vragen aan bod komen.

De eerste studie (Leenheer en Bijmolt 2005a) bestaat uit een vragenlijstonderzoek onder 180 detaillisten aangevuld met expertmetingen. Een onderzoek vanuit het perspectief van de aanbieder biedt inzicht in de loyaliteitsprogramma-adoptie van detaillisten (Figuur 1, blok (1)). Daarnaast kan het een meer compleet beeld geven van de effectiviteit van loyaliteitsprogramma's dan met consumentendata mogelijk is, bijvoorbeeld met betrekking tot de ontwikkeling van klantkennis (Figuur 1, blok (3)). In de tweede studie (Leenheer et al. 2004) staat het effect van loyaliteitsprogramma's op gedragsloyaliteit centraal (blok (3) + design). Voor de empirische analyse gebruiken we paneldata van 1909 Nederlandse huishoudens met hun koopgedrag in supermarkten en hun deelname aan de zeven loyaliteitsprogramma's in deze sector. De studie houdt expliciet rekening met de tweezijdige causaliteit tussen deelname aan een loyaliteits-

programma (blok (2)) en gedragsloyaliteit (blok (3)). De studie voert ook een winstgevendheidsanalyse uit (blok (4)). De derde studie (Leenheer en Bijmolt 2005b) bestudeert de effecten van een beloning gegeven als onderdeel van een loyaliteitsprogramma (blok (3) + design). Daarbij wordt gebruik gemaakt van de database van een specifiek loyaliteitsprogramma. De drie studies worden nu achtereenvolgens gepresenteerd (deel 2-4 van dit paper), waarna we afsluiten met conclusies (deel 5).

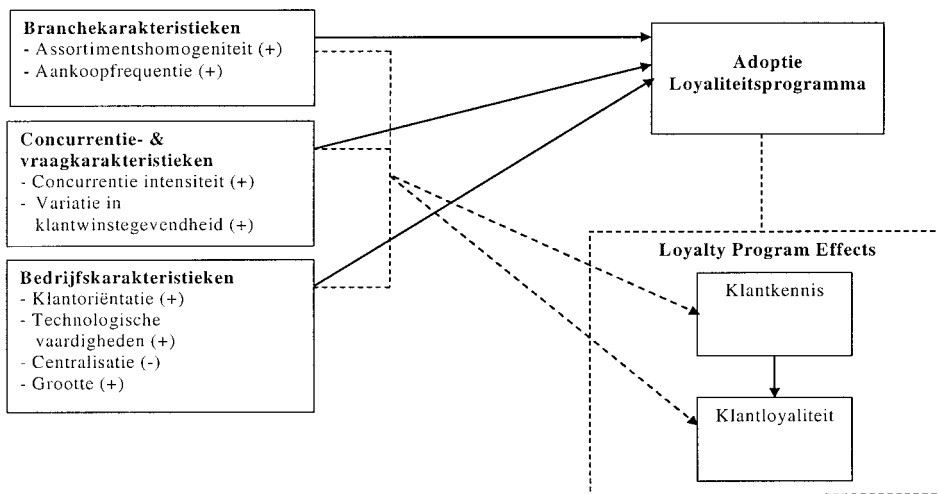
2. LOYALITEITSPROGRAMMA'S VANUIT HET PERSPECTIEF VAN DE DETAILLIST

2.1. Onderzoeksmodel

Hoewel loyaliteitsprogramma's een enorme vlucht hebben genomen in de Nederlandse detailhandel, spelen ze niet in iedere sector een even grote rol. En zelfs in sectoren waar loyaliteitsprogramma's wijdverspreid zijn, zijn er verschillende spelers die ze niet gebruiken. De bestaande literatuur (zie Leenheer (2004) voor een overzicht) heeft als gemene deler dat het onderzoek is gericht op de effecten van bestaande loyaliteitsprogramma's. Het is echter onduidelijk wat detaillisten er in de eerste plaats toe aanzet om loyaliteitsprogramma's te introduceren.

Een loyaliteitsprogramma kent twee belangrijke baten (positieve effecten) voor een detaillist: een verbetering van klantkennis en een verhoogde klantloyaliteit. Beter klantkennis kan worden verkregen door data-analyse van de (aankoop)gegevens van klantenkaarthouders. De verhoging van klantloyaliteit kan worden gerealiseerd met de marketingacties die als onderdeel van het programma worden uitgevoerd (zoals prijs-promoties, spaarprogramma's). Een detaillist zal bij het nemen van de adoptiebeslissing de kosten en baten van een loyaliteitsprogramma moeten inschatten. In navolging op Voss en Seiders (2003) nemen we aan dat een detaillist deze inschatting maakt op basis van drie categorieën van antecedenten: branchekarakteristieken, concurrentie- & vraagkarakteristieken en bedrijfskarakteristieken (zie Figuur 2). Op basis van bestaande literatuur op het gebied van loyaliteitsprogramma's, relatiemarketing en adopties zijn een aantal specifieke karakteristieken geselecteerd. Hun verwachte invloed zal nu kort worden toegelicht.

Assortimentshomogeniteit en aankoopfrequentie zijn de relevante branchekarakteristieken, weergegeven in Figuur 2. We verwachten allereerst dat de mate waarin de assortimenten van de verschillende detaillisten in een branche op elkaar lijken een positieve invloed heeft op loyaliteitsprogramma-adoptie. Heterogene assortimenten zijn beter toegesneden op specifieke consumentsegmenten en kunnen daardoor als relationele binding en natuurlijke switchbarrière fungeren. Bij homogene assortimenten zijn klanten echter meer geneigd van aanbieder te wisselen. Een loyaliteitsprogramma kan dan switchgedrag verminderen (Zhang et al. 2000). Daarnaast verwachten we dat de gemiddelde aankoopfrequentie in een branche loyaliteitsprogramma-adoptie positief beïnvloedt. Klanten zullen bij een lage aankoopfrequentie minder snel een klantenkaart adopteren en is hun loyaliteit naar het programma lager (Noble en Philips 2004; Yi en Jeon 2003). Het ontberen van frequent contact middels transacties bemoeilijkt zodoende het opbouwen van een klantrelatie (Dwyer et al. 1987). Bovendien zijn de baten van het verkrijgen van klantspecifieke data klein als klanten slechts infrequent aankopen doen.



Figuur 2. Onderzoeksmodel voor de adoptie van loyaliteitsprogramma's door detaillisten.

Daarnaast zijn concurrentie- & vraagkarakteristieken van belang, zie Figuur 2. Een hoge concurrentie-intensiteit heeft naar verwachting een positieve invloed op de innovativiteit van detaillisten en dus op de adoptie van loyaliteitsprogramma's. Ook zullen grote verschillen in de winstegevendheid van klanten de adoptie stimuleren. Een loyaliteitsprogramma is namelijk een geschikt instrument om tussen klanten te differentiëren: 1) tussen klantenkaarthouders en niet-klantenkaarthouders en 2) tussen klantenkaarthouders onderling.

Ten slotte zullen ook de bedrijfskarakteristieken een rol spelen bij de adoptiebeslissing volgens Figuur 2. In de eerste plaats verwachten we dat ondernemingen met een hoge mate van klantoriëntatie vaker een loyaliteitsprogramma adopteren, omdat deze oriëntatie nauw verwant is aan de mogelijkheden van een loyaliteitsprogramma. Voorts zullen ondernemingen die over goede technologische vaardigheden beschikken beter de mogelijkheden van de verkregen data kunnen benutten en dus een grotere verbetering van klantkennis kunnen realiseren. Centraal geleide ondernemingen zullen minder snel tot loyaliteitsprogramma-adoptie overgaan. In dergelijke ondernemingen is de betrokkenheid en toewijding van het personeel lager, hetgeen de succesvolle implementatie van innovaties belemmert (Damanpour 1991). Ten slotte zullen voor grote ondernemingen vanwege schaalvoordelen de kosten van een loyaliteitsprogramma lager zijn.

2.2. Onderzoeksmethode

Om het model te toetsen is een onderzoek uitgevoerd in 15 verschillende branches van de detailhandel, zoals supermarkten, warenhuizen, kledingwinkels en benzinestations. We maken daarbij gebruik van twee databronnen: 1) expertmetingen en 2) vragenlijsten onder detaillisten. We gebruiken experts voor het meten van de branchekarakteristieken, aangezien een correcte inschatting hiervan een breed perspectief vereist. Het expertpanel bestaat uit vijf academici en vijf praktijkmensen met een brede ervaring in de detailhandel. Iedere expert beoordeelt de mate waarin assortimenten van aanbieders verschillen en de gemiddelde aankoopfrequentie van een consument voor ieder van de

15 branches. In maart 2003 zijn 418 Nederlandse ketens in de detailhandel benaderd voor deelname aan het onderzoek. Daarbij verleenden 180 ketens hun medewerking (43.1% response), van wie 37% (67 ketens) een loyaliteitsprogramma heeft. De 180 deelnemers werd gevraagd naar concurrentie- en vraagkarakteristieken, bedrijfskarakteristieken en loyaliteitsprogramma-adoptie. De 67 detaillisten met een loyaliteitsprogramma is gevraagd de effecten van het programma op klantkennis en klantloyaliteit te beoordelen. Behalve loyaliteitsprogramma-adoptie (dichtome variabele), zijn alle variabelen met behulp van multi-item likertschalen gemeten, welke zo veel mogelijk zijn ontleend aan bestaande schalen.

2.3. Resultaten

We schatten het adoptiemodel met een random effects probit model. Dit model is geschikt om dichotome afhankelijke variabelen (0/1) te analyseren, in ons geval het niet/wel adopteren van een loyaliteitsprogramma. De interpretatie van beta-coëfficiënten is soortgelijk als voor regressieanalyse. Een positief (negatief) teken duidt daarbij op een positieve (negatieve) relatie van de onafhankelijke variabele met loyaliteitsprogramma-adoptie. Verder behoren detaillisten van nature tot een bepaalde branche. Dit wordt voornamelijk bepaald door het productassortiment en ondersteund door branchespecifieke organisaties, tijdschriften en congressen (Voss en Seiders 2003). Het model schat daarom een storingsterm op zowel brancheniveau als bedrijfsniveau.

Het model voorspelt 71.7% van de cases correct. Beide sectorkarakteristieken zijn van invloed op de loyaliteitsprogramma-adoptie. Zoals verwacht, stimuleren zowel assortimentshomogeniteit ($\beta = .452, p = .01$) als aankoopfrequentie ($\beta = .570, p = .05$) de loyaliteitsprogramma-adoptie van detaillisten. Verder vinden we dat concurrentie-intensiteit de adoptie positief beïnvloedt, hoewel het effect slechts indicatief is ($\beta = .118, p = .10$). Conform het onderzoeksmodel, heeft de diversiteit in klantwinstgevendheid een positief effect op adoptie dat ook indicatief is ($\beta = .111, p = .06$). Met betrekking tot de bedrijfskarakteristieken zien we een gemengd beeld. Klantoriëntatie heeft een positieve significante invloed op loyaliteitsprogramma-adoptie ($\beta = .290, p = .02$). Daarnaast lijken gecentraliseerde bedrijven minder snel geneigd een programma te adopteren ($\beta = -.100, p = .10$). We vinden echter dat zowel technologische vaardigheden als bedrijfsgrootte niet van invloed zijn op de adoptiebeslissing van detaillisten ($p > .10$).

De detaillisten met een loyaliteitsprogramma blijken deze vooral te gebruiken voor spaarprogramma's, directe prijspromoties en het sturen van persoonlijke post. Van andere elementen zoals loterijen, kredietfaciliteiten of speciale evenementen voor klantenkaarthouders wordt slechts sporadisch gebruik gemaakt. De meeste programma's maken geen gebruik van samenwerking met andere bedrijven (single-vendor programma's). Het merendeel van de detaillisten beoordeelt zijn programma positief, hoewel er duidelijke onderlinge verschillen bestaan. Om te bepalen in hoeverre deze verschillen samenhangen met branchekarakteristieken, concurrentie- en vraagkarakteristieken en bedrijfskarakteristieken, schatten we twee random-effects regressiemodellen met achtereenvolgens het effect op klantkennis en het effect op klantloyaliteit als de afhankelijke variabelen [2]. In het model voor klantloyaliteit nemen we ook klantkennis als onafhankelijke variabele op (zie Figuur 2).

Het model voor klantloyaliteit verklaart 25,5% van de variantie en het model voor klantkennis 18,6%. Zoals verwacht heeft klantkennis een positieve invloed op klant-

loyaliteit ($\beta = .265, p < .001$). Verder heeft technologische kennis een positieve invloed op klantkennis ($\beta = .295, p = .04$), terwijl dit antecedent niet werd meegewogen bij de adoptiebeslissing. Daarnaast heeft centralisatie een positief indicatief effect op klantkennis ($\beta = .453, p = .09$) en klantloyaliteit negatief ($\beta = -.313, p = .05$). De andere branche-, concurrentie-, vraag- en bedrijfskarakteristieken beïnvloeden de effectiviteit van het loyaliteitsprogramma niet. Dit duidt erop dat de detaillisten deze karakteristieken op een juiste manier meewegen bij hun adoptiebeslissing.

De gepresenteerde studie laat zien dat de beslissing van een detaillist om een loyaliteitsprogramma te adopteren voor een belangrijk deel wordt bepaald door sectorkarakteristieken. Sectoren die worden gekenmerkt door homogene assortimenten en hoge aankoopfrequenties zijn het meest geschikt voor loyaliteitsprogramma's. In het volgende hoofdstuk zullen we de effectiviteit van loyaliteitsprogramma's in een dergelijke sector (supermarktsector) analyseren. Daarnaast beïnvloeden concurrentie- en vraagkarakteristieken (concurrentie-intensiteit en variatie in klantwinstgevendheid) de baten van een loyaliteitsprogramma. Echter, niet voor iedere onderneming in een soortgelijke marktsituatie is een loyaliteitsprogramma even geschikt. Van belang zijn de oriëntatie (klantoriëntatie), bedrijfsstructuur (decentralisatie) en vaardigheden (technologische vaardigheden) van de onderneming. Vooral de laatste worden onvoldoende meegenomen in de adoptiebeslissing; hierdoor blijven verkregen loyaliteitsprogramma-data vaak onderbenut.

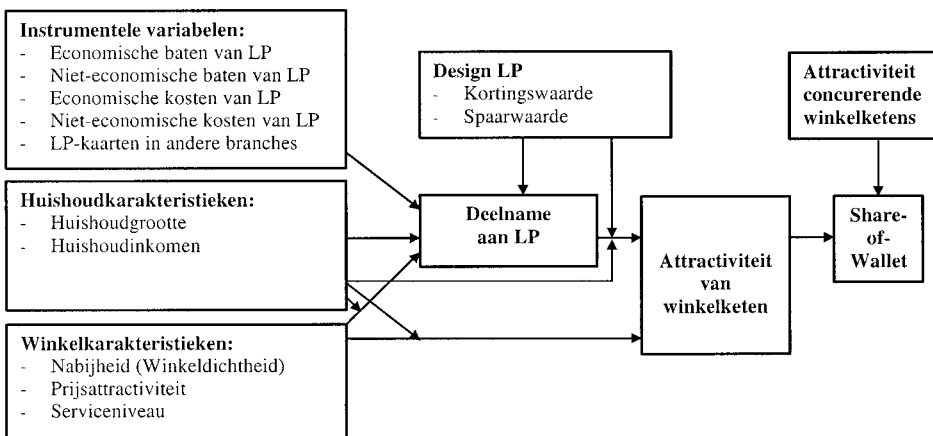
3. DE EFFECTIVITEIT VAN LOYALITEITSPROGRAMMA'S: HET ZELFSELECTIEPROBLEEM

3.1. Inleiding

Om de verschillen in effectiviteit van een groot aantal verschillende loyaliteitsprogramma's te analyseren, is een effectiviteitsrapportage door detaillisten geschikt (deel 2 van dit paper). Voor de exacte kwantificering van het effect van een programma op gedragsloyaliteit is het beter gebruik te maken van consumentendata. De vraag is in hoeverre de deelname aan een loyaliteitsprogramma (praktisch gezegd: het gebruik van een klantenkaart) bijdraagt aan klantloyaliteit. Vaak wordt daarbij geanalyseerd of deelnemers aan het programma loyaler gedrag vertonen dan klanten die niet deelnemen (bijv. Bolton et al. 2000; Mägi 2003). Er treedt daarbij echter een belangrijk causaliteits- of zelfselectieprobleem op. Men mag namelijk verwachten dat klanten die reeds loyaal zijn sterker geneigd zijn lid te worden van het loyaliteitsprogramma, aangezien zij meer baten (zoals beloningen) aan het programma zullen onttelen. Voor Nederlandse supermarkten met loyaliteitsprogramma's geldt dat hun klantenkaarthouders een gemiddelde share-of-wallet hebben van 36% en de niet-klantenkaarthouders een share-of-wallet van 7% (data van deze studie). Dit verschil moet deels worden toegeschreven aan zelfselectie van klanten en wellicht deels aan de effectiviteit van het programma. Het navolgende project presenteert een methodiek om met cross-sectionele data de effectiviteit van een loyaliteitsprogramma te kunnen bepalen en deze te scheiden van het zelfselectie-effect. Hoewel de noodzaak van een dergelijke analyse wordt onderkend (Bolton et al. 2004), is hierover nog geen bestaande literatuur. We maken voor deze studie gebruik van marktbrede consumentendata in de supermarktsector en bestuderen de census van 7 loyaliteitsprogramma's in deze sector.

3.2. Een model voor het meten van de effectiviteit van klantenkaarten

Figuur 3 geeft het onderzoeksmodel van de studie weer. De componenten zullen hieronder worden besproken. We kiezen voor share-of-wallet (SOW) als de centrale maat voor gedragsloyaliteit; deze maat is geschikt voor markten waarin consumenten regelmatig aankopen doen bij verschillende aanbieders, zoals de supermarktsector. Hoe de consument zijn/haar aankopen spreidt over de verschillende supermarktketens hangt af van de attractiviteit van de verschillende ketens. De share-of-wallet hangt positief samen met de attractiviteit van de eigen keten en negatief met de attractiviteit van de concurrerende aanbieders. De attractiviteit van een supermarkt wordt naar verwachting positief beïnvloed door de nabijheid, de prijsattractiviteit en het serviceniveau. Daarnaast stimuleert mogelijk ook deelname aan het loyaliteitsprogramma (vanaf nu: LP-deelname) de attractiviteit van de winkelketen.



Figuur 3. De effecten van loyaliteitsprogramma's op gedragsloyaliteit.

Zoals gezegd worden consumenten niet willekeurig deelnemer aan het programma. Bij hun adoptiebeslissing wegen zij de verwachte baten en kosten tegen elkaar af, welke deels samenhangen met de attractie van de keten zelf. Omdat loyale klanten de meeste baten ontlenden aan het programma, zijn klanten die al loyaal zijn het meest tot deelname geneigd. Bij analyse van het effect van LP-deelname op share-of-wallet treedt er econometrisch gezien een endogeniteitsprobleem op. Relevant daarbij is dat de variabele LP-deelname en de storingsterm in een dergelijk model positief zijn gecorreleerd. Dit zal ertoe leiden dat het effect wordt overschat. Het opnemen van controlevariabelen zoals nabijheid van de winkel, prijsattractiviteit en service verkleinen dit probleem maar lossen het niet op (Franses 2005). Een oplossing ligt in het volgen van een 2SLS-procedure (two-stage least squares). Dat houdt in dat we eerst een model schatten voor LP-deelname. In dit model dient een aantal instrumentele variabelen te worden opgenomen die exogeen zijn, dat wil zeggen variabelen die van invloed zijn op LP-deelname maar die geen directe invloed hebben op share-of-wallet. De voorspellingen op basis van dit model zullen als verklarende variabele in het SOW-model worden ingebracht (en niet de geobserveerde LP-deelname).

De gepercipieerde kosten en baten van LP-deelname zijn deels programmaspecifiek. Voor een deel verschillen consumenten echter ook op persoonsniveau in hun waardering voor loyaliteitsprogramma's. Deze attitudes ten aanzien van loyaliteitsprogramma's in het algemeen gebruiken we als exogene variabelen voor het LP-model. In de eerste plaats gaat het om de economische baten die loyaliteitsprogramma's bieden. Daarnaast kennen loyaliteitsprogramma's ook niet-economische baten, zoals het gevoel om als vaste klant te worden herkend. De economische kosten van LP-deelname zoals lidmaatschapsbijdrage zijn over het algemeen verwaarloosbaar. Wel kunnen er niet-economische kosten zijn, bijvoorbeeld de angst om privacy te verliezen met het weggeven van persoonsinformatie via de klantenkaart (Phelps et al. 2001). Ten slotte kan deelname aan loyaliteitsprogramma's in andere sectoren ook als exogene variabele worden opgenomen, aangezien zij een indicator zijn voor de attitude voor loyaliteitsprogramma's in het algemeen. Onze procedure vereist dat we tevens alle huishoud- en winkelkarakteristieken in het model voor LP-deelname opnemen, maar het betreft hier geen exogene variabelen.

Ten slotte nemen we ook design van het programma op in het onderzoeksmodel en wel op twee plaatsen. In de eerste plaats kan design de deelname aan het loyaliteitsprogramma beïnvloeden en ten tweede kan het design invloed hebben op de effecten van LP-deelname op SOW (moderator). De twee meest belangrijke designaspecten voor loyaliteitsprogramma's in de detailhandel zijn prijskortingen voor klantenkaarthouders en spaarprogramma's.

3.3. Data

Voor onze empirische analyse maken we gebruik van paneldata (GfK Panelservices Benelux) met het aankoopgedrag van 1909 huishoudens in Nederlandse supermarkten gedurende twee jaar (juli 1998-juni 2000). In de sector bestaan in die periode 7 verschillende loyaliteitsprogramma's. Zij verschaffen alle prijskortingen en een spaarprogramma, maar gebruiken die niet met dezelfde intensiteit. We definiëren kortingswaarde als het gemiddelde kortingspercentage verstrekt aan klantenkaarthouders op hun aankopen en spaarwaarde als percentage van het aankoopbedrag dat de consument in spaarbeloningen terug krijgt. Voor de meting van promotiewaarde maken we gebruik van metingen van de Consumentenbond onder een groot aantal klantenkaarthouders. Spaarwaarde achterhalen we via websites en managers, waarbij we veronderstellen dat klanten de punten ook daadwerkelijk inwisselen. Deze data ondersteunen de operationalisatie van de constructen in Figuur 3.

In het GfK-panel geven huishoudens jaarlijks via een vragenlijst aan welke klantenkaarten van supermarkten zij bezitten. Ongeveer 88% van de huishoudens heeft tenminste één klantenkaart. Veel huishoudens bezitten meer dan één klantenkaart; het gemiddeld aantal klantenkaarten per huishouden is 1.68. Op winkelniveau is de penetratie van klantenkaarten gemiddeld 64%. Voor de schatting van ons model gebruiken we alleen de eerste jaarlijkse periode en selecteren we alleen de 1307 huishoudens die hun klantenkaartportefeuille niet hebben gewijzigd gedurende de observatieperiode. De resterende data worden gebruikt voor modelvalidatie.

De huishoudens geven in de vragenlijst verder aan hoeveel klantenkaarten van benzinestations zij bezitten. Daarnaast geven zij via een serie stellingen aan wat hun houding is ten opzichte van klantenkaartprogramma's in het algemeen. Met factoranalyse extraheren we drie factoren die zijn te benoemen als economische baten van loyaliteitspro-

gramma's, niet-economische baten van loyaliteitsprogramma's en privacyangst. Ten slotte is de huishoudgrootte en het maandelijks inkomen bekend.

Voor de operationalisatie van de winkelkarakteristieken maken we gebruik van twee databronnen. De locaties van supermarktvestigingen ontleen we aan Elsevier Business Information. We definiëren winkeldichtheid als het aantal vestigingen van een supermarktketen gedeeld door het totaal aantal supermarktvestigingen in de provincie waarin een panelhuishouden woont. Voorts maken we gebruik van exitinterviews onder 8000 supermarktbezoekers die hun supermarkt beoordelen op acht dimensies. Met factoranalyse op de gemiddelde waardes per supermarkt extraheren we twee factoren: prijs-attractiviteit en service.

3.4. Methode

Voor het modelleren van share-of-wallet is het noodzakelijk een model op te stellen dat schattingen oplevert die altijd liggen tussen 0 en 1 en waarbij de som van de share-of-wallets per huishouden optelt tot 1. We kiezen daarom voor een attractiemodel (Hanssens et al. 2001). Het idee is dat de share-of-wallet (SOW) van een huishouden i in supermarkt s wordt bepaald door de relatieve attractie van de supermarkt (vergelijking 1). We specificeren daarbij de attractiefunctie als een multinomiaal logit model, waarbij de attractie exponentieel afhangt van een set van variabelen X_{jis} . De set X_{jis} bestaat uit winkelkarakteristieken en LP-deelname, waarbij huishoud- en designkarakteristieken als interacties worden opgenomen.

$$(1) \quad SOW_{is} = \frac{A_{is}}{\sum_{s=1}^S A_{is}} ; A_{is} = \exp. (X_{jis}\beta).$$

Met de methode van log-centering (Nakanishi en Cooper 1982), verkrijgen we een log-lineaire relatie die met regressieanalyse is te schatten. Een probleem bij dit model is echter dat share-of-wallet in een groot aantal gevallen gelijk is aan 0 (een huishouden bezoekt een bepaalde supermarkt gedurende het jaar geen enkele keer). Een ander probleem is dat het nu gepresenteerde model niet vrij is van het eerder besproken endogeneiteitsprobleem. Om aan beide zaken tegemoet te komen volgen we de procedure voor type-II Tobit modellen met endogene variabelen (Wooldridge 2002). Voor de technische details van dit model wordt de lezer verwezen naar Leenheer et al. (2004).

3.5. Resultaten

Allereerst schatten we een model voor LP-deelname. Dit model heeft een $R^2 = .148$ en alle vier instrumentele variabelen zijn significant. Economische baten van loyaliteitsprogramma's hebben een positieve invloed op LP-deelname ($\beta = .061$, $p < .001$), en hetzelfde geldt voor niet-economische baten ($\beta = .038$, $p < .001$). Privacyangst heeft een negatief effect ($\beta = .020$, $p = .006$), wat inhoudt dat privacy angst consument ervan kan weerhouden aan loyaliteitsprogramma's deel te nemen. Ten slotte hangt het aantal klantenkaarten van benzinstations positief samen met LP-deelname in supermarkten ($\beta = .074$, $p < .006$). We vinden tevens dat zowel spaarwaarde ($\beta = .070$, $p = .003$) als kortingswaarde ($\beta = .026$, $p = .001$) huishoudens stimuleren om aan een loyaliteitsprogramma deel te nemen. Het effect van spaarwaarde is significant groter dan dat van

kortingswaarde ($p < .01$). Ten slotte heeft prijsattractiviteit een negatief effect ($\beta = -.055, p < .01$) en de interactie tussen winkeldichtheid en huishoudinkomen een positief effect ($\beta = .089, p < .05$).

In de tweede stap schatten we het SOW-model. We vinden allereerst dat LP-deelname een positieve invloed heeft op supermarktattractie ($\beta = .479, p < .001$). Door deelname aan het loyaliteitsprogramma vindt een huishouden de supermarkt aantrekkelijker, wat vervolgens leidt tot een hogere share-of-wallet (zie vergelijking (1)). Er gaat geen significant modererend effect uit van kortingswaarde of spaarwaarde. We kunnen dus concluderen dat hogere beloningspercentages vooral LP-deelname stimuleren, maar geen effect hebben op klanten die al deelnemen aan het programma. Ook beide huishoudkarakteristieken modereren het LP-effect niet, hetgeen betekent dat het effect van LP-deelname op SOW niet wordt beïnvloed door het inkomen of de grootte van het huishouden. De winkelkarakteristieken winkeldichtheid ($\beta = 5.387, p < .01$) en prijsattractiviteit ($\beta = .355, p < .01$) leiden beide tot hogere share-of-wallets. Huishoudinkomen heeft een significant modererend effect op winkeldichtheid (+) en service (+). Huishoudens met hoge inkomens waarderen de nabijheid van winkels en goede service dus meer dan huishoudens met lage inkomens. Huishoudgrootte heeft een significant modererend effect op winkeldichtheid (?) en prijsattractiviteit (+). Grote huishoudens hechten minder aan de nabijheid van supermarkten maar meer aan een aantrekkelijke prijsstelling dan kleine huishoudens.

Tabel 1. Veranderingen in share-of-wallet en (netto) omzet per klant als gevolg van deelname aan loyaliteitsprogramma.

Supermarkt keten	Δ SOW*	Δ Omzet**	Δ Kosten**	Δ Netto Omzet**
LP1	.072	€ 430.20	€ 92.37	€ 337.83
LP2	.047	€ 272.67	€ 6.06	€ 266.61
LP3	.065	€ 368.92	€ 88.34	€ 280.58
LP4	.029	€ 176.11	€ 38.01	€ 138.10
LP5	.017	€ 109.93	€ 60.53	€ 49.40
LP6	.072	€ 402.73	€ 105.90	€ 296.83
LP7	.024	€ 155.82	€ 147.16	€ 8.66
Gemiddeld	.047	€ 273.77	€ 76.91	€ 196.86

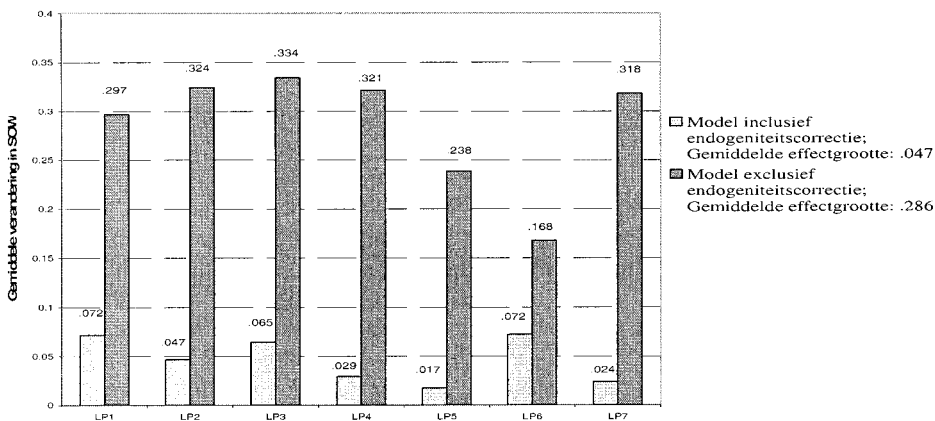
* Gemiddelde voorspelde verandering in SOW voor een klant van de supermarktketen

** Gemiddelde voorspelde verandering voor een klant van de supermarktketen in euro's per jaar

Doordat het SOW-model niet lineair is en betrekking heeft op attractie is de effectgrootte van LP-deelname niet meteen af te leiden (zie vergelijking (1)). Om hierin inzicht te verkrijgen berekenen we voor alle huishoudens in ieder van de zeven supermarktketens met LP het verschil in SOW tussen de situatie met en zonder LP-deelname volgens de modelvoorspelling. De resultaten zijn weergegeven in Tabel 1. De eerste kolom (Δ SOW) toont per supermarkt de gemiddelde verandering in SOW van een klant als gevolg van het loyaliteitsprogramma. Een loyaliteitsprogramma vergroot de share-of-wallet van een klant gemiddeld met 4.7 procentpunt, maar het effect varieert tussen 1.7 en 7.2 voor de verschillende loyaliteitsprogramma's. Vervolgens berekenen we het effect in termen van additionele omzet per klant per jaar (Δ Omzet). Na aftrek van de beloningskosten (de verleende kortings- en spaarwaarde) verkrijgen we het effect op de netto-omzet (Δ Netto Omzet). Alle zeven loyaliteitsprogramma's leiden tot positieve

netto-opbrengsten: de baten in termen van extra omzet zijn groter dan de kosten in termen van de beloningen die men weggeeft. Voor de loyaliteitsprogramma's met een klein effect op SOW (LP5 en LP7) is de toename in netto-omzet echter gering.

Ten slotte valideren we ons model door het te vergelijken met een naïef benchmarkmodel om vast te stellen of het inderdaad belangrijk is om met zelfselectie van loyaliteitsprogramma's rekening te houden. Het benchmarkmodel houdt geen rekening met zelfselectie, maar neemt rechtstreeks de geobserveerde LP-deelname als variabele in het SOW-model op. Het effect van LP-deelname is volgens dit model aanzienlijk hoger ($\beta_1 = 3.424, p < .001$). Figuur 4 toont de verschillen in de berekende effectgrootte voor het naïeve en ons complete model. Het naïeve model voorspelt een gemiddelde toename van 29 procentpunt, hetgeen ongeveer 6 keer zo veel is als ons model voorspelde.



Figuur 4. Veranderingen in share-of-wallet als gevolg van LP-deelname.

We vergelijken de actuele veranderingen in SOW van de consumenten die gedurende de observatieperiode een klantenkaart kregen of afstootten (de validatiesample) met de voorspelde verandering in SOW volgens beide modellen. We berekenen daarbij de MSE (mean squared error) en MAPE (mean absolute percentage error) (Van Dijk et al. 2004) en hoe lager de waarden hoe beter de modelvoorspelling is. Het complete model heeft een MSE = .014 en MAPE = 35.9, terwijl het benchmarkmodel een MSE = .254 en MAPE = 513.3 heeft. Het model met endogeniteitscorrectie voorspelt de veranderingen in SOW dus aanzienlijk beter. We concluderen dus dat het negeren van zelfselectie tot een behoorlijke overschatting van de effectiviteit van loyaliteitsprogramma's leidt.

In algemene zin toont deze studie aan dat loyaliteitsprogramma's een effectief instrument is om de share-of-wallet van deelnemers te vergroten. Dit geldt zelfs in een markt waar consumenten vaak verschillende klantenkaarten van concurrenten naast elkaar bezitten. Wel is het essentieel om bij de evaluatie van loyaliteitsprogramma's voor endogeniteit te corrigeren. Om consumenten tot LP-deelname te bewegen helpt het om een positieve basishouding ten aanzien van loyaliteitsprogramma's te creëren en aantrekkelijke prijs- en spaarbeloningen te geven. De beloningen die het programma geeft, worden in deze studie als kosten beschouwd. In de volgende studie laten we zien dat van beloningen ook een stimulerende werking kan uitgaan.

4. DE EFFECTIVITEIT VAN LOYALITEITSPROGRAMMA'S: DE POTENTIE VAN BELONINGEN

4.1. Inleiding

Spaarprogramma's zijn een krachtig instrument om loyaliteitsprikkelers te geven aan consumenten. Consumenten ontvangen spaarpunten als ze aankopen doen, maar kunnen deze pas verzilveren voor een beloning als ze een minimale spaardrempel hebben bereikt. Op deze manier worden consumenten aangezet tot herhalingsaankopen en grote transacties. Bestaande literatuur heeft aangetoond dat consumenten spaarpunten inderdaad meewegen in hun aankoopbeslissingen en soms zelfs zwaarder dan rationeel is (Hsee et al. 2003; Van Osselaer et al. 2004). Er is echter nauwelijks aandacht voor de effecten die spaarbeloningen hebben op consumentengedrag. De beloningen lijken vooral te worden beschouwd als een noodzakelijke kostenpost om consumenten tot sparen aan te zetten. De vraag is of dit terecht is. Ook van het verkrijgen van een beloning kan een stimulerende werking uitgaan die als opmaat geldt voor een nieuwe spaar-cyclus.

De derde studie focust op de effecten van spaarbeloningen. De centrale vraag is hoe spaarbeloningen aankoopgedrag beïnvloeden en hoe de effecten variëren over tijd en tussen consumenten. We zullen nu allereerst de specifieke empirische toepassing bespreken.

4.2. Data

We maken gebruik van data ontleend aan een loyaliteitsprogramma van een kledingde-taillist. Klanten kunnen kosteloos deelnemen aan het programma en ongeveer 80% van de omzet wordt door klantenkaarthouders gegenereerd. Het belangrijkste component van het loyaliteitsprogramma is het spaarsysteem. Een klant ontvangt een spaarpunt (met een waarde van € 0.25) voor iedere € 5 die zij/hij besteedt. Halfjaarlijks ontvangen klanten met 20 of meer spaarpunten een winkelvoucher met de waarde van de gespaarde punten. Het spaarsaldo wordt vervolgens op nul gezet. De winkelvoucher kan worden gebruikt op het gehele assortiment en is een half jaar geldig.

We beschikken over aankoopdata op consumentniveau gedurende bijna 3 jaar (2000-2002, week 36). De data bevatten tevens klantspecifieke informatie over het bezit en gebruik van winkelvouchers, de waarde van de vouchers en inschrijfdata. Tevens weten we op klantniveau in welke weken klanten direct mails en huis-aan-huis folders ontvangen. Voor onze analyse maken we gebruik van aankoopdata die precies vier spaar-cycli bestrijken (2000, week 35 – 2002, week 36); de voorgaande periode wordt gebruikt om een aantal variabelen te initialiseren. We selecteren de klantenkaarthouders die minstens één voucher hebben ontvangen gedurende de observatieperiode en minstens één aankoop hebben gedaan in de initialisatieperiode. De 45.996 geselecteerde klanten maken gemiddeld 4.16 transacties per jaar met een gemiddelde transactiegrootte van € 47.68.

Ongeveer de helft van de winkelvouchers wordt door klanten ingewisseld. De kans op verzilvering stijgt met de hoogte van het voucherbedrag en met het aantal direct mails dat de ontvanger in de inwisselperiode ontvangt. Verder is de kans op inwisselen aanzienlijk groter als de klant ook in de vorige periode een voucher inwisselde. De lidmaatschapsduur heeft een licht negatief effect. De vraag is of de voucher ook leidt tot

additionele aankopen. Voor de analyse hiervan maken we gebruik van een aselechte steekproef van 5000 klanten uit de totale geselecteerde set.

4.3. De effecten van een spaarbeloning

Om de invloed van de winkelvoucher te analyseren splitsen we de aankoopbeslissing van consumenten op in aankooptiming en transactiegrootte. Als een consument een winkelvoucher in bezit heeft, beïnvloedt dit mogelijk de timing van de volgende aankoop. Als de klant de voucher vervolgens bij een aankoop gebruikt, kan dit het aankoopbedrag beïnvloeden.

Voucherbezit verlaagt de prijs van een aankoop en kan zodoende leiden tot versnelde aankopen, switchgedrag, en toename van totale categoriebestedingen (Gupta 1988). In de laatste twee gevallen leidt dit tot extra aankopen bij de betreffende winkel, in het eerste geval worden de aankopen slechts vervroegd. Het feit dat de voucher als onderdeel van een spaarprogramma wordt uitgekeerd geeft het enkele speciale eigenschappen. Omdat consumenten expliciet voor de voucher hebben gespaard, kunnen zij het als hun eigen verdienste beschouwen. Enerzijds kunnen consumenten zich hierdoor minder schuldig voelen bij het doen van aankopen en zichzelf een beloning (extra aankoop) gunnen (Kivetz en Simonson 2002). Anderzijds kunnen consumenten die bewust moeite doen om de voucher te verwerven, ook meer berekenend gedrag vertonen waarbij zij de voucher gebruiken op een moment waarop ze sowieso een aankoop hadden gepland. In dat geval zal de aankoopbeslissing nauwelijks worden beïnvloed. Het is dus op voorhand niet zeker dat de voucher de timing van een aankoop stimuleert.

Hetzelfde geldt voor de grootte van de transactie waarbij de *voucher* wordt gebruikt. De voucher geeft namelijk geen prikkel om een grotere aankoop te doen, aangezien er sprake is van een vaste en niet van een percentuele verlaging van de prijs. Dergelijke promoties hebben een beperkte invloed op de transactiegrootte (Bawa 1996). Hoewel zelfbeloning de transactiegrootte zou kunnen stimuleren, kan berekenend gedrag zelfs tot verlaging van de transactiegrootte leiden (men probeert een aankoop te doen ter grootte van de voucher).

Voucherwaarde en de *verlooptijd* kunnen de relaties tussen voucherbezit en -gebruik en aankoopgedrag beïnvloeden. Uit onderzoek naar coupons blijkt dat deze vooral in de eerste weken na uitgifte tot additionele aankopen leiden (Inman en McAlister 1994). We gaan er ook voor vouchers vanuit dat mensen die de voucher niet snel gebruiken minder aangetrokken worden door het aanbod of berekenend gedrag vertonen. Literatuur over coupons toont aan dat er een positief effect uitgaat van de couponwaarde. Toch is het niet duidelijk dat hogere voucherwaardes leiden tot meer additionele aankopen, aangezien de grootste vouchers in het bezit zijn van de meest loyale klanten. Aangezien bestedingen een natuurlijk plafond kennen, leiden de vouchers bij deze loyale klanten wellicht juist niet tot veel extra aankopen.

Nadat een voucher is gebruikt kunnen er twee *post-voucher effecten* optreden: 1) een effect op toekomstig aankoopgedrag en 2) verandering in de reactie op toekomstige vouchers. Indien de voucher consumentenaankopen heeft gestimuleerd, zou vouchergebruik kunnen worden gevolgd door een langere periode tot de daaropvolgende aankoop en een kleinere transactiegrootte (post-voucher dip). Bovenstaande bespreking laat zien dat de voucher velerlei effecten kan genereren. Theoretisch gezien is de richting van die effecten vaak niet eenduidig en empirisch onderzoek zal hierover uitsluitsel moeten geven.

4.4. Analyse

Voor de empirische analyse gebruiken we voor de timing van aankoop een discreet hazard rate model (zie Vermunt 1996) en voor transactiegrootte wordt een lineair regressiemodel gebruikt [3]. De hazard rate is de conditionele kans dat een aankoop plaatsvindt in een betreffende week. De parameterschattingen zijn weergegeven in Tabel 2. De parameters voor de vierwekelijkse seizoensdummies zijn vanwege ruimte-overwegingen weggelaten. Als controlevariabelen hebben we ook de tijd sinds de vorige aankoop opgenomen (tevens in logaritmische vorm). Op deze manier controleren we voor het natuurlijke aankoopverloop. Daarnaast nemen we het aankoopverleden mee. Voor aankooptiming nemen we de fractie weken waarin een aankoop heeft plaatsgevonden in het afgelopen half jaar op. In het model voor transactiegrootte nemen we de transactiegrootte van de vorige aankoop op. Op deze manier controleren we voor heterogeniteit tussen klanten in hun normale aankoopgedrag.

Tabel 2. De invloed van de spaarbeloning op aankooptiming en aankoopbedrag.

	Aankooptiming		Aankoopbedrag	
	Parameter	<i>t</i> -waarde	Parameter	<i>t</i> -waarde
Constante	-2.880	-92.01***	43.066	38.92***
Tijd sinds vorige aankoop	-.009	-12.13***	.006	.21
LN (tijd sinds vorige aankoop)	.034	3.89***	1.561	4.91***
Lagged aankooptiming	4.349	70.76***		
Lagged transactiegrootte			4.215	8.88***
Voucher bezit	-.070	-1.64		
Voucher gebruik			-6.279	-2.48**
Moderatoren:				
Waarde	.015	7.44***	.842	6.54***
Tijd sinds toezending	-.005	-3.69***	-.424	-3.17***
Vorige voucher gebruikt	.242	10.41***	8.451	5.60***
Vorige voucher ontvangen, ongebruikt	.119	4.39***	7.055	3.16***
Lidmaatschapsduur	.000	.05	.002	.11
Voucher gebruikt bij vorige aankoop	-.074	-4.03***	.094	19.43***
Lidmaatschapsduur	-.0003	-1.83*	-.010	-1.76*
Direct mail	.368	29.72***	-.284	-.50
Huis-aan-huis folder	.372	25.29***	2.409	3.39***

Significantieniveaus: *** $p < .01$; ** $p < .05$, * $p < .10$ (tweezijdige testen).

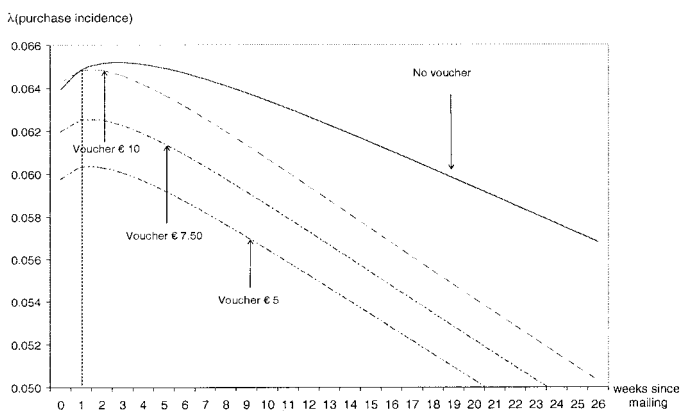
Het hoofdeffect voor voucherbezit is niet-significant ($\beta = .070$, $p > .10$) en voor vouchergebruik negatief ($\beta = -6.279$, $p < .05$). Dit hoofdeffect representeert echter het effect van een voucher met minimale waarde (?.5) in de eerste week na uitgifte voor een klant die de vorige voucher niet ontving. We nemen enkele moderators op waaruit blijkt dat de effectiviteit behoorlijk varieert tussen vouchers en klanten. Voor zowel aankooptiming ($\beta = .015$, $p < .01$) als transactiegrootte ($\beta = .842$, $p < .01$) geldt dat het effect van de voucher toeneemt naarmate de waarde stijgt. We nemen verder voucher-historie op met een categorische variabele bestaande uit drie niveaus: vorige voucher

niet ontvangen (basisniveau), vorige voucher ontvangen maar niet gebruikt en vorige voucher ontvangen en gebruikt. De vouchers blijken aanzienlijke effectiever te zijn als de klant ook de vorige voucher ontving (aankooptiming: ($\beta = .119, p < .01$); transactiegrootte: ($\beta = 7.055, p < .01$)) en zelfs nog effectiever als deze voucher ook nog werd ingewisseld (aankooptiming: ($\beta = .242, p < .01$); transactiegrootte: ($\beta = 8.451, p < .01$)). Verder blijkt de effectiviteit van de vouchers af te nemen gedurende de looptijd (aankooptiming: ($\beta = .005, p < .01$); transactiegrootte: ($\beta = -.424, p < .01$)). Dit houdt in dat een effectieve voucher na verloop van tijd ineffectief kan worden. Na het gebruik van een voucher wordt een aankoopdip waargenomen in de timing van de erop volgende aankoop ($\beta = -.074, p < .01$). De voucher stimuleert consumenten om sneller een aankoop te doen, maar na deze versnelling treedt een vertraging op waarbij de consument gedurende een langere tijd geen aankoop doet. Dergelijke dips zijn al vaak geconstateerd voor prijspromoties (post-promotie dips) (Van Heerde, Leeflang en Wittink 2000). Wel blijkt het transactiebedrag van de volgende aankoop iets hoger te liggen ($\beta = .094, p < .01$).

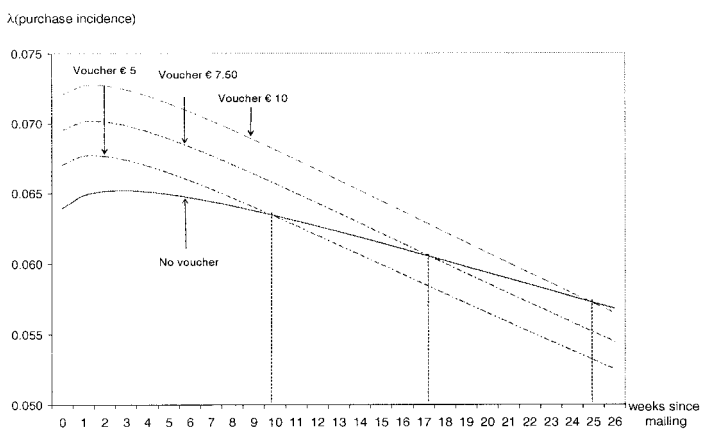
Figuur 5 laat het effect zien van voucherbezit op aankooptiming gedurende de verloop-tijd voor drie verschillende voucherwaardes en drie verschillende voucherhistories (zie Leenheer en Bijmolt 2005b voor technische details). Op de verticale as staat de hazard rate: de kans dat er een aankoop plaatsvindt gegeven dat er sinds week $t = 0$ geen aankoop meer heeft plaatsgevonden. Als benchmark is in de figuur ook de hazard rate aangegeven voor een vergelijkbare consument die geen voucher bezit (benchmarkcurve). Op deze manier is de effectiviteit van de voucher waar te nemen. De figuur toont de dominante invloed van voucherhistorie. In het geval dat de consument de vorige voucher niet heeft ontvangen is de huidige voucher alleen effectief gedurende de eerste twee weken en alleen voor de grootste voucher (€ 10). Alleen dan ligt de curve namelijk boven de benchmarkcurve. In de overige weken en voor lage voucherwaardes (€ 5 of € 7.50) is de aankoopkans voor een vergelijkbare consument zonder voucher hoger dan voor een consument met voucher. Daarentegen is de voucher voor een klant die de vorige voucher ook al ontving gedurende de volledige looptijd voor alle voucherwaardes effectief. De curves die de situatie voor een consument met voucher representeren liggen namelijk steeds boven de benchmarkcurve. Een middenpositie wordt ingenomen door de situatie waarin de vorige voucher is ontvangen maar niet gebruikt. Hier slaat het teken gedurende de looptijd om voor alle drie voucherwaardes. Deze vouchers zijn dus alleen gedurende de eerste weken effectief.

Voor transactiegrootte zijn soortgelijke figuren op te stellen. Ook daar geldt dat voor consumenten die de vorige voucher niet ontvingen, de huidige voucher niet effectief is. Voor consumenten die de vorige voucher ontvingen (en eventueel ook inwisselden) leidt de voucher tot grote transacties in de eerste weken na uitgifte. De vouchers met de grootste waarde zijn wederom het meest effectief.

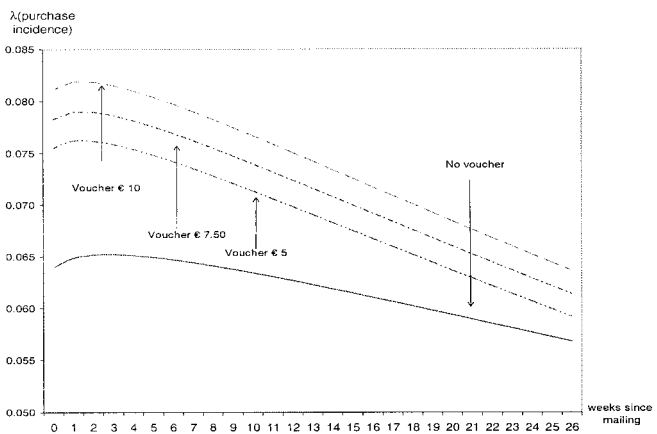
Deze studie naar winkelvouchers laat zien dat de beloningen in een spaarprogramma stimulerend kunnen zijn. Bestaande literatuur liet al zien dat consumenten zich bij hun koopbeslissing laten leiden door spaarpunten, maar onze analyses laten zien dat ook van de beloning een positieve werking op koopgedrag kan uitgaan. Dit geldt vooral voor grote vouchers, hetgeen pleit voor het zetten van voldoende hoge spaardrempels. Daarnaast zijn de vouchers vooral net na uitgifte effectief. De detaillist zou dus moeten overwegen de inwisselperiode kort te houden. Het duidelijkste effect gaat echter uit van de voucherhistorie: klanten reageren positiever als ze al eens eerder een beloning hebben gekregen. Dit pleit voor het hebben van lange-termijn programma's boven het houden van kortlopende spaaracties.



Figuur 5a Vorige voucher niet ontvangen



Figuur 5b Vorige voucher ontvangen, maar ongebruikt



Figuur 5c Vorige voucher ontvangen en gebruikt

Figuur 5. De relatie tussen voucherbezit en aankoopoptiming.

5. CONCLUSIES

Loyaliteitsprogramma's kunnen een aantrekkelijk instrument zijn voor het opbouwen van klantkennis en klantloyaliteit. Een studie onder 180 detaillisten laat zien dat dit vooral geldt in branches waarin consumenten met enige regelmaat aankopen doen en detaillisten weinig onderscheidend zijn met hun productassortiment (zoals benzinestations). Bij de adoptiebeslissing moet een detaillist de mogelijkheden en wenselijkheid van klantrelatiemanagement voor zijn specifieke onderneming meewegen. De informatie verkregen uit klantenkaarten biedt mogelijkheden om klantkennis te vergroten. Zonder de benodigde technologische vaardigheden is dit echter moeilijker te realiseren.

Voor het succes van een loyaliteitsprogramma is het essentieel dat klanten aan het programma deelnemen. De tweede studie toont aan dat een aantrekkelijk programmadesign consumenten hiertoe kan bewegen. Verder blijkt de adoptiebeslissing van consumenten deels te worden ingegeven door een algemene attitude ten aanzien van loyaliteitsprogramma's. Opvallend daarbij is dat de angst voor het verliezen van privacy nog altijd voor sommige consumenten een overweging is om van deelname af te zien; dit terwijl detaillisten vaak maar heel beperkt gebruik maken van de verkregen data.

Onderzoek in de supermarktbranche (studie 2) toont aan dat loyaliteitsprogramma's in staat zijn om de gedragsloyaliteit te verbeteren. Bij het bepalen van de effectiviteit kan het negeren van zelfselectie echter gemakkelijk leiden tot overschatting. Klanten worden namelijk niet willekeurig lid van een loyaliteitsprogramma: juist klanten die reeds loyaal zijn, voelen zich het meest aangetrokken tot het loyaliteitsprogramma. We laten ook zien dat consumenten over de tijd sterker op beloningen reageren (studie 3). Dit pleit voor langlopende programma's boven 'snelle' spaaracties.

Detaillisten dienen zich te realiseren dat de vormgeving van het programma belangrijk is (studie 2 en 3). Omdat differentiatie op basis van loyaliteitsprogramma-deelname zwak is in het geval van hoge penetratie, moet het programma zelf voldoende mogelijkheden bieden om tussen leden te differentiëren. Vooral van spaarprogramma's gaat een stimulerende werking uit. Men kan aanzienlijk op beloningskosten besparen als deze consumenten stimuleren om verdere aankopen te doen (zoals winkelvouchers). Een veelzijdig gebruik en optimale inzet van designelementen komt ook de winstgevendheid van het programma ten goede.

Het onderzoek naar loyaliteitsprogramma's is nog volop in ontwikkeling. Op het gebied van de optimalisering van het programmadesign is bijvoorbeeld nog veel onderzoek te doen. Daarbij is het gebruik van veldexperimenten een uitstekende manier om de validiteit en causaliteit van metingen te controleren. Ook over de winstgevendheid van loyaliteitsprogramma's is nog onvoldoende kennis, grotendeels omdat de beschikbare data daarvoor niet aanwezig zijn. Samenwerking tussen bedrijfsleven en academische onderzoekers kan vruchtbaar zijn om data (uit de praktijk) en analyse (vanuit universiteiten) samen te brengen.

NOTEN

1. Share-of-wallet meet de bestedingen bij een specifieke detaillist als fractie van de totale bestedingen in een sector; het begrip kent geen duidelijke vertaling in het Nederlands.
2. Omdat deze regressiemodellen truncated zijn (alleen bedrijven die een loyaliteitsprogramma hebben geadopteerd, rapporteren over de effectiviteit ervan), is het voor een onvertekende

schatting van de parameters noodzakelijk de Heckman two-step procedure te volgen en een zogenaamde Heckman's correctie term (ook wel inverse Mills ratio) op te nemen in de regressies (Verbeek 2000, p.208).

3. Beide modellen schatten random effects om rekening te houden met de panelstructuur van de data.

LITERATUUR

- Bawa, K., 1996. Influences on Consumer Response to Direct Mail Coupons: An Integrative Review. *Psychology & Marketing*, 13(2), 129-156.
- Bolton, R.N., P.K. Kannan en M.D. Bramlett, 2000. Implications of Loyalty Program Membership and Service Experiences for Customer Retention and Value. *Journal of the Academy of Marketing Science*, 28(1), 95-108.
- Bolton, R.N., K.N. Lemon en P.C. Verhoef, 2004. The Theoretical Underpinnings of Customer Asset Management: A Framework and Propositions for Future Research. *Journal of the Academy of Marketing Science*, 32(3), 271-293.
- Damanpour, F., 1991. Organizational Innovation: A Meta-Analysis of Effects of Determinants and Moderators. *Academy of Management Journal*, 34(3), 555-590.
- Dwyer, F.R., P.H. Schurr en S. Oh, 1987. Developing Buyer-Seller Relationships. *Journal of Marketing*, 51(2), 11-27.
- Franses, P.H., 2005. On the Use of Econometric Models for Policy Simulation in Marketing. *Journal of Marketing Research*, 42(1), 4-14.
- Gupta, S., 1988. Impact of Sales Promotions on When, What, and How Much to Buy. *Journal of Marketing Research*, 25(4), 342-355.
- Hanssens, D.M., L.J. Parsons en R.L. Schultz, 2001. *Market Response Models: Econometric and Time Series Analysis*. Boston: Kluwer Academic Publishers.
- Hsee, C.K., F. Yu, J. Zhang en Y. Zhang, 2003. Medium Maximization. *Journal of Consumer Research*, 30(1), 1-14.
- Inman, J.J. en L. McAlister, 1994. Do Coupon Expiration Dates Affect Consumer Behavior?. *Journal of Marketing Research*, 31(3), 423-428.
- Kivetz, R. en I. Simonson, 2002. Self-Control For The Righteous: Toward a Theory Of Pre-Commitment To Indulgence. *Journal of Consumer Research*, 29(2), 199-217.
- Leenheer, J., 2004. *The Adoption and Effectiveness of Loyalty Programs in Retailing*. Tilburg University. CentER Dissertation Series no.135.
- Leenheer, J., H.J. Van Heerde, T.H.A. Bijmolt en A. Smidts, 2004. Do Loyalty Programs Really Enhance Behavioral Loyalty? An Empirical Analysis Accounting for Self-Selecting Members. Working paper.
- Leenheer, J. en T.H.A. Bijmolt, 2005a. The Adoption and Effectiveness of Loyalty Programs: The Retailer's Perspective. Working paper.
- Leenheer, J. en T.H.A. Bijmolt, 2005b. The Dynamic Effects of Loyalty Program Rewards on Customer Purchases. Working paper.
- Mägi, A.W., 2003. Share of Wallet in Retailing: the Effects of Customer Satisfaction, Loyalty Cards and Shopper Characteristics. *Journal of Retailing*, 79(2), 97-106.
- Nakanishi, M. en L.C. Cooper, 1982. Simplified Estimation Procedures For MCI Models. *Marketing Science*, 1(3), 314-320.
- Noble, S.M. en J. Phillips, 2004. Relationship hindrance: why would consumers not want a relationship with a retailer? *Journal of Retailing*, 80(4), 289-303.
- Phelps, J., G. D'Souza en G. Nowak, 2001. Antecedents and Consequences of Consumer Privacy Concerns: An Empirical Investigation. *Journal of Interactive Marketing*, 15(4), 2-17.
- Van Dijk, A., H.J. Van Heerde, P.S.H. Leeflang en D.R. Wittink, 2004. Similarity-Based Spatial Methods to Estimate Shelf Space Elasticities. *Quantitative Marketing and Economics*, 2(3), 257-277.

- Van Heerde, H.J., P.S.H. Leeflang en D.R. Wittink, 2000. The Estimation of Pre- and Post-Promotion Dips with Store-Level Scanner Data. *Journal of Marketing Research*, 37(3), 383-395.
- Van Osselaer, S.M.J., J.W. Alba en P. Manchanda, 2004. Irrelevant Information and Mediated Intertemporal Choice. *Journal of Consumer Psychology*, 14(3), 257-270.
- Verbeek, M., 2000. *A Guide to Modern Econometrics*. Chichester: John Wiley & Sons, LTD.
- Vermunt, J., 1996. *Log-linear event history analysis: A general approach with missing data, latent variables, and unobserved heterogeneity*. Tilburg: Tilburg University Press.
- Voss, G.B. en K. Seiders, 2003. Exploring the Effect of Retail Sector and Firm Characteristics on Retail Price Promotion Strategy. *Journal of Retailing*, 79(1), 37-52.
- Yi, Y. en H. Jeon, 2003. Effects of Loyalty Programs on Value Perception, Program Loyalty, and Brand Loyalty. *Journal of the Academy of Marketing Science*, 31(3), 229-240.
- Wooldridge, J.M., 2002. *Econometric Analysis of Cross Section and Panel Data*. Cambridge, Massachusetts: The MIT Press.
- Zhang, Z.J., A. Krishna en S.K. Dhar, 2000. The Optimal Choice of Promotional Vehicles: Front-loaded or Rear-Loaded Incentives. *Management Science*, 46(3), 348-362.

4. *De interviewer, hulp of hindernis?*

J. VAN DER ZOUWEN¹

SAMENVATTING

Voor het verklaren van gedrag zijn subjectieve begrippen als meningen en houdingen onmisbaar. De meting van die niet-direct waarneembare begrippen vereist strikte standaardisatie van het proces van vraagbeantwoording. Als daarbij interviewers moeten worden ingeschakeld om er voor te zorgen dat er überhaupt een waarnemingsproces plaats vindt, dan kunnen diezelfde interviewers de standaardisatie belemmeren. Is de interviewer nu een onontbeerlijke *hulp* voor onderzoeker en respondent, of eerder een lastige *hindernis* bij het verkrijgen van niet-vertekende informatie en goed vergelijkbare antwoorden?

Ter beantwoording van die vraag zijn transcripten van survey interviews geanalyseerd; met name de wijze waarop interviewers inadequate antwoorden van respondenten proberen te herstellen. Op grond van deze analyse, en inzichten ontleend aan de cybernetica, kan geconcludeerd worden dat het niet effectief is om het herstelgedrag van interviewers sterk te standaardiseren.

1. WAAROM ONDERZOEK DOEN MET GESTANDAARDISEERDE VRAGENLIJSTEN?

Sociologen, politicologen en andere beoefenaren van de sociale wetenschappen zijn ambitieuze mensen. Ze willen het gedrag van mensen niet alleen *beschrijven*, ze willen het ook *verklaren*: waarom gedragen mensen zich zoals ze doen, waarom nemen ze juist déze beslissingen en waarom maken ze juist déze keuzen? Ditzelfde geldt voor marktonderzoekers.

Bij het zoeken naar die verklaringen komt men al gauw terecht bij vooroordelen, meningen, voorkeuren en houdingen. Kenmerkend voor dergelijke *subjectieve begrippen* is dat ze niet direct waarneembaar zijn [1]. Vrijwel de enige manier om erachter te komen wat iemands mening of houding is, is om hem of haar daarnaar te vragen.

Hetzelfde geldt voor gebeurtenissen en gedrag dat zich in het verleden heeft afgespeeld, en dat niet in documenten is vastgelegd. Ook dan zit er weinig anders op dan de persoon in kwestie daarnaar te vragen.

En als men dan toch aan het vragen stellen is, dan kan men ook “in één moeite door” vragen naar zaken die wél direct waarneembaar zijn of die wél zijn vastgelegd in documenten. Vandaar dat in de sociale wetenschappen de ondervraging verreweg de meest gebruikte waarnemingsmethode is (Presser, 1984; Saris & Van der Zouwen, 1999; Van den Berg et al., 2002).

¹ Dit is een bewerking van een rede, gehouden op 8 oktober 2004, ter gelegenheid van mijn afscheid van de Vrije Universiteit.

Nu kan men *verschillen* tussen mensen qua gedrag alleen verklaren door te wijzen op andere verschillen die er tussen die mensen bestaan. En *veranderingen* in het gedrag bij een bepaalde persoon kan men alleen maar verklaren door te wijzen op andere, eerdere, veranderingen bij diezelfde persoon.

Omdat subjectieve begrippen als meningen en houdingen alleen indirect waarneembaar zijn, is het van het grootste belang dat de ondervraagden allemaal de, op dezelfde wijze geformuleerde, vragen beantwoorden. En als ze op meerdere tijdstippen worden ondervraagd, dat ze dan telkens precies dezelfde vragen beantwoorden. Want anders weet de onderzoeker niet of een verschil tussen de antwoorden van twee respondenten het gevolg is van een verschil *qua mening*, of *qua vraagstelling*. Of dat een verandering in het antwoord van dezelfde respondent duidt op een gewijzigde mening, of het gevolg is van een gewijzigde vraagformulering.

De onderzoeker die variabelen behorend bij subjectieve begrippen wil meten moet als het ware “één vergelijking met twee onbekenden” oplossen. Dat kan alleen maar door één van die twee onbekenden constant te houden. Kortom, onderzoekers die subjectieve variabelen willen meten, zijn gedwongen om gebruik te maken van gestandaardiseerde vragenlijsten [2].

Bij een strikte standaardisatie van de vragenlijst hoort ook dat de vragen in dezelfde volgorde worden gesteld. En verder dat de vragen, voor zover mogelijk, gesloten geformuleerd zijn; dat wil zeggen dat de respondent niet zelf een antwoord formuleert, maar een keuze maakt uit door de onderzoeker aangeboden antwoordalternatieven.

2. WAAROM GESTANDAARDISEERDE VRAGENLIJSTEN IN INTERVIEWS AFNEMEN?

Dergelijke gestandaardiseerde vragenlijsten kan men op verschillende wijzen afnemen. De onderzoeker kan ze in een klaslokaal laten invullen, per post opsturen, of via het internet aanbieden en er vervolgens op hopen dat ze correct worden ingevuld en massaal teruggestuurd. Die werkwijze is betrekkelijk goedkoop. En als de vragenlijst met zorg is ontworpen kan dit tot bevredigende resultaten leiden [3].

In het verleden werden vragenlijsten vaak aangeboden door interviewers die de vragen oplazen en de antwoorden van de respondenten noteerden (Platt, 2002). Deze methode van vragenlijstafname is echter erg kostbaar geworden. Zeker als de interviews niet per telefoon worden afgenomen maar tijdens een bezoek bij de respondenten thuis, dus in een zogenaamd *persoonlijk interview*. Dat is voor de doorsnee onderzoeker een bijna onbetaalbare luxe geworden. Zo bleek onlangs dat de kosten van een gestandaardiseerd persoonlijk interview in Duitsland in de laatste twintig jaar zijn verdrievoudigd, tot meer dan 110 euro per afgenomen vraaggesprek (Koch, 2002).

Toch doen er zich situaties voor waarbij het wenselijk, of zelfs noodzakelijk is, dat de vragenlijst wordt afgenomen in een persoonlijk interview. Bijvoorbeeld als de respondenten de taal waarin de vragenlijst gesteld is slecht beheersen, als ze over weinig cognitieve vaardigheden beschikken, of als ze erg weinig gemotiveerd zijn om de vragenlijst te beantwoorden. En ook als de vragenlijst lang is en veel lastig te beantwoorden vragen bevat, of bestaat uit ingewikkelde series vervolgvragen, die gericht zijn op de reconstructie van iemands levensloop of sociaal netwerk.

Kortom, *als* het invullen van de vragenlijst zoveel problemen oproept dat er iemand aanwezig moet zijn om de respondent te helpen bij het uitvoeren van diens taak, *dan* is

het persoonlijk interview, al dan niet computergestuurd, de aangewezen waarnemingsmethode.

3. DE INTERVIEWER ALS BEDREIGING VAN HET STANDAARDISATIE PRINCIPE

Nu is het inschakelen van interviewers niet alleen duur, het staat ook op gespannen voet met het ideaal van de gestandaardiseerde ondervraging.

Immers, de loutere aanwezigheid van een tweede persoon die de vragen stelt en naar de antwoorden luistert kan invloed hebben op de antwoordkeuzen van de respondent. Extreme, en maatschappelijk ongewenste meningen worden minder vaak geuit en sociaal onwenselijk gedrag wordt minder vaak toegegeven dan bij een schriftelijke vragenlijst (De Leeuw, 1992; Van der Zouwen & De Leeuw, 1990, 1991). Het lastige probleem voor de onderzoeker is dat die vertekening van de antwoorden in een meer sociaal wenselijke richting niet voor alle respondenten even sterk is. Respondenten verschillen namelijk van elkaar in de mate waarin ze de goedkeuring van de interviewer zoeken (Gove, 1982; Stocké, 2004).

Bovendien, *als* men interviewers inschakelt zal men *meerdere* interviewers moeten inschakelen. Want zelfs bij een minimale steekproef van 400 personen heeft men toch minstens 10 interviewers nodig. Die interviewers verschillen van elkaar; in hun uiterlijk en spraak, maar ook in hun bejegening van de respondent. En die verschillen tussen interviewers vormen een niet te onderschatten foutenbron (Groves, 1989).

Tenslotte is er nog een factor die de standaardisatie bedreigt. Die factor betreft de taak van de interviewers. Een deel van hun takenpakket, de *respondent – onafhankelijke taken*, moeten ze bij iedere respondent vervullen, zoals het oplezen van de vraag en het aanbieden van de antwoordalternatieven.

Maar een ander deel van hun taak is afhankelijk van wat de respondent doet: de interviewer hoeft alleen een toelichting op de vraag te geven indien de respondent daarom vraagt. En de interviewer hoeft alleen maar door te vragen als het antwoord van de respondent daartoe aanleiding geeft.

Kortom, in het ene interview moet de interviewer ander gedrag vertonen dan in het andere interview. De uitvoering van deze *respondent – afhankelijke* taken bedreigt zodoende ook de standaardisatie van het waarnemingsproces.

4. DE INTERVIEWER: HULP OF HINDERNIS?

Met zevenmijlslaarzen zijn we naar het kernprobleem van dit hoofdstuk gelopen: Voor het verklaren van gedrag gebruiken sociale wetenschappers subjectieve begrippen als meningen en houdingen. De meting van de bij die niet-direct waarneembare begrippen behorende variabelen, vereist strikte standaardisatie van het proces van vraagbeantwoording. Maar als daarbij interviewers moeten worden ingeschakeld om ervoor te zorgen dat er überhaupt een waarnemingsproces plaats vindt, dan kunnen diezelfde interviewers de standaardisatie belemmeren.

De vraag dringt zich dan ook op of de interviewer gezien moet worden als een onontbeerlijke *hulp* voor onderzoeker en respondent, of eerder als een lastige *hindernis* bij het verkrijgen van niet-vertekende informatie en goed vergelijkbare antwoorden. Van daar de titel: “De interviewer, hulp of hindernis?”

5. HET IMPLICIETE ANTWOORD UIT DE ONDERZOEKSPRAKTIJK

Onderzoekers hebben de neiging om zich de vraag naar de invloed van de interviewer helemaal niet te stellen. Of ze gaan er gemakshalve maar van uit dat de interviewer hooguit een neutraal verlengstuk van de door hen opgestelde vragenlijst is. Onder die aanname kan de onderzoeker de ingevulde vragenlijsten, afkomstig van verschillende interviewers, eenvoudig samenvoegen tot één gegevensbestand en hoeft informatie omtrent de interviewer bij de analyse van dat bestand verder geen rol te spelen. Wel zo gemakkelijk.

Echter, als de interviewer er “niet toe doet”, zou er geen verschil mogen bestaan tussen de antwoorden die door de ene interviewer zijn verzameld en die van de andere interviewer. Als er wél zo’n verschil blijkt te zijn dan kan men nog denken dat dit komt omdat de ene interviewer andere respondenten ondervroeg dan de andere. Maar als de respondenten – zoals het eigenlijk behoort – louter op basis van toeval zijn toegewezen aan de interviewers, dan *kan* zo’n verklaring niet kloppen. Dan moet er kennelijk een of ander systematisch *interviewer effect* hebben plaats gevonden.

Soms gaat de negatieve invloed van de interviewer op de kwaliteit van de gegevens nog veel verder. Dat is het geval bij de zogenaamde *interviewerfraude*, een onderwerp waar men in de wereld van het survey-onderzoek liever niet over spreekt, laat staan over publiceert.

Bij interviewerfraude gaat het om het verschijnsel dat sommige interviewers delen van de vragenlijst overslaan en zelf invullen, bijvoorbeeld op een geschikt moment in hun eigen woning [4]. Op deze wijze kan de interviewer de tijd die het afnemen van de vragenlijst kost aanzienlijk bekorten, of zelfs de voor de respondent bedoelde beloning in eigen zak steken. Dergelijke, door deze interviewers verzonden antwoorden, verstoren natuurlijk in hoge mate de kwaliteit van het onderzoeksmateriaal [5].

6. HET OPENEN VAN DE BLACK BOX VAN HET INTERVIEW

Het gedrag van de interviewers lijkt dus heel wat minder probleemloos te zijn dan sommige onderzoekers veronderstellen of hopen. Maar we hebben hier alleen nog met *vermoedens* te maken.

De precieze oorzaak van het interviewer effect is nog niet aangegeven en de verdenking van interviewer fraude is met de voorhanden gegevens vaak moeilijk te bewijzen.

Om na te kunnen gaan in hoeverre de interviewer een hulp dan wel een hindernis is moeten we weten wat de interviewer *feitelijk* doet tijdens het interview:

- In hoeverre gedraagt de interviewer zich tijdens het interview overeenkomstig de door de onderzoeker gegeven *instructies*?
- In hoeverre is de interviewer alert op *problemen* die de respondent heeft met het begrijpen van de vraag en bij het zoeken van een adequaat antwoord op de vraag?
- In hoeverre is de interviewer in staat om zó door te vragen dat de verkregen informatie *meer relevant en adequaat* wordt, zonder vertekend te worden door suggesties van de interviewer?
- En tenslotte: Hoe zorgvuldig is de interviewer bij het *noteren* en het correct coderen van de door de respondent gegeven antwoorden?

De enige bruikbare methode om dergelijke onderzoeksvragen te beantwoorden is door de ‘black box’ [6] van het interview open te maken en te registreren wat interviewer en

respondent feitelijk doen en zeggen tijdens het interview.

In de praktijk gebeurt dat meestal door de interviews op geluidsband op te nemen, en woordelijk uit te schrijven tot interviewprotocollen. Die protocollen worden vervolgens geanalyseerd met een codeer systeem [7].

7. HOE CORRECT WORDEN DE VRAGEN GESTELD DOOR DE INTERVIEWERS?

Een basis vereiste voor een gestandaardiseerd interview is dat de vragen uit de vragenlijst aan alle respondenten op dezelfde wijze gesteld worden. Dat klinkt zó triviaal dat men zou verwachten dat dit ook inderdaad gebeurt.

De werkelijkheid is echter anders: het komt regelmatig voor dat men bij het afluisteren van interviews, de interviewer een andere tekst hoort uitspreken dan die er in de vragenlijst staat. Meestal gaat het om onschuldige variaties die de inhoud van de vraag intact laten. Maar het komt ook voor dat de interviewer de vraag zo herformuleert dat die, ook inhoudelijk gezien, afwijkt van de vraag uit de vragenlijst (Brenner, 1982; Van der Zouwen & Smit, 2004a; Van der Zouwen et al. 2005).

Als de respondent die *gewijzigde* vraag beantwoordt door een van de aangeboden antwoordalternatieven te noemen, ontstaat op dat moment irrelevante informatie. Omdat de onderzoeker niet kan weten dat dit antwoord van de respondent betrekking heeft op een *andere* dan de bedoelde vraag, kan de onderzoeker tot verkeerde conclusies komen. In dat geval is de interviewer geen hulp, maar een *hindernis* voor de onderzoeker.

Hoe komt het nu dat interviewers soms de vraagtekst wijzigen?

In de eerste plaats kan men natuurlijk denken aan onoplettendheid, gebrekkige motivatie, of zelfs incompetentie van de interviewer.

Maar men kan een deel van de ‘schuld’ ook zoeken in de tekst van de vragenlijst zelf. Sommige vragen, die ‘op papier’ duidelijk lijken, kunnen problemen geven als ze voorgelezen moeten worden. Bijvoorbeeld omdat de toelichting op een begrip uit de vraag pas gegeven wordt nadat de vraag zelf gesteld is, en sommige respondenten al beginnen te antwoorden voordat ze de toelichting hebben gehoord (Houtkoop-Steenstra, 2000, ch. 5). Of omdat een vraag *of* men tevreden is met iets, beantwoord blijkt te moeten worden door een keuze te maken uit meerdere *gradaties* van tevredenheid (Van der Zouwen & Smit, 2004a).

Sommige interviewers gaan tijdens het veldwerk de vraagtekst verbeteren, bijvoorbeeld door de toelichting in de vraagtekst zelf op te nemen. Of door niet te vragen “*Bent u tevreden met..?*” maar “*Hoe tevreden bent u met ..? Andere interviewers brengen die verbeteringen niet aan waardoor verschillende respondenten verschillend geformuleerde vragen krijgen voorgelegd. En daarmee wordt de vergelijkbaarheid van de antwoorden aangetast.*

Het komt ook voor dat de respondent inmiddels al zoveel over zichzelf verteld heeft, dat de te stellen vraag eigenlijk al beantwoord is; of feitelijk niet van toepassing blijkt.

De interviewer is dan gedwongen om te improviseren en ter plekke een vraagtekst te bedenken die de betekenis van de vraag handhaaft en die de respondent toont dat er naar diens eerder gegeven informatie wel degelijk geluisterd is [8]. Als een interviewer daarin slaagt is die interviewer duidelijk als hulp voor de onderzoeker opgetreden.

8. INADEQUATE ANTWOORDEN EN DE REACTIE DAAROP VAN DE INTERVIEWER

Idealiter volgt op de vraag van de interviewer, de antwoordkeuze van de respondent. In de praktijk komt die ideale vraag – antwoord sequentie minder vaak voor dan onderzoekers hopen:

- Bijvoorbeeld omdat de respondent niet een antwoord geeft, maar verzoekt om een *toelichting* op de vraag.
- Of omdat niet de respondent, maar een *derde aanwezige*, de vraag begint te beantwoorden.
- Of omdat de respondent *weigert* de vraag te beantwoorden of zegt het antwoord op de vraag *niet te weten*.
- Of de respondent beantwoordt feitelijk een *andere* vraag dan de gestelde.

Door dergelijke reacties van de respondent dreigt de vraag – antwoord sequentie, te ontsporen. Met als gevolg dat er geen antwoord of slechts een onvolledig, meerduidig of irrelevant antwoord wordt opgeleverd. Het is dan ook de taak van de interviewer om de “trein weer op de rails te zetten”. Anders gezegd: om in te grijpen en te proberen de sequentie te herstellen [9]. Ik zal daar enkele voorbeelden van geven.

Onderzoekers kunnen vragen laten stellen die moeilijk te beantwoorden zijn. De respondent kan daarop reageren met “dat weet ik niet”. Maar het moeten geven van zo’n reactie is niet bevorderlijk voor iemands zelfwaardering. Aantrekkelijker is het om gewoon een andere, verwante, vraag te beantwoorden waarop men het antwoord wél kan geven (Belson, 1981). Zo antwoordde een respondent op de vraag of hij verwachtte dat zijn inkomen de komende twee jaar zal toenemen met “Ja, dat hoop ik wel”. Een oplettende interviewer zal in zo’n geval niet meteen de antwoordmogelijkheid “ja” aankruisen maar de respondent erop wijzen dat het bij deze vraag niet gaat om wat men *hoopt*, maar om wat men *verwacht*.

Dat klinkt triviale dan het is. Want bij de training is de interviewer op het hart gedrukt om een goede verstandhouding met de respondent op te bouwen. Terwijl hier de interviewer de respondent erop moet wijzen dat hij slechts de makkelijke vraag naar de *inkomenswens* beantwoord heeft, en dat hij alsnog een antwoord moet geven op de veel lastiger vraag naar de *inkomensverwachting*. Dat vereist de nodige tact van de interviewer.

9. HOE ‘VRIJE’ ANTWOORDEN TE HERSTELLEN?

Bij gesloten vragen hoeft de respondent alleen maar dát antwoordalternatief te noemen, dat het best past bij zijn situatie, of zijn mening het beste weergeeft. Die taak lijkt eenvoudig. Desondanks gebeurt het dikwijls dat de respondent een antwoord geeft dat wel slaat op de vraag, maar niet gelijk is aan een van de antwoordalternatieven (Dijkstra, 2002a; Ongena, 2005). Op dat moment heeft de interviewer de taak om dat ‘vrije’ antwoord om te zetten in één van de in de vragenlijst genoemde antwoordalternatieven: en dat in overleg met de respondent.

De uitvoering van deze hersteltaak vergt ook de nodige tact van de interviewer. Want feitelijk moet de interviewer de respondent erop wijzen dat diens ‘vrije’ antwoord wel relevant is, maar niet adequaat – althans niet vanuit het perspectief van de onderzoeker. En dat de respondent de beantwoording van de vraag dus nog maar eens moet overdoen.

Dat verzoek komt de verstandhouding tussen interviewer en respondent natuurlijk niet ten goede. Vandaar dat sommige interviewers maar zelf het ‘vrije’ antwoord van de respondent ‘vertalen’ in een van de officiële antwoordalternatieven en dat in de vragenlijst aankruisen of intypen [10]. De interviewer corrigeert de respondent dus zonder dat die dit in de gaten heeft. Dat bespaart tijd, en het voorkomt irritatie bij de respondent. Maar het druist wel in tegen het basisprincipe van het interview, namelijk dat de *respondent* het uiteindelijke antwoord moet geven, en niet de interviewer.

Auteurs van handboeken over het gestandaardiseerde interview zijn het erover eens dat het zojuist genoemde ‘kiezen’ voor de respondent door de interviewer geen correct interviewgedrag is. Maar ze verschillen van mening over wat een interviewer dan wél moet doen:

- Moet de interviewer *alle* antwoordalternatieven herhalen en de respondent vragen alsnog daaruit een keuze te maken (Brenner, 1982; Fowler & Mangione, 1990) [11]?
- Of is het voldoende als in tweede instantie slechts die *paar* alternatieven genoemd worden die gelet op het ‘vrije’ antwoord van de respondent in principe in aanmerking komen (Viterna & Maynard, 2002)?
- Of kan de interviewer volstaan met het aan de respondent ter goedkeuring voorleggen van juist dat *ene* antwoordalternatief dat naar de mening van de interviewer het ‘vrije’ antwoord van de respondent het dichtst benadert (de ‘verification probe’): “Bedoelt u ...?” (Moore, 2000).

Voor alle drie manieren van doorvragen zijn goede argumenten aan te voeren. Het opnieuw opsommen van alle alternatieven maakt de kans op beïnvloeding door de interviewer het kleinst, maar het komt wel als onnatuurlijk en zelfs nogal ‘bot’ over; tenslotte heeft de respondent de vraag in eerste instantie wel degelijk beantwoord.

Het alleen ter goedkeuring voorleggen van het alternatief dat het best lijkt te passen bij het eerder gegeven ‘vrije’ antwoord klinkt veel natuurlijker en belast de verstandhouding tussen respondent en interviewer veel minder. Maar de kans op beïnvloeding door de interviewer is daarbij veel groter. Tenslotte is de vraag van de interviewer (“Bedoelt u ...?”) een lichtelijk suggestieve vraag. Of die suggestie een negatieve invloed heeft op de kwaliteit van de verkregen gegevens hangt in hoge mate af van vaardigheid van de interviewer om bij het ‘vrije’ antwoord van de respondent het best passende antwoordalternatief te zoeken.

10. IN HOEVERRE SLAGEN DE HERSTEL POGINGEN VAN DE INTERVIEWER?

We hebben gezien dat de onderzoeker – omwille van de vergelijkbaarheid van de antwoorden – het vraaggedrag van de interviewer en het antwoordgedrag van de respondent zo precies mogelijk voorschrijft.

We hebben ook gezien dat respondenten, in meer gevallen dan de onderzoeker lief is, zich niet aan die voorschriften houden; waardoor de *ideale* vraag – antwoord sequentie

verandert in een *problematische*. Er is dan een probleem en de interviewer moet dat zien op te lossen door een herstellpoging te doen.

Om na te gaan hoe vaak in een concreet onderzoek zo'n herstellpoging correct wordt uitgevoerd en slaagt, hebben Jan Smit en ik een analyse gemaakt van 201 interviews. Die interviews werden gehouden in het kader van het LASA onderzoek naar zelfstandigheid en welzijn van ouderen (zie Van der Zouwen & Smit, 2004a en 2004b).

In dit interview werden een aantal vragen gesteld inzake het inkomen; wat de hoogte is van dat inkomen, of er de laatste drie jaar een inkomensachteruitgang heeft plaats gevonden (en zo ja, hoeveel), hoe tevreden men is met het inkomen, en of men verwacht dat het inkomen in de komende twee jaar voor- of achteruit zal gaan [12].

Voor deze vier vragen hebben we, aan de hand van uitgeschreven interviewprotocollen, nagegaan in hoeverre het antwoord van de respondent op de gestelde vraag adequaat was of niet, en in hoeverre de interviewer op correcte wijze de vragen en doorvragen stelde.

In totaal hebben we dus vier keer 201, oftewel 804 vraag – antwoord sequenties geanalyseerd. In de tabel staan de voornaamste uitkomsten.

Tabel 1. Resultaten van de pogingen tot herstel van inadequate initiële antwoorden.

Van de 804 initiële antwoorden waren er

472 adequaat (59%)

332 inadequate (41%)

Op de 332 inadequate initiële antwoorden volgde

246 maal een herstellpoging (75%)

86 maal geen herstellpoging (25%)

Van de 246 herstellpogingen

slaagden er 212 (86%)

faalden er 34 (14%)

Van de 246 herstellpogingen werden er

185 wel correct uitgevoerd (75%)

61 niet correct uitgevoerd (25%)

Negen en vijftig procent van die sequenties leverden meteen een adequaat antwoord op, doordat de respondent één van de aangeboden antwoordalternatieven koos. In die gevallen kon de interviewer dat antwoord noteren en direct overgaan naar de volgende vraag. Echter, in 41 procent van de gevallen was het in eerste instantie gegeven antwoord niet adequaat. Dat betekent dat in 332 van de 804 geanalyseerde sequenties de interviewer een herstellpoging moest ondernemen.

Er is goed nieuws en er is slecht nieuws. Om met het goede nieuws te beginnen: in driekwart van die 332 gevallen deed de interviewer wat hij of zij moest doen, namelijk de respondent verzoeken om alsnog een adequaat antwoord te geven.

En als zo'n herstellpoging werd gedaan dan had die in 86 procent van de gevallen succes: de respondent gaf alsnog een adequaat antwoord. Door toedoen van de interviewers steeg het aantal adequate antwoorden van de respondenten daarmee van 472 tot 688, ofwel een verbetering met 46 procent. Hier fungeerde de interviewer dus duidelijk als hulp voor de onderzoeker.

Maar er is ook slecht nieuws. In één kwart van de gevallen waarin de interviewer een herstelpoging had moeten ondernemen, vond die in werkelijkheid niet plaats. De interviewer accepteerde het inadequate antwoord zonder meer, bedacht een antwoordcode, en ging over naar de volgende vraag. Dergelijke interviewers kan men moeilijk een hulp van de onderzoeker noemen.

Bovendien, in de 246 gevallen waarin *wel* een herstelpoging gedaan werd, gebeurde dat in een kwart van de gevallen op incorrecte wijze, meestal doordat de interviewer de respondent een antwoord suggereerde, een suggestie die gewoonlijk gretig door de respondent werd beaamd. Dergelijke hulpvaardigheid van de interviewers is een hindernis op de weg naar geldige, niet-vertekende informatie.

11. WAT IS DE BESTE ‘HERSTEL’ METHODE?

Door ons codeerwerk weten we waar in welke interviews, de interviewers een correcte en geslaagde herstelpoging hebben ondernomen. Alle reden om die interviewfragmenten nader te bekijken, want van die fragmenten zouden we kunnen leren wat de beste manier is om een herstelpoging te ondernemen.

Eerlijk gezegd viel het resultaat van die zoektocht in eerste instantie erg tegen. Want er bleken niet één, of enkele beste manieren te bestaan, maar tientallen. Of een herstelmethode goed werkt blijkt af te hangen van de aard van de vraag, van de aard van de inadequatheid van het antwoord van de respondent, maar ook van de creativiteit van de interviewer. Dus moest het streven naar een overzichtelijk lijstje met ‘best practices’ opgegeven worden, en werd duidelijk dat het ideaal van de volledige standaardisatie van het interview niet vol te houden is.

12. HET INTERVIEW ALS GESTUURD COMMUNICATIE PROCES

“Achteraf is het makkelijk praten”. Dat blijkt ook hier weer. Want bij nader inzien bedacht ik mij dat het helemaal zo gek nog niet is dat juist bij het ‘herstellen’ van afwijkingen van de ideale vraag – antwoord sequentie, standaardisatie niet werkt. Om deze stelling toe te lichten moet ik een uitstapje maken op het terrein van de cybernetica, de wetenschap inzake sturing en communicatie (Wiener, 1965).

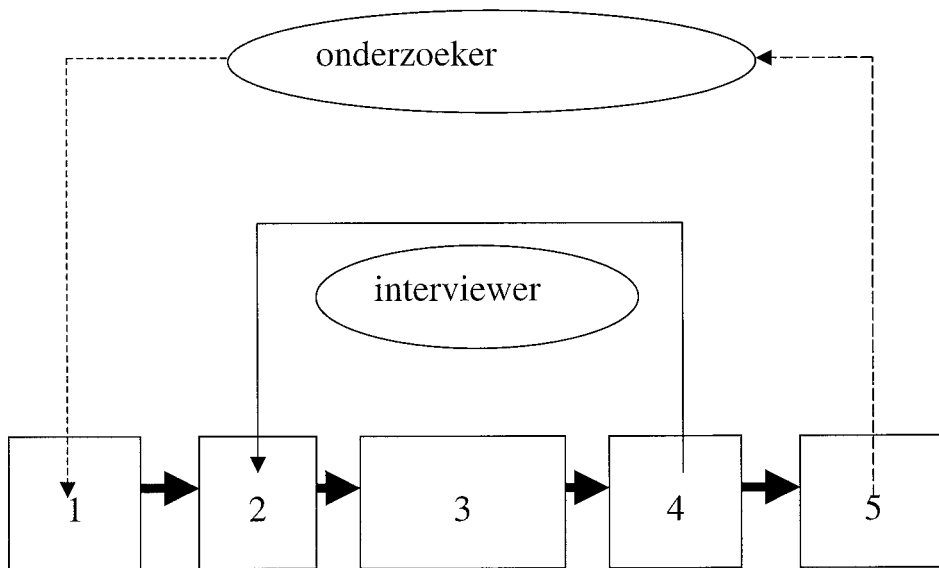
We kunnen het interview beschrijven als een gestuurd communicatieproces, waarbij drie actoren zijn betrokken, elk met hun eigen takenpakket (Van der Zouwen, Dijkstra & Van de Bovenkamp, 1986; Van der Zouwen, 2001; Van der Zouwen & Smit, 2005).

In de eerste plaats is er de *onderzoeker* die de vragenlijst opstelt, de interviewers instrueert en de door hen toegekende codes analyseert (zie Figuur 1.).

In de tweede plaats is er de *interviewer* die de vragen uit de vragenlijst, voorlegt aan de respondent, en de gegeven antwoorden beoordeelt. Adequate antwoorden worden gecodeerd terwijl bij inadequate antwoorden de interviewer een herstelpoging dient te ondernemen.

In het midden van dit schema bevindt zich de derde actor, de *respondent* die de vraag, via een intern cognitief proces, interpreteert en beantwoordt.

De activiteiten van deze drie actoren zijn elk op een bepaald doel gericht. Het doel van de onderzoeker is het verkrijgen van vergelijkbare, relevante, en niet-vertekende gegevens omtrent de respondenten. Om dat doel te bereiken moet hij het interviewproces



1. Door onderzoeker opgestelde vraag uit vragenlijst en instructies aan interviewers
2. Door interviewer gestelde (door)vraag
3. Cognitief proces bij respondent
4. Door respondent gegeven antwoord
5. Door interviewer toegekende code

Figuur 1. Het interview als gestuurd communicatie proces.

sturen. Nu worden er in de cybernetica twee typen sturing onderscheiden, namelijk sturing zónder, en sturing mét terugkoppeling.

Idealiter zou er in het gestandaardiseerde interview alleen sprake hoeven te zijn van sturing zónder terugkoppeling. Dan verloopt het proces regelrecht van fase 1 naar fase 5: De door de onderzoeker geformuleerde vraag (fase 1) wordt op overeenkomstige wijze gesteld door de interviewer (fase 2), en leidt tot een cognitief proces in het hoofd van de respondent (fase 3), resulterend in een adequaat antwoord (fase 4), dat door de interviewer correct wordt gecodeerd (fase 5) en daarna doorgegeven aan de onderzoeker.

Maar soms gaat er iets mis, bijvoorbeeld in het cognitieve proces bij de respondent, waardoor het antwoord niet adequaat is, en moet de interviewer een herstelpoging wagen. Dan is er sprake van sturing mét terugkoppeling, in dit geval van fase 4 terug naar fase 2.

Nu zegt de eerste hoofdwet van de cybernetica dat sturing met terugkoppeling alleen maar effectief is als er voldoende variëteit van handelingsalternatieven is (Ashby, 1956/1964). Die variëteit moet minstens zo groot zijn als de variëteit in de verstoringen die bijgestuurd moeten worden.

Anders gezegd, *als* de interviewer moet bijsturen, dus een herstelpoging moet ondernemen, *dan* moet hij ook over een uitgebreid gedragsrepertoire kunnen beschikken.

Er zijn echter onderzoekers die, met een beroep op de noodzakelijke standaardisering, het gedrag van interviewers ook bij herstelpogingen sterk inperken. Zij miskennen

daarmee het fundamentele verschil tussen sturing zónder, en sturing mét terugkoppeling. Een voorbeeld van zo'n inperking van het gedragsrepertoire van de interviewer is dat een inadequaat antwoord van de respondent, ongeacht de aard van dat antwoord, alleen maar hersteld mag worden door opnieuw alle antwoordalternatieven op te sommen. Een ander voorbeeld is dat de interviewer, als de respondent vraagt wat met een bepaald woord in de vraag wordt bedoeld, alleen maar mag reageren met "wat u er zelf onder verstaat". Daarmee belemmeren deze onderzoekers hun interviewers om als een effectieve hulp op te kunnen treden.

13. CONCLUSIE: DE WEDERKERIGE RELATIE TUSSEN ONDERZOEKER EN INTERVIEWER

Deze voorbeelden wijzen erop dat de relatie tussen de onderzoeker en de interviewer een *wederkerige* relatie is. Immers, we zagen dat de *interviewer* kan optreden als hulp, maar ook als hindernis voor het bereiken van het doel van de onderzoeker, namelijk het verkrijgen van correcte en vergelijkbare antwoorden van de respondenten. Maar, omgekeerd, zagen we ook dat de *onderzoeker* de interviewer kan helpen, of hinderen, bij het bereiken van diens doel, namelijk het op efficiënte wijze, en in een goede verstandhouding met de respondent, afnemen van het interview.

Daarom dient de onderzoeker zich bij het opstellen van de vragenlijst in te leven in de lastige taak die interviewers en respondenten bij bepaalde vragen hebben. De gebruikersvriendelijkheid van de vragenlijst dient grondig getest te worden in proefinterviews met respondenten die tot de onderzoekspopulatie behoren; dus niet alleen bij begripvolle collega's. En tijdens het eigenlijke veldwerk dient regelmatig – liefst via het beluisteren van geluidsopnamen – gecheckt te worden of de vragenlijst inderdaad zo wordt afgenomen als dat door de onderzoeker bedoeld was.

Die activiteiten zijn tijdrovend en duur. Maar gelet op de hoge kosten van het persoonlijk interview is het beknibben daarop te karakteriseren als "penny wise and pound foolish". Het persoonlijke interview is er voor het afnemen van vragenlijsten onder lastige omstandigheden. Maar dan moet het ook voor die lastige omstandigheden geschikt gemaakt zijn (De Leeuw, 2000).

En tenslotte de interviewer. In de literatuur over het gestandaardiseerde interview wordt de interviewer vooral beschreven als een mogelijke foutenbron, die met zijn afwijken van de vraagformulering, met zijn niet of juist suggestief doorvragen, of zelfs door het plegen van fraude een hoge hindernis vormt voor het bereiken van goede kwaliteit van de gegevens.

Het zou mooi zijn als die literatuur werd aangevuld met artikelen waarin werd beschreven hoe tactvol interviewers kunnen optreden bij het herstellen van inadequate antwoorden en hoe creatief bij het opvangen van tekortkomingen in de vragenlijst.

NOTEN

1. De term *subjectieve begrippen* die ik hier gebruik is ongeveer synoniem met de term *subjective phenomena* van Turner en Martin (1984) en verwant aan de term *attitudinal information*, zoals gebruikt door Sudman en Bradburn (1974).
2. Een wat uitgebreider argumentatie voor de standaardisatie van vragenlijsten heb ik elders gegeven (Van der Zouwen, 2002). Zie verder Maynard & Schaeffer (2002).

3. Met name Don Dillman (1978 en 2000) heeft erg veel gedaan voor de ‘upgrading’ van de schriftelijke vragenlijst, en – recentelijk – het internet survey.
4. Voorbeelden van ‘interviewer cheating’ zijn o.m. te vinden in Schreiner, et al. (1988), Biemer & Stokes (1989), Harrison & Krauss (2002) en Eyerma et al. (2004).
5. Een opmerkelijk voorbeeld van interviewerfraude is recentelijk gerapporteerd door Matschinger et. al (forthcoming). Door het tijdens het veldwerk in toenemende mate overslaan van screening vragen inzake bepaalde typen psychische problemen, respectievelijk het negeren van positieve antwoorden daarop, wisten de interviewers de interviewtijd steeds meer te bekorten: “The fraction of positively answered screening questions within each month of the study dropped from 80% to only 7.3% in the last month”. Het gevolg van deze fraude was dat dit kostbare face to face survey volstrekt onbruikbare, want sterk vertekende, resultaten heeft opgeleverd.
6. In mijn inaugurele oratie (Van der Zouwen, 1977) heb ik het interview als waarnemingsmethode geanalyseerd, als ware het interview in zijn geheel een black box. In deze afscheidsrede wil ik laten zien wat de mogelijkheden zijn om die black box – gedeeltelijk – te openen.
7. Een elegante procedure voor het verwerken en analyseren van interviewprotocollen (SEQUENCE) is ontwikkeld door Wil Dijkstra. Zie Dijkstra (2002a en 2002b) en voor een recente toepassing Ongena, 2005.
8. In Houtkoop-Steenstra (1995) wordt dit proces van ‘meeting both ends’ door de interviewer, beschreven.
9. In de conversatie analyse worden dergelijke herstelactiviteiten vanouds aangeduid met ‘repair’. Zie bijvoorbeeld Houtkoop-Steenstra (2000, pp 29-33), Moore & Maynard (2002) en Van der Zouwen & Smit (2004b).
10. Dit inadequate interviewergedrag werd eerder aangeduid met de term ‘choosing’ (Dijkstra & Van der Zouwen, 1988).
11. Deze werkwijze wordt ook geleerd aan de interviewers die de opleiding tot marktonderzoek interviewer volgen (persoonlijke mededeling van een van de opleiders, de heer Gert Schouten).
12. Deze vier vragen zijn gekozen omdat dergelijke inkomensvragen in veel surveys gesteld worden, en omdat uit voorgaand onderzoek (Van der Zouwen & Smit, 2004a) bekend is dat ze tot veel problematische sequenties leiden.

LITERATUUR

- Ashby, W.R., (1956/1964). *An Introduction to Cybernetics*. London: Methuen & Co.
- Belson, W.A., (1981). *The Design and Understanding of Survey Questions*. Aldershot: Gower.
- Berg, H. van den, M. Mehciz, H. Houtkoop-Steenstra & B. Holleman. (2002). *Opinie-maken of opinie-meten: De rol van stellingvragen in markt- en opinieonderzoek*. Amsterdam: Stichting voor Culturele Studies.
- Biemer, P.P. & S.L. Stokes (1989), The Optimal Design of Quality Control Samples to Detect Interviewer Cheating. *Journal of Official Statistics*, 5(1), 23-40.
- Brenner, M. (1982), “Response Effects of “Role-restricted” Characteristics of the Interviewer”. In: W. Dijkstra & J. van der Zouwen (eds.), *Response Behaviour in the Survey-interview*. London: Academic Press, pp. 131-165.
- Dijkstra, W. (2002a), *Methodologie als luis in de pels*. Inaugurele oratie Vrije Universiteit.
- Dijkstra, W. (2002b), “Transcribing, Coding, and Analyzing Verbal Interactions in Survey Interviews”. In: D.W. Maynard, H. Houtkoop-Steenstra, N.C. Schaeffer & J. van der Zouwen (eds.), *Standardization and Tacit Knowledge; Interaction and Practice in the Survey Interview*. New York: John Wiley.
- Dijkstra, W. & J. van der Zouwen (1988), “Types of inadequate interviewer behavior in survey-interviews”. In: W.E. Saris & I.N. Gallhofer (eds.), *Sociometric Research, Volume 1: Data Collection and Scaling*. Basingstoke: Macmillan. pp. 24-35.

- Dillman, D.A. (1978), *Mail and Telephone Surveys; The Total Design Method*. New York: John Wiley.
- Dillmann, D.A. (2000), *Mail and Internet Surveys; The Tailored Design Method. Second Edition*. New York: John Wiley.
- Eyerman, J., P. Biemer, C. Hill & C. Ellis (2004), "Statistical and Technical Methods for the Detection and Prevention of Interviewer Fraud in Data Collection". Paper presented at the Sixth International Conference on Logic and Methodology, August 16-20, 2004, Amsterdam, The Netherlands.
- Fowler Jr., F.J. & T.W. Mangione (1990), *Standardized Survey Interviewing*. Newbury Park: Sage.
- Gove, W.R. (1982), "Systematic Response Bias and Characteristics of the Respondent". In: W. Dijkstra & J. van der Zouwen (eds.), *Response Behaviour in the Survey Interview*. London: Academic Press, pp. 167-187.
- Groves, R.M. (1989), *Survey Errors and Survey Costs*. New York: John Wiley.
- Harrison, D.E. & S.I. Krauss (2002), "Interviewer Cheating: Implications for Research on Entrepreneurship in Africa". *Journal of Developmental Entrepreneurship*, 7(3), 319-330.
- Houtkoop-Steenstra, H. (1995), "Meeting both ends. Between standardization and recipient design in telephone survey interviews". In: P. ten Have & G. Psathas (eds.), *Situated order: studies in the organization of talk and embodied activities*. Washinton, DC: University Press of America, pp. 91-107.
- Houtkoop-Steenstra, H. (2000), *Interaction and the Standardized Survey Interview; The living questionnaire*. Cambridge: Cambridge University Press.
- Koch, A. (2002), "20 Jahre Feldarbeit im ALLBUSS: Ein Blick in die Black Box". *ZUMA-Nachrichten* 51, Jrg. 26, November 2002, 9-37.
- Leeuw, E.D. de (1992), *Data Quality in Mail, Telephone and Face to Face Surveys*. Proefschrift Vrije Universiteit, Amsterdam: TT-Publikaties.
- Leeuw, E.D. de (2000). Methodologische uitdagingen voor het nieuwe millennium; Van het 19-de eeuwse sociaal survey naar het World Wide Web. In A.E. Bronner et al (eds), *Recente ontwikkelingen in het marktonderzoek. Jaarboek van de Nederlandse Vereniging voor Marktonderzoekers en Informatie Management*, 2000,23-34. Haarlem: De Vrieseborch.
- Matschinger, H., S. Bernert & M.C. Angermeyer (forthcoming), An Analysis of Interviewer Effects on Screening Questions in a Computer Assisted Personal Mental Health Interview. *Journal of Official Statistics*, forthcoming.
- Maynard, D.W. & N.C. Schaeffer (2002), "Standardization and Its Discontents". In: D.W. Maynard, H. Houtkoop-Steenstra, N.C. Schaeffer & J. van der Zouwen (eds.), *Standardization and Tacit Knowledge: Interaction and Practice in the Survey Interview*. New York: John Wiley, pp. 3-45.
- Moore, R.J. (2000), "Differences in probing practices at two U.S. survey centers". In: J. Blasius, J. Hox, E. de Leeuw & P. Schmidt (eds.), *Social science methodology in the new millenium*. Proceedings of the Fifth International Conference on Logic and Methodology, held at Cologne, Germany, October 3-6, 2000. Opladen: Leske & Budrich, CD-ROM.
- Moore, R.J. & D.W. Maynard (2002), Achieving understanding in the standardized survey interview: repair sequences. In: D.W. Maynard, H. Houtkoop-Steenstra, N.C. Schaeffer & J. van der Zouwen (eds.), *Standardization and tacit knowledge: interaction and practice in the survey interview*. New York: John Wiley, pp. 281-311.
- Ongena, Y.P (2005), *Interviewer and Respondent Interaction in Survey Interviews*. Proefschrift Vrije Universiteit, Amsterdam.
- Platt, J. (2002), "The History of the Interview". In: J.F. Gubrium & J.A. Holstein (eds.), *Handbook of Interview Research; Context & Method*. Thousand Oaks: Sage.
- Presser, S. (1984), "The use of survey data in basic research in the social sciences". In: C.F. Turner & E. Martin (eds.), *Surveying Subjective Phenomena*. New York: Russell Sage Foundation.

- Saris, W.E. & J. van der Zouwen (1999), *Feasibility Study Concerning a Research Program on the Quality of Survey Measurement*. Report to the 'Gebiedsbestuur' MAGW. Amsterdam.
- Schreiner, I., K. Pennie & J. Newbrough (1988), "Interviewer Falsification in Census Bureau Surveys". American Statistical Association, *Proceedings of the Section of Survey Research Methods*, pp. 491-496.
- Stocké, V. (2004), "The Interdependence of Determinants for the Strength and Direction of Social Desirability Bias in Racial Attitude Surveys". Report No. 04-14 SonderforschungsBereich 504. Universität Mannheim, Mannheim.
- Sudman, S. & N.M. Bradburn (1974), *Response Effects in Surveys; A Review and Synthesis*. Chicago: Aldine.
- Turner, C.F. & E. Martin (eds.)(1984), *Surveying Subjective Phenomena*. New York: Russell Sage Foundation.
- Viterna, J.S. & D.W. Maynard (2002), "How Uniform is Standardization? Variation Within and Across Survey Research Centers Regarding Protocols for Interviewing." In: D.W. Maynard, H. Houtkoop-Steenstra, N.C. Schaeffer & J. van der Zouwen (Eds.), *Standardization and Tacit Knowledge: Interaction and Practice in the Survey Interview*. New York: John Wiley, pp. 365-397.
- Wiener, N. (1965), *Cybernetics: or Control and Communication in the Animal and the Machine*. Second Edition. Cambridge, MA: MIT-Press.
- Zouwen, J. van der (1977), *Het interview ondervraagd; naar een model van het interview-proces*. Inaugurele oratie VU. Alphen aan den Rijn: Samsom.
- Zouwen, J. van der (2001), "Cybernetics and Interviewing". *Kybernetes* 30(9/10), 1064-1071.
- Zouwen, J. van der (2002), "Why Study Interaction in the Survey Interview? Response from a Survey Researcher". In: D.W. Maynard, H. Houtkoop-Steenstra, N.C. Schaeffer & J. van der Zouwen (eds.), *Standardization and Tacit Knowledge: Interaction and Practice in the Survey Interview*. New York: John Wiley, pp. 47-65.
- Zouwen, J. van der, W. Dijkstra & J. van de Bovenkamp (1986), "The control of interaction processes in survey interviews". In: F. Geyer & J. van der Zouwen (eds.), *Sociocybernetic Paradoxes: Observation, Control and Evolution of Self-steering Systems*. London: Sage, pp. 55-63.
- Zouwen, J. van der & E.D. de Leeuw (1990 en 1991), "The Relationship between Mode of Administration and Quality of Data in Survey Research". *Bulletin de Méthodologie Sociologique* Nr. 29, pp. 3-14 (First Part) en Nr. 31 (Final Part) pp. 49-60.
- Zouwen, J. van der & J.H. Smit (2004a), "Evaluating Survey Questions by Analyzing Patterns of Behavior Codes and Question-Answer Sequences: A Diagnostic Approach". In: S. Presser, J.M. Rothgeb, M.P. Couper, J.T. Lessler, E. Martin, J. Martin & E. Singer (eds.), *Methods for Testing and Evaluating Survey Questionnaires*. New York: John Wiley, pp. 109-130.
- Zouwen, J. van der & J.H. Smit (2004b), "The Effectiveness of Repair Strategies of Interviewers in Survey Interviews". In: C. van Dijkum, J. Blasius & B. van Hilten (eds.), *Recent Developments and Applications in Social Research Methodology*. Proceedings of the Sixth International Conference on Logic and Methodology, August 17-20, 2004, Amsterdam. Amsterdam: SISWO, CD-ROM.
- Zouwen, J. van der & J.H. Smit (2005), "Control processes in survey interviews: a cybernetic approach". *Kybernetes* 34(5), 602-616.
- Zouwen, J. van der, J.H. Smit & M.H.L. van der Horst (2005), "Expectations of Interviewers as Determinants of Deviant Interviewer Behavior". Paper presented at the First Conference of the European Association for Survey Research, Barcelona, 18-22 July, 2005.

5. *De levende of digitale interviewer: Een vergelijking tussen CASI, CAPI en CATI*

A.E. BRONNER en T. KUIJLEN

SAMENVATTING

Eén van de kernactiviteiten binnen het marktonderzoek is dataverzameling via ondervraging. Een drietal ontwikkelingen heeft deze activiteit de laatste jaren sterk beïnvloed: lager wordende respons, sterk stijgende interviewer-kosten en het groeiende idee dat de respondent als een “echte” klant moet worden behandeld. Deze ontwikkelingen hebben geleid tot andere steekproeven (van vers/ad hoc naar access panels) en andere methoden van dataverzameling (van CATI/CAPI naar CASI). In dit artikel staan de effecten van de methoden van dataverzameling centraal. Bij de CASI situatie zonder interviewer worden eerder sociaal onwenselijke gedragingen toegegeven. De keerzijde is dat de respondenten via CASI minder opscheppen over sociaal wenselijk gedrag. CASI leidt ook tot extreme antwoordposities, meer afgewogen oordelen en het meer spontaan noemen van bijvoorbeeld merken. Al met al blijkt de aan- of afwezigheid van de interviewer wel degelijk invloed te hebben op de antwoorden. Bij samenvoeging van data die via verschillende methoden zijn verzameld is voorzichtigheid geboden.

1. INLEIDING: ONTWIKKELINGEN DIE DATAVERZAMELING BEINVLOEDEN

Eén van de kernactiviteiten binnen het marktonderzoek is dataverzameling via ondervraging. Aan respondenten worden vragen gesteld, er worden antwoorden gegeven en de antwoorden worden vastgelegd. En hoewel er een tendens is om dataverzameling een lagere status toe te kennen dan data-analyse (Morton-Williams, 1993) vormt dataverzameling onmiskenbaar de basis. Terugblikkend heeft een drietal factoren deze activiteit sterk beïnvloed. Toenemende non-respons, sterk stijgende kosten en het besef dat de respondent als klant behandeld moet worden zijn de drie factoren die de keuzen bij de dataverzameling gedurende de laatste tien jaar hebben bepaald. Wij zullen deze factoren achtereenvolgens toelichten. Allereerst de toenemende non-respons.

Nederland ontwikkelde zich in de jaren negentig tot een land met lage responspercentages. De non-respons werd hoger dan in vergelijkbare landen en nam snel toe (de Heer, 1999). In telefonisch onderzoek zakte in 10 jaar het responspercentage bij een verse steekproef van 60% naar 30%. En bij het face-to-face uitgevoerde Nationale Kiezers Onderzoek zakte de respons van 70% in de jaren zeventig tot 50% in de jaren negentig. Maar er bleek geen sprake van een vaste groep die nooit aan enquêtes meedoet of juist wel altijd meedoet. Schattingen zijn wel gedaan dat 15% van de populatie altijd mee

doet, 15% nooit en dat de resterende 70% het laat afhangen van onderwerp, situatie, beschikbare tijd. Dat werd recentelijk bevestigd door Ineke Stoop in haar proefschrift (Stoop, 2005) die aantoont dat zowel de bereikbaarheid als bereidwilligheid kunnen worden verhoogd. In het AVO 1999 werd bij het eerste bezoek 48% van de huishoudens thuis getroffen. Na vier contactpogingen was 92% bereikt en na 8 pogingen 98%. Ook het opnieuw benaderen van weigeraars levert succes op, zo toonde Stoop aan. In de literatuur wordt dit laatste 'refusal conversion' genoemd. Heel wat mensen blijken alsnog 'ja' te zeggen als een andere interviewer aan de deur komt of als dezelfde interviewer nog een keer langskomt op een mogelijk geschikter moment. Vergelijking met andere landen leert dat juist die 'refusal conversion' in Nederland succes heeft. En dat verleidde Stoop tot de volgende stelling: "Nederland stond bekend als Europees kampioen non-respons. In het European Social Survey (ESS) 2002/2003 scoort Nederland echter bovengemiddeld en komt dicht bij de 70% respons. Nederland verdient op grond van het ESS de titel: Europees kampioen overhalen van weigeraars". Het feit dat deze conversie mogelijk is wijst er ook weer op dat er geen kern van vaste weigeraars is.

Een tweede belangrijke beïnvloedende factor van de dataverzameling zijn de kosten. Door verandering van sociale wetgeving werden de enquêtekosten veel hoger en gecombineerd met de toenemende reiskosten maakte dit de inzet van enquêteurs een kostbare zaak. "The shift towards Internet surveys got an unintended push last year when the costs of face-to-face interviews rose by 30%, due to new labour laws affecting interviewers" (Havermans, 2003, p.6). Omdat bovendien het vak interviewen niet als een hoge status activiteit werd beschouwd verminderde ook het aantal geschikte kandidaten. Vrouwen bijvoorbeeld die in de 70-er en 80-er jaren parttime interviewen een aantrekkelijk alternatief vonden kozen in de 90-er jaren steeds meer voor vastere banen.

De derde ontwikkeling is dat de marktonderzoekwereld de respondenten steeds meer als een soort klant zijn gaan zien. En klanten hebben wensen en die moet je respecteren. Er wordt nu tevoren gevraagd of mensen bereid zijn op gezette tijden aan onderzoek mee te doen, respondenten worden in staat gesteld de vragen te beantwoorden op een tijdstip dat hen schikt, een adequate beloning voor de inspanning wordt gegeven en de privacy wordt gerespecteerd. Dit nieuwe denken wordt wel 'permission research' genoemd. Ook leidde deze nieuwe opvatting van de respondent als klant tot de mogelijkheid dat een respondent zelf kan kiezen hoe hij of zij ondervraagd wilde worden. De ene respondent preferert telefonisch, de andere face-to-face en een derde wil het liefst de vragenlijst zelf achter de pc invullen. Zo ontstonden "multi-mode" onderzoeken dat wil zeggen onderzoeken waarin verschillende methoden van dataverzameling parallel worden gebruikt.

Deze ontwikkelingen leidden in de eerste plaats tot de opzet van access panels door onderzoeksbureaus. Stoop (2005, p.41) omschrijft deze panels als volgt: "Access panels are groups of people who have agreed to regularly participate in surveys run by a specific organization, generally a market research organization, and are the modern equivalent of mail panels". Groot voordeel is dat leden van zo een panel aangeven op gezette tijden aan onderzoek mee te willen doen hetgeen dus de non-respons reduceert. Ook is informatie over die respondenten beschikbaar zodat socio-demografische gegevens niet telkens gevraagd behoeven te worden en zodat respondenten met bepaalde kenmerken

getrokken kunnen worden (lezers van blad X, gebruikers product Y) zonder kosten verblindende screening. De kwaliteit van zo een access panel wordt sterk bepaald door de werving van de participanten en het panelonderhoud. Een panel benadert het meest een random steekproef als mensen in ‘verse’ random steekproeven aan het eind van het vraaggesprek wordt gevraagd lid te willen worden van het panel. Een slechte benadering is mensen bijvoorbeeld via internet werven of vrijwillig lid laten worden. In het eerste geval bestaat het gevaar voor oververtegenwoordiging van ‘computerfreaks’, in het tweede geval voor ‘beroepsrespondenten’ die lid zijn van vele panels.

Bij adequaat panelonderzoek moeten we denken aan panelverversing, voorkomen dat een groep ‘beroepsrespondenten’ overblijft, niet te frequent benaderen, niet twee keer voor hetzelfde onderwerp benaderen.

In principe kunnen de leden van een dergelijk access panel met alle methoden van dataverzameling worden benaderd. Maar in de praktijk is de combinatie van access panel met CASI (Computer Assisted Self Interviewing) populair. In de terminologie van Bronner, Tchaoussoglou, Ross (2003) gaat het meer precies om CASI-IA, dat wil zeggen “Interviewer Absent”. CASI kan namelijk ook plaats vinden door de respondent onder het toezicht van de enquêteur die helpt als er problemen zijn. Deze vorm van CASI wordt aangeduid als CASI-IP (Interviewer Present). En deze CASI-IA methode kan op twee manieren plaatsvinden: off-line of respondenten ontvangen een e-mail invitatie met een URL naar de vragenlijst die op een website staat. Vaak wordt de combinatie access panel/CASI aangeduid als internetpanels. In het vervolg van dit artikel bedoelen wij met CASI telkens CASI-IA.

De combinatie access panel/CASI speelt in op de drie ontwikkelingen die in het begin van deze paragraaf zijn beschreven. De respons is hoog binnen het panel, met beloning wel 80%, de kosten zijn laag omdat er geen enquêteurs aan te pas hoeven te komen en de respondent wordt behandeld als klant omdat hij/zij een beloning krijgt en de vragenlijst op een zelf gekozen moment kan invullen. Mosmann (1998, p.44) spreekt van een verschuiving: “the interview situation shifts from a ‘man-to-man’ to a ‘man-to-machine-to-man’ relationship.” Omdat in de ogen van de marktonderzoeker de combinatie access panel/CASI veel problemen oplost is in veel longitudinale projecten een overstap gemaakt naar deze methodiek. Aan de andere kant zijn veel bedrijven en overheidsinstellingen zich juist bewust geworden dat onderzoek geen korte termijn maar een lange termijn zaak is. Benchmarking en tracking impliceren dat gegevens over een langere termijn op vergelijkbare basis worden verzameld. Maar deze recente methode wisseling bemoeilijkt de vergelijking van gegevens over tijd. Vaak gaan onderzoeksbureaus verschillen tussen scores in oude en nieuwe situatie naast elkaar zetten. Maar probleem bij deze vergelijking is dat twee elementen zijn veranderd: (a) de steekproef (van vers/ad hoc naar panel) en (b) de methode van dataverzameling (van CATI of CAPI naar CASI). En bij eventuele verschillen tussen oude en nieuwe cijfers weten wij dus niet of die veroorzaakt worden door de steekproefprocedure of methode van dataverzameling (modaliteit).

Om een beter zicht op het modaliteiteffect te verkrijgen hebben wij samen met TNS NIPO een onderzoek opgezet waarin steekproefeffect en modaliteiteffect zijn te onderscheiden. Er zijn drie volledig gematchte steekproeven uit het access panel van TNS NIPO getrokken en elk van deze steekproeven is voor dezelfde vragen met een andere

methode van dataverzameling benaderd. Zo weten wij dat eventuele verschillen niet zijn toe te schrijven aan de steekproef maar aan de modaliteit. Er zijn drie methoden van dataverzameling vergeleken: CASI (Computer Assisted Self Interviewing), CATI (Computer Assisted Telephone Interviewing) en CAPI (Computer Assisted Personal Interviewing dat wil zeggen face-to-face). In de volgende paragraaf zullen wij schetsen welke verschillen op grond van de literatuur zijn te verwachten. Daarna gaan wij nader in op de details van de onderzoeksopzet en uitvoering, gevolgd door een beschrijving van de resultaten.

2. WELKE VERSCHILLEN TUSSEN CASI, CATI, CAPI ZIJN TE VERWACHTEN?

In CASI neemt de pc de rol van de interviewer over, respondenten lezen zelf de vragen op het scherm en voeren de antwoorden in door codes in te tikken of verbale antwoorden op open vragen te geven. Op deze wijze worden de voordelen van de klassieke schriftelijke enquête gecombineerd met de mogelijkheid van complexe routings die in schriftelijk onderzoek niet mogelijk waren. Uit de literatuur rond dataverzameling kunnen we een aantal mogelijke effecten afleiden. Bij deze verwachtingen gaat het om verschillen tussen CASI versus CATI/CAPI.

2.1. Meer privacy en anonimiteit

De respondenten ervaren bij CASI meer privacy en anonimiteit zodat zij geneigd zijn meer over zichzelf prijs te geven en er minder sprake is van een sociale wenselijkheid vertekening (de Leeuw et al., 2000).

Omdat men eerlijker en openhartiger is treden twee effecten op: (a) minder geaccepteerd sociaal gedrag zoals veel alcohol drinken of drugs gebruiken wordt meer toegegeven en (b) aan de andere kant wordt sociaal gewaardeerd gedrag zoals veel sporten of veel culturele belangstelling hebben minder naar voren gebracht. De Leeuw et al. (2000, p.2) zeggen nog dat behalve dit aspect van vertrouwelijkheid ook het volgende speelt: “the use of a computer itself may enhance the feeling of privacy. After an answer is given, the answer disappears from the screen, while an answer that is written down remains on the paper for everyone to see”. Beide aspecten van minder sociale wenselijkheid (meer rapporteren van sociaal onwenselijk gedrag, minder rapporteren van sociaal wenselijk gedrag) zullen wij kort afzonderlijk behandelen.

2.2.1. Via CASI meer rapportage van sociaal onwenselijk gedrag

De algemene bevinding in de literatuur is dat hoe gevoeliger de gevraagde informatie des te minder de sociale wenselijkheid vertekening bij CASI. Veel onderzoek is gedaan op het terrein van de geestelijke gezondheidszorg.

Epstein et al. (2001) concluderen dat respondenten meer ‘mental health symptoms’ rapporteren als zij via CASI worden ondervraagd. En voorts zeggen zij (p.543): “Similar results have been reported when comparing CASI to other forms of survey administration with regard to sensitive topics, such as male-male sex, injection drug use, drug use in general, and number of sexual partners among females”. En Newman et al. (2002, p. 294) concluderen: “comparisons of CASI with face-to-face interviewing have concluded that subjects completing computer interviews disclose more socially undesirable

attitudes, facts and behaviours”. Poynter en Conely (2003, p.175-176) constateren grote verschillen bij vragen over verkeersgevaarlijk gedrag. Op de vraag of men de losse mobiele telefoon onder het rijden gebruikt antwoordde in de UK via CASI 37% dat wel eens te doen, terwijl dat percentage via CATI gemeten op 21% lag.

2.1.2. Via CASI minder rapportage van sociaal wenselijk gedrag

De keerzijde van de medaille is dat respondenten via CASI minder ‘opscheppen of overdrijven’ als het gaat om sociaal wenselijk gedrag. Respondenten willen zichzelf in een bepaald licht presenteren. Men toont hogere scores op zelfvertrouwen in een situatie met de enquêteur, doet meer aan sport, gebruikt vaker een pc (Mosmann, 1998, p.39). Dit verschijnsel wordt ook wel ‘social apprehension’ of ‘evaluation apprehension’ genoemd. Bronner, Tchaoussoglou en Ross (2003) rapporteren een onderzoek dat uitgevoerd is om sociale wenselijkheid te onderzoeken in relatie tot gerapporteerde mediaconsumptie. In een studie die NOM samen met TNS NIPO uitvoerde is gekozen voor een opzet waarbij *dezelfde* respondenten twee keer zijn benaderd, de eerste keer met methode x en de tweede keer met methode y. Eventuele verschillen kunnen aldus niet door steekproeffluctuaties worden bepaald.

De volgende procedure is gevolgd:

- aan ongeveer 500 respondenten is via CASI gevraagd welke tijdschriften en dagbladen zij lezen. Zij komen uit het access panel en zijn ondervraagd door TNS NIPO
- enkele weken later zijn dezelfde respondenten benaderd met dezelfde vragen, maar nu via CATI en door een ander bureau namelijk Veldkamp
- de respondenten werd niet verteld dat er een relatie was tussen de twee metingen
- omdat de paar weken verschil effect zouden kunnen hebben is parallel aan de CATI meting een verse CASI steekproef ondervraagd. De bereiksniveaus van de verse CASI steekproef en de eerste CASI steekproef uit het experiment verschilden niet significant van elkaar zodat we kunnen concluderen dat het is geoorloofd in het experiment de CASI en CATI meting te vergelijken. Er hebben zich in de tussenliggende weken geen ontwikkelingen voorgedaan. Nu volgen de conclusies. Bij dagbladen en bijvoorbeeld computerbladen zien we geen verschillen, maar bij opiniebladen en zakenbladen was er wel een duidelijk verschil: via CATI met een enquêteur hebben deze bladen duidelijk hogere bereikcijfers dan via CASI zonder een enquêteur. Bronner et al. (2003, p. 129) concluderen: “to an interviewer the respondents want to state that they read news weeklies and business magazines because that sounds more ‘comme il faut’.”

2.2. Via CASI minder item non-respons, meer antwoorden op open vragen, minder gebruik extreme categorieën

Interviewers, zeker als zij op stukloon betaald werden betaald, hadden nogal eens de neiging in zeer hoog tempo door de vragenlijst te gaan. In CASI kan de respondent alle tijd nemen om rustig over de antwoorden na te denken. Daardoor is er ook minder item non-respons (de Leeuw et al. 2000).

Oosterveld en Willems (2003) rapporteren een groot verschil bij open vragen. CASI levert in vergelijking tot CATI meer antwoorden en meer gedetailleerde antwoorden op. Allison en O’Konis (2002, p.91/92) namen waar dat CATI respondenten meer de neiging hebben extreme categorieën op een schaal te gebruiken (helemaal eens of aan de

andere kant helemaal oneens). Terwijl CASI respondenten meer gebruik maken van de minder extreme categorieën. Dillman et al. (2001) vonden hetzelfde in hun studie en 'recommend great care in the use of classic scales with anchored points (e.g. where 1 equals strongly agree and 5 equals strongly disagree)'.

2.3. Via CASI toonmateriaal en visuele stimuli benutten

Vergeleken met CATI kunnen bij CASI wel afbeeldingen worden getoond. In media onderzoek (Bronner et al. 2003) blijkt dat het tonen van logo's of reclame-uitingen leidt tot hogere herkenning ('recognition'). En bij leesgedrag kan CATI leiden tot titelverwarring omdat geen logo's kunnen worden getoond (bijv. verschillen tussen Kampioen en de Waterkampioen).

2.4. Via CASI minder tijdsdruk, meer deliberatietijd

Zoals ook al bij punt 2 naar voren is gebracht kan de respondent bij CASI de vragen in een zelf gekozen tijdbestek beantwoorden. Dat levert meer deliberatietijd op, die men kan benutten door dieper in het geheugen te graven. Dat impliceert meer meningen (minder 'weet niet').

Ook kan men spontaan dan meer merken noemen of zich meer mediaconsumptie momenten herinneren. "This way of interviewing captures more casual reading, exactly the kind most likely to get passed over in a time-rushed interview" (Lindner, Mallett, 1995). In de CASI procedure zijn mensen meer bereid cognitieve inspanning te leveren en zij tonen een grotere bereidheid over een antwoord na te denken. Mosmann (1998) concludeert dat interviews zonder enquêteur en zonder tijdsdruk hogere cijfers qua spontane merkbekendheid opleveren.

Een CATI gebruiker "would be horrified to see how forgetful consumers had become about his product. And he would presumably hardly find it comforting that this same 'amnesia' was also affecting his competitors" (Mosmann, 1998, p.26).

Al deze verwachtingen uit de literatuur hebben ons gestuurd bij het opstellen van de vragenlijst die benut is voor de vergelijking van CASI, CATI en CAPI.

Meer specifiek hebben de volgende hypothesen ons geleid. Als we CASI (zonder enquêteur) aan de ene kant vergelijken met CATI en CAPI aan de andere kant (met enquêteur) dan zullen we:

- H1: via CASI hogere percentages bij sociaal onwenselijk gedrag vinden
- H2: via CASI lagere percentages bij sociaal wenselijk gedrag vinden
- H3: via CASI minder item non-respons en weet niet vinden en minder gebruik van extreme categorieën op een rating-scale tegenkomen
- H4: via CASI hogere herkenningsscores vinden door het gebruik van toonmateriaal
- H5: via CASI meer oordelen en meer genuanceerde oordelen verkrijgen.

3. ONDERZOEKSOPZET

Zoals al eerder gesteld in dit artikel was een belangrijke voorwaarde voor onze opzet dat steekproeffecten onderscheiden kunnen worden van modaliteitseffecten.

Om deze effecten van elkaar te kunnen scheiden staan twee wegen open:

- (a) *dezelfde* respondenten worden met verschillende methoden benaderd
- (b) gematchte steekproeven met *verschillende* respondenten worden met verschillende methoden benaderd.

De eerste aanpak is gevolgd in de eerder beschreven studie van Bronner et al. (2003), maar daar ging het om de vergelijking tussen twee modaliteiten namelijk CASI en CATI. En dat achten wij haalbaar bij dezelfde respondenten. Maar in de voorliggende test gaat het om drie modaliteiten en dat lijkt ons te veel om bij één respondent toe te passen. Na de tweede keer verwachten wij dat wantrouwen ontstaat ('alweer diezelfde vragen?') of doorziet men de opzet. Daarom hebben wij ervoor gekozen om voor een verschiltest van de modaliteiten te kiezen voor drie gematchte steekproeven. Dat impliceert dat de steekproeven volledig vergelijkbaar moeten zijn.

Uit het access panel van TNS NIPO zijn op dezelfde wijze drie parallel steekproeven getrokken uit pc bezittende huishoudens (de officiële naam van het access panel is TNS NIPObase). De database omvat in 2005 ruim 200.000 potentiële respondenten. In de database is veel informatie over alle respondenten beschikbaar, bijvoorbeeld socio-demografische kenmerken, waarde-oriëntatie (WIN), informatie over mediaconsumptie, merkgebruik en bezittingen. Omdat zoveel gegevens vooraf bekend zijn, kan vooraf uitvoerig gecontroleerd worden of de steekproeven niet van elkaar afwijken.

Aldus is een optimale matching gerealiseerd.

Het streven was drie steekproeven van elk 300 waarnemingen te realiseren. De uiteindelijk netto aantallen zijn:

- CASI n = 283
- CAPI n = 287
- CATI n = 273

Het veldwerk is uitgevoerd in 2003.

Bij CASI hebben de respondenten op een zelf gekozen moment achter de pc de vragen beantwoord. Bij CATI werd de respondent gebeld vanaf de centrale telefonische unit van TNS NIPO in Amsterdam. En bij CAPI gingen enquêteurs gespreid over het land thuis langs bij de respondenten, stelden mondeling de vragen vanaf het scherm en voerden de antwoorden direct op de pc in.

De drie steekproeven zijn zoveel mogelijk gematched getrokken en kleine verschillen zijn via weging gecorrigeerd. Uiteindelijk zijn de steekproeven op de volgende kenmerken gelijk: geslacht, leeftijd, sociale klasse, regio, bruto jaarinkomen van huishoudens, levensfase, hoogst genoten opleiding hoofdkostwinner, op dit moment werkzaam, beroep hoofdkostwinner.

In de vragenlijst hebben wij een variëteit aan vragen opgenomen deels variërend naar onderwerp (media, consumptiegedrag, maatschappelijk gedrag), deels naar kennis/hou-

ding/gedrag, deels naar antwoord-mogelijkheden (open, gesloten vragen).
In de volgende paragraaf beschrijven wij de resultaten van de uitgevoerde analyses.

4. DE RESULTATEN

Bij de beantwoording van enquête- of interviewvragen kunnen factoren optreden die kunnen leiden tot over- of onderrepresentatie van penetratiecijfers, ontwijken van een eerlijk antwoord, etc. Voorbeelden van dergelijke factoren zijn sociale wenselijkheid, geheugeneffecten en antwoordtendenties. In paragraaf 2 is op basis van de literatuur een aantal verwachtingen gepresenteerd. Deze verwachtingen betreffen met name de afwijkingen van CASI (zonder enquêteur) ten opzichte van CATI en CAPI (met enquêteur). De resultaten groeperen wij in vier onderdelen:

- a. sociaal onwenselijk en wenselijk gedrag
- b. schaaleffecten zoals weet niet, gebruik extreme categorieën
- c. effect van het gebruik van visuele stimuli
- d. effect van deliberatietijd

De resultaten zijn weergegeven in de tabellen 1 tot en met 7. De toelichting op de betekenis van de variabelen in de tabellen 1 tot en met 7 is opgenomen in de appendix. Voor de analyses is gebruik gemaakt van Anova en incidenteel van tabelanalyse.

4.1. Sociaal onwenselijk gedrag

Omdat bij CASI de respondent in een betrekkelijk anonieme situatie vragen kan beantwoorden en geen telefonisch of face-to-face contact heeft met de interviewer zal sociale wenselijkheid een geringere rol spelen. Daardoor zullen sociaal onwenselijke antwoorden bij CASI hoger zijn dan bij de andere interview instrumenten (zie hypothese 1).

In tabel 1 is te zien dat respondenten zoals verwacht bij CASI significant meer verkeersovertredingen rapporteren dan bij CAPI en CATI. Dit is in overeenstemming met het eerder geciteerde onderzoek van Poynter en Conely (2003). De variabele Sumv1(31_1 to v31_6) is een somscore over een aantal overtredingen zoals te hard rijden, bellen tijdens het rijden, voetgangers geen voorrang verlenen.

Als het gaat om de rapportage van ziektes, sum(v37_1 to v38_1), de somscore over het lijden ziektes als epilepsie, astma, reuma blijkt het gerapporteerde aantal ziektes in tegenstelling tot de verwachting bij CATI het hoogst te zijn en bij CASI het laagst. Blijkbaar speelt bij het vermelden aan welke ziektes men lijdt sociale wenselijkheid geen rol.

In tabel 2 zijn de percentages weergegeven die respondenten op de vragen antwoorden als wordt gevraagd naar een bepaald gedrag of situatie, o.a.:

- kijken naar GTST, NOS Journaal
- gebruik marihuana
- mate waarin men kan rondkomen van het gezinsinkomen. (comfortabel, kan rondkomen, moeilijk rond te komen, heel moeilijk rond te komen, wil niet zeggen)

Zoals verwacht is het aantal respondenten dat zegt GTST te kijken significant groter bij CASI dan bij CAPI en CATI. Ook hier lijkt op te gaan dat naarmate de fysieke afstand

tot een interviewer groter is, en de invloed van sociale wenselijkheid kleiner, men meer geneigd is het echte kijkgedrag te melden. Zoals verwacht zijn er geen instrumentverschillen met betrekking tot de mate waarin men zegt naar het NOS Journaal te kijken. Het percentage respondenten dat moeilijk tot zeer moeilijk zegt rond te kunnen komen is zoals verwacht het hoogst bij CASI. Dat geeft men blijkbaar gemakkelijker toe in een situatie zonder enquêteur. Eveneens conform verwachtingen is het percentage mensen dat zegt comfortabel rond te kunnen komen met het huidige inkomen het hoogst bij CAPI. In het gebruik van marihuana zijn geen verschillen te rapporteren.

Tabel 1. Verschillen tussen CASI, CAPI en CATI (ANOVA).

Gedrag gevoelig voor sociale wenselijkheid	Gemiddelden				df	F ratio	Sign.
	CASI	CAPI	CATI	Totaal			p<
sumv(31_1 to v31_6) wets-overtredingen	2,09	1,77	1,47	1,78	2	15,81	0,00
sum(v38_1 to v38_7) ziektes	0,34	0,41	0,54	0,43	2	5,42	0,01

Tabel 2. Verschillen tussen CASI, CAPI en CATI (chikwadraat toetsen).

Gedrag gevoelig voor sociale wenselijkheid	Percentages				df	X ²	Sign.
	CASI	CAPI	CATI	Totaal			P <
Gebruikt u weleens marihuana	1,4%	3,9%	4,3%	3,2%	2	4,33	0,09
Kijkt u wel eens naar GTST	43,8%	25,4%	31,8%	33,6%	2	21,72	0,00
Kijkt u naar het NOS journaal	84,7%	80,4%	83,2%	82,7%	3	1,86	0,39
Hoe rondkomen van inkomen? (% Comfortabel)	29,9%	47,7%	37,5%	38,4%	8	27,36	0,00
Hoe rondkomen van inkomen? (% Moeilijk tot zeer moeilijk)	22,8%	13,6%	14,4%	16,9 %	8	22,43	0,00

4.2. Sociaal wenselijk gedrag

De keerzijde van de medaille is dat te verwachten is dat CASI juist lagere scores oplevert bij sociaal zeer wenselijk gedrag. Evaluatie gevoeligheid betreft de neiging van mensen zich beter voor te doen of beter of “slimmer” te willen lijken dan ze zijn wanneer ze zich geëvalueerd weten, zoals in interviewsituaties.

Bij vragen naar wie verantwoordelijk is voor aankoopbeslissingen in het gezin kan dit verschijnsel zich voordoen.

Er is een aantal vragen beschikbaar waarin gevraagd wordt wie de belangrijkste beslisser is voor aankopen zoals de auto, apparatuur, kleding, voeding. De antwoordmogelijkheden op deze vragen zijn “IK” (de geïnterviewde), de partner of beiden in gelijke mate. Voor deze vragen is een somscore berekend voor het aantal keren dat een respondent zichzelf (“IK”) aanmerkte als de belangrijkste beslisser.

Omdat het geslacht van de respondent van invloed zal zijn op de rol die men zichzelf toekent bij aankoopbeslissingen zijn separaat de gemiddelde somscores (ANOVA)

berekend voor sexe, de drie interview instrumenten en de interactie tussen de interviewinstrumenten.

Tabel 3. ANOVA (Voor toetsing zie tabel 4).

Gemiddelden				
Wie is belangrijkste beslisser	CASI	CAPI	CATI	Totaal
Mannen: gemiddeld aantal "IK" antwoorden	2,41	2,51	2,43	2,45
Vrouwen: gemiddeld aantal "IK" antwoorden	2,54	3,15	3,12	2,93
Totaal	2,47	2,83	2,77	2,69

In bovenstaande tabel 3 is te zien dat overall CAPI het gemiddeld hoogste aantal 'Ik' antwoorden oplevert en dat vrouwen zich vaker als de belangrijkste beslisser zien. Zowel het interviewinstrument als sexe hebben een significant effect op het aantal keren dat respondenten zich als belangrijkste beslisser zien. Daarnaast blijkt er ook een significant interactie-effect van instrument* sexe te zijn. Dit is weergegeven in figuur 1 en tabel 4. Hieruit blijkt dat het 'evaluation apprehension effect' voor mannen te verwaarlozen is.

Tabel 4. ANOVA: Tests voor Tussen-Subjecten Effecten.

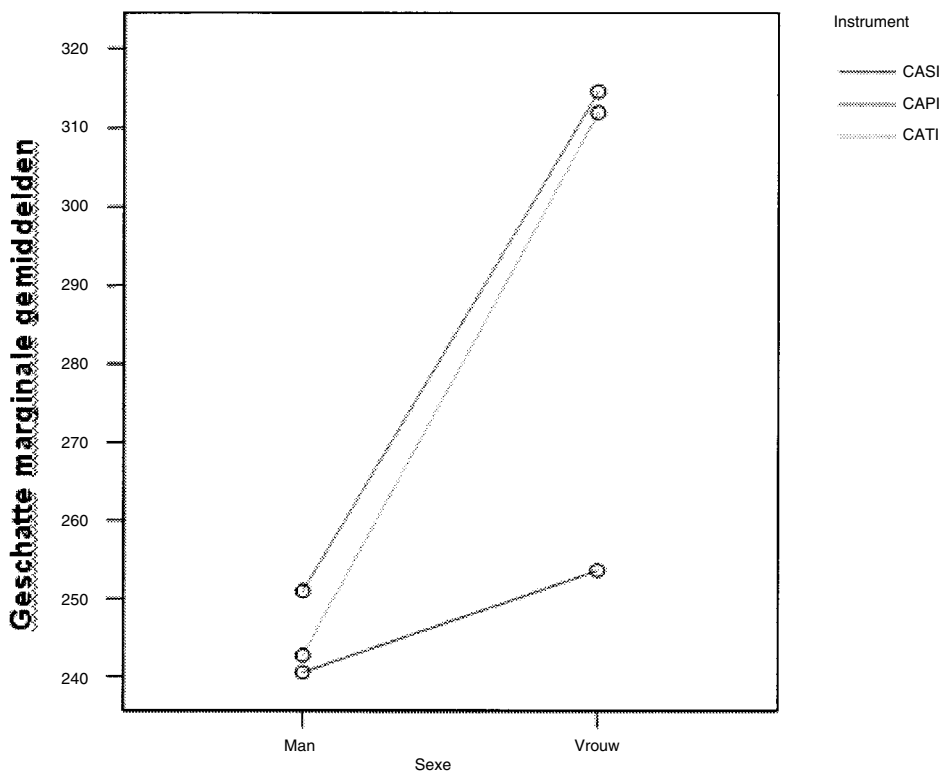
Afhankelijke Variable: SUM (V24_1 TO V24_6) EQ 1 (= de geïnterviewde)					Sign.
Bron	Type III Kwadratensom	df	Gemiddelde kwadratensom	F	P<
Gecorrigeerd Model	60,55	5	12,11	8,89	0,00
Intercept	4550,47	1	4550,47	3339,79	0,84
Type	16,29	2	8,15	5,98	0,02
Sexe	37,19	1	37,19	27,29	0,04
Type * sexe	10,63	2	5,32	3,90	0,01
Fout	858,38	630	1,36		
Totaal	5479,00	636			
Gecorrigeerd Totaal	918,92	635			

Er zijn geen significante verschillen tussen de interviewinstrumenten voor mannen. Maar voor vrouwen blijkt de rol die zij zichzelf toemeten als belangrijkste beslisser voor aankopen in het gezin bij CASI significant lager te zijn dan bij CAPI en CATI (zie fig. 1).

4.3. Schaaffecten

Bij de beantwoording van bijvoorbeeld kortere of langere series van stellingen (met 5 of 7 punts Likert schalen lopend van 1=oneens tot 5 of 7 = zeer mee eens), semantische differentiaal vragen (met categorie schalen vaak lopend van 1 tot 7), rapportcijfers voor merken of producten treden vaak respons-tendenties op, zoals:

- het overwegend gebruik van de extreme antwoordcategorieën,
- het overwegend gebruik van de middenposities van antwoordschalen
- het veelvuldig gebruik van de "weet niet" of "geen oordeel" categorieën.



Figuur 1. Interactie effect van interviewinstrument en sexe.

Verondersteld wordt dat het voorkomen van dergelijke antwoordtendenties samenhangt met het gebruikte interviewinstrument.

Naar verwachting zullen respondenten het gebruik van extremen hanteren om snel op vragen te kunnen reageren. Dit zal zich vooral voordoen bij instrumenten waar de ervaren tijdsdruk groter is, zoals bij CAPI en CATI. Feitelijk gebruiken respondenten deze categorieschaal dan als een simpele ja-nee schaal.

De in deze analyse betrokken vragen zijn:

- Een semantische differentiaal bestaande uit 13 vragen met 7-puntscategorieschalen voor een groot biermerk. Over deze 13 vragen is het gebruik van de score 1 óf 7 gesommeerd in de variabele v52_xtreme. Voor de middenposities is het aantal keren geteld dat de middenpositie (score 4) is gekozen en gesommeerd in V52_midpos.
- Een achttal stellingen, zoals “winkelen is echt een ontspanning”, “ik praat graag met anderen”, etc, met 5-punts “eens-oneens” categorieschalen. Over deze 8 vragen is het gebruik van de score 1 óf 5 gesommeerd in de variabele v8_xtreme. Voor de middenposities is het aantal keren geteld dat de middenpositie (score 3) is gekozen en gesommeerd in V8_midpos
- Rapportcijfers voor een zestal bekende Nederlandse merken. Over deze 6 vragen is het gebruik van de score 1 of 10 gesommeerd in de variabele v6_xtreme. Voor de middenposities is het aantal keren geteld dat de middenpositie (score 5 óf 6) is geko-

zen en gesommeerd in V6_midpos. Voor deze 6 vragen is ook het aantal keren geteld dat de “weet niet” categorie is gekozen en gesommeerd in V6som_wn.

- Zeven vragen m.b.t. verkeersovertredingen (te hard rijden, bellen onder het rijden, etc). Voor deze vragen is het aantal keren geteld dat de “weet niet” categorie is gekozen en gesommeerd in somv31_WN.
- Voor zeven rapportcijfers over tv programma’s cijfers zoals (GTST, NOS Journaal, Breekijzer, etc. is het aantal “weet niet” antwoorden gesommeerd in V11som_wn.
- Van een drietal vragen over de *rechtvaardiging van de oorlog in Irak, de aanleg van de Betuwelijn* en het *vertrouwen in de Nederlandse regering* is het aantal keren dat “weet niet” of “geen mening” is gekozen, gesommeerd in vsum_gewetensvragen_WN.

In tabel 5 is te zien dat het gebruik van de extreme antwoordposities het laagst is bij CASI en afwisselend hoger is bij CAPI en CATI. Het gebruik van de middenposities op de hierboven vermelde vragen komt vaker voor bij CASI.

Ook het gebruik van de ‘weet niet’ antwoordmogelijkheid komt vaker voor bij CASI. Kennelijk staan respondenten bij CASI minder onder druk snel een antwoord te geven of een antwoord te geven terwijl ze eigenlijk geen mening hebben.

Tabel 5. Verschillen tussen CASI, CAPI en CATI (ANOVA).

	Gemiddelden			Sign.			
	CASI	CAPI	CATI	Totaal	df	F	p <
Extreme scores							
v52_xtreme	2,54	3,15	3,49	3,06	2	7,21	0,00
V6_xtreme	2,16	2,63	2,01	2,26	2	8,04	0,00
V8_xtreme	0,14	0,09	0,14	0,13	2	1,07	0,35
all_xtreme	4,84	5,87	5,63	5,44	2	5,42	0,01
Scores op de middenposities							
V6_midpos	1,92	1,42	1,25	1,53	2	21,91	0,35
V8_midpos	1,62	1,65	1,78	1,68	2	1,41	0,24
v52_midpos	3,24	2,97	1,99	2,73	2	16,55	0,00
allmidpoint	6,78	6,03	5,01	5,94	2	19,01	0,00
Som weet niet antwoorden							
V6som_wn (uitspraken, oa. Winkelen)	0,01	0,01	0,00	0,01	2	1,84	0,16
V11som_wn (cijfers tv programma’s)	0,58	0,43	0,21	0,40	2	17,63	0,00
somv31_WN (overtredingen)	0,05	0,07	0,04	0,05	2	0,67	0,51
sum_gewetensvragen_WN (betuwe, iraq, rondkomen, pc)	0,14	0,10	0,04	0,09	2	5,79	0,00
wn_all (WN= weet niet)	0,71	0,59	0,26	0,52	2	14,59	0,00

4.4. Effect gebruik visuele stimuli en antwoordtijd

Het gebruik van visuele stimuli kan ertoe leiden dat bijvoorbeeld herinnering van reclame bij CASI hoger zal zijn (bij CASI kunnen beelden getoond worden die overeenkomen met de beelden die tijdens televisiereclame zijn getoond). Bovendien zal de beschikbare antwoordtijd er naar verwachting toe leiden (bij CASI kan de respondent

immers het eigen tempo bepalen) dat de respondent meer tijd heeft om het geheugen te raadplegen, hetgeen tot hogere herinneringscijfers kan leiden. In tabel 6 is het aantal respondenten weergegeven dat claimt de Rexona spot (deze spot hebben wij als enige voorbeeldspot in het onderzoek opgenomen) gezien te hebben. De herinnering in de geholpen en niet geholpen conditie is het hoogst voor CASI.

Tabel 6. Resultaten tabel analyses.

Rexona spot	Gemiddelden			Sign.			
	CASI	CAPI	CATI	Totaal	df	X ²	P <
Rexonaspot gezien (ja)v13	87,9%	83,6%	79,7%	83,7%	2	7,01	0,00
Heeft u deze Rexona tv-spot de afgelopen weken gezien? V19	74,7%	47,5%	48,9%	57,0%	4	99,17	0,00

4.5. Effect deliberatietijd

De beschikbare denktijd zal er naar verwachting toe leiden (bij CASI kan de respondent immers het eigen tempo bepalen) dat vragen die enig nadenken behoeven (bijvoorbeeld vragen met betrekking tot politieke of maatschappelijke onderwerpen) de respondent vaker een (afgewogen) oordeel heeft en bijvoorbeeld minder vaak ‘weet niet’ of ‘geen oordeel’ zal kiezen. Vragen waarvoor de antwoord bedenktijd naar verwachting een rol zal spelen (‘gewetensvragen’) zijn:

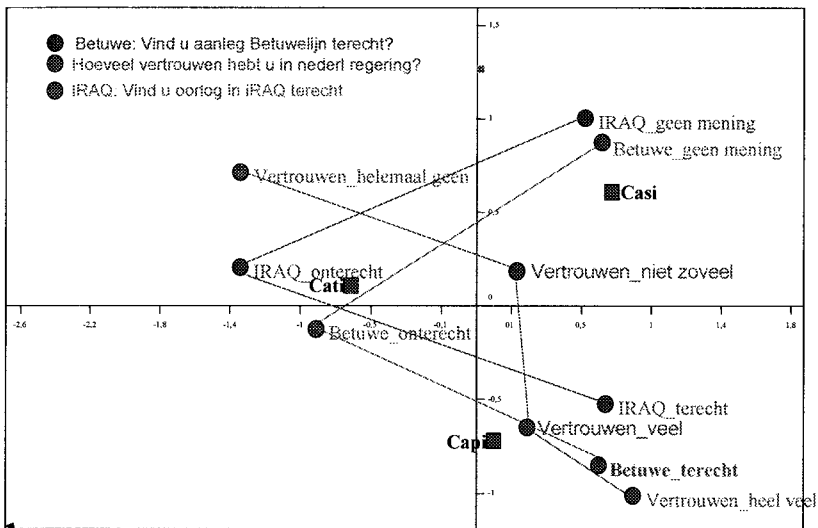
- vindt U de oorlog in Irak terecht? (terecht, onterecht, geen mening)
- vindt u de aanleg van de Betuwelijn terecht? (terecht, onterecht, geen mening)
- hoeveel vertrouwen heeft u in de Nederlandse regering? (heel veel, veel, vertrouwen, niet zoveel, helemaal geen, weet niet).

In tabel 7 is duidelijk dat bij het gebruik van CASI men aanzienlijk minder vaak ‘onterecht’ kiest voor de oorlog in Irak en de aanleg van de Betuwelijn en minder vaak “groot vertrouwen” zegt te hebben in de Nederlandse regering.

Tabel 7. Resultaten tabel analyses.

Gewetensvragen	Gemiddelden				Sign.			
	CASI	CAPI	CATI	Total	df	X ²	P <	
Oorlog in Irak terecht?								
% Onterecht->	29,9%	40,4%	42,9%	37,7%	6	20,98	0,00	
Aanleg Betuwelijn terecht?								
% Onterecht->	48,9%	56,7%	60,5%	55,4%	6	27,21	0,00	
Vertrouwen Nederlandse regering (% op de scores 1+2= groot vertrouwen)	28,3%	34,5%	34,8%	32,2%	2	20,69	0,01	

De samenhang tussen de drie interviewinstrumenten en de 3 hiervoor besproken ‘gewetensvragen’ is gevisualiseerd in onderstaande multiple correspondentie plot. Geconcludeerd kan worden dat CASI zoals in de voorgaande tabel relatief minder ‘onterecht’ antwoorden oplevert evenals minder vaak het antwoord: “groot vertrouwen in de



Figuur 2. Homals plot voor 3 gewetensvragen en de 3 interviewinstrumenten.

Nederlandse regering”. Met CATI worden meer ‘onterecht’ antwoorden verkregen evenals het antwoord “helemaal geen vertrouwen in de Nederlandse regering”. Tegen de verwachting in levert CASI echter vaker ‘geen mening’ antwoorden op. Blijkbaar leidt meer ‘bedenktheid’ ertoe dat men genuanceerder gaat denken of dat men minder vaak een uitgesproken oordeel heeft. CAPI levert meer ‘terecht’ antwoorden op evenals een groter aantal respondenten met ‘heel veel vertrouwen’ in de Nederlandse regering.

5. CONCLUSIES

Eén van de kernactiviteiten binnen het marktonderzoek is dataverzameling via onder-vraging. Een drietal factoren heeft deze activiteit de laatste jaren sterk beïnvloed. Lage respons percentages, sterk stijgende interviewers kosten en het groeiende idee dat de respondent als een ‘echte’ klant moet worden behandeld hebben een verschuiving in methoden van dataverzameling teweeg gebracht. Deze ontwikkelingen leidden in de eerste plaats tot de opzet van access panels door onderzoeksbureaus. Leden van zo’n panel geven aan op gezette tijden aan onderzoek mee te willen doen. Informatie over die respondenten is beschikbaar zodat sociodemografische gegevens niet telkens gevraagd behoeven te worden en zodat respondenten met bepaalde kenmerken getrokken kunnen worden (lezers van blad x, gebruikers van product y) zonder kostenverslindende screening. In principe kunnen de leden van een dergelijk access panel met alle methoden van dataverzameling worden benaderd. Maar in de praktijk is de combinatie van access panel met CASI (Computer Assisted Self Interviewing) populair. Respondenten vullen op een zelf gekozen moment de vragenlijst achter de pc in. In longitudinale projecten voor overheid en bedrijfsleven is vaak de laatste paar jaar een overstap gemaakt van CATI (Computer Assisted Telephone Interviewing) of CAPI (Computer Assisted Personal Interviewing) naar CASI. Van een interview situatie met enquêteur naar een interview situatie zonder enquêteur. Maar deze recente methode wisseling

bemoeilijkt een vergelijking over tijd. Vaak zetten onderzoeksbureaus verschillen tussen oude en nieuwe scores naast elkaar. Maar probleem is bij deze vergelijking dat twee elementen zijn veranderd: de steekproef (van ad hoc naar panel) en de methode van dataverzameling (van CATI/CAPI naar CASI). En bij eventuele verschillen tussen oude en nieuwe cijfers weten wij dus niet of die veroorzaakt worden door de steekproefprocedure of methode van dataverzameling (modaliteit). Vandaar dat wij samen met TNS NIPO een onderzoek hebben opgezet waarin het modaliteiteffect is te onderscheiden. Op basis van de literatuur is een aantal verwachtingen geformuleerd over de verschillen tussen CASI aan de ene kant en CATI/CAPI aan de andere kant. In zijn algemeenheid ondersteunen de resultaten van onze analyses deze verwachtingen.

De respondenten ervaren bij CASI meer privacy en anonimiteit zodat zij geneigd zijn eerlijker over sociaal onwenselijke gedragingen te rapporteren, met andere woorden zij geven eerder bepaalde gedragingen toe. De keerzijde van de medaille is dat zij via CASI minder 'opscheppen' over sociaal zeer wenselijk gedrag. Voor beide verschijnselen vonden wij in onze data ondersteuning. Er worden via CASI bijvoorbeeld significant meer verkeersovertredingen gerapporteerd zoals te hard rijden, bellen tijdens het rijden, voetgangers geen voorrang verlenen. En het omgekeerde vinden wij bij vragen over de invloed in de 'gezinspraak' als het om productkeuzen in het gezin gaat. Voor vrouwen blijkt de rol die zij zichzelf toemeten als belangrijkste beslisser voor aankopen in het gezin bij CASI significant lager te zijn dan bij CATI en CAPI.

Wat betreft schaafeffecten ondersteunt ons onderzoek de bevinding uit de literatuur dat bij CASI minder extreme antwoordposities worden gebruikt. Respondenten hanteren het gebruik van extremen om snel op vragen te kunnen reageren. Dit zal zich vooral voordoen bij instrumenten waar de ervaren tijdsdruk groter is, zoals bij CAPI en CATI. Het gebruik van visuele stimuli zoals de mogelijkheid om binnen CASI reclamespots te laten zien verhoogt de herkenningsscores ('recognition') van reclame.

De beschikbare deliberatietijd binnen CASI leidt ertoe dat mensen dieper in het geheugen kunnen graven en bijvoorbeeld meer merken spontaan kunnen noemen, maar blijkt ook te leiden tot meer afgewogen oordelen over maatschappelijke problemen.

Al met al blijkt de aanwezigheid of afwezigheid van een interviewer wel degelijk invloed op de antwoorden op bepaalde vragen te hebben. Dat impliceert dat voorzichtigheid is geboden bij de vergelijking over tijd van gegevens die met verschillende methoden van datacollectie zijn verzameld. Ook is eerder in dit artikel gesignaleerd dat de respondent steeds meer als 'klant' wordt behandeld. En dat leidt er in sommige onderzoeksopzetten toe dat aan hem/haar de keuze wordt geboden zelf de methode van dataverzameling te kiezen via welke hij of zij ondervraagd wil worden (mixed mode of multi-mode, zie ook De Leeuw, 2005). Ook bij een dergelijke opzet is bij de samenvoeging van data die via verschillende methoden zijn verzameld voorzichtigheid geboden!

Appendix

Gedrag gevoelig voor sociale wenselijkheid

Bij bepaalde vragen speelt naar verwachting sociale wenselijkheid een rol bij beantwoording.

Het betreft een aantal vragen, zoals marihuana-gebruik, te hard rijden, bellen onder het rijden, rechts inhalen, etc. De “ja” antwoorden op de vragen (v31) zijn gesommeerd

Ook de vragen m.b.t. het lijden aan bepaalde ziektes (V 38) zijn gesommeerd

Somscore (Vragen v31_1 tot en met v31_6 indien ja) **begaan van wetsovertredingen**

Somscore (v38_1 to v38_7) **lijden aan bepaalde ziektes**

Antwoord tendenties

Voor vragen die onderdeel zijn van bepaalde schalen mbt opinies, semantische differentiaal, schoolcijfers en 8 stellingen (zoals ik vind winkelen leuk, ik praat graag met anderen, etc) is over de categorieschalen een sommatie gemaakt van het aantal keren dat een extreme antwoord positie werd gekozen, dus een sommatie van het aantal enen of het vijven, tien en of zevens (1=zeer goed tot 5 zeer slecht, rapportcijfers 1 tot 10, semantische differentiaal met scores van 1 tot 7).

Zo is ook een sommatie gemaakt over de middenposities (te weten 3 bij 5-punt schalen en 5/6 bij de 10-puntschalen en 4 bij de 7-puntschalen (semantische differentiaal)

Extreme scores

Som extremen over 13 semantische differentiaal vragen voor een groot biermerk (v52_xtreme)

Som extremen over 8 stellingen (V6_xtreme)

Som extremen over rapportcijfers van zestal bekende Nederlandse merken (V8_xtreme)

all_xtreme is de somscore van alle extreme antwoorden (v52_xtreme + v6_xtreme + v8_xtreme)

Scores op de middenposities

Som middenposities over 13 semantische differentiaal vragen voor een groot biermerk v52_midpos

Som middenposities over 8 stellingen V6_midpos

Som middenposities over rapportcijfers van zestal bekende Nederlandse merken V8_midpos

Allmidpoint is all_xtreme is de somscore van alle middenpositie antwoorden

Som “weet niet” antwoorden

Voor een aantal vragen is bovendien nagegaan of het gebruik van de “weet niet” categorie afhankelijk is van het gebruikte interviewinstrument.

V6som_wn (som aantal keren “weet niet” m.b.t. 8 uitspraken, o.a. winkelen)

V11som_wn (som aantal keren “weet niet” cijfers over 7 tv programma’s)

somv31_WN (som van het aantal keren “weet niet” bij de 7 vragen m.b.t. wetsovertredingen)

sum_gewetensvragen_WN (som van het aantal keren weet niet bij de zg. gewetensvragen, zie hierna)

wn_all (is som van v6som_wn; v11som_wn; somv31_wn; sum_gewetensvragen_wn)

Rexona spot

Rexonaspot gezien (ja)v13

Heeft u deze Rexona tv-spot de afgelopen weken gezien? V19

(Gewogen) Somscores

factorscore(v52_i, Heinekenvragen)

sum(v6_1 to v6_ > 3) stellingen winkelen

v6_1_8/8 allscores 6=mis uitspraken winkelen etc

sum(v8_1 to v6_ > 6) som van aantal keren dat men een voldoende geeft aan 6 bekende merken

sum(v8_1 to v6_ < 6) som van aantal keren dat men een onvoldoende geeft aan 6 bekende merken

sum(v10_1 to v10_6 = ja) somscore over het aantal (max 6) televisieprogramma's waarnaar men kijkt

sum(v31_1 to v31_6 = ja) som van aantal keren dat men een wetsovertreding begaat (max 8)

sum(v3_1 to v3_8) som van aantal keren (max 8) dat men bepaalde gelegenheid (cafe, etc) bezoekt

sum(v4_1 to v4_6) som aantal keren dat men bepaalde activiteit (eco-product kopen, etc) verricht

sumv(v41_1 to v41_9) Orange KPN

Gewetens vragen

Oorlog in Iraq terecht?

Aanleg betuwelijn terecht

Hoe rondkomen van inkomen? Comfortabel

Hoe rondkomen van inkomen? Moeilijk

Vertrouwen Nederlandse regering :groot (score 1+2)

Wie is belangrijkste beslisser in gezin?

Over een aantal product aankopen is nagegaan wie zegt de belangrijkste beslisser te zijn. Het betreft producten die traditioneel tot het beslissingsdomein van de vrouw behoren, zoals kinderkleding, supermarktinkopen, etc en aankopen die traditioneel tot het beslissingsdomein van de man behoren, zoals aanschaf auto, hobbygereedschap, etc. Over de verschillende producten is een sommatie gemaakt van het aantal keren dat de geïnterviewde zichzelf (**IK**) als belangrijkste beslisser aanmerkt.

sum(v24_1 to v24_6 eq 1) aantal keren dat geïnterviewde zichzelf als belangrijkste beslisser zag ("IK")

LITERATUUR

- Allison, J. & C. O'Konis (2002). If given the choice, *Quirk's Marketing Research Review*, July/Aug issue, 20.
- Bronner, F., C.Tchaoussoglou, R.Ross (2003). The virtual interviewer, *Worldwide Readership Research Symposium*, Cambridge, Mass., 121-130.
- Dillman, D.A., G.Phelps, R.Tortora, K.Swift, J.Kohrell, J.Berck (2001). Response rate measurement differences in mixed mode surveys using mail, telephone, interactive voice response and internet. *AAPOR Annual Conference*, Montreal.
- Epstein, J., P.Barker, L.Kroutil (2001). Mode effects in self-reported mental health data. *Public Opinion Quarterly*, 65, 529-549.
- Havermans, J. (2003). The strength is in online research. *Research World*, 1, 4-6.
- Heer, W. de (1999). International response trends, results of an international survey. *Journal of Official Statistics*, 15, 2, 129-142.
- Leeuw, E. de, J.Hox, S.Kef, M.van Hattum (2000). Computer assisted self-interviewing tailored for special populations: a guide on how to overcome the problems of special interviews and sensitive topics. *Proceedings of the 5th Int. Conf. on Logic and Methodology*, Cologne.
- Leeuw, E. de (2005). To mix or not to mix data collection modes in surveys. *Journal of Official Statistics*, 21.2, 233-255.
- Lindner, G. & D.Mallett (1995). The interviewer effect on readership levels, 7th *Worldwide Readership Research Symposium*, Berlin, 209-217.
- Macer, T. (2003). We seek them here, we seek them there. In: R.Banks et al. (eds.) *Survey and Statistical Computing IV, the impact of technology on the survey process*. Association for survey computing.
- Morton-Williams, J. (1993). *Interviewer approaches*. Aldershot: Dartmouth Publishing.
- Mosmann, H. (1998). Finding the facts: response patterns PAPI-CAPI-CATI. *INRA Deutschland*.
- Newman, J., D. Des Jarlais, C.Turner, J.Gribble, P.Cooley, D.Paone (2002). The differential effects of face-to-face and computer interview modes. *American Journal of Public Health*, 92, 2, 294-297.
- Oosterveld, P. & P.Willems (2003). Two modalities: one answer. *ESOMAR Technovate Conference*, Cannes.
- Poynter, R. & P. Comely (2003). Beyond online panels. *ESOMAR Technovate Conference*, Cannes.
- Stoop, I.A.L. (2005). The hunt for the last respondent: nonresponse in sample surveys. Den Haag: SCP.

6. Landenvergelijkende surveys: *Vergelijking tussen landen en tussen surveys*

I.A.L. STOOP

SAMENVATTING

Landenvergelijkende surveys verschillen van nationale surveys door het doel – het vergelijken van landen – en omdat er sprake is van verschillende talen, verschillende steekproefkaders, verschillende onderzoeksorganisaties en verschillende onderzoekstradities. Deze verschillen bemoeilijken de vergelijkbaarheid van resultaten. Bij het opzetten en analyseren van landenvergelijkende surveys wordt men al gauw geconfronteerd met de vraag of en in hoeverre landen eigenlijk wel vergelijkbaar zijn, hoe vergelijkbaarheid wordt gedefinieerd en of een survey binnen verschillende landen op een vergelijkbare manier kan worden opgezet en uitgevoerd. Om deze vragen te kunnen beantwoorden is kennis over de nationale context en nationale onderzoekstradities noodzakelijk. Meer nog dan nationale surveys vereisen landenvergelijkende surveys uitvoerige documentatie. De hoofdstuk geeft een overzicht van de belangrijkste problemen die zich voor doen bij de opzet van landenvergelijkende surveys, en laat zien hoe geprobeerd is deze op te lossen in het European Social Survey.

1. INTRODUCTIE

In Nederland worden veel gegevens verzameld in het kader van landenvergelijkende surveys. De resultaten van deze surveys zijn niet alleen van belang voor onderzoekers en beleidsmakers die uitspraken willen doen over verschillen binnen Europa, of Europa met de VS willen vergelijken. Internationaal onderzoek wordt steeds belangrijker in een tijd van globalisering en Europese integratie. Zeker nu Europees beleid zich niet alleen meer met kolen en staal, maar ook met discriminatie, kinderopvang en de afstemming van werk en gezinsleven bezighoudt, vormt de natiestaat niet meer de vanzelfsprekende afbakening voor onderzoek en is het steeds noodzakelijker over de grens te kijken. Europa wordt dan het object van onderzoek, of regio's binnen Europa, of de grote steden in Nederland en Duitsland. Ook als de onderzoeksvraag zich op een concreet Nederlands probleem richt, kunnen gegevens uit andere landen inzicht bieden. Hoe verloopt de integratie van minderheden in naburige landen, wat zijn de ervaringen met 'nieuw leren' over de grens, is de kloof tussen burgers en politiek in Nederland groter dan in andere landen en welke oorzaken en verklaringen worden daar genoemd? Europa is een natuurlijk laboratorium, en hoe beter gegevens uit verschillende landen vergelijkbaar zijn, hoe meer ze kunnen bijdragen aan het beantwoorden van belangrijke nationale en cross-nationale vragen.

Dit hoofdstuk gaat over landenvergelijkend onderzoek onder leden van de bevolking,

dus niet onder bedrijven of over systemen of gebaseerd op macro-gegevens. Gegevens over personen kunnen afkomstig zijn uit registraties en administraties of uit surveys. De nadruk ligt hier op surveys. Verder kan er bij landenvergelijkend onderzoek sprake zijn van ex-post harmonisatie, waarbij de bron minder relevant is en gegevens achteraf vergelijkbaar worden gemaakt, en van geharmoniseerde dataverzameling, waar de opzet en vraagstelling – en hopelijk ook de resultaten – zo goed mogelijk vergelijkbaar zijn. Deze laatste procedure wordt vooral gevolgd in onderzoek naar meningen, attitudes en waarden, waarbij de wijze van vraagstelling en formulering van vragen grote effecten kunnen hebben op de resulterende antwoorden (zie bijvoorbeeld De Leeuw, 1992, en Saris, 2003) en harmonisatie achteraf niet mogelijk.

In de volgende paragrafen wordt ingegaan op de problemen die onlosmakelijk verbonden zijn met landenvergelijkend onderzoek. Deze problemen zijn met een stringente methodologische aanpak voor een deel op te lossen, maar voor een deel zal men er mee moeten leren leven. Des te belangrijker is dat duidelijk is waar die problemen optreden en hoe ze de resultaten kunnen beïnvloeden. Paragraaf 2 bevat een overzicht van een aantal landenvergelijkende surveys, en par. 3 van de belangrijkste problemen die optreden bij het streven naar optimale vergelijkbaarheid. In paragraaf 4 wordt ingegaan op methodologische problemen rond steekproeftrekking en nonrespons. In paragraaf 5 wordt ten slotte geprobeerd de besproken problemen enigszins in perspectief te zetten. Het nieuwe European Social Survey krijgt speciale aandacht, omdat in dit onderzoek expliciet optimale vergelijkbaarheid en methodologische vernieuwing wordt nagestreefd.

2. LANDENVERGELIJKEND ONDERZOEK

Landenvergelijkend onderzoek is niet een trend van de laatste jaren en niet specifiek voor Europa. Zo kent Zuid-Amerika de Latinobarómetro en Afrika de Afrobarometer. In dit hoofdstuk staat Europees vergelijkend onderzoek centraal. Een van de belangrijkste organisaties op dit terrein is Eurostat, het statistisch bureau van de EU, dat al bijna 50 jaar een grote hoeveelheid gegevens op sociaal en economisch terrein verzamelt en bewerkt, waarmee een toenemend aantal landen kan worden vergeleken en ontwikkelingen kunnen worden gevolgd (De Michelis and Chantraine, 2003). Zeker nu de statistieken van Eurostat vrijelijk voor iedereen toegankelijk zijn, is hiermee een schat aan informatie beschikbaar. Eurostat verzamelt zelf niet of nauwelijks gegevens, maar krijgt deze aangeleverd door de nationale bureaus voor de statistiek, zoals het CBS in Nederland. Hoe de gegevens tot stand komen, en welke bronnen hierbij worden gebruikt, is minder van belang dan harmonisatie achteraf. Sommige landen verzamelen de statistieken vooral via surveys, andere vooral via registraties. Een mooi voorbeeld van uiteenlopende werkwijzen in verschillende landen zijn de volkstellingen, waar in sommige landen duizenden volkstellers op pad gaan, terwijl de volkstellingstatistiek in Nederland op basis van registraties wordt samengesteld (Schulte Nordholt et al., 2004).

Lang niet alle gegevens waarover men zou willen beschikken, zijn in registraties opgenomen. De Enquête Beroepsbevolking die door de statistische bureaus van de EU-landen wordt uitgevoerd, bevat bijvoorbeeld vragen over aspecten van werk en werk zoeken die alleen aan de respondent zelf kunnen worden gesteld. Deze vragen worden gesteld in face-to-face of telefonische surveys. Ook in het nieuwe *Survey on Income*

and Living conditions van de EU (EU-SILC) is in ieder geval een deel van de vragen alleen via surveys onder de bevolking te beantwoorden. Verder is EUROSTAT nauw betrokken bij het Harmonized European Time Use Survey (HETUS). Naast deze landenvergelijkende statistische surveys bestaat een groot aantal internationale onderzoeken op verschillende deeltherreinen, zoals het International Crime Victim Survey (ICVS), het Survey of Health, Ageing and Retirement in Europe (SHARE), de Trends in Maths and Science Study (TIMSS), het Programme for International Student Assessment (PISA), de Progress in International Reading Literacy Study (PIRLS) and het International Adult Literacy Survey (IALS) (zie ook het overzicht van weblinks aan het eind van dit hoofdstuk). In alle gevallen is het in deze onderzoeken nodig om respondenten in verschillende landen persoonlijk te bevragen, en in alle gevallen worstelt men met de vergelijkbaarheid van onderzoek en resultaten. In de meeste gevallen is methodologische informatie over de opzet, uitvoering en kwaliteit van deze surveys, en van de surveys die in paragraaf 2 aan de orde komen, beschikbaar op het internet.

O'Shea et al. (2004) vergelijken vier internationale studies waarin naar attitudes, waarden en opinies wordt gevraagd. De drie eerste studies hebben een respectabele geschiedenis: de Eurobarometer surveys bestaan nu ongeveer 30 jaar, de European Value Study (EVS) 25 jaar en het International Social Survey Programme (ISSP) 20 jaar. De Eurobarometer surveys hebben als doel de Europese Commissie informatie te leveren over *“the topic most dear to its heart: the process of European unification, as reflected in the attitudes and beliefs of the people in the EU member states”* (Saris & Kaase, 1997:9). Behalve de Commissie maakt ook de wetenschap veel gebruik van de Eurobarometers, onder andere omdat in de loop van de jaren veel verschillende onderwerpen aan de orde zijn gekomen. Na afronding en rapportage door de Commissie worden de gegevens beschikbaar gesteld voor analyse door derden. De European Value Study, waarvoor iedere negen jaar gegevens worden verzameld, is een tweede gevestigde en waardevolle reeks surveys. Uit de EVS zijn ook de vijfjaarlijkse World Value Surveys voortgekomen. De EVS richt zich op het meten van veranderingen in morele en sociale waarden, met nadruk op christelijke waarden en alternatieve wijzen van zingeven (zie Halman, 2001). Een derde vergelijkend onderzoek met een lange geschiedenis is het ISSP dat in 1985 werd opgericht door vier instellingen in Groot-Britannië, de Verenigde Staten, Duitsland en Australië. Het bestaat nu uit een groep van 39 instellingen die voor jaarlijkse peilingen een relatief korte vragenlijst opstellen die onregelmatig wordt herhaald. De ISSP kent een informele structuur en een wat losse regelgeving, voortkomend uit het feit dat het *“came into being without much serious planning, having emerged as a vague idea during an impromptu meeting”* en zich sindsdien *“somewhat haphazardly”* heeft ontwikkeld (Davis & Jowell, 1989:3). Volgens Kish (1994:178) is het ISSP meer een *“cross-national collaboration of national studies”* dan een werkelijk cross-nationaal onderzoek, omdat er geen centrale financiering aanwezig is en de vragenlijst in veel landen als aanhangsel bij een ander onderzoek is opgenomen. In Nederland is het ISSP jarenlang een schriftelijke drop-off vragenlijst geweest bij het onderzoek Culturele Veranderingen in Nederland van het Sociaal en Cultureel Planbureau. Vanaf de peiling van 2003 wordt het ISSP in Nederland uitgevoerd door de Vrije Universiteit.

Een nieuwe loot aan de stam is het tweejaarlijkse European Social Survey (ESS). Voor de ESS zijn in 2002 en 2004 in meer dan 20 Europese landen gegevens verzameld. Het survey wordt centraal aangestuurd met behulp van financiering van de EU en de Euro-

pean Science Foundation. Het veldwerk en de nationale coördinatie worden gefinancierd door de deelnemende landen. Het ESS is opgezet omdat de ruimte voor aanvullende vraagstellingen, de doelstelling en de invalshoek van de drie onderzoeken uit de vorige paragraaf beperkt is. Het bestaat uit een face-to-face onderzoek met een vast deel van ca. een half uur vragenlijst, deels gewijd aan achtergrondkenmerken en deels aan vragen op het terrein van mediagebruik, politieke houdingen, waarden, etc. die in iedere peiling herhaald worden. Daarnaast is er een half uur beschikbaar voor vragenmodules waarvoor door internationale groepen onderzoekers in competitie een voorstel kan worden ingediend. In de peiling van 2002 gingen deze roterende modules over de houding ten opzichte van immigranten en over burgerschap, betrokkenheid en democratie; in 2004 over economische moraal, de houding ten opzichte van medicijngebruik en gezondheid, en werk, gezin en welzijn.

Het ESS kent niet alleen een inhoudelijke doelstelling. Bij de eerdergenoemde, al langer bestaande onderzoeken lag van origine minder de nadruk op methodologische kwaliteit en op vergelijkbaarheid van de opzet van de surveys in de verschillende landen. Als gevolg hiervan is het mogelijk dat verschillen in wijzen van afname van het onderzoek, verschillen in de wijze van steekproeftrekken en verschillen in de behaalde respons de vergelijkbaarheid van nationale peilingen ernstig kunnen beperken. Bovendien is de kwaliteit van survey-onderzoek niet in alle landen even hoog (zie o.a. De Heer, 2000), en blijkt de kwaliteit van landenvergelijkend onderzoek vaak lager dan die van nationaal onderzoek. Het ESS heeft als expliciete methodologische doelstelling een cross-nationaal onderzoek op te zetten dat beantwoordt aan de hoogste standaarden van nationaal onderzoek, en zelfs om de bestaande methodologie van survey-onderzoek te verbeteren. Dit gebeurt door middel van centrale aansturing, centrale specificaties voor de uitvoering van het onderzoek, nauwgezette monitoring en de inzet van inhoudelijke en methodologische experts op verschillende terreinen. Daarnaast worden de surveys en het veldwerk uitgebreid gedocumenteerd. Het gecombineerde bestand van de eerste peiling is sinds september 2003 vrijelijk beschikbaar voor gebruik door derden; het ESS kent inmiddels (augustus 2005) bijna 6000 geregistreerde gebruikers en is de bron van bijna 100 artikelen, papers, boeken en presentaties. De tweede peiling is sinds september 2005 beschikbaar.

3. STREVEN NAAR KWALITEIT EN VERGELIJKBAARHEID

In landenvergelijkende surveys kan men nooit overal identiek te werk gaan, al is het maar omdat talen verschillen. Daarom spreekt men vaak over equivalentie van methoden, vraagstelling en uitvoering. Equivalentie is echter een ongrijpbaar begrip: Johnson (1998) komt tot 52 betekenissen. Daarom wordt vaak gesproken van het streven naar optimale vergelijkbaarheid. Harkness et al. (2003) laten zien waarom optimale vergelijkbaarheid zo moeilijk te bereiken is en waarom landenvergelijkend onderzoek zoveel gecompliceerder is dan nationaal onderzoek. In deze paragraaf wordt aandacht besteed aan vier soorten problemen bij landenvergelijkend onderzoek. Ten eerste kunnen verschillen tussen landen veroorzaken dat het object van onderzoek verschilt. Deze *conceptuele* verschillen brengen de noodzaak met zich mee om voor hetzelfde achterliggende verschijnsel in verschillende landen verschillende indicatoren te gebruiken. Ten tweede loop je bij het stellen van vragen aan tegen *linguïstische* problemen: taalver-

schillen, schaalverschillen en vertalingsproblemen. Ten derde botst de wens standaardisatie van dataverzameling en veldwerk in verschillende landen op verschillen in de *surveypraktijk*. En ten vierde zijn er *methodologische* problemen rond populatie, steekproef en respons. Op deze laatste aspecten wordt in paragraaf 4 ingegaan.

De conceptuele verschillen komen tot uiting in het feit dat dezelfde variabelen of vragen in verschillende landen of culturen een verschillende betekenis kunnen hebben. Braun en Mohler (2003: 111) bespreken de vergelijkbaarheid van achtergrondvariabelen. Neem bijvoorbeeld de simpele vraag naar burgerlijke staat. In een aantal landen is of was echtscheiding niet mogelijk. De categorie 'gescheiden' heeft daar dan ook weinig betekenis; 'gescheiden van tafel en bed' zou in de vragenlijst dezelfde functie kunnen vervullen. Behalve dat dergelijke verschillen in objectieve omstandigheden aanpassing van de vraagstelling nodig maken (het is bijvoorbeeld nodig te vragen naar burgerlijke staat en daarnaast naar vormen van samenleven) betekent het bestaan van dergelijke verschillen ook dat dezelfde objectieve omstandigheden niet overal als indicatoren van hetzelfde concept kunnen worden beschouwd. De relatie tussen gescheiden zijn en geluk, tevredenheid en sociale isolatie zal in landen waar men niet of nauwelijks kan scheiden niet bestaan. Daar zal men andere indicatoren moeten vinden. Vergelijkbare effecten kunnen optreden bij woonomstandigheden of arbeidsomstandigheden. Wonen in een flat in het centrum van een grote stad is in Nederland vaak een indicator van lager welzijn en lagere welvaart. De bewoners van luxe appartementen in het centrum van Londen of Parijs zullen zich hier niet in kunnen herkennen. Parttime werk wordt in Nederland als een verworvenheid gezien en is normaal, vooral bij vrouwen. In andere landen wordt parttime werk gezien als deelwerkeloosheid en is het hebben van een parttime baan iets negatiefs. Nog gecompliceerder wordt het bij vragen over religie. Het begrip kerkgenootschap bestaat niet bij hindoes, het begrip god bestaat niet bij boeddhisten, en het verband tussen godsdienst en staat is in sommige landen erg hecht, terwijl andere landen een strikte scheiding betrachten.

Het European Social Survey bevatte in 2002 een module met vragen over houdingen ten opzichte van immigranten. Los van het feit dat de houding in verschillende landen kan verschillen, blijkt dat het begrip 'immigranten' overal een andere connotatie heeft. In Nederland wonen veel allochtonen en asielzoekers. Veel van origine Turkse of Marokkaanse bewoners van Nederland zijn hier geboren; Antillianen hebben de Nederlandse nationaliteit. Dit zijn formeel geen immigranten, maar deze groepen spelen wel een belangrijke rol bij het tot stand komen van meningen over 'buitenlanders' (wat het ook niet zijn). In het Verenigd Koninkrijk spreken immigranten van origine wel vaak Engels, in Duitsland hebben personen van Duitse afkomst ten alle tijden het recht terug te keren, Joden kunnen vrijelijk naar Israël emigreren, etc. Immigranten zijn in verschillende landen duidelijk andere groepen. Een manier om dit probleem op te lossen is vooraf duidelijk te expliciteren wat met een bepaalde reeks vragen getracht wordt te meten. Gaat het om immigranten, of buitenlanders, of mensen met een andere religie of een andere huidskleur? In het ESS is geprobeerd de term immigrant te vermijden door te vragen naar 'mensen die in uw land komen wonen' en die afkomstig zijn uit armere resp. rijkere landen binnen resp. buiten Europa, en die van hetzelfde resp. een ander ras of etnische groep zijn. Om welke groepen het gaat zal in ieder land verschillen, en de vragen bleken ook niet in ieder land even makkelijk te beantwoorden.

In het European Social Survey is voorafgaand aan het maken van de vragenlijst aan een groep inhoudelijke experts verzocht op verschillende terreinen (sociaal-structurele kenmerken, mediagebruik, waarden, politieke betrokkenheid, welzijn, sociale uitsluiting, religie) aan te geven wat de theoretische achtergronden van begrippen zijn, welke indicatoren in de loop van de tijd en in verschillende surveys gebruikt zijn om de concepten te meten, en wat in de praktijk de resultaten waren. Op grond van deze rapporten zijn nieuwe vragen ontwikkeld. Ook bij de roterende modules (zie paragraaf 2) is aan de onderzoeksgroepen verzocht niet zozeer een lijst met vragen in te leveren, maar een theoretische achtergrondstudie en een overzicht van concepten en indicatoren. In beide gevallen is nagegaan of de resulterende vragen in alle deelnemende landen dezelfde betekenis hadden, al is dat niet eenvoudig als een vragenlijst wordt afgenomen in zulke uiteenlopende landen als Griekenland, Finland en Oekraïne. Een andere manier om contextverschillen niet zozeer op te lossen als wel zichtbaar te maken is het aanvullen van survey gegevens met contextgegevens. In de ESS is dit gebeurd door geaggregeerde informatie te verzamelen en beschikbaar te stellen over landen (van bevolkingsopbouw tot BNP), te verwijzen naar bestaande informatie over structuren en systemen (kiessystemen, schoolsoorten) en informatie te verzamelen over belangrijke gebeurtenissen die plaatsvonden gedurende het veldwerk en die mogelijk effect hadden op de antwoorden op de vragenlijst. Deze gebeurtenissen zijn opgeslagen in een afzonderlijke, algemene *event* databank.

Zelfs als duidelijk is wat er gevraagd moet worden, blijft het probleem van de *linguïstische* vergelijkbaarheid. Het is niet makkelijk om in verschillende landen, met verschillende culturele gebruiken en vooral verschillende talen hetzelfde te vragen. Letterlijke vertalingen noch *back-translation* (heen- en terugvertalen) zorgen voor optimale vergelijkbaarheid (denk maar aan het begrip immigranten). Harkness (1999) heeft laten zien dat equivalentie van betekenis tussen landen veel meer is dan een kwestie van (goed) vertalen. In het ESS wordt een gestructureerd vertalingsproces voorgeschreven, waarbij in de bronvragenlijst annotaties zijn opgenomen die het doel van de vraag beschrijven, per land meer dan een vertaler wordt ingezet en verschillen tussen vertalers worden besproken en beoordeeld, en waar aanvullend overleg plaatsvindt tussen landen waarin dezelfde taal wordt gesproken, zoals België en Nederland, of Duitsland, Oostenrijk en Zwitserland. Dat laatste overleg heeft vooral ten doel onnodige verschillen te vermijden. Zelfs bij deze strikte procedure zijn fouten in de vertaling opgetreden (zoals in de gevallen waar *wealthy* vertaald werd met het equivalent van *healthy*).

Vergelijkbaarheid van vragen is meer dan goed vertalen van de vraagtekst. Het voorbeeld van responschalen (Harkness, 2003) kan dit illustreren. Scores op een schaal van 0 tot 10 heeft in Nederland de simpele betekenis van rapportcijfers. In andere landen worden op rapporten echter volstrekt andere schalen gehanteerd, zodat een schaal van 0 tot 10 daar veel minder vanzelfsprekend is. Binnen veel landen is er een traditie ontstaan van eigen responschalen en bijbehorende semantische categorieën. Men kan dus kiezen: of onvergelijkbare, vertrouwde schalen, of vergelijkbare, ‘vreemde’ schalen. In de ESS is gekozen voor identieke schalen, en identieke (ofwel zoveel mogelijk vergelijkbare) labels in alle landen. Een mogelijkheid om het effect van deze keuze te testen is met behulp van zogenaamde *multitrait-multimethod* (MTMM) studies (Saris, 2003). In de ESS is een dergelijk MTMM-design geïmplementeerd en wordt aan dezelfde respondent dezelfde vraag tweemaal (met enige tussenpoos) gesteld met ver-

schillende responschalen. Op deze manier kan binnen ieder land het methode-effect van responschalen worden geïsoleerd van de ware score op een bepaalde kenmerk (*trait*).

Tot zover is ingegaan op het zoveel mogelijk vergelijkbaar maken van de vragenlijst. Bij het afnemen van de vragenlijst kunnen weer verschillen ontstaan die samenhangen met de *surveypraktijk* in verschillende landen, ofwel met de wijze van interviewen. Of vragen schriftelijk, telefonisch, mondeling of via internet worden gepresenteerd, kan invloed hebben op de uitkomsten van een survey (De Leeuw, 1992; Saris & Kaase, 1997:168-9). Bij mondelinge en telefonische surveys kan de interviewer de respondent aanwijzingen geven, maar is er ook een grotere kans op sociaal-wenselijke antwoorden. Bij telefonisch onderzoek zijn geen toonkaarten mogelijk, bij computergestuurde vraagstelling zijn er legio manieren die anders ontbreken om foto's en filmpjes op te nemen en de volgorde van de vraagstelling aan te passen. Holbrook et al. (2003) kwamen op grond van een vergelijking van (lange) telefonische en face-to-face surveys tot de conclusie dat de eerste goedkoper zijn, sneller kunnen worden uitgevoerd en beter volgens vaste richtlijnen kunnen worden afgenomen omdat supervisie van interviewers simpeler is, maar in vrijwel alle gevallen ook tot een hogere nonrespons leiden. Item nonrespons is in het algemeen hoger als de respondenten zelf een vragenlijst invullen (Kalgraff Skjak en Harkness, 2003: 183).

Los van welke dataverzamelmethode, of *mode*, nu 'beter' is, lijken de geconstateerde verschillen ervoor te pleiten om in landenvergelijkende surveys overal dezelfde methode te gebruiken. Kallgraf Skjak en Harkness (2003: 191) geven aan dat deze *one mode fits all* benadering ook nadelen heeft. Het belangrijkste nadeel is dat de optimale methode per land kan verschillen. Face-to-face interviewen is ongebruikelijk geworden in Noord-Europese landen en Zwitserland. Dat kan betekenen dat het in deze landen onmogelijk is om ervaren interviewers te vinden. Schriftelijk enquêteren is moeilijk in landen met een relatief hoog analfabetisme, internet-enquêteren in landen met een lage internet-penetratie en telefonisch enquêteren in landen waar veel mensen geen telefoon, of alleen een mobieltje, of een geheim nummer hebben. Behalve landen kunnen ook personen een geprefereerde mode van interviewen hebben. Een mixed-mode benadering kan daarom leiden tot een hogere respons (Dillman, 1999; Snijkers en Luppens, 2000), maar ook tot verschillen tussen personen die een gevolg zijn van de geprefereerde methode. Bij de ESS worden in samenwerking met Gallup Europe experimenten uitgevoerd om het effect van verschillende modes te testen, mode-invariante vragen te ontwerpen of methoden te ontwikkelen voor het corrigeren van mode-effecten (Roberts et al., 2005).

Uit bovenstaand overzicht blijkt op hoeveel manieren men in landenvergelijkend onderzoek vergelijkbaarheid nastreeft en hoe moeilijk dit daadwerkelijk te bereiken is. Niet alleen zijn er conceptuele en linguïstische barrières en verschillen tussen gebruikelijke dataverzamelmethode, ook blijkt dat er bij het trekken van een steekproef en het behalen van een hoge respons grote verschillen tussen landen te bestaan.

4. POPULATIE, STEEKPROEF EN NONRESPONS

Een strikt identieke opzet van een survey is niet altijd de beste manier om optimale vergelijkbaarheid te realiseren, omdat een identieke opzet niet in ieder land of iedere cultuur even effectief hoeft te zijn (zie bovenstaande voorbeelden van interview *mode* en responschalen). Het is echter steeds van belang na te gaan wanneer met een nationaal optimale aanpak zo goed mogelijk hetzelfde doel wordt nagestreefd, of wanneer er sprake is van onnodige, versturende verschillen tussen landen of uitvoerende onderzoeksorganisaties (zie De Heer, 2000). Een standaard omschrijving van doelpopulatie, random steekproeftrekking en responsberekening komt de kwaliteit van landenvergelijkend onderzoek ten goede, maar verschillen in de samenstelling van de populatie, aanwezige steekproefkaders en gerealiseerde respons zijn niet te vermijden.

Zelfs als de doelpopulatie precies gedefinieerd is, bijvoorbeeld alle inwoners van een land van 15 jaar en ouder in zelfstandige huishoudens, kan deze definitie in verschillende landen anders uitwerken. Ten eerste blijkt het in de praktijk lastig om nationale deelnemers ervan te overtuigen de onder- en bovengrens aan de leeftijd die normaal gehanteerd wordt, op te geven, zeker omdat voor het interviewen van de jongsten in veel landen toestemming van de ouders nodig is, en bij de alleroudsten nonrespons om fysieke en psychische redenen vaak hoog is. Maar zelfs als men zich aan de omschrijving houdt, treden er praktische problemen rond vergelijkbaarheid op. Wat te doen in een land als Luxemburg, waar veel inwoners voor kortere perioden verblijven: wanneer is men een inwoner? En wat zijn de consequenties van het feit dat Luxemburg geen universiteit heeft, en dus ook geen studenten kent (die wonen tijdelijk in een ander land). Verschillende gewoonten van ouderenzorg kunnen leiden tot substantiële verschillen in het aantal thuiswonende ouderen tussen landen. En verschillen in regels rond dienstplicht kunnen leiden tot aanzienlijke verschillen in het aantal jongeren (of jonge mannen) dat deel uitmaakt van een zelfstandig huishouden. Bij de identificatie van de doelpopulatie blijkt dus al hoe moeilijker landen vergelijkbaar zijn.

Een ander cruciaal aspect van vergelijkbaarheid heeft te maken met het trekken van de steekproef. In veel landenvergelijkend onderzoek is er – zeker in de beginfase – gebruik gemaakt van quota steekproeven. Couper en De Leeuw (2003) laten zien dat deze gewoonte ook nu nog niet volledig is uitgestorven. Om een goed beeld te kunnen geven van de populatie zijn random steekproeven, waarin ieder lid van de populatie een bekende kans groter dan 0 heeft om getrokken te worden, te prefereren. In veel onderzoek wordt in enigerlei fase substitutie toegestaan, dat wil zeggen het vervangen van een onwillig huishouden door een naburig huishouden, of het vervangen van de te interviewen persoon door een ander lid van het gezin. Bij de EVS was in 19 van de 32 deelnemende landen in enige fase van het veldwerk sprake van substitutie (Halman, 2001), en dit kwam voor in de helft van de ISSP landen (Park and Jowell, 1996). Substitutie heeft echter verschillende nadelen (zie ook O'Shea et al., 2004). Ten eerste worden interviewers aangemoedigd om niet te veel energie te besteden aan 'moeilijke gevallen' als ze dezen mogen vervangen door bereidwilliger personen. Ten tweede zijn responscijfers moeilijk te berekenen als nonrespondenten door andere vervangen mogen worden. Responscijfers hebben dan nog maar weinig betekenis, en, wat erger is, de effecten van nonrespons zijn dan ook veel moeilijker te bepalen en te corrigeren. Het grootste nadeel van substitutie is dat hierdoor een oververtegenwoordiging van

makkelijke respondenten wordt gecreëerd, terwijl er aanwijzingen zijn dat deze verschillen van moeilijk bereikbare, of minder bereidwillige respondenten (zie Stoop, 2005). Als de responsgeneigdheid in verschillende landen verschilt, zal er ook in verschillende mate sprake zijn van substitutie, wat nog een extra nadeel is. In verband met vergelijkbaarheid en hoge surveykwaliteit is het dus van belang dat in ieder land een random steekproef wordt getrokken, waarbij substitutie niet is toegestaan.

Idealiter zou er bij een Europees onderzoek één (gestratificeerde) steekproef getrokken moeten kunnen worden. Helaas bestaat er geen Europees bevolkingsregister dat als steekproefkader kan worden gebruikt. In de meeste landen bestaat zelfs geen bevolkingsregister waaruit simpel een steekproef kan worden getrokken. Dat betekent dat in het ene land een personensteekproef uit het bevolkingsregister kan worden getrokken, in het ander land een adressensteekproef wordt gebruikt, dat soms telefoonboeken als steekproefkader dienen en soms kiesregisters en soms combinaties van verschillende kaders. Gelukkig heeft Kish (1994) aangetoond dat verschillende steekproefkaders vergelijkbare resultaten kunnen opleveren, mits er maar sprake is van random steekproeven zonder substitutie. Bij de ESS (zie Häder en Gabler, 2003) is een expert panel verantwoordelijk voor de equivalentie van de steekproeven van de verschillende landen. Een steekproefdesign mag pas worden geïmplementeerd nadat het goedgekeurd is door dit panel. Steekproefexperts helpen landen bij de constructie van een goed steekproefkader, het ontwikkelen van een steekproefprocedure en het berekenen van design-effecten, die optreden bij ongelijke kansen van personen om getrokken te worden (bijvoorbeeld bij adressteekproeven, waar leden van grote huishoudens een kleinere kans hebben om in de steekproef te komen) en bij getrapte steekproeven waarbij er sprake kan zijn van clustereffecten. Op deze wijze wordt de kans dat verschillen in steekproeftrekking de resultaten van een survey beïnvloeden, geminimaliseerd.

Een van de grootste problemen van landenvergelijkend onderzoek is nonrespons, en in het bijzonder verschillen in responspercentage en in samenstelling van de respons tussen landen. Nonrespons (Groves, 1989) kan tot vertekening leiden, en een grote variatie in respons kan tot verschillen in vertekening leiden. O'Shea et al. (2004) beschrijven zeer grote verschillen in responspercentages bij het ISSP (56% in Letland tot 94% in Bulgarije, waar waarschijnlijk substitutie werd toegestaan) en bij de EVS (13% in Spanje en 95% in Slowakije). Om deze verschillen te voorkomen werd in het ESS een streefpercentage van 70% respons voorgeschreven. Voor sommige landen was dat niet bijzonder hoog; voor Nederland leek dat bijna onhaalbaar. Uit veel onderzoek is immers gebleken dat Nederland op het terrein van respons slechte resultaten boekt (De Heer, 1999; De Heer, 2000; De Leeuw en De Heer, 2002).

Een ander probleem is dat responscijfers vaak op zeer verschillende wijze worden berekend en dat daardoor vergelijking van respons op verschillende onderzoeken en in verschillende landen een hopeloze taak is (zie Schnell; 1997; Smith, 2002). Couper en De Leeuw (2003: 165) concluderen na vergelijking van de respons in verschillende landen op verschillende surveys: *"It should be noted that the quality of the response information provided varied; often few details beyond basic response rate information were provided in the reports. This lack of documentation makes it difficult for users of the data to make judgments about the quality and comparability of these cross-national surveys"*. Soms wordt respons alleen berekend op basis van de mensen die daadwerke-

lijk bereikt zijn, soms is er sprake van substitutie, soms worden groepen waarvan de respons naar verwachting laag zal zijn bij voorbaat uitgesloten, etc. De responscijfers van de ESS zijn uniek omdat met behulp van deze gegevens veel exacter dan in andere surveys, waar het design verschilt en de respons op uiteenlopende manier wordt berekend, verschillen tussen landen zijn aan te geven.

Tabel 1. Responspercentages European Social Survey 2002/2003.

	respons		respons		respons
Griekenland	80,0	Nederland	67,9	Verenigd Koninkrijk	55,5
Finland	73,2	Denemarken	67,6	Spanje	53,2
Polen	73,2	Noorwegen	65,0	Luxemburg	43,9
Israël	71,0	Ierland	64,5	Italië	43,7
Slovenië	70,5	Oostenrijk	60,4	Tsjechië	43,3
Hongarije	69,9	België	59,2	Frankrijk	43,1
Zweden	69,5	Duitsland	57,1	Zwitserland	33,5
Portugal	68,8				

Bron: http://ess.nsd.uib.no/2003_Fworksummary.jsp

Tabel 1 geeft een overzicht van de responspercentages in de eerste peiling van het ESS. In deze tabel vallen twee dingen op. Ten eerste slaat Nederland in het ESS helemaal geen gek figuur. De hoge respons in Nederland werd deels gerealiseerd door het herbenaderen van het merendeel van de initiële weigeraars en het alsnog overhalen van een groot deel van hen, iets wat in de andere deelnemende landen niet werd geprobeerd of niet lukte (zie Stoop, 2005: 266). Het tweede dat opvalt is dat, ondanks het streefpercentage van 70%, er ook bij de ESS een zeer grote spreiding is in responspercentages, van 33,5% in Zwitserland tot 80% in Griekenland. Deze spreiding heeft meer betekenis dan de eerder beschreven verschillen tussen landen, omdat de afhandeling van de steekproef, de wijze van enquêteren en de berekening van de respons in de ESS volgens strikte regels verliep.

Het is niet eenvoudig te achterhalen waarom in de ESS de respons in sommige landen zoveel lager lag dan in andere. De nationale kosten van het survey en de getrooste inspanningen zijn niet voldoende om de verschillen te verklaren. Overigens heeft in de tweede ronde van het ESS Zwitserland, na veel overleg met externe experts en de veldwerkorganisatie en o.a. door extra beloningen voor de respondenten, spectaculaire verbeteringen geboekt en is de respons gestegen van 33% naar 48%.

5. OPTIMALE VERGELIJKBAARHEID EN TRANSPARANTIE

In het ESS is met veel inspanning geprobeerd een zo goed mogelijk vergelijkbaar survey op te zetten, en ook daar blijkt vertalen moeilijk (en worden fouten gemaakt), ook daar blijkt dat dezelfde wijze van ondervragen niet in ieder land optimaal is, en ook daar blijkt de respons zeer uiteen te lopen. Betekent dit alles dat landenvergelijkend onderzoek niet mogelijk is? Als het antwoord op deze vraag bevestigend zou luiden, zou je ook moeten concluderen dat nationaal onderzoek niet mogelijk is: ook bij nationale surveys treden vergelijkbare culturele verschillen op tussen groepen respondenten

als bij landenvergelijkende surveys, alleen wordt dat vaak niet beseft. Responscijfers verschillen soms zeer aanzienlijk in grote steden en op het platteland. Ook in een klein en relatief homogeen land als Nederland kan men voor verschillende groepen verschillende optimale modes van vraagstelling hebben. De formulering van vragen zou idealiter moeten worden aangepast aan achtergrond, opleiding en taalgebruik van respondenten. Een extra complicatie is dat er sprake kan zijn van interactie, als vragen in een landenvergelijkend survey wel leiden tot optimaal vergelijkbare antwoorden van hoger opgeleiden, maar niet van lager opgeleiden. Cross-culturele vergelijkbaarheid binnen landen is dus net zo goed een probleem als vergelijkbaarheid tussen landen (Scheuch, 1989:155). Het verschil met landenvergelijkend onderzoek is dat er bij nationaal onderzoek sprake is van één taal (of slechts een beperkt aantal), meestal van een veldwerkbureau, één steekproefdesign en uniforme systeemkenmerken (onderwijs, gezondheidszorg, verkiezingen).

Landenvergelijkend onderzoek is moeilijk, maar de gesignaleerde problemen zijn geen reden om dan maar op te houden onderzoek te doen, zeker niet als de volgende drie aanbevelingen worden opgevolgd. *Ten eerste* is het aan te bevelen een identieke aanpak te hanteren waar dat mogelijk is, en een specifieke aanpak waar dit noodzakelijk is. Verschillen in onderzoeksopzet zouden niet veroorzaakt moeten worden door verschillende gebruiken en gewoonten binnen de deelnemende landen of bij uitvoerende onderzoeksorganisaties, maar beredeneerd moeten worden toegelaten (zie De Heer, 2000; O'Shea et al., 2004). *Ten tweede* komt men er niet met het uitvaardigen van centrale richtlijnen. Betrokkenheid van en overleg met nationale experts, ondersteuning en begeleiding van nationale coördinatoren, maar ook monitoring en controle van nationaal veldwerk is noodzakelijk. *Ten slotte* is het van groot belang goed te documenteren wat daadwerkelijk gebeurd is. Dat betekent dat niet alleen de databestanden beschikbaar moeten worden gesteld voor verder onderzoek, maar ook de vraagtekst, responscijfers, veldwerkprocedures en steekproefdesigns. Daarnaast zouden problemen met de vertaling, afwijkingen in de vragenlijst, afwijkingen in de voorgeschreven procedure, kortweg alle ontdekte afwijkingen moeten worden beschreven en actief onder de aandacht van de gebruikers moeten worden gebracht. Bij de ESS zijn alle documentatie en alle gesignaleerde afwijkingen vastgelegd op de data-website van het Noorse dataarchief met als paradoxaal resultaat dat bij de ESS centraal voorgeschreven procedures en strikte standaarden gepaard gaan met een zeer grote, zichtbare hoeveelheid afwijkingen. Hoe meer documentatie, hoe meer fouten lijkt het wel.

Een combinatie van strikte standaarden en centrale specificaties, nauwe samenwerking met nationale vertegenwoordigers, grote transparantie en volledige documentatie is de beste garantie voor optimale vergelijkbaarheid en kan tevens voorkomen dat een onderzoeker niet meer kan nagaan of er nu sprake is van verschillen tussen landen of verschillen tussen surveys.

LITERATUUR

Braun, Michael and Peter Ph. Mohler, 2003. Background Variables. In: Janet A. Harkness, Fons J.R. van de Vijver, Peter Ph. Mohler (eds.). *Cross-Cultural Survey Methods*. New York: Wiley, pp. 101-115.

- Couper, Mick P. and Edith D. de Leeuw, 2003. Nonresponse in Cross-Cultural and Cross-National Surveys. In: Janet A. Harkness, Fons J.R. van de Vijver, Peter Ph. Mohler (eds.). *Cross-Cultural Survey Methods*. New York: Wiley, pp. 157-177.
- Davis, J., P.Ph. Mohler and T.W. Smith, 1994. Nationwide General Social Surveys. In: I. Borg and P.Ph. Mohler (eds.) *Trends and Perspectives in Empirical Social Research*. Berlin: Walter de Gruyter & Co.
- De Michelis, Alberto and Alain Chantraine, 2003. *Memoirs of Eurostat. Fifty years serving Europe*. Luxembourg: Office for Official Publications of the European Communities.
- Dillman, Don A., 1999. *Mail and Internet Surveys: The Tailored Design Method*. New York: Wiley.
- Groves, Robert M., 1989. *Survey Errors and Survey Costs*. New York: John Wiley & Sons, Inc.
- Häder, Sabine and Siegfried Gabler, 2003. Sampling and Estimation. In: Janet A. Harkness, Fons J.R. van de Vijver, Peter Ph. Mohler (eds.). *Cross-Cultural Survey Methods*. New York: Wiley, pp. 117-134.
- Halman, L., 2001. *The European Values Study: A Third Wave*. Source Book of the 1999/2000 European Values Study Surveys. Tilburg: EVS, WORC, Tilburg University.
- Harkness, J., 1999. In pursuit of quality: issues for cross-national survey research in *International Journal of Social Research Methodology*, Vol. 2, No. 2, pp. 125-140.
- Harkness, Janet, 2003. Questionnaire Translation. In: Janet A. Harkness, Fons J.R. van de Vijver, Peter Ph. Mohler (eds.). *Cross-Cultural Survey Methods*. New York: Wiley, pp. 35-56.
- Harkness, Janet A., Fons J.R. van de Vijver, Peter Ph. Mohler (eds.), 2003. *Cross-Cultural Survey Methods*. New York: Wiley.
- Heer, Wim de, 1999. International Response Trends, Results of an International Survey. *Journal of Official Statistics*, Vol. 15, No 2, pp. 129-142.
- Heer, Wim de, 2000. Survey Practices in European Countries. In: Siobhán Carey (ed.) *Measuring Adult Literacy. The International Adult Literacy Survey (IALS) in the European context*. London: Office for National Statistics, pp. 43-64.
- Holbrook, A. L., M. Green and J.A. Krosnick, 2003. Telephone vs. face-to-face interviewing of national probability samples with long questionnaires: Comparisons of respondent satisficing and social desirability response bias. *Public Opinion Quarterly*, Vol. 67, pp. 79-125.
- Johnson, T., 1998. Approaches to Equivalence in Cross-Cultural and Cross-National Survey Research. In: J.A. Harkness, ZUMA-Nachrichten Spezial No. 3. *Cross-Cultural Survey Equivalence*. Mannheim: ZUMA.
- Jowell, R., 1998. How Comparative is Comparative Research? *American Behavioural Scientist*, Vol. 42, No. 2, October 1998, pp. 168-177.
- Kallgraff Skjåk, Knut and Janet Harkness, 2003. Data Collection Methods. In: Janet A. Harkness, Fons J.R. van de Vijver, Peter Ph. Mohler (eds.). *Cross-Cultural Survey Methods*. New York: Wiley, pp. 179-193.
- Kish, L., 1994. Multipopulation Survey Designs: Five Types with Seven Shared Aspects. *International Statistical Review*, Vol. 62, No. 2, pp. 167-186.
- Leeuw, Edith D. de, 1992. *Data Quality in Mail, Telephone and Face to Face Surveys*. Amsterdam: TT Publikaties.
- Leeuw, Edith de, and Wim de Heer, 2002. Trends in Household Survey Nonresponse: A Longitudinal and International Comparison. In: Robert M. Groves, Don A. Dillman, John L. Eltinge and Roderick J.A. Little (eds.) *Survey Nonresponse*. New York: Wiley, pp. 41-54.
- Lynn, Peter, Lilli Japac and Lars Lyberg, 2005. *What's so Special About Cross-national Surveys?* Paper presented at the 3rd CSDI workshop, Madrid, March 2005.
- O'Shea, R., C. Bryson and R. Jowell, 2004. *Comparative Attitudinal Research in Europe*. London: ESS.
- Park, A. and R. Jowell, 1996. *Consistencies and differences in a cross-national survey*. The International Social Survey Programme 1995. London : SCPR.
- Roberts C., P. Lynn, R. Manchin and I. Agnes, 2005. *Methodological advances on the ESS: A*

- mixed mode future?* Paper presented at 60th AAPOR Annual Conference, Miami Beach, Florida, USA.
- Saris, Willem E., 2003. Multitrait-Multimethod Studies. In: Janet A. Harkness, Fons J.R. van de Vijver, Peter Ph. Mohler (eds.). *Cross-Cultural Survey Methods*. New York: Wiley, pp. 265-274.
- Saris, W. E., and M. Kaase (eds.), 1997. *Eurobarometer: Measurement Instruments for Opinions in Europe*. ZUMA, Mannheim.
- Scheuch, E. K., 1989. Theoretical Implications of Comparative Survey Research: Why the Wheel of Cross-cultural Methodology Keeps on Being Reinvented. *International Sociology*, Vol. 4, No. 2, pp. 147-167.
- Schnell, Rainer, 1997. *Nonresponse in Bevölkerungsumfragen, Ausmaß, Entwicklung und Ursachen*. Opladen: Leske und Budrich.
- Schulte Nordholt, Eric, Marijke Hartgers and Rita Gircour, 2004. *The Dutch Virtual Census of 2001. Analysis and Methodology*. Voorburg: Statistics Netherlands.
- Smith, Tom. W., 2002. Developing Nonresponse Standards. In: Robert M. Groves, Don A. Dillman, John L. Eltinge and Roderick J.A. Little (eds.) *Survey Nonresponse*. New York: Wiley, pp. 27-40.
- Snijkers, Ger and Martin Luppens, 2000. The best of two worlds: total design method and new Kontiv design. An operational model to improve respondent co-operation. *Netherlands Official Statistics*, Vol. 3, pp. 4-10.
- Stoop, Ineke A.L., 2005. *The Hunt for the Last Respondent*. Den Haag: Sociaal en Cultureel Planbureau.
- Vijver, F. J. R. van de, 2003. Bias and Substantive Analysis. In: Janet A. Harkness, Fons J.R. van de Vijver, Peter Ph. Mohler (eds.). *Cross-Cultural Survey Methods*. New York: Wiley, pp. 143-155.

LINKS NAAR WEBSITES

- Afrobarometer: <http://www.afrobarometer.org/>
- Eurobarometer: http://www.gesis.org/en/data_service/eurobarometer/
- Eurostat: <http://europa.eu.int/comm/eurostat/>
- European Social Survey: <http://www.europeansocialsurvey.org/>
- European Social Survey (data-archieff): <http://ess.nsd.uib.no/>
- European Value Study: <http://www.europeanvalues.nl/>
- International Adult Literacy Survey: <http://www.statcan.ca/english/Dli/Data/Ftp/ials.htm>
- International Social Survey Programme: <http://www.issp.org>
- International Time Use Survey: <http://www.iser.essex.ac.uk/mtus/>
- Latinobarómetro: <http://www.latinobarometro.org/>
- Programme for International Student Assessment: <http://www.pisa.oecd.org/>
- Survey of Health, Ageing and Retirement in Europe (SHARE): <http://www.share-project.org>
- Trends in Maths and Science Study (TIMSS) and Progress in International Reading Literacy Study (PIRLS): <http://timss.bc.edu/>
- World Values Survey: <http://www.worldvaluessurvey.com>

7. *Meten op markten*

P.S.H. LEEFLANG

SAMENVATTING

Door intensief te meten op markten en goede gegevens zijn er in de afgelopen decennia tal van verbanden gevonden tussen marktinstrumenten en responsvariabelen. Analyse van deze analyses leidt tot generalisaties over de werking van markten. Een aantal van deze generalisaties wordt in dit artikel besproken.

1. INTRO¹

In 2008 bestaat de Koninklijke Nederlandse Academie van Wetenschappen (KNAW) 200 jaar. Het instituut is opgericht door koning Lodewijk Napoleon ten tijde van zijn koningsschap van Holland (1806-1810). De KNAW is werkzaam op het gebied van wetenschappelijk onderzoek. Zij bevordert de uitwisseling van gedachten en informatie tussen haar (circa 200) leden onderling en tussen deze leden en andere beoefenaren van de wetenschap en wetenschappelijke organisaties. De KNAW beoordeelt de kwaliteit van wetenschappelijk onderzoek. Tevens is zij een koepelorganisatie voor overwegend fundamenteel-wetenschappelijke onderzoeksinstituten en informatieverzorgende instituten. Voorbeelden van deze instituten zijn het Meertens instituut (onderzoek en documentatie van de Nederlandse taalcultuur), het Rathenau instituut, het Nederlands Instituut voor Hersenonderzoek (NIH) en het Nederlandse Instituut voor Oorlogsdocumentatie (NIOD). Verder adviseert de KNAW de Minister van Onderwijs en Wetenschap.

Leden van de KNAW worden geacht om tenminste éénmaal een lezing te verzorgen voor de andere leden. In de bijna 200-jarige geschiedenis van de KNAW is er in 2004 voor het eerst een lezing gegeven op het terrein van Marketing en Marktonderzoek. Deze lezing is gepubliceerd in de reeks Mededelingen der KNAW (Leeflang, 2005). In de bijdrage voor het Jaarboek zal ik een verkorte en aangepaste versie van deze lezing geven. Dit artikel gaat over de mogelijkheden die er zijn om op markten te meten. Deze mogelijkheden zijn de laatste decennia sterk toegenomen door het beschikbaar komen van grote hoeveelheden gegevens over vraag en aanbod en de ontwikkeling en toepassing van methoden en technieken van marktonderzoek, waarmee informatie aan de gegevens ontleend kan worden. Metingen op markten en bewerkingen van gegevens hebben mogelijkheden gecreëerd om tot generalisaties over marktwerking te komen. Deze generalisaties vormen, samen met de ontwikkelde methoden en technieken van marktonderzoek, de kern van marketing als veld van wetenschappelijk onderzoek: mar-

keting science. In dit artikel zullen we op deze generalisaties ingaan. Tevens zullen we bespreken welke waarde de kennis omtrent wetmatigheden op markten zou kunnen hebben voor de bedrijfspraktijk.

2. MARKETING SCIENCE

In 1926 beweerde John Wanamaker, een warenhuis magnaat uit Philadelphia², 'ik weet dat de helft van mijn reclame-uitgaven verspilld zijn en niets opleveren, alleen ik weet niet welke helft'.

Anno 2004 wordt deze uitspraak in veel organisaties nog als adagium gebruikt. Dit is nauwelijks voor te stellen, gegeven de kennis die zowel over de werking van reclame als over het meten van de effecten van reclameboodschappen gegeneerd is.

In veel economieboeken worden uitvoerige beschouwingen gewijd aan vraag- en aanbodfuncties en het al dan niet in evenwicht zijn van markten.³ Vraag en aanbod worden daarbij gerelateerd aan de prijs van een product. De relatie tussen vraag en/of aanbod en de prijs per eenheid product kan empirisch worden bepaald. Aan de hand van deze relatie en met behulp van enige andere informatie kan, in principe, de optimale prijs worden bepaald. Er zijn echter weinig organisaties bekend waar dit op een formele wijze ook zo gebeurt. Vaak neemt men zijn toevlucht tot ervaringsregels of baseert men zijn beslissingen op intuïtie. Men gaat dan voorbij aan de ruime mogelijkheden die er zijn om op markten te meten en deze metingen als uitgangspunt te nemen voor het specificeren van vraag- en aanbodrelaties. Tevens gaat men voorbij aan de mogelijkheden om gebruik te maken van *kennis* die in marketing science ontwikkeld is. Deze kennis is gebaseerd op analyses van analyses: meta-analyses. Meta-analyses leiden tot generalisaties over de werking van markten.

Marketing science heeft betrekking op de ontwikkeling van methoden en technieken om het gedrag van vragers en aanbieders te beschrijven en te verklaren en om het toepassen van deze methoden en technieken. Daarnaast worden concepten, te toetsen hypothesen en theorieën (T) ontwikkeld. Na ontwikkeling worden deze theorieën vaak empirisch (E) getoetst. Vanaf 1980 is er steeds meer aandacht gekomen voor het formuleren van kennis over hoe en onder welke omstandigheden de marktinstrumenten welke effecten hebben.

De inhoud van marketing science wordt bepaald door uitspraken die alleen zinvol zijn als er empirische evidentie bestaat. Marketing science sluit dan ook goed aan bij het logisch positivisme (Koster, 1991). De gegeneerde kennis wordt sterk door inductie bepaald. Daarbij bestaat er een voorkeur om empirische patronen vast te stellen die tot generalisaties en uiteindelijk tot een theorie leiden. Dit wordt wel de ET-benadering genoemd (E: empirie, T: theorie).⁴ Dit in tegenstelling tot de TE-benadering, die men in marketing als onderzoeksveld ook aantreft, waar men vaak geïsoleerd van de empirie een theorie ontwikkelt en deze dan empirisch toetst. Naarmate meer empirische toetsing plaatsvindt, zal meer inhoud aan theorie (T) gegeven kunnen worden en die theorie vormt dan een betere basis voor empirische toetsing (E). Zo zal na enige tijd de vraag welke van de twee benaderingen TE of ET de beste mogelijkheden biedt voor het ontwikkelen van marketing science, overbodig worden. De volgorde is of ETETET... of TETETE... hetgeen na enige tijd tot convergentie moet leiden.

Bij de bespreking van generalisaties zullen zowel de resultaten van meta-analyses als van situaties die wel tot convergerende resultaten leiden maar die op een beperkt aantal

producten/merken betrekking hebben aan de orde komen. Het onderscheid geven we aan met de symbolen ETETET respectievelijk ET.

3. GENERALISATIES: BEGRIPSBEPALING

Het meten op markten vindt zo'n vijftig jaar op enigszins grote schaal plaats. De eerder genoemde data-explosie is als gevolg van scanning zo'n vijftien jaar geleden ingezet. Dit heeft ertoe geleid dat er tal van analyses op veel verschillende markten zijn gedaan. Door statistische analyse van de resultaten van de individuele studies, kunnen door 'meta-analyses' generalisaties, wetmatigheden op markten worden geformuleerd (Wold, 1986). In de meta-analyses die in marketing science zijn uitgevoerd ligt de nadruk op het generaliseren van elasticiteiten en het bepalen van patronen in het koopgedrag. Ook zijn er meta-analyses verricht die de snelheid van acceptatie van nieuwe producten in kaart brengen. De generalisaties die door meta-analyses worden verkregen leiden niet tot één getal (parameter), bijvoorbeeld één gemiddelde of één elasticiteit. In empirische generalisaties wordt een bepaalde verdeling van de parameter verklaard uit andere variabelen, die de variatie in de parameter bepalen. Zo zullen prijselasticiteiten voor verschillende merken op verschillende markten niet allemaal dezelfde waarde bezitten. De prijselasticiteiten zullen systematisch van elkaar verschillen. In meta-analyses probeert men deze systematiek te achterhalen en te verklaren (Farley, Lehmann, Sawyer, 1995). De variabelen die hierbij gebruikt worden, worden *moderators* genoemd.

In deze bijdrage zullen we enkele belangrijke generalisaties bespreken. Sommige generalisaties zijn gebaseerd op meta-analyses die op hun beurt zijn gebaseerd op honderden individuele analyses. In deze analyses is het ETETET-proces geconvergeerd. In andere empirische studies hebben de uitkomsten van empirische studies geleid tot theorievorming (ET). De kennis die gebaseerd is op de ET-studies zal onder andere omstandigheden, dat wil vaak zeggen, andere markten in andere landen, maar ook gedurende andere tijdperiodes aangevuld en gevalideerd moeten worden.

We zullen achtereenvolgens enkele wetmatigheden in het koopgedrag, het concurrentiegedrag, de diffusie van nieuwe producten/merken, de invloed van verkrijgbaarheid (distributie), de prijsgevoeligheid, reclame en promoties bespreken.⁵ Daarbij gaat de meeste aandacht uit naar generalisaties die betrekking hebben op markten waar de afnemers finale consumenten zijn en waar niet-duurzame consumptiegoederen worden aangeboden. Markten met organisaties als afnemers zijn én onderling heterogener dan consumentenmarkten, én zijn minder onderzocht, mede omdat er minder marktonderzoeksgegevens beschikbaar zijn die het mogelijk maken om tot generalisaties te komen.⁶ De markten zijn vaak zo specifiek dat de resultaten van effectstudies niet gepubliceerd mogen worden.⁷

4. CONSUMENTENGEDRAG

Door een combinatie van deductie en inductie heeft men een min of meer uitgekristalliseerd beeld gekregen van de variabelen die de processen waaruit het consumentengedrag is opgebouwd bepalen. We verwijzen in dit verband naar het model van Blackwell, Miniard en Engel (2001).

Andere generalisaties hebben betrekking op de bepaling van die variabelen die tot een herhaalde aankoop leiden op basis van een relatie tussen vrager en aanbieder. Uit meta-analyses blijkt dat het vertrouwen van de afnemer in de aanbieder (*trust*) en de wederzijdse binding van de klant met de aanbieder (*commitment*) variabelen zijn, die bepalend zijn voor een relatie (Morgan, Hunt, 1994). De mate van tevredenheid over de relatie (*satisfaction*) bepaalt het vertrouwen en dit op zijn beurt het *commitment* in een relatie. Dit is trouwens ook het resultaat van een meta-analyse naar de factoren die van invloed zijn op relaties tussen ondernemingen in het distributiekanaal (Geyskens, Steenkamp, Kumar, 1999).

Opmerkelijke generalisaties over de samenstelling van de vraag naar een merk zijn in de loop der jaren geformuleerd door de Britse hoogleraar A.S.C. Ehrenberg. Hij heeft laten zien dat de verdeling van het aantal gekochte eenheden van een merk/product goed beschreven kan worden door een Negatief Binomiale Verdeling (NBD) of een *Logarithmic Series Distribution* (LSD) (Ehrenberg, 1959). Dezelfde statistische verdeling (maar dan met andere parameters) beschrijft de verdeling van de gemiddelde aankoopfrequentie van consumenten die tot herhalingsaankopen overgaan (Ehrenberg, 1972).

Uncles, Ehrenberg en Hammond (1995) laten zien dat merken met een hoog marktaandeel niet alleen meer kopers hebben dan merken met een laag marktaandeel, maar dat de kopers van de kleine merken minder kopen dan kopers van grote merken. Deze wetmatigheid staat bekend als *double jeopardy* ('dubbel gevaar'). Fader en Schmittlein (1993) vonden bovendien dat merken met een hoog marktaandeel een hoge mate van merktrouw bezitten.

5. CONCURRENTENGEDRAG

We zullen nu stilstaan bij generalisaties die in marketing science met betrekking tot het concurrentiegedrag geformuleerd zijn. Als een aanbieder de prijs voor merk A verlaagt, dan volgen reacties van andere merken. Veelal gebeurt dit met hetzelfde marktinstrument (hier prijs): *simple competitive reactions*.⁸ Een aanbieder kan reageren met een ander instrument dan dat waarmee zijn reactie wordt uitgelokt. Merk B kan zijn inspanningen vergroten om zijn verkrijgbaarheid te verhogen, of de reclame van merk B kan geïntensiveerd worden: *multiple competitive reactions*.

Uit een analyse van verschillende analyses⁹ blijkt dat:

1. de frequentie van *single competitive reactions* hoger is dan de frequentie waarmee elk van de combinaties van *multiple competitive reactions* zich voordoen;
2. er vaker sprake is van overreacties (merk B reageert op de acties van merk A terwijl dit niet nodig is om zijn verkopen of marktaandeel te beschermen) dan van onderreacties (merk B zou moeten reageren maar doet dit niet);
3. er reacties zijn die én nodig én effectief zijn. Deze reacties zorgen ervoor dat het verlies aan verkopen van merk B als gevolg van de actie van merk A door de reactie van merk B weer (voor een deel) wordt teruggewonnen. Deze reacties gaan veelal met verlies gepaard.

Deze reacties worden voornamelijk bestudeerd op oligopolide markten. Uit een meta-analyse van Dekimpe en Hanssens (1995) blijkt bovendien dat er weinig merken zijn die na de introductie- en groeiperiode van het merk een permanente verandering in het

marktaandeel kennen. Op veel markten treffen we een lange-termijn evenwicht aan in termen van marktaandelen. De relatieve posities van merken worden vaak alleen maar tijdelijk verstoord door acties van aanbieders. Deze resultaten gaan vooral op voor markten waarbij de introductie- en groeiperiode van de echte merken die op zo'n markt verschijnen voorbij is en de markt voor het eerst naar een stabiel evenwicht convergeert. Dit evenwicht blijft op de lange termijn vaak bestaan. Dit impliceert dat vanuit een macrostandpunt gezien, veel acties en reacties verspillingen zijn van aanbieders. De recente prijsoorlogen tussen de grote levensmiddelenlijsten in Nederland (2003-2005) bevestigen dit.

Evenwichten kunnen evenwel verstoord worden door nieuwkomers op markten die of een sterk ander product aanbieden, en/of concurreren met een permanent lagere prijs, en/of andere distributievormen hanteren (Shankar, Carpenter, Krishnamurthi, 1998). Zo heeft de introductie van drie commerciële zenders, Veronica, SBS6 en TMF in augustus 1995, het medialandschap permanent veranderd. De aandelen van reclameinkomsten van de commerciële en niet-commerciële zenders die zich richten op de Nederlandse markt, liggen na 1995 op een ander niveau dan voor 1995.¹⁰ Een ander voorbeeld: het permanent lage prijsniveau van Aldi en Lidl gekoppeld aan het grote assortiment eigen en B-merken dat zij voeren, heeft uiteindelijk geleid tot permanente verandering van de marktaandelen van met name de niet-geïntegreerde, zelfstandige detailhandel. Deze marktaandelen zijn tot een permanent lager niveau gedaald.

Internet biedt mogelijkheden om de (hierna nog te bespreken) sterke relatie tussen marktaandeel en beschikbaarheid (distributie) te doorbreken.¹¹ Ook dit kan leiden tot een permanente verandering van het marktaandeel van de aanbieder die internet gebruikt.

Alhoewel permanente veranderingen in marktaandelen niet veel voorkomen, evolueren de *verkopen* (in eenheden) van veel merken wel. De verkopen van een merk zijn per definitie gelijk aan het marktaandeel vermenigvuldigd met de vraag naar de productcategorie (primaire vraag). Als de verkopen evolueren en de marktaandelen niet, dan betekent dit dat de over merken geaggregeerde *vraag naar het product* in veel gevallen wel evolueert.

De evolutie van de *verkopen* op categorieniveau (primaire vraag) blijkt frequenter voor te komen dan op merkniveau. Wanneer één of meer merken toch in verkopen evolueren (maar dat komt relatief weinig voor), evolueren per definitie de verkopen op productcategorie-niveau. Dekimpe en Hanssens hebben met behulp van diverse moderators deze resultaten verder ontleed. De evolutie in verkopen was, zo bleek, het sterkst in de jaren zeventig van de vorige eeuw. Het percentage markten dat een permanente groei in de verkopen op productklasse-niveau vertoont, is groter in de Verenigde Staten en Canada dan in Europa. Voor meer generalisaties over concurrentiegedrag verwijzen we naar de special issue van *Marketing Science* vol. 24, nr.1.

6. GENERALISATIES OVER NIEUWE PRODUCTEN

We zullen enkele van de wetmatigheden die met betrekking tot productbeslissingen geformuleerd zijn bespreken. We beginnen met enige aandacht voor de entree van nieuwe producten op markten.

Er zijn veel studies die de relatie tussen de entreevolgorde op een markt en de lange-termijn waarde van het marktaandeel hebben bepaald. Markten worden gecreëerd door

succesvolle pioniers. Van de niet-succesvolle pioniers wordt na enige tijd niet veel meer vernomen en er worden geen gegevens vastgelegd. Doorgaans hebben geslaagde innovators het hoogste marktaandeel (Robinson, Fornell, 1985). Zelfs na decennia zijn de pioniers van toen nog marktleiders.¹² Het merk dat als tweede een markt betreedt, heeft een lange-termijn marktaandeel dat lager is dan dat van de pioniers; hetzelfde gaat op voor merken die nog later geïntroduceerd worden. In diverse meta-studies heeft men deze relatie geobjectiveerd. Stel dat men als x -de merk een markt betreedt, dan is y , het lange-termijn marktaandeel van dat merk ten opzichte van het merk van de innovator, gelijk aan: $y = 1/\sqrt{x}$. Het merk dat bijvoorbeeld als vierde ($x=4$) op de markt geïntroduceerd wordt heeft een marktaandeel dat circa de helft is van de pionier ($y=1/2$). In drie meta-analyses waarin gegevens over consumentengoederen en geneesmiddelen onderzocht worden is de relatie $y = 1/\sqrt{x}$ bepaald; zie tabel 1.

Tabel 1. Relatie tussen entreevolgorde en marktaandeel ten opzichte van het merk dat als eerste de markt betreedt.

	Consumentengoederen		Geneesmiddelen (maagzuurremmers)
	Urban et.al (1986)	Kalyanaram en Urban (1992)	Berndt et.al (1995)
1	1.02	1.00	1.00
2	0.71	0.76	0.70
3	0.58	0.64	0.57
4	0.51	0.57	0.49
5	0.45	0.53	0.44
6	0.41	0.49	0.40

Bron: Kalyanaram, Robinson, Urban (1995)

De verschillen in marktaandelen tussen innovator, volgers en late toetreders worden gedurende de levenscyclus van het product kleiner (Brown, Lattin, 1994; Huff, Robinson, 1994). De verschillen tussen het marktaandeel van een pionier en de marktaandelen van volgers worden niet alleen bepaald door de entreevolgorde, maar ook door de tijd dat een pionier alleen op de markt is (*time in market*). Is hij lange tijd monopolist dan zal het verschil in marktaandeel tussen de pionier en de eerste volger groter zijn dan wanneer het monopolie slechts korte tijd kan bestaan (Brown, Lattin, 1994). *Time in market* is een voorbeeld van een *moderator*.

De resultaten uit de studies die in tabel 1 vermeld staan zijn gebaseerd op succesvolle pioniers. Wanneer alle merken die ooit op een bepaalde markt geïntroduceerd zijn in de analyse worden betrokken zijn die uitkomsten anders. Golder en Tellis (1993) nemen in hun analyse wel alle merken die ooit geïntroduceerd zijn in beschouwing. Golder en Tellis (1993) vinden een hoog percentage mislukkingen van merken van pioniers en concluderen dat het voordeel van pioniers vaak een legende is. Met behulp van de gegevens uit hun database (PIMS data), concluderen zij dat het percentage mislukkingen 47% is.

Eerder hebben we gesteld dat veel generalisaties gebaseerd zijn op metingen op markten waar niet-duurzame consumptiegoederen aangeboden worden. Een belangrijke uit-

zondering hierop wordt gevormd door de kennis die gegenereerd is op het terrein van de adoptie en diffusie van (nieuwe) duurzame consumptiegoederen (Mahajan, Muller, Bass, 1995). De generalisaties hebben met name betrekking op het verloop van de diffusiecurven en welke variabelen dit verloop beïnvloeden.

Tellis, Stremersch en Yin (2003) hebben de *takeoff* in de verkopen van 137 producten, in 10 categorieën duurzame consumptiegoederen, in 16 Europese landen verklaard. De *takeoff* is de snelle groei, na de introductieperiode, van een nieuw product. De tijd tot dat de *takeoff* plaatsvindt wordt beïnvloed door het land waar het product geïntroduceerd wordt, hetgeen samenhangt met de cultuur (risicoaversie, ambitieniveaus), met de welvaart en de media-intensiteit in de verschillende landen. Daarnaast zijn er verschillen tussen de verschillende productcategorieën. De moderatoren land en productcategorie bepalen, met andere woorden, de lengte van de tijd tot de *takeoff*.

We geven enkele andere generalisaties met betrekking tot nieuwe en bestaande producten in tabel 2 weer.

Tabel 2. Enkele wetmatigheden met betrekking tot productbeslissingen.

-
1. De diffusie van een nieuw product wordt bepaald door externe invloeden zoals reclame, direct mail, inzet van vertegenwoordigers en interne invloeden zoals mond-tot-mond reclame. De externe invloed is bijna altijd veel kleiner dan de invloed die afnemers onderling op elkaar hebben (Sultan, Farley, Lehmann 1990) (ETETET).
 2. Verbeteringen in de kwaliteit van een product hebben een positieve invloed op de verkopen. Verdere verbeteringen in de kwaliteit van een product leiden tot afnemende meeropbrengsten (Lambin, 1976) (ET).
 3. De perceptie van de kwaliteit van een product door afnemers, heeft een negatief verband met de hoogte van het marktaandeel (Hellofs, Jacobson, 1999) (ETETET).
-

7. DISTRIBUTIE

Diverse studies en observaties¹³ geven aan dat er een nauwe relatie bestaat tussen het marktaandeel van een merk en het aantal plaatsen/distributiepunten/winkels waar het merk verkrijgbaar is (numerieke/gewogen distributie). Een hogere gewogen distributie leidt tot een hoger marktaandeel, waarbij veelal sprake is van een convex verloop. Het omgekeerde doet zich ook voor: een hoger marktaandeel leidt tot een hogere gewogen distributie. Het ontwarren van de wederzijdse relaties tussen marktaandeel en numerieke/gewogen distributie staat nog in de kinderschoenen. We worden hier geconfronteerd met endogeniteitsproblemen, waardoor we andere dan de gebruikelijke schattingsmethoden zullen moeten toepassen.¹⁴

Een dergelijk probleem doet zich ook voor wanneer we de relatie tussen de ruimte die producten in het schap in een winkel innemen en de verkopen van deze producten willen bepalen. Recent is men er in geslaagd om met behulp van methoden die in de ruimtelijke econometrie ontwikkeld zijn, schapelasticiteiten te bepalen.¹⁵ Deze liggen gemiddeld rond de 0,2%. Een verhoging van de schapruimte van een artikel met 10% zal tot een verhoging van de verkopen van dat artikel leiden met 2%. Schapelasticiteiten met deze waarde zijn ook gevonden in andere studies die voornamelijk met behulp van experimenten verricht werden.¹⁶

8. PRIJS

De gevoeligheid van de vraag en het aanbod voor prijsfluctuaties wordt doorgaans gemeten met behulp van prijselasticiteiten. We maken een onderscheid in studies waarin de prijselasticiteit op het niveau van de productklasse (primaire vraagniveau) wordt bepaald en studies waarin prijselasticiteiten op merkniveau vastgesteld zijn.

De prijselasticiteiten op het niveau van de primaire vraag zijn, in absolute zin, beduidend lager dan op merkniveau. We zien een grote spreiding van prijselasticiteiten op het niveau van de productklasse. Wel liggen alle elasticiteiten tussen nul en min één (Pagoulatos, Sorensen, 1986). Ontbijtproducten, dierenvoeding en siroop bezitten prijselasticiteiten die dicht bij nul liggen. Vis, visproducten in blik en sigaren hebben prijselasticiteiten met een (absolute) waarde van rond 0,7. Welke moderatoren verklaren nu deze verschillen? Pagoulatos en Sorensen (1986) vonden dat productklassen waar veel reclame wordt gemaakt en/of die hoge research- en ontwikkelingskosten hebben, lagere prijselasticiteiten bezitten dan productklassen waar dat niet het geval is. Zij vonden bovendien dat producten die aan bedrijven worden geleverd (industriële producten en diensten), minder prijselastisch zijn dan finale consumentengoederen.

Dat de prijselasticiteiten op merkniveau (absoluut) veel hoger zijn dan die op productklassenniveau, valt te verklaren uit het feit dat een initiële prijsverandering vaak gevolgd wordt door prijsveranderingen van concurrenten. Daardoor wordt het initiële effect op de vraag voor een deel gecompenseerd (zie hiervoor: concurrentiegedrag).

In een in 1988 gepubliceerde studie, vond Tellis (1988) een gemiddelde prijselasticiteit van -1,76. Recent vonden Bijmolt, Van Heerde en Pieters (2005) een gemiddelde prijselasticiteit van -2,62 op basis van een meta-analyse waarin 1851 verschillende prijselasticiteiten uit 81 studies geanalyseerd werden. De verschillen tussen de prijselasticiteiten kunnen verklaard worden door de volgende moderatoren:

1. het type goederen: duurzame consumptiegoederen hebben hogere prijselasticiteiten dan levensmiddelen en deze zijn (althans in de Verenigde Staten) hoger dan prijselasticiteiten voor geneesmiddelen;
2. prijselasticiteiten nemen in absolute waarde toe in de tijd wanneer de verkopen (en niet marktaandeel) als afhankelijke variabele wordt gebruikt;
3. inflatie leidt tot hogere prijselasticiteiten;
4. elasticiteiten op variëteit-niveau (Stock Keeping Units: sku-niveau) zijn (absoluut) hoger dan elasticiteiten op merkniveau (en die waren hoger dan elasticiteiten op productklasse-niveau).

Kruislingse elasticiteiten zijn op merkniveau frequent gemeten. Uit een meta-analyse van Sethuraman, Srinivasan en Kim (1999) blijkt dat de absolute waarden van deze elasticiteiten lager zijn dan 'eigen' elasticiteiten. Ongeveer 70 procent van de (absolute waarde van de) geschatte elasticiteiten ligt tussen de 0 en 1 en 15 procent ligt tussen 1 en 2. De gemiddelde kruislingse elasticiteit bedraagt 0,52. Wanneer merken meer op elkaar lijken, zijn de kruislingse elasticiteiten hoger.

9. RECLAME

Een verhoging van de reclame-uitgaven leidt (bijna) nooit tot een permanente verandering van het koopgedrag (Dekimpe, Hanssens, 1999). De zogenaamde *persistence* van

reclame is met andere woorden gelijk aan nul. Uitzonderingen zijn onder meer de effecten van verhogingen van het reclamebudget van Crest tandpasta waarin benadrukt werd dat Crest fluor bevatte¹⁷ en de verhoging van het budget voor Gorbatsjov-wodka toen de gelijknamige Russische leider sterk in de belangstelling kwam te staan (Simon, 1997).

Op korte en (middel)lange termijn kan reclame tot geringe effecten in de verkopen of het marktaandeel leiden. De korte-termijn reclame-elasticiteiten liggen tussen nul en 0,2 (Lambin, 1976; Leone, Schultz, 1980; Assmus, Farley, Lehmann, 1984; Vakratsas, Ambler, 1999). Reclame heeft met name effect als dit verbonden is met iets nieuws. Zo hebben nieuwe producten hogere reclame-elasticiteiten (0,25) dan bestaande producten (0,05). Andersom: wanneer een prijsstijging ondersteund wordt door reclame, dan is de prijselasticiteit (veel) hoger (in absolute waarde) dan wanneer de prijsverandering niet gecommuniceerd wordt (Bemmar, Mouchoux, 1991; Kaul, Wittink, 1995; Van Heerde, Leeflang, Wittink, 2000, 2001, 2004). Wanneer er gekozen wordt voor het communiceren van andere elementen dan de prijs (non-price advertising) dan neemt de gevoeligheid voor prijsveranderingen af (Kaul, Wittink, 1995). Het communiceren over nieuwe verpakkingen, productmodificaties, nieuwe variëteiten en nieuwe distributiepunten waar een merk verkrijgbaar is, leidt veelal tot significante (van nul verschillende) effecten van de reclame op de verkopen.

Het kan enige tijd duren voordat de effecten van reclame in de tijd volledig weggeëbd zijn. De tijd dat 90 procent van het totale effect gerealiseerd is (90% *duration interval*) ligt tussen de drie en negen maanden (Clarke, 1976). Voor televisiereclame zijn evenwel 90%-*duration intervals* van tussen de twee en drie jaar gevonden.¹⁸

Hogere reclame-uitgaven leiden al snel tot afnemende meeropbrengsten (Little, 1979). Wanneer consumenten driemaal geconfronteerd zijn geweest met een reclameboodschap, neemt de aandacht voor de boodschap af. Reclamemakers kunnen zich beter richten op het bereiken van meer mensen dan op het herhaald bereiken van potentiële consumenten (Deighton, Henderson, Neslin, 1994).¹⁹

Alhoewel reclame effect kan hebben op de verkopen op het niveau van het merk, zijn de effecten van merkreclame op het niveau van de primaire vraag nihil. Voor veel productcategorieën is het primaire vraagniveau een gegeven. Dit niveau kan niet door reclame voor één of meer merken beïnvloed worden.

Naast kennis die wordt gegenereerd met behulp van stimulus-response modellen en modellen die via een bepaalde sequentie of hiërarchie aangeven hoe reclame kan werken, is er recent aandacht voor de effecten van de inhoud van reclameboodschappen. Deze effecten kunnen tegenwoordig goed gemeten worden door de aandacht van reclameboodschappen te meten met behulp van oogbewegingen.²⁰ Zo blijkt dat 'humor' en het geven van weinig informatie in een televisiecommercial een positief effect hebben op de aandacht voor de commercial.²¹ Voor uitingen van advertenties in 'print' (dagbladen, tijdschriften) wordt de aandacht met name verkregen door de afbeelding, het plaatje en de foto die in de advertentie is opgenomen. Het tweede belangrijke element van een advertentie, de grootte van de oppervlakte van de tekst, heeft een rechtstreekse positieve relatie met de aandacht voor een reclame-uiting. Het derde element, de aandacht voor het merk (merknaam/symbool), heeft een positieve relatie met de aandacht voor de afbeelding en de tekst (Pieters, Wedel, 2004).

10. PROMOTIES

Promoties zijn tijdelijke wijzigingen in de marketingmix. Leiden deze wijzigingen nu tot wijzigingen in de verkopen? Het verhaal wordt wellicht eentonig, want evenals bij de effecten van reclame moeten we concluderen dat promoties in het algemeen niet tot permanente wijzigingen in de vraag leiden, noch op merkniveau, noch op het niveau van de primaire vraag (Nijs, Dekimpe, Steenkamp, Hanssens, 2001). De korte termijn-effecten van promoties zijn hoog. Bij reclame komen we elasticiteiten tussen nul en 0,25 tegen, bij prijskortingen die door reclame binnen en buiten de winkel ondersteund wordt, ontmoeten we elasticiteiten van -2 tot zelfs -10. Dit betekent dat een (ondersteunde) korting van 20/25 procent wel tot een afzetstijging (in eenheden) van 250 procent kan leiden.²² In de praktijk zijn de prijskortingen vaak te laag of te hoog, zijn of worden niet ondersteund door reclame binnen en buiten de winkel, vinden te vaak plaats (frequentie te hoog) en worden ingezet als reactie op het gedrag van concurrenten terwijl dit niet nodig is. Het is dan ook niet verwonderlijk dat slechts 16 procent van de promoties die gericht zijn op de detailhandel en 11 procent van de consumentenpromoties winstgevend zijn.²³

Wanneer ondernemers gebruikmaken van wetmatigheden aangaande de werking van promoties, dan kunnen de eerder genoemde winstpercentages aanzienlijk verhoogd worden. De wetmatigheden staan beknopt beschreven in onderstaande opsomming:

1. Promoties leiden niet tot permanente veranderingen in de verkopen, noch op merkniveau, noch op productcategorieniveau (Nijs, Dekimpe, Hanssens, Steenkamp, 2001).
2. De korte termijneffecten van (prijs)promoties zijn hoog: elasticiteiten zijn in absolute waarde groter dan 2. (Wittink, Addona, Hawkes, Porter, 1988).
3. Elasticiteiten van prijspromoties zijn hoger dan prijselasticiteiten (Blattberg, Briesch, Fox, 1995).
4. Prijspromotie-elasticiteiten zijn hoger voor merken in productcategorieën met weinig merken, korte tussenaankooptijden, een hogere penetratie van het product en voor productcategorieën waarvoor makkelijk voorraad te houden is (Narasimhan, Neslin, Sen, 1996).
5. Tijdelijke aandacht voor een merk door (tijdelijke) reclame buiten de winkel (*feature-ring*²⁴ in dagbladen en huis-aan-huis bladen) en binnen de winkel (*displays*) heeft een groot effect op de verkopen (Blattberg, Briesch, Fox, 1995).
6. Prijskortingen die ondersteund worden door *feature advertising* en/ of een *display* hebben meer effect dan niet-ondersteunde prijskortingen (Bemmar, Mouchoux, 1991; Van Heerde, Leeflang, Wittink, 2000, 2001, 2004). (ET).²⁵
7. Diepe promoties (hoge kortingen) leiden tot een (steeds) meer inelastische vraag (Foekens, Leeflang, Wittink, 1999). (ET).
8. Des te frequenter een promotie wordt gehouden, des te inelastischer wordt de vraag naar het merk (Blattberg, Briesch, Fox, 1995; Foekens, Leeflang, Wittink, 1999). Hetzelfde geldt voor de vraag naar de productcategorie (Raju, 1992).
9. Promotionele kortingen die lager zijn dan 10 procent leveren niet meer verkopen op dan wanneer er geen korting wordt gegeven. Kortingen boven de 25 procent leveren niet meer verkopen dan kortingen rond de 25 procent (Van Heerde, Leeflang, Wittink, 2001). (ET).
10. De additionele vraag naar een merk als gevolg van een prijspromotie bestaat voor: 1/3 uit merksubstitutie;

- 1/3 uit voorraadvorming en anticipatie op promoties;
 1/3 uit meerverkopen op categorieniveau. (Van Heerde, Gupta, Wittink, 2003; Van Heerde, Leeflang, Wittink, 2004). (ET).
11. Een promotie voor één variëteit van een merk (sku) leidt tot kannibalisatie van andere variëteiten van hetzelfde merk en wel tot 80 procent (Kalyanam, Putler, 1997; Van Heerde, Leeflang, Wittink, 2004). (ET).
 12. Een belangrijk deel van de meerverkopen op categorieniveau in een winkel als gevolg van een promotie, komt voort uit winkelsubstitutie (Van Heerde, Leeflang, Wittink, 2004). (ET).
 13. Des te hoger het marktaandeel van een merk, des te minder gevoelig is dat merk voor een prijskorting (Blattberg, Briesch, Fox, 1995).
 14. De promotie voor een merk in een productcategorie i heeft slechts beperkte effecten op de verkopen in andere productcategorieën $j \neq i$ (Van Dijk, Leeflang, Parreño Selva, Wittink, 2004; Wedel, Zhang, 2004). (ET).

11. ANDERE WETMATIGHEDEN

Naast de gegeven voorbeelden wordt er hard gewerkt om ook op tal van andere terreinen te kunnen formuleren *how marketing works*. Zo wordt er heel wat onderzoek verricht naar merkextensies. Dat wil zeggen welke nieuwe producten kunnen onder dezelfde, bestaande merknaam op de markt gebracht worden en welke merken zijn beter af met een nieuwe naam (Aaker, Keller, 1992; Bottomley, Holden, 2001; Czellar, 2003). Iets dergelijks doet zich voor wanneer twee merken samen als het ware een nieuw 'merk' gaan vormen (Dell met Intel processors). Dit wordt *co-branding* genoemd. Wanneer wordt dit wel geapprecieerd door afnemers en wanneer minder en wat zijn de effecten op de constituerende merken (*spill-over effects*)?²⁶

In ander onderzoek vraagt men zich af onder welke omstandigheden het zinvol is om een internetkanaal toe te voegen aan de bestaande directe en indirecte kanalen. 'Zinvol' wordt dan vertaald in termen van bijdrage tot de winst, de cashflow of de reactie op de beurskoers.²⁷

12. TOEKOMSTIG ONDERZOEK

Het meten op markten heeft voornamelijk plaatsgevonden op markten van frequent aangekochte consumentengoederen. De kennis op markten waar duurzame consumptiegoederen worden aangeboden is aanzienlijk minder. Hetzelfde geldt voor markten waar diensten worden geleverd aan finale consumenten. Kennis in de vorm van elasticiteiten, de spreiding van elasticiteiten die verklaard kan worden door moderators en werking is beperkt in de business-to-business marketing (industriële marketing). Toekomstig onderzoek kan gericht zijn op het genereren van generalisaties op deze veelal heterogene markten, waar relatief weinig gegevens voor handen zijn.

Op markten met veel gegevens zijn tal van nieuwe onderzoeken te entameren. Zo is er nog relatief weinig bekend over de rol die internet speelt of gaat spelen in het consumentengedrag.²⁸ Bij het meten van effecten van inspanningen op markten worden veelal maatstaven als verkopen, winst, rendement of contributiemarge gebruikt. We zien dat er steeds meer situaties komen waarbij de effecten van inspanningen op variabelen als klanttevredenheid, klantbehoud (*customer retention*) en *customer share* (deel van de

consumptie dat voor rekening van een bepaalde aanbieder komt) gemeten worden (Verhoef, 2003). Ook het meten van de waarde die klanten voor organisaties bezitten, staat steeds meer in de belangstelling (Bolton, Lemon, Verhoef, 2004).

Er zijn tal van technische problemen op te lossen die bij het modelleren van markten een rol spelen. Te denken valt aan aggregatievraagstukken zoals 'Hoe aggregeren we individuele vraagrelaties tot vraagrelaties op winkelniveau en op marktniveau?'²⁹ Een ander aggregatievraagstuk heeft betrekking op de definiëring van vraageenheden. Modelleren we de vraag naar een *stock keeping unit* (sku) zoals een flesje Fanta, citroensmaak in een plastic fles van 500 cl, of modelleren we de vraag naar het merk Fanta?

Een ander probleem waar we mee worstelen is het endogeniteitsvraagstuk.³⁰ Hoe kunnen we de relaties tussen endogene variabelen op een zo goed mogelijke wijze ontwarren zodat we parameters schatten die geen *bias* bezitten? Veel van het toekomstig onderzoek zal verder gedomineerd worden door het toepassen van de modernere methoden van econometrisch- en psychometrisch onderzoek.

De resultaten van het onderzoek dat we hiervoor beschreven kunnen van belang zijn voor nieuwe afnemers. Daarbij denken we aan beleidsmakers bij de nationale en Europese overheid. Vraagstukken als 'Wat is het effect van het verbod van reclame voor bepaalde producten (geneesmiddelen,³¹ rookwaren, alcoholhoudende dranken, cosmetica, speelgoed, enz.)?' of 'Wat zijn de gevolgen van accijnsveranderingen?' kunnen met behulp van marketing science benaderd worden. Ook kartelpolitieke uitspraken, onder meer over het toestaan van een fusie of alliantie tussen ondernemingen en het gevolg ervan, kunnen met behulp van modellen uit marketing science worden ondersteund.

Ook zien we dat modellen steeds vaker een rol gaan spelen in onderzoek naar de waarde van merken. Zeker daar waar het bepalen van de schade aan de waarde van merken in het geding is, hebben deze modellen in ieder geval hun waarde bewezen.³²

Samengevat, door te meten op markten, en zeker op markten met veel gegevens, gaan we meer weten over marktwerking. Van koopmanschap, via het marketing management wordt marketing daarmee een meer volwassen veld van wetenschappelijk onderzoek. De gegeneerde kennis en de methoden die we gebruiken om hiertoe te komen en die bekend staan als marketing science zullen zowel de marketing manager als de wetenschap zelf kunnen dienen.

NOTEN

1. Dank ben ik verschuldigd aan Dr. J. van Doorn, Prof. dr. J.C. Hoekstra, Dr. F. van Perlo-ten Kleij en Prof. dr. P.C. Verhoef (Fac. der Economische Wetenschappen, Rijksuniversiteit Groningen) die zich de moeite gaven een vorige versie van dit stuk van commentaar te voorzien. Veel dank ben ik verschuldigd aan mijn vriend en collega Prof. dr. D.R. Wittink (School of Management, Yale University, New Haven, Conn. en Fac. der Economische Wetenschappen, Rijksuniversiteit Groningen) met wie ik in de afgelopen 15 jaar intensief mocht samenwerken. Zijn plotselinge overlijden op 4 juni 2005 betekent helaas een abrupt einde van deze samenwerking.
2. Dalrymple, G. (1990, p. 28)
3. Katz, Rosen (1991, Ch. 10); Mas-Colell, Whinston, Green (1995, Ch. 10).
4. Ehrenberg (1994).
5. Zie ook Koster, Leeflang (1992), Leeflang, Wittink, Wedel, Naert (2000, pp. 28-32), Hanssens, Parsons, Schultz (2001, pp. 319-356) en diverse onderdelen uit Leeflang (2003).
6. Eliashberg, Lilien, Kim (1995).

7. Zie ook Brand (1993) en Brand, Leeftang (1994) voor een overzicht.
8. Lambin, Naert, Bultez (1975).
9. Leeftang, Wittink (1992, 1996); Steenkamp, Nijs, Hanssens, Dekimpe (2005).
10. Kornelis, Dekimpe, Leeftang (2005).
11. Lynch Jr., Arieli (2000); Geyskens, Gielens, Dekimpe (2002).
12. Leeftang (1994, p. 134).
13. Reibstein, Farris (1995).
14. Wansbeek, Wedel (1999), Ebbes (2004).
15. Van Dijk, Van Heerde, Leeftang, Wittink (2004).
16. Curhan (1972); Heinsbroek (1977); Bultez, Naert (1988); Hoch, Drèze, Purk (1994).
17. Wichern, Jones (1977).
18. Lodish et. al (1995a, 1995b).
19. Voor een overzicht van andere wetmatigheden met betrekking tot de werking van reclame zie Vakratsas, Ambler (1999).
20. Pieters (1998); Rosbergen (1998); Leeftang, Pieters (1999); Wedel, Pieters (2000); Pieters, Warlop, Wedel (2002); Pieters, Wedel (2004).
21. Woltman Elpers (2003); Woltman Elpers, Pieters, Wedel (2003).
22. Wittink, Adonna, Hawkes, Porter (1988); Van Heerde, Leeftang, Wittink (2001, 2004).
23. Abraham, Lodish (1989); Campo (1998); Smit (1999).
24. Featuring: reclame waarin een detaillist aandacht vestigt op een merk dat veelal in de aanbieding is. Hierbij worden huis-aan-huis bladen, folders en dagbladen als media ingezet. Displays: extra aandacht voor een merk door middel van borden in de winkel, extra reclameuitingen in de winkel, enz.
25. (ET) betekent dat de wetmatigheden gebaseerd zijn op een beperkt aantal studies. De andere in de opsomming gemeende wetmatigheden zijn gebaseerd op een groot aantal studies c.q. studies waarin een groot aantal verschillende productcategorieën zijn betudeerd (zie ook paragraaf 2).
26. Zie Helmig, Huber, Leeftang (2005).
27. Geyskens, Gielens, Dekimpe (2002).
28. Zie bijv. de special issue van *Marketing Science*, vol. 19, nr. 1, 2000.
29. Foekens, Leeftang, Wittink (1994); Christen, Gupta, Staelin, Porter, Wittink (1997).
30. Zie ook Wansbeek, Wedel (1999); Ebbes (2004).
31. Zie Leeftang, Wittink (2002); Leeftang, Wieringa, Wittink, Wolttil (2004), Leeftang, Wieringa, Wittink (2005).
32. Zie Hanssens, Leeftang, Wittink (2005).

LITERATUUR

- Aaker, D.A. & K.L. Keller (1992), 'The Effects of Sequential Introduction of Brand Extensions', *Journal of Marketing Research*, vol. 29, pp.35-51.
- Abraham, M.M. & L.M. Lodish (1989), 'Fact-Based Strategies for Managing Advertising and Promotion Dollars: Lessons from Single Source Data', Working Paper # 89-006, Marketing Department, The Wharton School of the University of Pennsylvania.
- Assmus, G., J.U. Farley & D.R. Lehmann (1984), 'How Advertising Affects Sales: Meta Analysis of Econometric Results', *Journal of Marketing Research*, vol. 21, pp. 153-158.
- Bemmar, A.C. & D. Mouchoux (1991), 'Measuring the Short Term Effect of In-store Promotion and Retail Advertising on Brand Sales: A Factorial Experiment', *Journal of Marketing Research*, vol. 28, pp. 202-214.
- Berndt, E.R., L. Bui, D.R. Reiley & G.L. Urban (1995), 'Information, Marketing, and Pricing in the U.S. Antilucer Drug Market', *The American Economic Review*, vol. 85, pp. 100-105.
- Bijmolt T.H.A., H.J. van Heerde & R.G.M. Pieters (2005), 'Determinants of Price Elasticity:

- New Empirical Generalizations', *Journal of Marketing Research*, vol. 42, pp. 141-156.
- Blackwell, R.D., P.W. Miniard & J.F. Engel (2001), *Consumer Behavior*, The Dryden Press, Orlando, 9th International Edition.
- Blattberg, R.C., R. Briesch & E.J. Fox (1995), 'How Promotions Work', *Marketing Science*, vol. 14, pp. G122-G132.
- Bolton, R. N., K.N. Lemon & P.C. Verhoef, 'The Theoretical Underpinnings of Customer Asset Management: A Framework and Propositions for Future Research', *Journal of the Academy of Marketing Science*, vol. 32, pp. 271-293.
- Bottomley, P.A. & S.J.S. Holden (2001), 'Do We Really Know How Consumers Evaluate Brand Extensions? Empirical Generalizations Based on Secondary Analysis of Eight Studies', *Journal of Marketing Research*, vol. 38, pp. 494-500.
- Brand, M.J. & P.S.H. Leeflang (1994), 'Research on Modeling Industrial Markets' in Laurent, G., G.L. Lilien & B. Pras (eds.), *Research Traditions in Marketing*, Kluwer Publishers, Boston, pp. 231-261.
- Brand, M.J. (1993), 'Effectiveness of the Industrial Marketingmix: An Assessment Through Simulation of the Organizational Buying Process', Unpublished Ph.D. thesis, University of Groningen.
- Brown, C.L. & J.M. Lattin, (1994), 'Investigating the Relationship between Time in Market and Pioneering Advantage', *Management Science*, vol. 40, pp. 1361-1369.
- Bultez, A., P.A. Naert (1988), 'S.H.A.R.P.: Shelf Allocation for Retailers' Profit', *Marketing Science*, vol. 7, pp. 211-231.
- Campo, K. (1998), 'Promoties: Effectief op Korte Termijn, Destructief op Lange Termijn? Een Overzicht van het Marketingdebat', *Economisch en Sociaal Tijdschrift*, nr. 2, pp. 175-220.
- Christen, M., Sachin Gupta, J.C. Porter, R. Staelin, D.R. Wittink (1997), 'Using Market-Level Data to Understand Promotion Effects in a Nonlinear Model', *Journal of Marketing Research*, vol. 34, pp. 322-334.
- Clarke, D.G. (1976), 'Econometric Measurement of the Duration of Advertising Effect on Sales', *Journal of Marketing Research*, vol. 13, pp. 345-357.
- Curhan, R.C. (1972), 'The Relationship between Shelf Space and Unit Sales in Supermarkets', *Journal of Marketing Research*, vol. 9, pp. 406-412.
- Czellar, S. (2003), 'Consumer Attitude toward Brand Extensions: An Integrative Model and Research Propositions', *International Journal of Research in Marketing*, vol. 20, pp. 97-115.
- Dalrymple, G. (1990). *Advertising in America: The First Two Hundred Years*. Harry N. Abrams, New York.
- Deighton, J., C.M. Henderson & S.A. Neslin (1994), 'The Effects of Advertising on Brand Switching and Repeat Purchasing', *Journal of Marketing Research*, vol. 31, pp. 28-43.
- Dekimpe, M.G., & D.M. Hanssens (1999), 'Sustained Spending and Persistent Repsonse: A New Look at Long-Term Marketing Profitability', *Journal of Marketing Research*, vol. 36, pp. 397-412.
- Dijk A. van, H.J. van Heerde, P.S.H. Leeflang & D.R. Wittink (2004), 'Similarity-Based Spatial Methods to Estimate Shelf Space Elasticities', *Quantitative Marketing and Economics*, vol. 2, pp. 257-277.
- Dijk, A. van, P.S.H. Leeflang, J. Parreño Selva & D.R. Wittink (2004), 'Sales Bump Decomposition to Infer Cross-Category Effects Using Daily Store Data', research paper, University of Groningen.
- Ebbes, P. (2004), 'Latent Instrumental Variables, a New Approach to Solve for Endogeneity', Unpublished Ph. D. thesis, University of Groningen, the Netherlands.
- Ehrenberg, A.S.C. (1959), 'The Pattern of Consumer Purchases', *Applied Statistics*, vol. 8, pp. 26-41.
- Ehrenberg, A.S.C. (1972), *Repeat Buying, Theory and Applications*, North-Holland Publishing Company, Amsterdam.
- Ehrenberg, A.S.C. (1994), 'Theory or Well-based Results: Which Comes First?' in Laurent, G.,

- G.L. Lilien & B. Pras (eds.), *Research Traditions in Marketing*, Kluwer Academic Publishers, Boston, pp. 79-108.
- Eliashberg, J., G.L. Lilien & N. Kim (1995), 'Searching for Generalizations in Business Marketing Negotiations', *Marketing Science*, vol. 14, pp. G47-G60.
- Fader, P.S. & D.C. Schmittlein (1993), 'Excess Behavioral Loyalty for High Share Brands: Deviations from the Dirichlet Model for Repeat Purchasing', *Journal of Marketing Research*, vol. 30, pp. 478-493.
- Farley, J.U., D.R. Lehmann & A. Sawyer (1995), 'Empirical Marketing Generalizations Using Meta-Analysis', *Marketing Science*, vol. 14, pp. G36-G46.
- Foekens, E.W., P.S.H. Leeflang & D.R. Wittink (1994), 'A Comparison and an Exploration of the Forecasting Accuracy of a Loglinear Model at Different Levels of Aggregation', *International Journal of Forecasting*, vol. 10, pp. 245-261.
- Foekens, E.W., P.S.H. Leeflang & D.R. Wittink (1999), 'Varying Parameter Models to Accommodate Dynamic Promotion Effects', *Journal of Econometrics*, vol. 89, pp. 249-268.
- Geyskens, I., J.B.E.M. Steenkamp & N. Kumar (1999), 'A Meta-Analysis of Satisfaction in Marketing Channel Relationships', *Journal of Marketing Research*, vol. 36, pp. 223-238.
- Geyskens, I., K. Gielens & M.G. Dekimpe (2002), 'The Market Valuation of Internet Channel Additions', *Journal of Marketing*, vol. 66, pp. 102-119.
- Golder, P.N. & G.J. Tellis (1993), 'Pioneer Advantage: Marketing Logic or Marketing Legend?', *Journal of Marketing Research*, vol. 30, pp. 158-170.
- Hanssens, D.M., L.J. Parsons & R.L. Schultz (2001), *Market Response Models: Econometric and Time Series Analysis*, 2nd ed., Kluwer Academic Publishers, Boston.
- Hanssens, D.M., P.S.H. Leeflang & D.R. Wittink (2005), 'Market Response Models and Marketing Practice', *Applied Stochastic Models for Business and Industry*, vol. 21, nr 4/5, 423-434.
- Heerde, H. van, S. Gupta & D.R. Wittink (2003), 'Is 75% of the Sales Promotion Bump due to Brand Switching? No Only 33% Is', *Journal of Marketing Research*, vol. 40, pp. 481-491.
- Heerde, H. van, P.S.H. Leeflang & D.R. Wittink (2000), 'The Estimation of Pre- and Postpromotion Dips with Store-Level Scanner Data', *Journal of Marketing Research*, vol. 37, pp. 383-396.
- Heerde, H. van, P.S.H. Leeflang & D.R. Wittink (2001), 'Semiparametric Analysis to Estimate the Deal Effect Curve', *Journal of Marketing Research*, vol. 38, pp. 197-215.
- Heerde, H. van, P.S.H. Leeflang & D.R. Wittink (2004), 'Decomposing the Sales Promotion Bump with Store Data', *Marketing Science*, vol. 23, pp. 317-334.
- Heinsbroek, H. (1977), 'De Maximalisatie van de Schapruimte-toewijzing in de Supermarkt', Unpublished Ph.D. thesis, Free University of Brussels, Belgium.
- Hellofs, L.L. & R. Jacobson (1999), 'Market Share and Customers' Perceptions of Quality: When Can Firms Grow Their Way to Higher Versus Lower Quality?', *Journal of Marketing*, vol. 63, pp. 16-25.
- Helmig, B., J.A. Huber & P.S.H. Leeflang (2005), 'Co-branding: The State-of-the-Art', Research Paper, Department of Business Administration, University of Fribourg, Switzerland.
- Hoch, S.J., X. Drèze & M.E. Purk (1994), 'EDLP, Hi-Lo, and Margin Arithmetic', *Journal of Marketing*, vol. 58, pp. 16-27.
- Hoekstra, J.C., P.S.H. Leeflang & D.R. Wittink (1999), 'The Customer Concept', *Journal of Market Focused Management*, vol. 4, pp. 43-75.
- Huff, L.C. & W.T. Robinson (1994), 'Note: The Impact of Leadtime and Years of Competitive Rivalry on Pioneer Market Share Advantages', *Management Science*, vol. 40, pp. 1370-1377.
- Kalyanam, K. & D.S. Putler (1997), 'Incorporating Demographic Variables in Brand Choice Models: An Indivisible Alternatives Framework', *Marketing Science*, vol. 16, pp. 166-181.
- Kalyanaram, G. & G.L. Urban (1992), 'Dynamic Effects of the Order of Entry on Market Share, Trial Penetration, and Repeat Purchases for Frequently Purchased Consumer Goods', *Marketing Science*, vol. 11, pp. 235-250.
- Kalyanaram, G., W.T. Robinson & G.L. Urban (1995), 'Order of Entry: Established Empirical

- Generalizations, Emerging Empirical Generalizations and Future Research', *Marketing Science*, vol. 14, pp. G212-G221.
- Katz, M.L. & H.S. Rosen (1991), *Microeconomics*, Irwin, Boston.
- Kaul, A. & D.R. Wittink (1995), 'Empirical Generalizations about the Impact of Advertising on Price Sensitivity and Price', *Marketing Science*, vol. 14, pp. G151-G160.
- Kornelis, M., M.G. Dekimpe & P.S.H. Leeftang (2005), 'Are Late Entrants in Television Advertising Markets Breaking News?', Research Paper, University of Groningen.
- Koster, J.M.D. (1991), *Grondslagen van de Marketingwetenschap*, Stenfert Kroese, Leiden.
- Koster, J.M.D. & P.S.H. Leeftang (1992), 'Over Generalisaties en Wetmatige Uitspraken in de Marketingwetenschap', *MAB*, jrg. 66, pp. 239-253.
- Lambin, J.J. (1976), *Advertising, Competition and Market Conduct in Oligopoly over Time*, North-Holland Publishing Company, Amsterdam.
- Lambin, J.J., P.A. Naert & A.V. Bultez (1975), 'Optimal Marketing Behavior in Oligopoly', *European Economic Review*, vol. 6, pp. 105-128.
- Leeftang, P.S.H. (2005), *Meten op Markten*, Koninklijke Nederlandse Academie van Wetenschappen, Amsterdam.
- Leeftang, P.S.H. (1994), *Probleemgebied Marketing II; De Marktinstrumenten*, Stenfert Kroese, Houten.
- Leeftang, P.S.H. (2003), *Marketing*, Stenfert Kroese, Groningen.
- Leeftang, P.S.H. & R.G.M. Pieters (1999), 'Management van de Aandacht voor Merken', *Holland Management Review*, nr. 65, pp. 74-84.
- Leeftang, P.S.H., J.E. Wieringa en D.R. Wittink (2005), "Medicijnen, Marketing en Modellen", *Jaarboek 2005, Ontwikkelingen in het Marktonderzoek*, pp. 149-163.
- Leeftang, P.S.H., J.E. Wieringa, D.R. Wittink & M. Woltil (2004), 'De Invloed op Promotie en Prijs op de Afzet van Geneesmiddelen', *ESB*, pp. 272-274.
- Leeftang, P.S.H. & D.R. Wittink (1992), 'Diagnosing Competitive Reactions Using (Aggregated) Scanner Data', *International Journal of Research in Marketing*, vol. 9, pp. 39-57.
- Leeftang, P.S.H. & D.R. Wittink (1996), 'Competitive Reaction Versus Consumer Response: Do Managers Overreact?' *International Journal of Research in Marketing*, vol. 13, pp. 103-119.
- Leeftang, P.S.H. & D.R. Wittink (2002), 'Promotie van Geneesmiddelen', *ESB*, pp. 228-231.
- Leeftang, P.S.H., D.R. Wittink, M. Wedel & P.A. Naert (2000), *Building Models for Marketing Decisions*, Kluwer Academic Publishers, Boston.
- Leone, R.P. & R.L. Schultz (1980), 'A Study of Marketing Generalizations', *Journal of Marketing*, vol. 44, January, pp. 10-18.
- Little, J.D.C. (1979), 'Aggregate Advertising Models: The State of the Art', *Operations Research*, vol. 27, pp. 629-667.
- Lodish, L.M., M.M. Abraham, S. Kalmenson, J. Livelsberger, B. Lubetkin, B. Richardson & M.E. Stevens (1995a), 'How T.V. Advertising Works: A Meta Analysis of 389 Real World Split Cable T.V. Advertising Experiments', *Journal of Marketing Research*, vol. 32, pp. 125-139.
- Lodish, L.M., M.M. Abraham, J. Livelsberger, B. Lubetkin, B. Richardson & M.E. Stevens (1995b), 'A Summary of Fifty-five In-Market Experimental Estimates of the Long-Term Effect of T.V. Advertising', *Marketing Science*, vol. 14, pp. G133-G140.
- Lynch Jr., J.G. & D. Arieli (2000), 'Wine Online: Search Costs Affect Competition on Price, Quality, and Distribution', *Marketing Science*, vol. 19, pp. 83-93.
- Mahajan, V., E. Muller & F.M. Bass (1995), 'Diffusion of New Products: Empirical Generalizations and Managerial Use', *Marketing Science*, vol. 14, pp. G79-G89.
- Mas-Colell, A., M.D. Whinston & J.R. Green (1995), *Microeconomic Theory*, Oxford University Press, Oxford.
- Morgan, R.M. & S.D. Hunt (1994), 'The Commitment-Trust Theory of Relationship Marketing', *Journal of Marketing*, vol. 58, July, pp. 20-38.
- Narasimhan, C., S.A. Neslin & S.K. Sen (1996), 'Promotional Elasticities and Category Charac-

- teristics', *Journal of Marketing*, vol. 60, pp. 17-20.
- Nijs, V.R., M.G. Dekimpe, J.B.E.M. Steenkamp & D.M. Hanssens (2001), 'The Category-Demand Effects of Price Promotions', *Marketing Science*, vol. 20, pp. 1-22.
- Pagoulatos, E. & R. Sorensen (1986), 'What Determines the Elasticity of Industry Demand?', *International Journal of Industrial Organization*, vol. 4, pp. 237-250.
- Pieters, R.G.M. (1998), 'Een Beetje Invite en Heel Veel Marieke', *Tijdschrift voor Marketing*, maart, pp. 16-19.
- Pieters, R.G.M., L. Warlop & M. Wedel (2002), 'Breaking Through the Clutter: Benefits of Advertising Originality and Familiarity for Brand Attention and Memory', *Management Science*, vol. 48, pp. 765-782.
- Pieters, R. & M. Wedel (2004), 'Attention Capture and Transfer in Advertising: Brand, Pictorial, and Text-Size Effects', *Journal of Marketing*, vol. 68, pp. 36-50.
- Putler, D.S. (1992), 'Incorporation Reference Price Effects Into a Theory of Consumer Choice', *Marketing Science*, vol. 11, pp. 278-309.
- Raju, J.S. (1992), 'The Effect of Price Promotions on Variability in Product Category Sales', *Marketing Science*, vol. 11, pp. 207-220.
- Reibstein, D.J. & P.W. Farris (1995), 'Market Share and Distribution: A Generalization, a Speculation, and Some Implications', *Marketing Science*, vol. 14, pp. G190-G202.
- Robinson, W.T. & C. Fornell (1985), 'Sources of Market Pioneer Advantages in Consumer Goods Industries', *Journal of Marketing Research*, vol. 22, pp. 419-436.
- Rosbergen, E. (1998), 'Assessing Visual Attention to Print Advertising through Statistical Analysis of Eye-Movement Data', Unpublished Ph.D. thesis, University of Groningen.
- Sethuraman, R., V. Srinivasan & D. Kim (1999), 'Asymmetric and Neighborhood Cross-Price Effects: Some Empirical Generalizations', *Marketing Science*, vol. 18, pp. 23-42.
- Shankar, V., G.S. Carpenter & L. Krishnamurthi (1998), 'Late Mover Advantage: How Innovative Late Entrants Outsell Pioneers', *Journal of Marketing Research*, vol. 35, pp. 54-70.
- Simon, H. (1997), 'Hysteresis in Marketing – A New Phenomenon?', *Sloan Management Review*, vol. 38, pp. 39-49.
- Smit, R. (1999), 'Promoties Langs de Meetlat', *Tijdschrift voor Marketing*, april, pp. 34-40.
- Steenkamp, J.B.E.M., V.R. Nijs, D.M. Hanssens & M.G. Dekimpe (2005), 'Competitive Reactions to Advertising and Promotion Attacks', *Marketing Science*, vol. 24, pp. 35-54.
- Sultan, F., J.U. Farley & D.R. Lehmann (1990), 'A Meta-Analysis of Applications of Diffusion Models', *Journal of Marketing Research*, vol. 27, pp. 70-77.
- Tellis, G.J. (1988), 'The Price Elasticity of Selective Demand: A Meta-Analysis of Econometric Models of Sales', *Journal of Marketing Research*, vol. 25, pp. 331-341.
- Tellis, G.J., S. Stremersch & E. Yin (2003), 'The International Takeoff of New Products, The Role of Economics, Culture, and Country Innovativeness', *Marketing Science*, vol. 22, pp. 188-208.
- Uncles, M., A.S.C. Ehrenberg & K. Hammond (1995), 'Patterns of Buyer Behavior: Regularities, Models and Extensions', *Marketing Science*, vol. 14, pp. G71-G78.
- Urban, G.L., T. Carter, S. Gaskin & Z. Mucha (1986), 'Market Share Rewards to Pioneering Brands: An Empirical Analysis and Strategic Implications', *Management Science*, vol. 32, pp. 645-659.
- Vakratsas, D. & T. Ambler (1999), 'How Advertising Works: What do We Really Know', *Journal of Marketing*, vol. 63, January, pp. 26-43.
- Verhoef, P.C. (2003), 'Understanding the Effect of Customer Relationship Management Efforts on Customer Retention and Customer Share Development', *Journal of Marketing*, vol. 67, pp. 30-45.
- Wansbeek, T.W. & M. Wedel (1999), 'Marketing and Economics: Editor's Introduction', *Journal of Econometrics*, vol. 189, pp. 1-14.
- Wedel, W. & R. Pieters (2000), 'Eye Fixations on Advertisements and Memory for Brands: A Model and Findings', *Marketing Science*, vol. 19, pp. 297-312.

- Wedel, M. & J. Zhang (2004), 'Analyzing Brand Competition across Categories', *Journal of Marketing Research*, vol. 40, pp.448-456.
- Wichern, D.W. & R.H. Jones (1977), 'Assessing the Impact of Market Disturbances Using Intervention Analysis', *Management Science*, vol. 24, pp. 329-337.
- Wittink, D.R., M.J. Addona, W.J. Hawkes & J.C. Porter (1988), 'SCAN*PRO: The Estimation, Validation and Use of Promotional Effects Based on Scanner Data', Internal Paper, Cornell University.
- Wold, F.M. (1986), *Meta Analysis: Quantitative Methods for Research Synthesis*, California Sage Press, Newburg Park.
- Woltman Elpers, J.L.C.M. (2003), 'Consumers' Moment-to-Moment Processing of Television Commercials', Unpublished Ph.D. thesis, University of Groningen.
- Woltman Elpers, J.L.C.M., R.G.M. Pieters & M. Wedel (2003), 'Why Do Consumers Stop Viewing Television Commercials? Two Experiments on the Influence of Moment-to-Moment Entertainment and Information Value', *Journal of Marketing Research*, vol. 40, pp. 437-453.

8. *Marketing van producten die passen bij de genetische structuur van het individu*

A. RONTELTAP, S. WENDEL, H.C.M. VAN TRIJP en B.G.C. DELLAERT

SAMENVATTING

Recente ontwikkelingen in de levenswetenschappen (life sciences), en dan met name op het gebied van inzichten in de menselijke genen, zorgen ervoor dat bedrijven mogelijkheden krijgen om producten en diensten te ontwikkelen die afgestemd zijn op individuele fysieke behoeften van de consument. Het leveren van deze producten en diensten stelt de marketing in deze bedrijven voor een aantal nieuwe uitdagingen.

In dit artikel zetten de auteurs deze uitdagingen uiteen en worden hypothesen voor consumentenacceptatie van diensten op basis van dergelijke inzichten uit de levenswetenschappen geformuleerd en getoetst. De resultaten laten onder andere zien dat het van belang is deze diensten als integrale systemen te analyseren, en ze niet op afzonderlijke kenmerken alleen te beoordelen.

1. INLEIDING

De levenswetenschappen (of “life sciences”) maken op dit moment een sterke ontwikkeling door, met de potentie om belangrijke sectoren van de economie zoals de gezondheidszorg, farmaceutische industrie en de voedselproductie te hervormen (Enriquez en Goldberg 2000). Met name de recente ontrafeling van het menselijke genoom (IHGSC 2001; Venter et al. 2001) heeft een nieuwe impuls gegeven aan de kennis over hoe verschillende individuen reageren op verschillende varianten in voeding en medicijnen. De verwachting is dat deze ontwikkelingen de waarde creatie in de farmaceutische en voedingsindustrie drastisch zal veranderen (bijv. Muller en Kersten 2003; Ordovas en Corella 2004). Bedrijven krijgen nu de mogelijkheid om relevante verschillen in genetische constitutie van consumenten in kaart te brengen en producten en diensten te leveren die afgestemd zijn op de verschillende genetische behoeften van individuen of segmenten met vergelijkbare genetische constitutie (bijv. German, Yerezian en Watzke 2004). In medische en farmaceutische toepassingen wordt deze werkwijze al toegepast, en ook binnen de voedingsindustrie ontstaat er meer aandacht voor. Het Instituut voor de Toekomst (Cain en Smid 2003) voorspelt dat in 2010 een derde van de Amerikaanse bevolking gebruik zal maken van op het individu afgestemde voeding en de eerste op het menselijk genoom gebaseerde commerciële adviesbureaus hebben de markt al betreden. Een voorbeeld is het in de V.S. gevestigde bedrijf Sciona (www.sciona.com), dat consumenten gebaseerd op hun individuele genetische informatie, persoonlijk advies geeft over dieet en levensstijl.

Het leveren van producten die op maat gemaakt zijn naar de genetische samenstelling

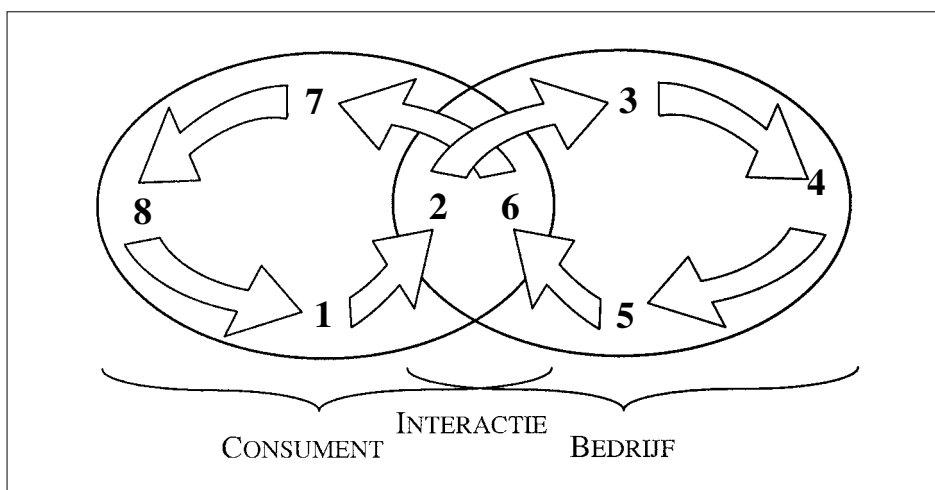
van het individu stelt de marketing functie in bedrijven voor een aantal nieuwe uitdagingen. In dit artikel analyseren we deze uitdagingen en ontwikkelen en toetsen we hypothesen over de consumentenacceptatie van producten en diensten die zijn gebaseerd op individuele genetische informatie en die tot stand zijn gekomen op basis van recente inzichten uit de levenswetenschappen (in het vervolg van het artikel duiden we dergelijk systemen aan met de term “life science systemen”). We beargumenteren dat life science systemen noodzakelijkerwijs uit meerdere fasen bestaan, en een structuur vereisen met een aantal consument-bedrijf interacties. We ontwikkelen een conceptueel model dat bestaat uit acht verschillende fasen en hypothetiseren dat consumenten het totaal van deze acht fasen in beschouwing nemen bij het evalueren van life science systemen. We hypothetiseren verder dat de effecten van elk van deze fasen op consumentenvoorkeuren gemedieerd wordt door het consumentenoordeel over een kleiner aantal onderliggende meer abstracte voordelen (“benefits”) welke de consument aan het systeem ontleent. We veronderstellen dat de set van traditionele beoordelingsdimensies van technologische innovaties: gemak, nut en plezier, in het geval van life science systemen eveneens van toepassing zijn, maar uitgebreid moeten worden met de dimensies controle en privacy (Parthasarathy, 2004; Rabino, 2003). Tenslotte onderzoeken we of de beoordeling van life science systemen afhangt van de vraag of het systeem gebruikt wordt als onderdeel van een zelf-zorg proces door het individu, of dat het systeem wordt aanbevolen binnen de medische sector, bijvoorbeeld door een advies van de huisarts (bijv. Stearns et al. 2000). Eerder onderzoek heeft aangetoond dat medicijnkeuzes door consumenten verschillend zijn wanneer deze producten voorgeschreven zijn door een arts dan wanneer de producten vrijelijk in de winkel beschikbaar zijn (bijv. Hoy, 1994; Trussel et al. 1993). We verwachten dat consumentenvoorkeuren voor life science systemen ook op een dergelijke wijze kunnen verschillen.

2. MARKETING VAN PRODUCTEN DIE PASSEN BIJ DE GENETISCHE STRUCTUUR VAN HET INDIVIDU

Om producten die passen bij de genetische structuur van het individu te kunnen produceren (e.g. German, Yeretzián en Watzke, 2004) is het nodig dat bedrijven de genetische constitutie van consumenten kunnen vaststellen. Hoewel de moderne ontwikkelingen in productie en communicatie technologie bedrijven steeds meer in staat stellen om hun producten en diensten te verpersoonlijken (bijv. Steckel et al. 2005), is het verwerven van genetische informatie aanzienlijk meer ingrijpend dan het direct of indirect vaststellen van consumentenvoorkeuren (Rabino, 2003; Parthasarathy, 2004). Voor consumenten is het beschikbaar stellen van gedetailleerde fysieke informatie (bijv. een bloedmonster) aan een bedrijf een ingrijpende stap in het interactieproces. Bovendien is deze interactie niet éénmalig, omdat een geregelde interactie tussen consument en bedrijf noodzakelijk is om de effectiviteit van de op life science gebaseerde producten te monitoren en zondig de productkenmerken aan te passen. In de nu volgende paragrafen beschrijven we de consumentenbeoordeling van deze interactiemechanismen in meer detail.

3. MEER-FASE INTERACTIEMECHANISMEN

In eerder onderzoek zijn verschillende fasen in het uitwisselingsproces tussen consumenten en bedrijven onderscheiden. Een tweetal analyses is m.n. relevant voor onze huidige studie. Murthi en Sakar (2003) suggereren drie fasen: (1) de leerfase (het bedrijf verzamelt data van de consumenten), (2) de afstemmingsfase (gebruikmakend van de informatie ontwikkelt het bedrijf een persoonlijk aanbod), en (3) de evaluatiefase (de effectiviteit van de afstemming wordt beoordeeld door de consument). In een andere studie stelden Ives en Mason (1990) een soortgelijke structuur voor toen zij hun “customer service life cycle”-concept introduceerden. Ze onderscheidde de volgende fasen: (1) vereisten (formuleren van productvereisten), (2) acquisitie (aankoop van het product), (3) bezit (gebruik van het product) en (4) uittrekking (reflectie op het product en het proces).



Figuur 1. Consument-bedrijf interacties bij marketing van producten gebaseerd op individuele genetische samenstelling.

Voortbouwend op deze modellen stellen wij een uitgebreider acht-fase proces voor voor life science systemen (Figuur 1). In de eerste fase (*informatie*) genereert de consument spezifische und möglich gefühlige information (bijv. ziektehistorie oder genetische constitution). In der folgenden fase (*communicatie*) wird diese information via eine physische (bijv. gesundheitszentrum oder ziekenhuis) oder digitale (internet) interface beschikbaar gesteld aan de aanbodketen. In fase 3 (*informatie behandeling*) wordt de information ontvangen und verwerkt in de keten. Deze information zal vervolgens ingevoerd worden in een beslissingstool (fase 4: *ontwerp*) van een bedrijf. Gebaseerd op de information und der beslissingstool zal de aanbodketen vervolgens een persoonlijk aanbod opstellen gericht op de specifieke behoeften van der consument (fase 5: *productie*). In fase 6 (*levering*) wordt dit persoonlijk product (oder der persoonlijke dienst) geleverd aan der consument via eine bepaalde gebruikersinterface (bijv. het verzenden van producten naar huis, advies via email). Der consument ontvangt het product und implementeert dit in het dagelijks leven (fase 7: *gebruik*). Op basis van der implementatie zal der consument vervolgens der toegevoegde waarde verifiëren (fase 8: *evaluatie*).

Bovengenoemde fasen kunnen geclassificeerd worden in drie domeinen: 1) het consumentendomein, 2) de informatie-uitwisseling tussen consument en bedrijf (interactie), en 3) het bedrijfsdomein (Zie Figuur 1). Deze drie domeinen in het uitwisselingsproces tussen consumenten en producenten maken gezamenlijk de creatie van verpersoonlijkte aanbevelingen in de levenswetenschappen mogelijk en we verwachten dat de verschillende fasen in deze domeinen de consumentenbeoordeling van de systemen voor levenswetenschappen zullen aansturen.

H1: Consumentenvoorkeuren voor life science systemen worden bepaald door een gezamenlijke evaluatie van de *vereiste acties door de consument, de kenmerken van de informatie-uitwisseling en de acties van het bedrijf in het systeem.*

4. DE MEDIERENDE ROL VAN BENEFITS

We verwachten dat verschillende life science systemen verschillende benefits kunnen opleveren voor consumenten (d.w.z. verschillende onderliggende voordelen die een consument verwacht van het gebruik, zie Gutman 1982, Myers 1976). Eerdere onderzoeken laten zien dat consumenten concrete eigenschappen van een product abstraheren naar een beperkter aantal gepercipieerde benefits van een product (bijv. Gutman 1982, Ter Hofstede, Audenaert, Steenkamp en Wedel, 1998), die vervolgens de consument leiden tot een beslissing (bijv. de keuze van een product). In onderzoek naar de adoptie van nieuwe technologie is een dergelijke structuur toegepast om consumentenevaluaties van complexe nieuwe technologieën te voorspellen (bijv. Davis, Bagozzi en Warshaw 1989, Rogers, 2003). We verwachten een vergelijkbare structuur voor de consumenten evaluatie van kenmerken van een life science systeem.

H2: De effecten van concrete kenmerken van de verschillende fasen van een life science systeem op de voorkeuren van de consument voor het systeem, worden gemedieerd door de gepercipieerde benefits van het systeem.

5. BENEFITS VAN EEN SYSTEEM VOOR LEVENSWETENSCHAPPEN

Er is weinig bekend over hoe consumenten complexe innovaties op het gebied van levenswetenschappen beoordelen. Om ons begrip op dit terrein verder te ontwikkelen putten we uit diverse onderzoeksstromen die gepercipieerde benefits identificeren van nieuwe technologie en waarvan bekend is dat ze de adoptie van deze technologie door consumenten bepalen (bijv. Davis, Bagozzi en Warshaw, 189, Dabholkar en Bagozzi 2002; Rogers 2003).

Rogers (2003) beschrijft waarom de door de consument waargenomen eigenschappen van een innovatie cruciaal zijn voor de adoptiebeslissing van de consument. Verschillende onderzoekers hebben voortgebouwd op dit model. Een prominent model, afkomstig uit het veld van informatie technologie, is het Technologie Acceptatie Model (TAM) (bijv. Davis, 1989), dat de houding van de consument en de intentie tot gedrag met betrekking tot een bepaald systeem verbindt. Het TAM is afgeleid uit de theorie van beredeneerd gedrag van Fishbein en Ajzen (TRA), dat het voorgenomen

gedrag van consumenten weergeeft door onderscheid te maken tussen opvattingen, houdingen, intenties en gedrag van consumenten (Ajzen en Fishbein, 1980).

Deze modellen lijken bruikbaar voor ons toepassingsgebied omdat consument-bedrijf interacties in onze context tevens sterk afhangen van technologie, zowel wat betreft het type product dat wordt ontwikkeld (bijv. voedingssupplementen) als het gebruik van informatietechnologie om de interactie te ondersteunen. Eerder onderzoek leverde drie belangrijke kosten-baten overwegingen op: gemak, nut en plezier. Voor de specifieke toepassingen op life science systemen, stellen we voor deze drie kernoverwegingen uit te breiden met de zorgen die consumenten kunnen hebben over privacy en controle.

5.1. Gemak en nut

De invloed van “gemak” en “nut” als belangrijke consumentenpercepties wordt ondersteund door het technologie acceptatie model (bijv. Davis, 1989; Davis, Bagozzi en Warshaw, 1989). Deze twee constructen zijn geoperationaliseerd en uitgewerkt, veelal in de context van het internet. Bijvoorbeeld, Chau en Hu (2002) identificeerden “gemak” en “nut” als twee cruciale antecedenten van de acceptatie door artsen van telemedicine software en verscheidene andere onderzoekers onderzochten online koopgedrag van consumenten (O’Cass en Fenech, 2003), klantacceptatie van internetbankieren (Suh en Han 2002), acceptatie door consumenten van internet (Moon en Kim 2002) en motivatie van consumenten om deel te nemen in online winkelen (Childers et al. 2001). In onze studie definiëren we “waargenomen gemak” als “de mate waarin een persoon gelooft dat gebruik van een life science systeem zonder moeite zal zijn” en “waargenomen nut” als “de mate waarin een persoon gelooft dat gebruik van een life science systeem gunstig zal zijn voor zijn of haar gezondheid”.

H3a: Groter waargenomen gemak verhoogt de voorkeur van een consument voor een life science systeem.

H3b: Grotere waargenomen nut verhoogt de voorkeur van een consument voor een life science systeem

5.2. Plezier

“Plezier” is een bijkomende determinant die de houding van consumenten ten opzichte van het gebruik van een technologie beïnvloedt (bijv. Dabholkar en Bagozzi 2002; van der Heijden, 2003). Onderzoek van Dabholkar en Bagozzi (2002) laat zien dat “plezier” een belangrijke factor is in het beïnvloeden van houdingen van consumenten ten opzichte van het gebruik van op technologie gebaseerde selfservice. Dit resultaat werd bevestigd door Van der Heijden (2003) in de context van websites. In onze studie definiëren we “plezier” als “de mate waarin het gebruik van een systeem van levenswetenschappen waargenomen wordt als een aangename ervaring”.

H4: Groter waargenomen plezier verhoogt de voorkeur van consumenten voor een life science systeem .

5.3. Privacy en Controle

Percepties van privacy en controle zijn twee doorslaggevende aspecten in de context van de levenswetenschappen (Parthasarathy 2004; Rabin 2003). In de Verenigde Staten

zijn consumenten voornamelijk bezorgd over hun privacy als het gaat om gezondheidsgerelateerde diensten en zijn ze in het algemeen onwillig om persoonlijke informatie te verschaffen, voornamelijk door een op technologie gebaseerd medium zoals het internet (Economist, 2005). Er zijn verschillende redenen waarom consumenten terughoudend kunnen zijn in het beschikbaar stellen van genetische informatie: Genetische informatie is erg gedetailleerd, de informatie kan tot in het oneindige bewaard worden en de informatie laat niet alleen het testen van de huidige toestand van het individu toe maar ook het specificeren van de *toekomstige* risico's van het individu. Bovendien kunnen meerdere verschillende andere partijen belang hebben bij de verwerven van de genetische informatie van een persoon, bijvoorbeeld huidige of toekomstige familie, organisaties waarmee individuen economische overeenkomsten hebben (werknemers, verzekeringsmaatschappijen, banken) of overheidsorganisaties (bijv. gezondheidsinstanties) (Takala en Gylling 2000). Dit lijstje illustreert de uiteenlopende economische, sociale en psychologische implicaties van de onthulling van genetische informatie. Beveiliging speelt ook een belangrijke rol bij de beleving privacy. Vijayasathy (2004) stelt dat beveiliging een direct effect heeft op de houding van consumenten ten opzichte van online winkelen. Dit stemt overeen met O'Cass en Fenech (2003) die voorstellen dat "waargenomen veiligheid" een web-ervaring is die indirect invloed heeft op de houding van consumenten ten opzichte van winkelen op het internet. We definiëren daarom privacy als "de mate waarin een life science systeem waargenomen wordt als relatief veilig voor het verschaffen van gevoelige persoonlijke informatie".

H5a: Grotere waargenomen privacy verhoogt de voorkeur van een consument voor een life science systeem

Zorgen van de consument over privacy zijn direct verbonden aan gevoelens van controle over zowel het proces als de uitkomst bij het gebruik van een specifieke technologie. "Waargenomen eigen controle" is een antecedent van intentie tot gedrag in het model van de theorie van gepland gedrag (TPB) (bijv. Ajzen, 1985; Ajzen en Madden, 1986). Dabholkar (1996) vond dat "controle" belangrijk is in de evaluatie van nieuwe op technologie gebaseerde zelfbedieningsopties. Ook Chau en Hu (2000) identificeren "waargenomen controle" als een relevante bepalende factor van gedragsintenties ten opzichte van telemedicine technologie. We definiëren "eigen controle" dus als "de mate waarin een systeem voor levenswetenschappen ervaren wordt als eenvoudig te controleren en gebruiken".

H5b: Grotere waargenomen eigen controle verhoogt de voorkeur van een consument voor een life science systeem.

5.4. *Het effect van zelf-zorg vs. medisch advies*

Verschillende onderzoekers hebben het effect belicht van *intrinsieke* versus *extrinsieke* motivatie op aankoopbeslissingen (Bloch en Richins 1983) en op technologieadoptie (Davis, Bagozzi en Warshaw, 1992). Eén van de grootste verschillen tussen deze twee motieven is een lagere doeloriëntatie en een sterkere nadruk op ervaringen bij intrinsiek gemotiveerd gedrag (Hoffman en Novak, 1996, Mathwick, Malhotra, and Rigdon, 2002). We verwachten dat dit onderscheid ook toepasbaar is in de context van ons onderzoek. Onze hypothese is dat consumenten die van plan zijn een systeem te gebrui-

ken als onderdeel van hun zelf-zorg activiteiten het plezier dat ze aan het gebruik ontle-
nen belangrijker vinden dan consumenten die het systeem te gebruiken op advies van
hun huisarts. In het laatste geval verwachten we dat het nut van het systeem belangrij-
ker is.

H6: Consumenten die een life science systeem gebruiken als onderdeel van hun zelf-
zorg activiteiten vinden (a) het plezier in het gebruik *belangrijker*, en (b) het nut
van het systeem *minder belangrijk*, dan consumenten die het systeem op een
medisch advies gebruiken.

6. METHODEN EN DATA

Nutrigenomics is één van de gebieden waar op genetische samenstelling gebaseerde
producten en aanbevelingen nu reeds in de markt beschikbaar beginnen te komen.
Nutrigenomics kan worden gedefinieerd als de “wetenschappelijke studie van de wijze
waarop specifieke genen reageren op nutriënten” – met andere woorden, het effect van
wat we eten op de manier waarop onze genen functioneren (Cain en Schmid, 2003).
Het Amerikaanse bedrijf Sciona (www.sciona.com) bijvoorbeeld heeft diagnostische
producten voor consumenten ontwikkeld en geïntroduceerd die op genetica zijn geba-
seerd. Een voorbeeld is het ‘Nutrition Screen’, een hulpmiddel dat online beschikbaar
is (www.mycellf.com) voor ongeveer \$200. Dit systeem vereist dat de consument thuis
enkele cellen uit de mond (wangslim) afneemt. Samen met een ingevulde vragenlijst
over eetgewoonten en levensstijl worden deze cellen opgestuurd naar het bedrijf. Twee
tot drie weken later ontvangt de consument een persoonlijk rapport van 40 pagina’s, dat
praktische aanbevelingen geeft voor verbetering zoals een advies voor voedingssupple-
menten, een persoonlijk dieet en een levensstijladvies, op basis van de genetische
samenstelling en de huidige levensstijl van de consument.

We testen onze hypotheses in de context van op het individu afgestemd voedseladvies
op basis van genetische informatie. De studie bestond uit drie fasen, waarvan de eerste
twee voorstudies betroffen en de derde fase de feitelijk hoofdstudie was. De eerste fase
richtte zich op het ontwikkelen van een kwalitatief begrip van de mogelijkheden van
nutrigenomics zoals gezien door de ogen van experts en consumenten. In deze fase zijn
gesprekken gevoerd met verschillende experts uit wetenschap en bedrijfsleven om zo
relevante fasen van het interactieproces te identificeren. Ook zijn consumenten geïnter-
viewd met als doel te beoordelen of de geïdentificeerde fasen begrepen werden door
consumenten. Gebaseerd op deze interviews zijn de fasen en labels geformuleerd zoals
gepresenteerd in Tabel 1.

De kwalitatieve interviews resulteerden tevens in de volgende omschrijving van de
zelf-zorg context versus medisch advies context:

1. Iemand die je kent heeft je verteld dat het mogelijk is persoonlijk voedingsadvies te
verkrijgen en je wilt dit graag proberen (zelf-zorg).
2. Je huisarts adviseerde je bij je gebruikelijke controle dat je je beter zou voelen als je
een persoonlijk voedingsadvies zou gebruiken (medisch advies).

De tweede fase richtte zich op betrouwbaarheid van de schalen en verfijning van de
schalen uit de literatuur. Hiervoor is gebruik gemaakt van een steekproef van 108
‘Bachelor’ en ‘Master’ studenten die elk enkele van de life science systeem kenmerken
en benefits beoordeelden in een korte vragenlijst. Elke respondent ontving een kleine

Tabel 1. Geselecteerde labels voor de evaluatietaak van de proefpersonen over life science systemen.

<i>Fase</i>	<i>Label</i>	<i>Fase</i>	<i>Label</i>
1	<i>Type persoonlijke informatie dat beschikbaar gesteld wordt</i> <ul style="list-style-type: none"> • Bloedsamenstelling • DNA/ genetische samenstelling • Voedselconsumptie gewoontes 	5	<i>Specificiteit van persoonlijk advies</i> <ul style="list-style-type: none"> • Niveau van ingrediënten • Niveau van productgroep • Niveau van speciale merkproducten
2	<i>Manier waarop informatie beschikbaar gesteld wordt</i> <ul style="list-style-type: none"> • Via een fitnessclub • Via het ziekenhuis • Via de huisarts 	6	<i>Manier waarop persoonlijk advies door consument wordt ontvangen</i> <ul style="list-style-type: none"> • Via de e-mail • Via een fitnessclub • Via de huisarts
3	<i>Manier waarop met informatie wordt omgegaan</i> <ul style="list-style-type: none"> • Volledig anoniem • Gedeeld tussen patiënt en huisarts • Beschikbaar voor commerciële voedingsmiddelenbedrijven 	7	<i>Gebruik</i> <ul style="list-style-type: none"> • Opgenomen in bestaande maaltijdpatronen • Specifieke producten toegevoegd aan gewone maaltijden • Klaarmaken van individuele, aangepaste maaltijden
4	<i>Opsteller van persoonlijk advies</i> <ul style="list-style-type: none"> • Commercieel voedingsmiddelenbedrijf • Verzekeringsmaatschappij • Voedingencentrum (overheid) 	8	<i>Evaluatiemogelijkheden voor effectiviteit van systeem</i> <ul style="list-style-type: none"> • Geen evaluatie mogelijk • Evaluatie mogelijk • Evaluatie verplicht

beloning voor het uitvoeren van de taak. Selectie van de factoren van life science systemen en de evaluatieconcepten van consumenten was willekeurig. Data werden verzameld door middel van een pen-en-papier taak en alle items zijn gemeten op 9-punts semantisch differentiaal schalen. De resultaten van de analyses bevestigden dat de meeste schaalitems goede tot zeer goede metingen opleverden. Gebaseerd op lage item-tot-totaal correlaties werden twee items uitgesloten van de originele schaal van nut. Onze resultaten lieten verder zien dat de toegevoegde evaluatieconcepten over privacy en gedragscontrole voldoende onderscheidend vermogen hadden om de voorgestelde modelstructuur te meten.

6.1. Steekproef en procedure van dataverzamelen voor de hoofdstudie

Data voor de hoofdstudie werden computergestuurd verzameld door een professioneel marktonderzoeksbureau onder 204 respondenten in een panel. Respondenten, woonachtig in Nederland en variërend in leeftijd van 18 tot 64, werden uitgenodigd in de centrale testfaciliteit. Elke respondent beoordeelde drie willekeurig gekozen scenario's uit een volledig 3⁸ factorieel experimenteel design. De scenario's waren volledige profielbeschrijvingen op basis van de acht fasen uit Figuur 1, elk met drie niveaus. De context variabele zelf-zorg vs. medisch advies werd gemanipuleerd door de taakinstructie. De helft van de proefpersonen werd gevraagd zich een zelf-zorg situatie voor de stellen, de andere helft werd geïnstrueerd zich de medische situatie voor te stellen. Proefpersonen werden willekeurig toegewezen aan één van deze situatievoorwaarden.

6.2. Benefits van life science systemen

Items om de benefits te meten waren gebaseerd op bestaande schalen en alle items zijn gescoord op 9-puntsschalen. *Gemak* werd gemeten met behulp van drie items ($\alpha = 0.90$). Dit systeem voor persoonlijk voedseladvies is (1) moeilijk te begrijpen – eenvoudig te begrijpen, (2) moeilijk om te leren gebruiken – makkelijk om te leren gebruiken, (3) moeilijk om te onthouden wat ik moet doen – makkelijk om te onthouden wat ik moet doen. *Nut* werd gemeten door drie items ($\alpha = 0.89$). Dit systeem voor persoonlijk voedseladvies is (1) niet bruikbaar – bruikbaar, (2) niet bruikbaar om mijn voeding te verbeteren – bruikbaar om mijn voeding te verbeteren, (3) niet van invloed op mijn aankopen – van invloed op mijn aankopen. *Plezier* werd gemeten door drie items ($\alpha = 0.95$). Het gebruik van dit systeem is (1) vervelend – interessant, (2) afstompend – onderhoudend, (3) niet prettig – prettig. *Privacy* beschouwingen werden gemeten door drie items ($\alpha = 0.94$). Tijdens het gebruik van dit systeem voel ik (1) mij onzeker over het geven van persoonlijke informatie – mij zeker over het geven van persoonlijke informatie, (2) mij onzeker over het geven van informatie over mijn gezondheid – mij zeker over het geven van informatie over mijn gezondheid, (3) dat het systeem niet veilig is – dat het systeem veilig is. *Controle* werd gemeten door drie items ($\alpha = 0.90$). Bij het gebruik van dit systeem zou ik (1) niet in staat zijn om het systeem te gebruiken om informatie te zoeken en voeding te kopen – wel in staat zijn het systeem te gebruiken om informatie te zoeken en voeding te kopen, (2) geen controle over het gebruik hebben – geheel controle over het gebruik hebben, (3) geen kennis hebben om het systeem te gebruiken – volledige kennis hebben om het systeem te gebruiken. *Voorkeur* werd gemeten door drie items ($\alpha = 0.98$). Twee semantisch differentiaal schalen. Deze bestonden uit de volgende schalen: Hoe waarschijnlijk is het dat u van dit bepaalde systeem van persoonlijk voedingsadvies gebruik zult gaan maken (1) onwaarschijnlijk – waarschijnlijk, (2) onmogelijk – mogelijk. Ook is een 9-punt Likert schaal gebruikt: Ik zou dit systeem voor voedingsadvies zeker gebruiken ((1) oneens, (2) eens).

7. RESULTATEN

7.1. Schaalvaliditeit

Eerst is beoordeeld of de items die gebruikt werden om benefits en voorkeuren te meten inderdaad verschillende benefits onderscheidde. Het één-factor-model paste slecht bij de data en werd duidelijk verworpen. Vervolgens schatten we het veronderstelde zes-factoren-model. Dit model paste goed bij de data: $\chi^2(120) = 382.79$ ($p < .001$), CFI = .97, NNFI = TLI = .96 and RMSEA = .061. Bovendien paste het zes-factoren-model significant beter bij de data dan het één-factor-model. Alle factorladingen waren significant en alle compleet gestandaardiseerde ladingen waren groter dan .56 en de gemiddelde compleet gestandaardiseerde lading was .85. Deze bevindingen ondersteunen de convergente validiteit van de items. Cronbach's alfa's waren .90, .89, .90, .94, .95 en .98 voor respectievelijk gemak, nut, controle, privacy, plezier en voorkeur. De constructen zijn vervolgens geoperationaliseerd als somscores van de items.

7.2. Toetsing van hypotheses

Tabel 2 toont de regressie analyse resultaten van drie fixed-effect modellen met de intentie tot gebruik van een life science systeem als afhankelijke variabele¹. Het eerste model gebruikt alleen benefits als verklarende variabelen, het tweede model alleen systeemkenmerken, en het derde model beide categorieën.

De eerste kolom van Tabel 2 laat zien dat percepties van consumenten over benefits goede voorspellers zijn van consumentenvoorkeuren voor een life science systeem ($R^2_{\text{adj}} = .721$). Groter gemak verhoogt ($\beta = -.014$, n.s.) de consumentenvoorkeur voor een life science systeem niet (Hypothese 3a verworpen), maar groter nut wel ($\beta = .196$, $p < .001$), hetgeen Hypothese 3b ondersteunt. Tevens is groter plezier de sterkste voorspeller van consumentenvoorkeur ($\beta = .542$, $p < .001$), wat Hypothese 4 ondersteunt. Daarnaast nemen consumentenvoorkeuren toe bij meer privacy ($\beta = .180$, $p < .001$), wat steun verleent aan Hypothese 5a, maar grotere gedragscontrole doet dat niet ($\beta = .027$, n.s.) (hypothese 5b wordt verworpen).

Uit de tweede kolom blijkt dat de kenmerken van het life science systeem ook een significant deel van de waardering van consumenten verklaren ($R^2_{\text{adj}} = .089$, $p < .001$). Daarbij dragen verschillende fasen op het niveau van vereiste consumentenacties (evaluatie), van de eigenschappen van interactieve uitwisselingsinterfaces (communicatie en levering) en van de acties van het bedrijf (omgaan met data, ontwikkeling en productie), significant ($p < .05$) bij aan de verklaring van consumentenvoorkeuren voor life science systemen. Consumenten geven de voorkeur aan de volgende kenmerken: communicatie via de huisarts, een systeem waarbij gegevens niet beschikbaar zijn voor commerciële bedrijven, een ketencoördinator die niet als verzekeringsbedrijf opereert, het aanleveren van gegevens via andere kanalen dan de fitnessclub, adviezen die niet zijn gespecificeerd in termen van merkproducten, en evaluatie met goede feedback. Zes van de acht fasen van het model dragen significant bij aan de verklaring van consumentenvoorkeuren, waarbij een zevende, informatie delen (ongunstig als DNA/genetische informatie beschikbaar wordt gesteld) significant is op het niveau van 10%. Samengevat leveren deze resultaten een sterke ondersteuning voor Hypothese 1.

Hypothese 2a stelt dat het effect van een life science systeem op consumentenvoorkeuren gemedieerd wordt door de benefits van het systeem. Volledige mediatie is aangetoond (bijv. Baron en Kenny, 1986) als de onafhankelijke variabelen (kenmerken van het systeem) significante effecten op de mediator (benefits van het systeem) uitoefenen, evenals op de afhankelijke variabele (voorkeur), maar het effect van de kenmerken van het systeem op de consumentenvoorkeuren verdwijnt als de mediërende variabelen opgenomen worden als co-variant. De resultaten in Tabel 2 ondersteunen hypothese 2a. De effecten van de kenmerken van het systeem zijn niet langer significant wanneer de benefits worden opgenomen in het model en er is volledige mediatie.

Hypothese 6 stelt dat de invloed van de benefits *plezier* en *nut* op consumentenvoorkeuren afhankelijk is van of het gebruik van het life science systeem gemotiveerd is vanuit zelfzorg of vanuit medische advies. De helft van de respondenten was geïnstrueerd dat gebruik van het systeem een aanbeveling van hun huisarts was en de andere helft dat het zelfmotivatie was. Om deze hypothese te toetsen is een additionele regressieanalyse uitgevoerd met waargenomen benefits en de interactie tussen waargenomen benefits en de situatie als onafhankelijke variabelen. De resultaten van deze analyse ondersteunen hypothese 6 en tonen dat het *effect van nut* groter is in de medische toepassing van het gebruik van een life science systeem ($p < .001$), en het *effect van ple-*

TABEL 2. Regressie resultaten.

<i>Benefits</i>	<i>Intentie tot gebruik</i>		
	<i>alleen benefits</i>	<i>alleen systeem- kenmerken</i>	<i>benefits en systeem- kenmerken</i>
Gemak	-0.014		-0.014
Nut	0.196 **		0.193 **
Plezier	0.542 **		0.534 **
Eigen controle	0.027		0.028
Privacy	0.180 **		0.179 **
<i>Kenmerken</i>			
Beschikbaar stellen: Basis= Fitness club			
Huisarts		0.159 **	-0.040
Ziekenhuis		0.093	-0.031
Type informatie: Basis=Bloedsamenstelling			
DNA/ genetische materiaal		-0.094	-0.043
Voedsel consumptie		0.048	-0.033
Omgaan met informatie: Basis= Anoniem			
Patiënt en huisarts		0.008	-0.024
Commercieel bedrijf		-0.147 *	-0.058
Opsteller advies: Basis= Voedingsbedrijf			
Verzekeraar		-0.196 **	-0.036
Voedingscentrum		-0.025	-0.009
Advies ontvangen: Basis = Email			
Fitness club		-0.176 **	-0.016
Huisarts		-0.071	-0.006
Specificiteit: Basis = Ingrediënten			
Productgroepen		0.077	-0.027
Productmerken		-0.125 *	-0.060
Gebruik: Basis = Bestaande maaltijd			
Toevoegen aan maaltijd		0.015	0.036
Individuele maaltijden		-0.002	0.025
Evaluatie: Basis= Geen evaluatie			
Evaluatie mogelijk		0.146 **	-0.004
Verplichte evaluatie		0.136 *	0.042
R ²	0.814	0.412	0.822
R ² adjusted	0.721	0.089	0.721

** significant p<.05

* significant p<.10

zier groter is in de toepassing van zelfzorggebruik van het life science systeem (p<0.05).²

8. CONCLUSIE EN DISCUSSIE

Maatwerk producten voor consumenten op basis van genetische samenstelling brengen een aantal uitdagingen voor marketing met zich mee. In dit artikel hebben we een theoretische structuur ontwikkeld en getoetst om deze uitdagingen te analyseren. In onze

empirische analyse vinden ondersteuning voor de vier kernconcepten van ons model, in het bijzonder dat bij de marketing van producten gebaseerd op genetische informatie: (1) een meer-fase interactiemechanisme vereist is, (2) het effect van de kenmerken van het systeem gemedieerd wordt door een beperkt aantal onderliggende benefits, (3) deze benefits naast gemak, nut en plezier ook privacy omvatten (maar niet controle), en (4) het relatieve belang van benefits en daardoor de acceptatie van het systeem afhangt van de motivering waardoor consumenten gebruik maken van een life science systeem.

Gezien het meer-fase karakter van het proces vinden we ondersteuning voor de rol van alle drie domeinen (de consument, de interactie en het bedrijf) aan de gepercipieerde benefits van, en de voorkeuren voor life science systemen. Zoals verwacht, is er nog steeds een aanmerkelijke terughoudendheid bij consumenten voor commerciële toepassingen van dit soort persoonlijke adviezen. Het zijn in het bijzonder de niveaus van “commerciële” kenmerken van de dienstverlening (bijv. gedeelde informatie met commerciële voedselbedrijven, kennissysteem in eigendom van een verzekeringsmaatschappij, advies geleverd als merkproducten en advies geleverd via de fitnessclub) die de aantrekkelijkheid voor consumenten van deze systemen het meeste schaden.

De analyse laat ook zien dat mediatie voor alle niveaus van kenmerken van het systeem optreedt. In overeenstemming met voorgaande studies vinden we dat consumentenvoorkeuren voor producten gebaseerd op genetische informatie grotendeels worden bepaald door de benefits nut, gemak en plezier. Gemak en nut kunnen niet helemaal gescheiden worden in de percepties van consumenten, hetgeen in overeenstemming is met voorgaande studies, waar de waardering voor het gemak, die van het nut bleek te beïnvloeden (Taylor en Todd, 1995; Kwon en Chidambaram, 2000). Plezier speelt ook een rol in consumentenvoorkeuren voor dit type dienstverlening (Dabholkar en Bagozzi, 2002). Echter, naast deze drie “gevestigde” benefits van consumenten zien we dat waargenomen privacy ook een belangrijke rol speelt in de consumentenvoorkeur. Deze perceptie lijkt zelfs een grotere rol te spelen in deze context dan in verschillende andere contexten (zoals internetbetalingen) (Hoffman, 1999). Dit stelt een interessante uitdaging aan dienstverlening op basis van genetica die per definitie impliceert dat echte één-op-één marketing plaatsvindt waar consumenten individueel worden geïdentificeerd en aangesproken. Omdat deze vertrouwelijke identificatie de sleutel tot het optimaliseren van het product is moet er wellicht een afweging gemaakt worden tussen nut en privacy: Minder privacy (bijv. persoonlijke identificatie) impliceert een kans op een groter nut (op maat gemaakte producten).

Concluderend kunnen we stellen dat genetische segmentatie in meerdere opzichten verschilt van de meer gebruikelijke segmentaties op basis van levensstijl. Levensstijlen zijn gebaseerd op fenotypen in de zin dat ze gemakkelijker van buitenaf waargenomen kunnen worden (of door specifieke vragen). Genotype is niet zichtbaar van buitenaf en vereist per definitie een meer intieme interactie. Zo is zelfclassificatie op basis van levensstijl (welke levensstijl past het best bij me) mogelijk, maar niet in het geval van genotype (welke andere genetische samenstellingen komen het dichtst bij die van mij). Ook noemenswaardig is de rol van de huisarts in de dienstverlening gebaseerd op life sciences. Op één lijn met voorgaande studies (bijv. Henneman et al., 2004) ondervinden we dat de huisarts een kernrol kan spelen in het bespreken van persoonlijke informatie. Anders dan bij andere diensten, steunt een op life sciences gebaseerde dienstverlening veel meer op bevestiging door de medische sector. Dit heeft vervolgens weer aanzienlijke implicaties voor de welke distributie- en communicatiekanalen het meest geschikt zijn voor dit type dienstverlening.

NOTEN

1. Het fixed-effect model is geïmplementeerd door het toevoegen van een dummy voor elke individuele respondent.
2. Merk op dat het hoofdeffect van de situatie niet beoordeeld kan worden in combinatie met respondent dummies, door multicollineariteit. De reden is dat de situatie een perfect lineaire combinatie van respondent dummies is. Daarom rapporteren we alleen het interactie-effect waarop de kernfocus ligt van hypothese 6.

LITERATUUR

- Ajzen, Icek (1985). "From intentions to actions: a theory of planned behavior". In: J. Kuhl and J. Beckman (Eds.). *Action control: from cognition to behavior* (pp. 11-39). Berlin: Springer-Verlag.
- Ajzen, Icek en Martin Fishbein (1980). "Understanding attitudes and predicting social behavior". Englewood Cliffs, NJ: Prentice-Hall.
- Ajzen, Icek en Ellen T.J. Madden (1986). "Prediction of goal directed behavior: Attitudes, intentions, and perceived behavioral control". *Journal of Experimental Social Psychology*, 22, 453-474.
- Baron, Reuben M. en David A. Kenny (1986). "The moderator-mediator variable distinction in social psychological research: Conceptual, strategic, and statistical considerations". *Journal of Personality & Social Psychology* 51(6), 1173-1182.
- Bloch, Peter H. en Marsha L. Richins (1983). "A theoretical model for the study of product importance perceptions". *Journal of Marketing* 47 (Summer), 69-81.
- Cain, Mary en Greg Schmid (2003). "From nutrigenomics science to personalized nutrition: the market in 2010". Maureen Davis en Charles Grosel (Eds.): Institute for the Future.
- Chau Patrick Y.K. en Paul J.-H. Hu (2002). "Investigating Healthcare Professionals' Decisions to Accept Telemedicine Technology: an Empirical Test of Competing Theories". *Information & Management* 39 (January, 4), 297-311.
- Childers, Terry L., Christopher L. Carr, Joann Peck en Stephen Carson (2001). "Hedonic and Utilitarian Motivations for Online Retail Shopping Behavior". *Journal of Retailing* 77 (Winter 3), 511-535.
- Dabholkar, Pratibha A. (1996). "Consumer Evaluations of New Technology-Based Self-Service Options: An Investigation of Alternative Models of Service Quality". *International Journal of Research in Marketing* 13, 29-51.
- Dabholkar, Pratibha A. en Richard P. Bagozzi (2002). "An Attitudinal Model of Technology-Based Self-Service: Moderating Effects of Consumer Traits and Situational Factors". *Journal of Academy of Marketing Science* 30 (3), 184-201.
- Davis, Fred D. (1989). "Perceived Usefulness, Perceived Use, and User Acceptance of Information Technology". *MIS Quarterly* 13 (3), 319-340.
- Davis, Fred D., Richard P. Bagozzi en Paul R. Warshaw (1989). "User Acceptance of Computer Technology: a Comparison of Two Theoretical Models". *Management Science* 35 (8), 982-1003.
- Davis, Fred D., Richard P. Bagozzi en Paul R. Warshaw (1992). "Extrinsic and intrinsic motivation to use computers in the workplace". *Journal of Applied Social Psychology* 22, 1111-1132.
- Economist, 4/30/2005. "The no-computer virus – IT in the health-care industry". Vol. 375 Issue 8424, p65, 3p, 1 graph, 2c.
- Enriquez, Juan en Ray A. Goldberg (2000). "Transforming Life, Transforming Business: The Life-Science Revolution". *Harvard Business Review* 78 (2), 94-104.
- German, J. Bruce, Chahan Yeretian en Heribert J. Watzke (2004). "Personalizing foods for health and preference". *Food Technology* 58 (12), 26-31.

- Gutman, Jonathan (1982). "A Means-End Model Based on Consumer Categorization Processes". *Journal of Marketing* 46 (Spring), 60-72.
- Henneman, Lidewij, Danielle R.M. Timmermans en Gerrit van der Wal (2004). "Public Experiences, Knowledge and Expectations about Medical Genetics and the Use of Genetic Information". *Community Genetics* 7, 33-43.
- Hoffman, Donna L., Thomas P. Novak en Marcos Peralta (1999). "Building Consumer Trust Online". *Communications of the ACM* 42 (4), 80-85.
- Hoffman, Donna L. en Thomas P. Novak (1996). "Marketing in Hypermedia Computer-Mediated Environments: Conceptual Foundations". *Journal of Marketing* 60(July), 50-68.
- Hoy, Mariea Grubbs (1994). "Switch Drugs Vis-à-Vis Rx and OTC: Policy, Marketing, and Research Considerations". *Journal of Public Policy & Marketing* 13(1), 85-96.
- IHGSC (International Human Genome Sequencing Consortium) (2001). "Initial sequencing and analysis of the human genome". *Nature* 409, 860-921.
- Ives, Blake en Richard O. Mason (1990). "Can Information Technology Revitalize your Customer Service?". *Academy of Management Executive* 4 (4), 52-69.
- Kwon, Hyosun.S. en Laku Chidambaram (2000). "A test of the Technology Acceptance Model – The case of cellular telephone adoption". in *Proceedings of the 33rd Hawaii International Conference on System Sciences* 1 (January 4-7), p.1013.
- Massoud, Mark, Harriet Ragozin, Greg Schmid en Leah Spalding (2001). "The future of nutrition: consumers engage with science". Charles Grosel en Julie Koyano (Eds.): *Institute for the future*.
- Mathwick,Charla, Naresh Malhotra en Edward Rigdon (2002). "The Effect of Dynamic Retail Experiences on Experiential Perceptions of Value: An Internet and Catalog Comparison". *Journal of Retailing* 78 (1), 51-62.
- Meuter, Matthew L., Mary Jo Bitner, Amy L. Ostrom en Stephen W. Brown (2005). "Choosing among alternative service delivery modes: an investigation of customer trial of self-service technologies". *Journal of Marketing* 69 (2), 61-83.
- Moon, Ji-Won en Young-Gul Kim (2001). "Extending the TAM for a World-Wide-Web Context". *Information & Management* 38 (4), 217-230.
- Müller, Michael en Sander Kersten (2003). "Nutrigenomics: Goals and Strategies". *Nature Reviews Genetics* 4, 315-322.
- Murthi, B.P.S. en Sumit Sarkar (2003). "The Role of the Management Sciences in Research on Personalization". *Management Science* 49 (October), 1344-1362.
- Myers, James H. (1976). "Benefit Structure Analysis: A New Tool for Product Planning". *Journal of Marketing* 40 (October), 23-32.
- Nightingale, Paul en Paul Martin (2004). "The myth of the biotech revolution". *Trends in Biotechnology* 22 (11), 564-569.
- Novak, Thomas P., Donna L. Hoffman en Adam Duhachek (2003). "The Influence of Goal-Directed and Experiential Activities on Online Flow Experiences". *Journal of Consumer Psychology* 13(1&2), 3-16.
- O'Cass, A. en T. Fenech (2003). "Web Retailing Adoption: Exploring the Nature of Internet Users Web Retailing Behavior". *Journal of Retailing and Consumer Services*, 10 (2), 81-95.
- Ordovas, Jose M. en Dolores Corella (2004). "Nutritional Genomics". *Annual Review of Human Genetics* 5, 71-118.
- Parthasarathy, Shobita (2004). "Regulating Risk: Defining Genetic Privacy in the United States and Britain". *Science, Technology & Human Values* 29, 332-352.
- Pine, Joseph en James Gilmore (1999). "The Experience Economy. Work is Theatre & Every Business a Stage". Boston, Harvard Business School Press.
- Rabino, Isaac (2003). "Genetic Testing and its Implications: Human Genetics Researchers Grapple with Ethical Issues". *Science, Technology & Human Values* 28, 365-402.
- Rogers, Everett M. (2003) *Diffusion of Innovations* (5 ed.). New York: Free Press.

- Stearns, Sally C., Bernard L. Shulamit, Sara B. Fasick, Robert Schwarz, Robert T. Konrad, Marcia G. Ory en Gordon H. DeFries (2000). "The Economic Implications of Self-Care: The Effect of Lifestyle, Functional Adaptations, and Medical Self-Care Among a National Sample of Medicare Beneficiaries". *American Journal of Public Health* 90 (10), 1608-1612.
- Steckel Joel, Russ Winer, Randy Bucklin, Benedict G.C. Dellaert, Xavier Drèze, Gerald Häubl, Sandy Jap, John Little, Tom Meyvis, Alan Montgomery en Arvind Rangaswamy (2005). "Choice in Interactive Environments". Forthcoming in *Marketing Letters*.
- Suh, Bomil en Ingoo Han (2002). "Effect of Trust on Customer Acceptance of the Internet Banking". *Electronic Commerce Research and Applications* 1 (Autumn-Winter), 247-263.
- Takala, Tuija en Heta A. Gylling (2000). "Who Should Know about Your Genetic Makeup and Why?". *Journal of Medical Ethics* 26, 171-174.
- Taylor, Shirley en Peter A. Todd (1995). "Understanding information technology usage: a test of competing models". *Information Systems Research* 6 (2), 144-76.
- Ter Hofstede, Frenkel, Anke Audenaert, Jan-Benedict E.M. Steenkamp en Michel Wedel (1998). "An Investigation into the Association Pattern Technique as a Quantitative Approach to Measuring Means-End Analysis". *International Journal of Research in Marketing* 15 (1), 37-50.
- Trussel, James, Felcicia Stewart, Malcolm Potts, Felicia Guest en Charlotte Ellertson (1993). "Should Oral Contraceptives Be Available without Prescription?". *American Journal of Public Health* 83 (8), 1094-1099.
- Van der Heijden, Hans (2003). "Factors influencing the Usage of Websites: the Case of a Generic Portal in The Netherlands". *Information & Management* 40, 541-549.
- Venter, J.Craig *et al* (2001), "The sequence of the human genome". *Science* 291 (5507), 1304-1351.
- Vijayasarathy, Leo R. (2004). "Predicting Consumer Intentions to Use On-line Shopping: The Case for an Augmented Technology Acceptance Model". *Information & Management* 41 (July, 6), 747-762.

9. *Publicatiedrift van Nederlandse Marketingwetenschappers*¹

P.C. VERHOEF² en S. STREMERSCHE

SAMENVATTING

De Nederlandse marketingwetenschappers zijn na de Amerikaanse marketingwetenschappers momenteel de meest productieve ter wereld, wanneer gekeken wordt naar het aantal publicaties in vijf toptijdschriften. Wanneer gecorrigeerd wordt voor het inwonersaantal is Nederland zelfs het meest productieve land. In Nederland scoren vooral de Erasmus Universiteit Rotterdam, de Rijksuniversiteit Groningen, en de Universiteit van Tilburg goed. Op individueel niveau zijn het vooral wetenschappers van deze drie universiteiten die tot de meest productieve Nederlandse marketingwetenschappers moeten worden gerekend. Tenslotte blijkt dat sommige Nederlandse universiteiten zich breed positioneren, terwijl andere zich specialiseren. We bespreken ook de thema's waarop de verschillende groepen zich focussen in hun onderzoek.

1. INLEIDING

Marketing is nog een relatief jonge wetenschappelijke discipline. De eerste marketingvakgroepen of secties aan Nederlandse universiteiten werden zo eind jaren '60 – begin jaren '70 van de vorige eeuw opgericht. Dit gebeurde vooral aan de verschillende economische faculteiten, zoals die van de Erasmus Universiteit Rotterdam, de Rijksuniversiteit Groningen en de Universiteit van Wageningen. Zo werd 40 jaar geleden in 1965 de leerstoel Marketing en Marktonderzoek opgericht aan de Universiteit van Wageningen, welke werd bezet door Thieu Meulenberg. Dat Wageningen hier vooropliep is niet zo verwonderlijk, omdat marketing als vakgebied zijn oorsprong heeft in de agrarische sector. De nadruk van veel van deze vakgroepen lag in de eerste jaren vooral op het verrichten van onderwijs op het gebied van marketing en marktonderzoek en het uitdragen van het marketingconcept richting het bedrijfsleven. Daarnaast werd er wel onderzoek gedaan door een beperkt aantal wetenschappers. Er heerste toen echter een sterke cultuur om in Nederlandse bladen, zoals het *Tijdschrift voor Marketing* (toen nog met de wetenschappelijke katern "Interface"), *ESB*, *Maandblad Accountancy en Bedrijfs-economie (MAB)* en het *Jaarboek voor Marktonderzoek*, te publiceren. Zo rapporteert

1. De auteurs danken de redacteur Prof.dr.ir. Ale Smidts voor de vele nuttige opmerkingen op een eerdere versie van dit artikel.
2. Correspondentieadres: Prof.dr. Peter C. Verhoef, Rijksuniversiteit Groningen, Faculteit Economie, WSN-413; Postbus 800, 9700 AV Groningen, Tel. 050 363 7320/ 7065, Fax 050 363 3720, E-mail: p.c.verhoef@rug.nl

Wierenga in 1981 dat 44% van alle publicaties van NL wetenschappers in marketing in de periode 1977-1980 Nederlandstalig zijn (Wierenga 1981). In de afgelopen twee decennia is daarin echter grote verandering gekomen. Wetenschappelijk onderzoek neemt een steeds prominentere plaats in bij vakgroepen marketing in Nederland, terwijl publiceren in hoogstaande internationale tijdschriften de norm is geworden.

De noodzaak voor wetenschappelijk onderzoek is dan ook nadrukkelijk aanwezig. Door een grotere nadruk op internationaal wetenschappelijk onderzoek in het algemeen binnen de universitaire wereld is het voor marketing als wetenschappelijke discipline van groot belang om ook wetenschappelijk onderzoek te verrichten. Alleen hierdoor kan diepgaande en vernieuwende kennis worden opgebouwd over verschillende marketingonderwerpen, zoals de werking van marketing mix instrumenten, de ontwikkeling en adoptie van nieuwe producten, klantloyaliteit, en marketingstrategie. Daarnaast kan men door gedegen onderzoek trachten generaliseerbare uitspraken te doen over bepaalde fenomenen (zie bijv. Bijmolt, Van Heerde en Pieters 2005; Leeflang 2004; Nijs et al. 2001; Tellis, Stremersch en Yin 2003). Bovendien moeten de economische- en bedrijfskundige faculteiten zich van de HEAO's en andere onderwijsinstellingen onderscheiden. Hoogstaand onderzoek en de verwerking van de inzichten van dit onderzoek in het onderwijs richting studenten is hierbij essentieel. Natuurlijk mogen ook meer "rationele" argumenten voor wetenschappelijk onderzoek, de wetenschappelijke nieuwsgierigheid en andere drijfveren van onderzoekers niet uit het oog worden verloren (zie bijv. Willemsen, Heijnen en Termeer 1998).

Wetenschappelijk marketingonderzoek wordt door de meeste marketinggroepen in Nederland als een primaire taak gezien. Er zou zelfs kunnen worden gesproken over een soort publicatiedrift van Nederlandse marketingwetenschappers. Deze publicatiedrift wordt zichtbaar in de resultaten van recent onderzoek van onze hand, waarin we in algemene zin gekeken hebben naar de globalisatie van wetenschappelijk marketingonderzoek en de consequenties daarvan (Stremersch en Verhoef 2005). Uit dit onderzoek blijkt dat momenteel Nederland na de VS het meest productieve land is, wanneer gekeken wordt naar publicaties in vijf toptijdschriften. In dit artikel willen we verder ingaan op de publicatiedrift van marketingwetenschappers verbonden aan Nederlandse universiteiten. De volgende vragen staan centraal:

- (1) Hoe ontwikkelt de productiviteit van marketingwetenschappers in de vijf internationale marketing toptijdschriften verbonden aan Nederlandse universiteiten zich over de tijd?
- (2) Hoe ontwikkelt de productiviteit in de vijf internationale marketingtoptijdschriften zich voor de verschillende universiteiten in Nederland?
- (3) Wie zijn de meest productieve marketingwetenschappers in deze vijf internationale marketing toptijdschriften over de jaren heen?
- (4) In welke deelgebieden van marketing doen de verschillende marketingvakgroepen van Nederlandse universiteiten voornamelijk onderzoek?

Het onderzoeken van de eigen productiviteit kan wel eens worden geïnterpreteerd als een soort navelstaren, waarin vooral de wetenschappers zelf geïnteresseerd zijn. Het is echter ook relevant voor mensen in de marketing- en marktonderzoekpraktijk. Immers het geeft inzicht in: (1) de wetenschappelijke productiviteit van de verschillende marketinggroepen van Nederlandse universiteiten en (2) de specialisaties van deze groepen. Bedrijven kunnen deze informatie gebruiken bij het zoeken van geschikte universiteiten voor contractonderzoek en de werving van nieuw personeel onder marketing studenten.

Dit artikel is als volgt opgebouwd. We zullen eerst ingaan op de achtergrond van het onderzoek. Vervolgens zullen de centrale onderzoeksvragen worden behandeld. We eindigen met een toekomstperspectief.

2. ACHTERGROND ONDERZOEK

In eerder vermeld onderzoek (Stremersch en Verhoef 2005) hebben we voor de periode van 1964-2002, de publicaties in vijf Engelstalige internationale marketingtijdschriften verzameld: *International Journal of Research in Marketing*, *Journal of Consumer Research*, *Journal of Marketing*, *Journal of Marketing Research*, en *Marketing Science*. Deze tijdschriften worden door de Nederlandse wetenschappelijke wereld in het algemeen als *de vijf toptijdschriften* binnen marketing beschouwd (Sterken en Leeflang 2000). Dit blijkt ook uit de lijsten van de VSNU en de gehanteerde lijsten van onderzoekscholen, zoals CENTER (Tilburg), ERIM (Erasmus), METEOR (Maastricht) en SOM (Groningen), waarin ieder van deze tijdschriften een hoge ranking hebben. De tijdschriften verschillen met betrekking tot hun focus en achtergrond. *Journal of Marketing* is het oudste tijdschrift van de American Marketing Association of AMA (opgericht 1936) en heeft van origine meer een focus op marketingtheorie-ontwikkeling en marketing management vraagstukken. *Journal of Marketing Research* (opgericht 1964) en *Marketing Science* (opgericht 1981) hebben beide een meer kwantitatieve focus en worden uitgegeven door respectievelijk de AMA en INFORMS (deze organisatie geeft tijdschriften uit op het gebied van Operations Research en Management Science). *Journal of Consumer Research* (opgericht 1974) van de Association of Consumer Research (ACR) richt zich vooral op consumentenonderzoek met daarin verschillende theoretische invalshoeken, zoals psychologie, sociologie en economie. *International Journal of Research in Marketing* (opgericht 1984) is een Europees tijdschrift en gelieerd aan de European Marketing Academy (EMAC). Deze vijf tijdschriften worden allen tot de kern van de marketingtijdschriften gerekend (Baumgartner en Pieters 2003).

Door de focus op deze vijf top marketingtijdschriften nemen we expliciet alleen publicaties in de top en de kern van de marketingdiscipline in beschouwing. Hierdoor worden publicaties van marketingwetenschappers in toptijdschriften buiten marketing (bijvoorbeeld *Management Science*, *Academy of Management Journal* en *Strategic Management Journal*)ⁱ, en publicaties in goede maar lager gerangschikte tijdschriften binnen marketing (bijvoorbeeld, *Marketing Letters*, *Journal of Consumer Psychology*, *Journal of the Academy of Marketing Science* en *Journal of Retailing*) niet meegeteld. Voor de periode 1964-2002 hebben we over alle artikelen in de vijf toptijdschriften gegevens, zoals de internationale achtergrond van de auteurs, verzameld. Hierbij zullen we ons in dit artikel vooral richten op de universiteit waaraan de auteur primair verbonden is tijdens publicatie van het artikel. Deze gegevens staan meestal vermeld in het artikel. Dit betekent dat in het onderzoek ook nadrukkelijk gesproken wordt over *wetenschappers verbonden aan Nederlandse universiteiten*. Deze wetenschappers kunnen dus een andere dan de Nederlandse nationaliteit (bijv. Belgische) hebben. Voor een uitgebreide beschrijving van deze database en de verzameling van deze data verwijzen we naar Stremersch en Verhoef (2005).

Tabel 1. Rangschikking van output volgens land van aanstelling, Top 10.

Land	Totaal		1964-1973		1974-1983		1984-1993		1994-2002		1994-1998		1999-2002	
	%	Rang	%	Rang	%	Rang	%	Rang	%	Rang	%	Rang	%	Rang
Verenigde Staten	85.5	1	92.9	1	89.8	1	86.4	1	78.0	1	78.7	1	77.2	1
Canada	3.9	2	2.6	2	5.2	2	4.1	2	3.3	3	3.7	3	2.8	3
Nederland	2.0	3	0.1	11	0.2	9	1.5	4	4.9	2	4.6	2	5.3	2
UK	1.7	4	2.5	3	1.1	3	1.8	3	1.9	4	1.9	4	1.9	6
Frankrijk	1.1	5	0.2	7	0.9	4	1.1	5	1.7	5	1.4	6	2.0	5
Israel	0.9	6	0.3	5	0.9	4	1.0	6	0.9	10	0.6	10	1.2	10
België	0.8	7	0.4	4	0.2	9	0.7	7	1.6	6	1.7	5	1.4	8
Australië	0.7	8	0.2	8	0.5	6	0.5	9	1.2	9	1.1	8	1.4	9
Duitsland	0.6	9	0.1	11	0.1	11	0.5	8	1.3	8	1.1	7	1.5	7
Hong Kong	0.4	10	0.0	16	0.0	17	0.0	20	1.4	7	0.9	9	2.0	4
Overige	2.4		0.7		1.1		2.4		3.8		4.3		3.3	

Bron: Stremersch en Verhoev (2005). Totaal aantal artikelen n = 9943

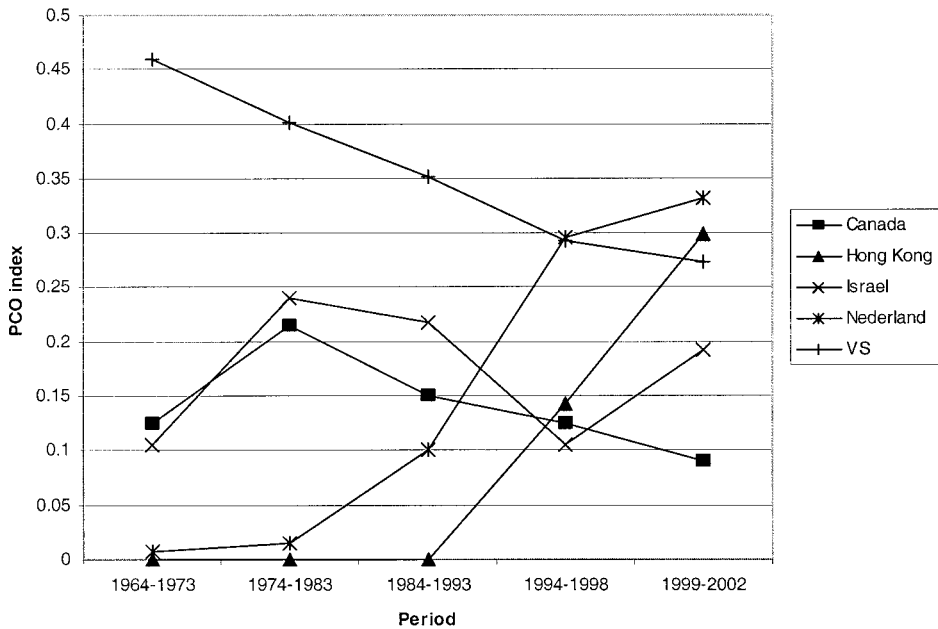
3. TOTALE PRODUCTIVITEIT OVER DE TIJD

In tabel 1 geven we een overzicht van de ontwikkeling van productiviteit van de verschillende landenⁱⁱ. Voor verschillende tijdsperioden wordt het percentage publicaties afkomstig van een land weergegeven met bijhorende rangorde. Hierbij beperken we ons tot de top 10 landen in de gehele bestudeerde periode. De Verenigde Staten neemt duidelijk een onbedreigde nummer 1 positie in. Marketing als vakgebied kent zijn oorsprong dan ook vooral in de Verenigde Staten. Dit blijkt onder andere uit de oorsprong van de in dit onderzoek bekeken tijdschriften. Daarnaast komen de internationaal bekende marketing- en marktonderzoektekstboeken, zoals *Marketing Management* van Philip Kotler, voornamelijk uit de Verenigde Staten. Het is dan ook niet verwonderlijk dat de meeste publicaties in de bestudeerde tijdschriften uit de Verenigde Staten komen en de Verenigde Staten de nummer 1 positie inneemt. Opvallend is echter wel, dat het percentage publicaties uit de Verenigde Staten over de tijd daalt. Dit wijst erop, dat marketing als wetenschap steeds internationaler wordt.

De positie van Nederland is over de tijd sterk verbeterd. In de periode van 1964-1984 heeft Nederland een zeer beperkt aandeel. In deze periode komt de productiviteit in de top marketingtijdschriften van een zeer beperkt aantal wetenschappers, zoals Berend Wierenga (nu Erasmus, toen werkzaam in Wageningen), Theo Verhallen (Tilburg), Fred van Raaij (nu Tilburg, toen Erasmus) en Peter Leeftang (Groningen) op gang (zie o.a. Leeftang en Reuyl 1984; Verhallen en Van Raaij 1982; Wierenga 1978; 1981; 1984). Vanaf 1984 neemt het aandeel echter sterk toe, met als resultaat een aandeel van 5.3% en de tweede positie in 1999-2002. In deze periode zijn er ook steeds meer wetenschappers verbonden aan Nederlandse universiteiten, die publiceren in de onderzochte tijdschriften. Hierbij dient wel te worden opgemerkt, dat vanaf 1984 *International Journal of Research in Marketing* wordt meegenomen in deze berekening. In dit tijdschrift van Europese origine hebben Europeanen en dus ook Nederlanders een relatief groot publicatieaandeel. Als we dit tijdschrift buiten beschouwing laten, zien we echter een zelfde ontwikkeling. De sterke positie van Nederland wordt nog verder onderstreept, wanneer de productiviteit per inwoner wordt bekeken. Dit cijfer is relevant omdat landen met veel inwoners, zoals de VS, relatief meer universiteiten en dus wetenschappers hebben en dus meer publicatiekracht hebben. Als we naar deze index kijken, blijkt dat Nederland het meest productieve land ter wereld is in de laatst bekeken periode (zie figuur 1). Deze index is berekend door het aantal publicaties in een bepaalde tijdsperiode van een land te delen door het aantal inwoners wonende in dat land in die tijdsperiode. Opvallend is verder de sterke stijging van productiviteit van landen uit Azië (Hong Kong).

3. PRODUCTIVITEIT PER UNIVERSITEIT

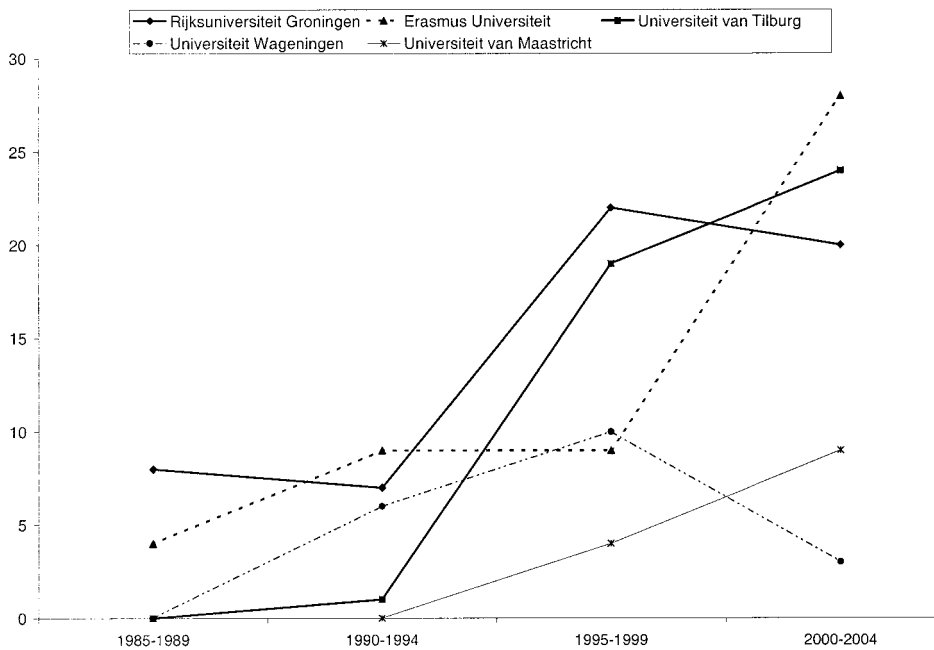
Het is vervolgens interessant om de productiviteit van de verschillende universiteiten te bekijken. Gezien de sterke stijging van het publicatieaandeel sinds 1984 bekijken we dit vanaf 1985, steeds per periode van 5 jaar. Om ook de meest recente ontwikkelingen in kaart te kunnen brengen hebben we bovendien data over de jaren 2003 en 2004 toegevoegd. Het analyseniveau is de universiteit. We maken dus geen onderscheid tussen verschillende groepen binnen een universiteit. Zo kunnen er bijvoorbeeld marketinggroepen bij een economische faculteit en een bedrijfskunde faculteit van dezelfde universiteit zijn.



Figuur 1. Productiviteit per land gecorrigeerd voor aantal inwoners (PCO Index).

De top 5 van universiteiten op het gebied van publicaties in deze bestudeerde perioden is: (1) Rijksuniversiteit Groningen, (2) Erasmus Universiteit Rotterdam, (3) Universiteit van Tilburg, (4) Universiteit van Wageningen en (5) Universiteit van Maastricht. Andere groepen, zoals die van de Vrije Universiteit, TU Delft en de Radboud Universiteit in Nijmegen, vallen duidelijk buiten de top 5ⁱⁱⁱ. Opvallend is wel de zesde plek van de Technische Universiteit Eindhoven, waar een kleine groep onderzoekers op het gebied van “Urban Planning” zich ook bezig houdt met vooral winkellocatie en de modelering van keuzegedrag van consumenten. De ontwikkeling van publicaties voor de top 5 wordt getoond in figuur 2. Deze figuur laat een aantal opvallende ontwikkelingen zien. In het algemeen stijgt de productiviteit van de meeste universiteiten over de tijd. De Universiteit van Tilburg heeft een zeer sterke stijging in output vanaf 1995. De Rijksuniversiteit Groningen neemt over de tijd gezien de meest dominante positie in, maar deze positie staat sinds 1998 onder druk. De productiviteit van de Erasmus Universiteit Rotterdam neemt duidelijk toe vanaf 1999, resulterend in een nummer 1 positie in de periode 2000-2004. Tenslotte vindt bij de Universiteit van Maastricht sinds 1995 een productiviteitstoename plaats. De productiviteit is echter vergeleken met die van de top 3 nog wel beperkt.

We houden bij deze rangschikking geen rekening met de grootte van de marketinggroepen van de verschillende universiteiten. Zo heeft Wageningen in vergelijking met de andere vier universiteiten een relatief kleine groep. Echter, het is erg moeilijk historische data over de precieze groeps grootte te verzamelen en aldus daarvoor te corrigeren.



Figuur 2. Evolutie van publicaties in toptijdschriften aan Nederlandse universiteiten (top 5).

4. PRODUCTIEVE WETENSCHAPPERS

Om inzicht te krijgen wie nu de meest productieve wetenschappers zijn in de bestudeerde periode van 1985-2004, is het aantal publicaties van Nederlandse wetenschappers verbonden aan Nederlandse universiteiten geteld^{iv}. Hierbij worden in de eerste kolom met de output alleen die publicaties geteld, waarbij de auteur aangaf op dat moment in Nederland te werken. In de tweede kolom met output wordt het totale aantal publicaties van de auteur vermeld. Publicaties van auteurs die op een bepaald moment niet verbonden zijn aan een Nederlandse universiteit maar aan een buitenlandse instelling, zijn hier dus in meegenomen. Volgens onze database hebben in deze periode ruim 80 auteurs verbonden aan Nederlandse universiteiten artikelen geschreven in de bestudeerde tijdschriften. De auteurs met een totale output van 4 of meer artikelen in de kerntijdschriften van marketing staan in tabel 2. Er ontstaat zo een top 8 van in totaal 16 auteurs.

De auteurs in deze tabel zijn bijna allemaal afkomstig van de top 3 universiteiten. De uitzonderingen zijn Hans van Trijp en Ko de Ruyter. Michel Wedel – tot voor kort verbonden aan de Rijksuniversiteit Groningen en nu verbonden aan University of Michigan – schreef 32 artikelen in deze topbladen en is daarmee de meest productieve marketingwetenschapper in Nederland in het bestudeerde tijdvak. Jan-Benedict Steenkamp (achtereenvolgens werkzaam in Wageningen, Leuven en Tilburg) staat tweede in de lijst met een totaal aantal publicaties van 30. Het vertrek van Michel Wedel naar een voltijdse aanstelling aan de University of Michigan en het aanstaande vertrek van Steenkamp naar de University of North Carolina (Chapel Hill), is dan ook een groot

Tabel 2. Top 10 auteurs (voorheen) primair verbonden aan NL universiteiten 1985-2004 op basis van publicaties in top 5 marketingtijdschriften.

Auteur	Momenteel verbonden aan	Promotie- jaar	Universiteit Promotie	Output in NL	Output totaal	Output zonder IJRM	Output per promotie- jaar
1. Michel Wedel	University of Michigan, Rijksuniversiteit Groningen	1990	Universiteit Wageningen	30	32	20	2.42
2. Jan-Benedict Steenkamp	Universiteit van Tilburg	1989	Universiteit Wageningen	15	30	16	2.00
3. Peter Leeftang	Rijksuniversiteit Groningen	1974	Erasmus Universiteit Rotterdam	16	16	4	0.85
4. Rik Pieters	Universiteit van Tilburg	1989	Universiteit Leiden	14	14	9	0.93
5. Tammo Bijmolt	Rijksuniversiteit Groningen	1996	Rijksuniversiteit Groningen	8	8	2	1.00
6. Stefan Stremersch	Erasmus Universiteit Rotterdam	2001	Universiteit van Tilburg	7	7	6	2.33
Berend Wierenga	Erasmus Universiteit Rotterdam	1974	Universiteit Wageningen	7	7	4	0.40
7. Harald van Heerde	Universiteit van Tilburg	1999	Rijksuniversiteit Groningen	5	5	5	1.00
Inge Geyskens	Universiteit van Tilburg	1998	Universiteit Leuven (B)	3	5	2	0.83
Gerrit van Bruggen	Erasmus Universiteit Rotterdam	1993	Erasmus Universiteit Rotterdam	5	5	3	0.45
Philip Hans Franses	Erasmus Universiteit Rotterdam	1991	Erasmus Universiteit Rotterdam	5	5	3	0.38
Els Gijbsbrechts	Universiteit van Tilburg	1984	UFSIA Antwerpen (B)	1	5	0	0.30
8. Katrijn Gielens	Erasmus Universiteit Rotterdam	1999	Universiteit Leuven (B)	4	4	2	0.80
Stijn van Osselaer	Erasmus Universiteit Rotterdam	1998	University of Florida (VS)	0	4	4	0.67
Hans van Trijp	Universiteit Wageningen	1995	Universiteit Wageningen	4	4	2	0.44
Ko de Ruyter	Universiteit van Maastricht	1993	Universiteit Twente	4	4	1	0.38

verlies voor de Nederlandse marketingwetenschap. Peter Leeflang en Rik Pieters volgen met een aantal van rond de 15. Uit de tabel valt ook de opkomst op van een aantal jonge onderzoekers, zoals Stefan Stremersch, Harald van Heerde en Katrijn Gielens. Dit wijst erop, dat er een aanwas is van nieuwe jonge onderzoekers.

In tabel 2 is niet alleen de huidige universiteit, waaraan de onderzoeker werkzaam is opgenomen, maar ook het jaar en de universiteit waaraan de onderzoeker gepromoveerd is. Opvallend is, dat 4 van de 16 onderzoekers gepromoveerd zijn aan de Universiteit van Wageningen. Aan de Erasmus Universiteit Rotterdam zijn 3 onderzoekers gepromoveerd.

Het is altijd erg moeilijk een rangschikking op te stellen, deze is geen uitzondering, wel in tegendeel. Gezien de huidige beloningsstructuur van universiteiten waarbij hoofdzakelijk het aantal publicaties in de vooraanstaande tijdschriften wordt geteld, is deze rangschikking dan ook een politiek erg gevoelige. Het is dan ook belangrijk om een aantal valkuilen bij de interpretatie van deze lijst te noemen.

Ten eerste is het totale aantal publicaties natuurlijk afhankelijk van de totale termijn waarover de onderzoeker heeft kunnen beschikken. Zo zijn sommige onderzoekers in het midden van de jaren '70 gepromoveerd (Leeflang en Wierenga), terwijl sommige andere erg recent (bijv. Gielens, Stremersch en Van Heerde) zijn gepromoveerd. De eersten hebben dan ook meer tijd gehad een groter aantal artikelen te publiceren. Een eenvoudige correctie hiervoor is de gemiddelde productiviteit per onderzoeker te berekenen. Hiervoor delen we de output door het aantal jaren, dat verstreken is tussen het promotiejaar van de onderzoeker en 2004. Voor de wetenschappers die eerder dan 1985 zijn gepromoveerd delen we de output door 20 (aangezien de bestudeerde publicatieperiode maximaal 1985-2004 omvat). Deze maatstaf is een benadering van de gemiddelde jaarlijkse productiviteit van een onderzoeker. Vijf onderzoekers hebben een productiviteit van 1 of meer publicaties per jaar: Michel Wedel (2.42), Stefan Stremersch (2.33), Jan-Benedict Steenkamp (2.00), Tammo Bijmolt (1.00) en Harald van Heerde (1.00). Natuurlijk is ook deze statistiek weer gevoelig voor kritiek. Zo evolueren wetenschappers gedurende hun carrière in productiviteit. Zo starten sommige onderzoekers erg traag, maar versnellen nadien. Andere onderzoekers starten zeer snel, maar vallen nadien terug. Terwijl een minderheid een relatief stabiele productiviteit heeft. Het arbitrair bestuderen van het gemiddelde tempo sinds promotie verhuult natuurlijk deze patronen. Ook is de onderzoeksfocus gericht op internationale toppublicaties in Nederland toch vrij recent (zie ook Sterken en Leeflang 2000). Daardoor scoren oudere onderzoekers ietwat lager in deze gecorrigeerde outputmaatstaf.

Ten tweede zou de lijst kunnen veranderen, wanneer het Europees georiënteerde tijdschrift *IJRM* niet wordt meegenomen. Om dit te bekijken is de output zonder *IJRM* berekend. Wanneer deze maatstaf wordt bekeken zou de top 2 hetzelfde blijven. Echter, daaronder zijn redelijk wat verschuivingen zichtbaar. Hierbij moet wel worden opgemerkt, dat pas sinds een aantal jaar het steeds vaker voorkomt dat wetenschappers geregeld in de vier van origine Amerikaanse tijdschriften publiceren.

Ten derde is zoals al eerder aangegeven een aantal toptijdschriften buiten marketing niet meegenomen. Hierdoor komen bepaalde marketingwetenschappers niet in de lijst voor, terwijl ze wel in hoogstaande niet-marketingtijdschriften hebben gepubliceerd.

Ondanks deze tekortkomingen, heeft de lijst wel enige "face-validity". Zo staan of stonden hoogleraren die hoog in onze lijst staan, zoals Michel Wedel, Jan-Benedict Steenkamp, Rik Pieters en Philip Hans Franses, ook hoog in de *Economist* Top 40.

5. SPECIALISATIES VAN UNIVERSITEITEN

5.1. Waar publiceert men?

Om een eerste indruk te krijgen van de specialisatie van de groepen, hebben we gekeken in welke tijdschriften marketingwetenschappers van de top 5 universiteiten vooral publiceren. Zoals aangegeven bij de beschrijving van de data hebben verschillende tijdschriften een iets andere positionering. In tabel 3 wordt de verdeling van de publicaties van 1984-2003 over de vijf meegenomen tijdschriften getoond. Procentueel scoort *International Journal of Research in Marketing* het hoogst bij alle vijf de groepen. Hiervoor zijn wellicht twee mogelijke oorzaken. Ten eerste wordt het *International Journal of Research in Marketing* internationaal meestal lager gerangschikt dan de andere 4 meegenomen tijdschriften (Baumgartner en Pieters 2004; Hult, Neese en Bas-haw 1997). Ten tweede, heeft *International Journal of Research in Marketing* sterk Nederlandse fundamente (waarbij Berend Wierenga (founding editor) en Jan-Benedict Steenkamp (editor) een sterke rol hebben gespeeld in de ontwikkeling van het tijdschrift).

Er zijn wel wat opvallende verschillen tussen de top 3 universiteiten. Zo heeft de Rijksuniversiteit Groningen relatief veel publicaties in de twee meest kwantitatieve tijdschriften: *Journal of Marketing Research* en *Marketing Science*. Dit wijst op de traditioneel kwantitatieve/econometrische focus van het marketing onderzoek van deze universiteit. Bij deze universiteit wordt er echter weinig aan consumentenonderzoek gedaan, gezien de lage score bij *Journal of Consumer Research*. De Erasmus Universiteit Rotterdam scoort relatief hoog bij *Journal of Marketing*. De percentages bij de *Journal of Marketing Research* en *Marketing Science* liggen echter dicht bij het gemiddelde. Bij *Journal of Consumer Research* wordt onder het gemiddelde gescoord. Over de bestudeerde periode lijkt de Erasmus Universiteit Rotterdam, dan ook zich wat meer te richten op zowel kwantitatief / econometrisch georiënteerd marketing onderzoek als onderzoek op het gebied van marketing management. Consumentenonderzoek krijgt minder aandacht. Op dit laatste veld scoort de Universiteit van Tilburg zeer goed. De andere percentages liggen echter dicht bij de totalen. Dit wijst erop, dat Tilburg in termen van wetenschappelijk marketingonderzoek, een brede onderzoeksgroep is. Ze doen kwantitatief georiënteerd onderzoek, onderzoek op het gebied van marketing theorie en marketing management en consumentenonderzoek. Opvallend is het hoge percentage bij *Journal of Marketing* voor de Universiteit van Maastricht. Hierbij moet echter wel gemeld worden, dat het totale aantal publicaties van Maastricht relatief klein is.

Tabel 3. Procentuele verdeling van publicaties over tijdschriften heen per universiteit.

	Journal of Marketing	Journal of Marketing Research	Marketing Science	Journal of Consumer Research	International Journal of Research in Marketing	Totaal aantal artikelen
Rijksuniversiteit Groningen	3.4%	32.2%	15.3%	3.4%	45.8%	57
Erasmus Universiteit Rotterdam	13.7%	23.5%	11.8%	3.9%	47.1%	50
Universiteit van Tilburg	9.1%	18.2%	11.4%	13.6%	47.7%	44
Universiteit Wageningen	5.3%	21.1%	5.3%	5.3%	63.2%	19
Universiteit van Maastricht	30.8%	7.7%	0.0%	0.0%	61.5%	13
Totaal	9.7%	23.7%	11.3%	5.9%	49.5%	183

5.2. Waarover publiceert men?

Om een verder inzicht te krijgen in de profilering van de verschillende groepen kijken we voor de top 5 universiteiten naar de inhoud van de artikelen. De artikelen worden op basis van hun titel en samenvatting toegewezen aan één bepaald thema. Zo wordt bijvoorbeeld een artikel dat een methode onderzoekt om merkbaarheid te meten alleen toegewezen aan de categorie marketing onderzoek en niet aan de categorie merkenbeleid (zie bijv. Bijmolt e.a. 1998). De resultaten van deze analyse staan in Tabel 4. In deze tabel staan per marketinggroep het percentage van output toegewezen aan een bepaald onderwerp.

Tabel 4. Inhoudelijke analyses van publicaties marketinggroepen (1985-2004) (n=183).

Thema	Erasmus					Gemiddeld
	Universiteit Rotterdam	Universiteit Groningen	Universiteit Maastricht	Universiteit Tilburg	Universiteit Wageningen	
Marketingstrategie	3.4%			2.1%		1.6%
Internationale Marketing		5.0%		4.2%	11.8%	3.6%
Innovatie, diffusie, adoptie	10.3%	1.7%		2.1%		4.2%
Customer Relationship Management/Direct Marketing	1.7%	6.7%	11.1%			3.1%
Marketingkanalen/ Business-to-Business Marketing	12.1%		11.1%	6.3%	5.9%	6.3%
Econometrische Marketingmodellen (effect van marketingmix)	13.8%	31.7%		10.4%	5.9%	17.2%
Marktonderzoektechnieken (segmentatie, conjoint, e.d.)	22.4%	43.3%		27.1%	47.1%	31.3%
Dienstenmarketing	3.4%		66.7%			4.2%
Reclame en Communicatie		8.3%		16.7%		6.8%
Merkenbeleid						
Marketing Beslissingsondersteuning	17.2%				5.9%	5.7%
Retailing	6.9%		11.1%	12.5%	5.9%	6.3%
Consumentengedrag	3.4%	3.3%		10.4%	17.6%	6.3%
Verkoopbeleid	5.2%					1.6%
Overig				8.3%		2.1%
Totaal	100%	100%	100%	100%	100%	100%

De Nederlandse marketingwetenschap is redelijk kwantitatief georiënteerd met veel onderzoek op het gebied van marktonderzoekstechnieken en econometrische marketingmodellen. Dit betreft ongeveer 50% van de publicaties. Opvallend is verder dat er minder aandacht is voor onderwerpen als marketingstrategie en merkenbeleid.

Tussen universiteiten zijn er duidelijke verschillen. De Erasmus Universiteit Rotterdam doet in vergelijking met de andere universiteiten relatief veel onderzoek op het gebied van innovatie, marketingkanalen en business-to-business marketing, marketing beslissingsondersteuning en verkoopbeleid. Echter ook aan deze universiteit is er een belangrijke onderzoeksstroming gericht op marktonderzoekstechnieken. De Rijksuniversiteit Groningen is het meest kwantitatief georiënteerd van alle groepen. Opvallend is verder de relatief hoge score op Customer Relationship Management (CRM) en Direct Marketing. Maastricht richt zich sterk (66%) op dienstenmarketing. Dit is een duidelijke focusstrategie aangezien het een specifiek (binnen Nederland handelen slechts 4% van alle publicaties over dienstenmarketing), maar wel belangrijk veld is in marketing (Vargo en Lush 2004). De Universiteit van Tilburg doet in vergelijking met andere groepen veel onderzoek op het gebied van reclame en communicatie, consumentenge-

drag, en retailing. De score op kwantitatieve methoden is ongeveer gelijk aan die van de Erasmus Universiteit Rotterdam. De Universiteit van Wageningen heeft ook relatief veel onderzoek gepubliceerd op het gebied van marktonderzoekstechnieken. Daarnaast scoort men relatief hoog op consumentengedrag.

Op basis van deze analyse kunnen we concluderen dat er duidelijke verschillen zijn tussen de groepen. De Erasmus Universiteit Rotterdam en de Universiteit Tilburg publiceren over een breed scala van onderwerpen en zijn dus marketing breed actief. De drie andere universiteiten hebben duidelijk een sterke focus in hun onderzoek.

Ook bij deze analyse zijn er wederom weer een aantal beperkingen. De belangrijkste is dat de analyse van de inhoud van het onderzoek alleen gebaseerd is op publicaties in de 5 top marketingtijdschriften. Deze tijdschriften zijn vooral geïnteresseerd in algemene methode- en theorieontwikkeling in marketing. Onderzoek over specifieke thema's met een meer interdisciplinair karakter, zoals bijv. marketing van voedingsmiddelen en in de agribusiness (Wageningen), zal daar minder snel gepubliceerd worden. Hierover wordt dan gepubliceerd in meer specifieke tijdschriften. Omdat deze tijdschriften niet zijn meegenomen, kunnen bepaalde onderzoeksonderwerpen niet of maar beperkt naar voren komen in de analyse.

5.3. Waar wil men over publiceren?

Aan bovenstaande analyses zit nog een nadeel verbonden. Men bekijkt de situatie in het verleden. De specialisatie van een universiteit heeft echter ook te maken met wat voor onderzoek men in de toekomst wil doen. Om de positionering van universiteiten verder te bekijken hebben we de websites van de verschillende vakgroepen bezocht. In feite wordt hiermee de vraag beantwoordt, waarover men wil publiceren. De inhoud van deze websites werd aangevuld en gecorrigeerd door interviews met vakgroepvoorzitters om zeker te weten dat de inhoud up-to-date is. Hierbij beperken we ons wederom tot de top 5 universiteiten. De resultaten van deze exercitie staan in tabel 5.

Tabel 5. Overzicht specialisaties van onderzoek van top 5 universiteiten.

Universiteit	Specialisaties
Erasmus Universiteit Rotterdam	Marketingmodellen, high tech marketing, life sciences marketing, innovatie, managerial decision-making in marketing, marketing management support systemen, neuroscience & marketing, consumer decision-making, sales en account management, and knowledge-based marketing
Rijksuniversiteit Groningen	Geavanceerde marketingmodellen (bijv.complexiteit, grote datasets), promoties, CRM, distributie en detailhandel, en marketing van geneesmiddelen
Universiteit van Maastricht	Marketing en management van diensten. markt leergedrag, interactieve marketing, relatiemarketing, logistiek en ketenbeheer, en merkenbeleid.
Universiteit Tilburg	Marketingstrategie, marketingmodellen en consumentengedrag
Universiteit van Wageningen	Marketing en consumentengedrag in de agri- en voedingsbranche

Wanneer we de specialisatie bekijken lijken de Erasmus Universiteit Rotterdam en de Universiteit van Tilburg het brede veld van marketingthema's te willen bestuderen. Ze doen onderzoek op alle deelgebieden binnen marketing: marketingstrategie (marketing besluitvorming bij bedrijven), consumentengedrag (besluitvorming bij consumenten) en marketingmodellen (kwantitatieve marketing). Dit heeft waarschijnlijk te maken met de grootte van de groep, waarbij beide veruit de grootste groepen zijn in voltijdse equivalenten binnen het docentenkorps. Beide groepen hebben echter ook specifieke thema's. Zo focust Tilburg o.a. op kwantitatieve promotiemodellen, terwijl Rotterdam een aantal unieke thema's heeft geclaimd: high-tech marketing, life sciences, marketing decision support systemen, sales & account management en neuroscience & marketing. De Rijksuniversiteit Groningen stelt data en kwantitatieve methoden in hun onderzoek centraal en past deze toe op verschillende thema's. Maastricht kiest voor een focus op diensten. Daar binnen lijkt men echter wel weer verschillende thema's te bestuderen, zoals interactieve marketing, relatiemarketing en logistiek. De agrarische focus van de Universiteit van Wageningen vertaalt zich ook in het marketingonderzoek, dat zich vooral richt op de agrarische bedrijven en voedingsmiddelen.

Naast de top 5 zijn er nog een aantal andere marketinggroepen. Deze groepen kiezen vaak voor een specifiek thema. Zo kiest de marketinggroep van de VU vooral voor strategische marketing, terwijl de marketinggroep van de Radboud Universiteit Nijmegen zich vooral richt op relatiemanagement. De marketinggroep van de Technische Universiteit Delft houdt zich vooral bezig met nieuwe productontwikkeling. De TU Twente (Faculteit Gedragwetenschappen) richt zich in hun onderzoek vooral op marketingcommunicatie met daarbij een duidelijke inbreng vanuit de sociale psychologie. De TU Eindhoven heeft een "urban planning groep" rond Harry Timmermans, die zich op het gebied van marketing bezig houdt met vraagstukken rond winkelkeuze en planning van winkels. Daarnaast is er bij de TU Eindhoven een aantal marketingwetenschappers verbonden aan de faculteit Technologie en Management. Hun onderzoek richt zich op het aanwenden van nieuwe technologieën binnen marketing (E-Commerce, M-Commerce, etc.) en nieuw productontwikkeling voor zowel goederen als diensten. Tot slot heeft Nyenrode een "Center for Supply Chain Management" met daarinbinnen aandacht voor marketingstrategie, CRM, account management, retailing en productontwikkeling.

Bij de analyses van de inhoudelijke focus van verschillende groepen dienen wel een aantal opmerkingen geplaatst te worden. Ten eerste, is de focus onder andere afhankelijk van specifieke personen. De onderzoeksagenda van een specifiek lid van de onderzoeksgroep kan ervoor zorgen, dat deze onderzoeksgroep relatief veel onderzoek doet op een bepaald gebied. Als deze onderzoeker vervolgens de betreffende universiteit verlaat, verzwakt de aandacht voor dat onderwerp. Ten tweede kunnen inhoudelijke positioneringen "politiek" bepaald zijn. Een faculteit kan bijvoorbeeld kiezen voor een bepaalde onderzoeksfocus, waarbinnen alle verschillende disciplines in ieder geval in hun positionering statements aan moeten voldoen. Dit zou in principe kunnen leiden tot een discrepantie tussen wat men zegt en wat men uiteindelijk onderzoekt. In onze analyses lijken deze verschillen niet zo duidelijk. De conclusies van de inhoudelijke analyse van de gepubliceerde artikelen komen gemiddeld genomen overeen met de positionering van de groepen op de websites. Desondanks blijft het bij de interpretatie van deze analyse belangrijk om zich te realiseren, dat al deze analyses de huidige status quo of zelfs die van het verleden aangeven. Zo kan het in de toekomst aantrekken van een nieuwe onderzoeker de inhoudelijke onderzoeksfocus duidelijk veranderen.

6. DISCUSSIE

In dit artikel is een overzicht gegeven van de wetenschappelijke output van marketingwetenschappers verbonden aan Nederlandse universiteiten. Onze resultaten laten duidelijk zien dat Nederlandse marketingwetenschappers op een hoog niveau onderzoek doen en zeer productief zijn. Nederland heeft een prominente plaats in de wetenschappelijke marketingwereld.

De centrale vraag is of dit niveau in de toekomst gehandhaafd blijft. Er zijn hiervoor een aantal mogelijke ontwikkelingen, die positief en negatief werken. Positief is het feit dat steeds meer marketinggroepen wetenschappelijk onderzoek belangrijk vinden en hun organisatie daar ook op inrichten. Hierbij kan gedacht worden aan de oprichting van onderzoeksscholen en een belonings- en bevorderingsstructuur waarbij het type en/of aantal publicaties een belangrijke rol spelen. Het vertrek van productieve wetenschappers naar het buitenland zou echter negatief kunnen werken. De salarisstructuur is duidelijk een uitdaging op dit vlak. Niet alleen blijft Nederland achter op de Verenigde Staten, maar in groeiende mate ook op landen als Hong Kong, Singapore, Frankrijk en het Verenigd Koninkrijk. De Verenigde Staten hebben voorlopig nog steeds de sterkste aantrekkingskracht, omdat daar nog steeds de grootste groep marketingwetenschappers zit en je dus als wetenschapper makkelijker toegang krijgt tot het dominante netwerk. Ons onderzoek laat zien dat deze toegang van groot belang is om impact van onderzoek te realiseren (Stremersch en Verhoef 2005). Aan de andere kant wordt de positie van de Verenigde Staten als belangrijke aantrekkingspool ook bedreigd. De voornaamste reden hiervoor zijn de toenemende "home land security" problemen, die in toenemende mate een druk zetten op migratiestromen van professoren en studenten. Sinds 9/11 worden door de Verenigde Staten veel visa-aanvragen uit het educatieve milieu geweigerd. Dit zorgt er onder meer ook voor dat een toenemend aantal wetenschappers van Europese origine terugkeert naar Europa. Of er een "brain-drain" dreigt voor de Nederlandse marketingwetenschap is dus een open vraag. Hoe universiteiten reageren, zal daarbij bepalend zijn. Op dit moment wordt volop gekozen voor twee oplossingen: (1) jonge onderzoekers sneller tot hoogleraar promoveren; (2) jonge hoogleraren sneller laten doorgroeien tot het maximale salaris. Maar hoe lang dit systeem volstaat, is op dit moment onzeker.

Ook is een kritische zelfreflectie over de marketingwetenschap op zijn plaats. Zo noteert Stremersch (2005) dat er een groeiende kloof is tussen de marketingwetenschap en de praktijk. Deze kloof wordt ook door andere wetenschappers gezien (zie o.a. Bemmaor en Franses 2005). Slechts een zeer beperkt aantal marketing mensen uit de praktijk lezen wetenschappelijke tijdschriften. Het dichten van deze kloof is een lastige zaak. Wetenschappers zullen vooral geïnteresseerd zijn in het verdiepen van kennis over onderwerpen en het publiceren over deze thema's. De praktijk wil vaak snelle oplossingen voor dagelijkse marketingproblemen (Leeflang 2005). Deze twee werelden lijken soms haast niet met elkaar verenigbaar. Goed wetenschappelijk onderzoek vergt nu eenmaal een diepgaande studie van de materie en een uitgebreide analyse van gegevens. Desondanks zijn er een aantal mogelijke oplossingen. Ten eerste zullen wetenschappers moeten trachten relevant en interessant onderzoek te doen. Onderzoekers zullen hierbij verder moeten gaan, dan alleen het verdiepen van theoretische concepten. Zo kan men focussen op interessante traditionele marketingonderwerpen, zoals de effectiviteit van marketinginstrumenten, en nieuwe onderzoeksvelden, zoals marketing en de medische wetenschap en marketing in een kenniseconomie. Onderzoek doen in

thema's die midden in het maatschappelijke debat staan zal de impact van de marketingwetenschap ook verhogen. Een probleem is echter wel dat maatschappelijke debatten vluchtig zijn. Wat nu in is, kan morgen al niet meer relevant zijn. Ten tweede zullen in een kenniseconomie ook Nederlandse marketeers zich continu moeten richten op het vermeerderen van hun marketingkennis. De link met de marketingwetenschap is dan essentieel. Marketingwetenschappers zullen dan echter wel hun onderzoeksresultaten toegankelijker moeten maken voor mensen uit de praktijk. Ten derde, kunnen samenwerkingsverbanden tussen universiteit en bedrijfsleven/overheid in instituten en onderzoekscentra zich vooral richten op de ontwikkeling van algemene kennis over bepaalde marketingfenomenen. Aan Nederlandse universiteiten zijn en worden initiatieven ontwikkeld om deze samenwerking een structureel karakter te geven.

Het dichten van de kloof tussen marketingwetenschap en praktijk, is niet alleen de verantwoordelijkheid van de wetenschap. De praktijk zal daadwerkelijk open moeten staan voor wetenschappelijk onderzoek en de toepassing daarvan. Dit betekent minder korte termijn denken en meer lange termijn denken gericht op gezamenlijke kennisontwikkeling. Dit betekent ook minder gericht op interessante cases en leuke gimmicks, maar ook meer gericht op inhoudelijke verdieping van marketing kennis. En het betekent, dat bedrijven ervoor open moeten staan om data en kennis te delen. Dit gebeurt al in samenwerkingsverbanden met grote onderzoeksbureaus, zoals GFK, TNS, IRI, Market Response en AC Nielsen, waarbij deze bureaus in grote onderzoeksprojecten data delen met onderzoekers (zie bijv. Nijs et al. 2001).

Zoals al aangegeven is dit onderzoek niet alleen relevant voor marketingwetenschappers, maar ook voor de marketing- en marktonderzoekpraktijk. Marktonderzoekbureaus krijgen met dit onderzoek beter inzicht in waar goed getrainde studenten in marketingonderzoek worden opgeleid. Daarnaast fungeren onderzoeksbureaus en marketingadviesbureaus als intermediairs bij de toepassing van nieuwe methoden, technieken en theorieën. Op basis van dit onderzoek kunnen zij bepalen met welke universiteiten zij mogelijk intensievere contacten willen hebben om zo toegang te hebben tot de nieuwste wetenschappelijke marketingkennis.

NOTEN

- i Verschillende marketingwetenschappers publiceren ook hoogstaand onderzoek in deze bladen (zie o.a. Leeflang, Foekens en Wittink 1999; Scheer, Kumar en Steenkamp 2003; Smidts, 1997).
- ii Als er in dit artikel gesproken wordt over productiviteit, dan wordt er gekeken naar auteur – artikel combinaties. Een artikel dat is geschreven door twee auteurs verbonden aan Nederlandse universiteiten telt dus twee keer mee voor Nederland.
- iii De exacte lijst met gegevens voor elke universiteit is beschikbaar bij de auteurs.
- iv Door de periode 1985-2004 te nemen, wordt een aantal publicaties van voor 1985 van langer actief zijnde wetenschappers niet meegenomen. De totale productiviteit van deze wetenschappers gedurende hun wetenschappelijke carrière kan dus duidelijk hoger liggen. Echter, de veranderingen in de lijst zijn beperkt. Zo zou Wierenga dan uitkomen op een output van 10, Leeflang op een output van 17 en Gijsbrechts op een output van 6. Bovendien zou dan Fred van Raaij (Tilburg) in de lijst voorkomen met een output van 4.

LITERATUUR

- Baumgartner, H. en R. Pieters (2003), The Structural influence of marketing journals: A citation analysis of discipline and its subareas over time, *Journal of Marketing*, 67 (2), 123-139
- Bijmolt, T.H.A., Van Heerde, H.J. en Pieters R.G.M. (2005), Determinants of price elasticity: New empirical generalizations, *Journal of Marketing Research*, 42 (2), 141-156
- Bijmolt, T.H.A., Wedel, M., Pieters R.G.M. en DeSarbo, W.S. (1998), Judgments of brand similarity, *International Journal of Research in Marketing*, 15(3), 249-268
- Bemmar A. en Franses, P.H. (2005), The diffusion of marketing science in the practitioners' community: Opening the black box, *Applied Stochastic Models in Business and Industry* (special issue on Bridging the Gap Between Academic Research in Marketing and Practitioners' Concerns) 21 (3), te verschijnen
- Hult, G.T.M., Neese, T.M. en Bashaw, R.E. (1997), Faculty perceptions of marketing journals, *Journal of Marketing Education*, 19 (1), 37-52
- Leeflang, P.S.H. en Reuyl J.C. (1984), On the predictive power of market share attraction models, *Journal of Marketing Research*, 21 (2), 211-215
- Leeflang, P.S.H., Foekens E.W. en Wittink, D.R. (1999), Varying parameter models to accommodate dynamic promotion effects, *Journal of Econometrics*, 89 (1-2): 249-268
- Leeflang, P.S.H. (2005), *Meten op Markten*, KNAW, Amsterdam
- Leeflang, P.S.H. (2004), Heeft marketing nog toekomst?, *Holland Management Review*, Januari-Februari, 74-79
- Nijs, V., DeKimpe, M.G., Steenkamp, J.B.E.M. en Hansens, D.M. (2001), The Category-demand effects of price promotions, *Marketing Science*, 20 (1), 1-22
- Scheer L.K., Kumar N. en Steenkamp J.B.E.M (2003), Reactions to perceived inequity in US and Dutch interorganizational relationships, *Academy of Management Journal*, 46 (3), 303-316
- Smidts A. (1997), The relationship between risk attitude and strength of preference: A test of intrinsic risk attitude, *Management Science*, 43 (3), 357-370
- Stremersch, S. en P.C. Verhoef (2005), The globalization of the marketing field: Does it help or hinder the field, *Marketing Science*, te verschijnen
- Stremersch, S. (2005), *Op zoek naar een publiek*, Oratie. Inaugural Adresses Series, Erasmus Research Institute of Management (ERIM), Rotterdam. Online beschikbaar via: <http://hdl.handle.net/1765/1945>
- Sterken, E. en P.S.H. Leeflang (2000), Een expertlijst van tijdschriften in de bedrijfswetenschappen, *Maandblad voor Accountancy en Bedrijfseconomie*, 74 (4), 73-85
- Tellis, G.J., Stremersch, S. en Yin, E. (2002), The international takeoff of new products: the role of economics, culture and country innovativeness, *Marketing Science*, 22 (2), 188-208
- Vargo, S.L en Lush, R.F. (2004), Evolving to a new dominant logic for marketing, *Journal of Marketing*, 68 (1), 1-17
- Verhallen, T.M.M.. en Van Raaij, W.F. (1982), Household behavior and the use of natural-gas for home heating, *Journal of Consumer Research*, 8 (3), 253-257
- Wierenga, B. (1978), Least-squares estimation method for linear learning model, *Journal of Marketing Research*, 15 (1), 145-153
- Wierenga, B. (1981), The development of marketing in the Netherlands, *Journal of Marketing*, 45 (4), 150-153
- Wierenga, B. (1984), Empirical test of the Lancaster characteristics model, *International Journal of Research in Marketing*, 1 (4), 263-293
- Willemsen, K., Heijnen G. en Termeer, G. (1998), *Het heilig vuur: De kern van het universitaire bestaan*, Wolters Noordhof, Groningen

10. Multi-niveau latente klasse analyse: Met een toepassing bij het simultaan clusteren van landen en consumenten

J.K. VERMUNT, T.H.A. BIJMOLT en L.J. PAAS

SAMENVATTING

In dit artikel bespreken we latente klasse analyse, een model gebaseerde clustertechniek. Speciale aandacht wordt besteed aan een recent voorgestelde uitbreiding van latente klasse analyse die gebruikt kan worden voor de analyse van multi-niveau gegevens, zoals gegevens van respondenten die gegroepeerd kunnen worden in postcodes, van klanten gegroepeerd in filialen of van consumenten gegroepeerd in landen. Na een introductie waarin we latente klasse analyse bespreken en vergelijken met andere clusteranalyse methoden, behandelen we het multi-niveau latente klasse model. Vervolgens illustreren we dit multi-niveau model met een toepassing waarin consumenten worden geclusterd op basis van hun portfolio aan financiële producten en landen worden geclusterd op basis van de verdeling van consumenten over de verschillende segmenten. We bespreken zowel de marketingimplicaties van de bevindingen als de bredere toepassing van het multi-niveau latent klasse model voor marktonderzoek.

1. INLEIDING

Consumenten verschillen sterk in preferenties voor producten en diensten en in reacties op reclames en veel andere voor marketing relevante variabelen. Eén van de doelstellingen bij marktonderzoek is dan ook het beschrijven van de heterogeniteit in de markt op basis van waargenomen of uitgesproken preferenties van consumenten (revealed versus stated preferences). Bij waargenomen preferenties gaat het bijvoorbeeld om informatie over het bezit of de aankoop van bepaalde producten, terwijl het bij uitgesproken preferenties meestal gaat om attitudes of gedragsintenties.

Afhankelijk van de toepassing zal men de heterogeniteit in preferenties beschouwen als een continu of juist als een discreet verschijnsel. In het eerste geval zal men gebruik maken van, bijvoorbeeld, factoranalyse of random-effects (hierarchical Bayes) regressietechnieken. Opsporen van discrete categorieën consumenten met verschillende preferenties (marktsegmenten) geschiedt over het algemeen met behulp van clusteranalytische technieken. In de academische marketing wereld is een debat gaande over welke van deze twee benaderingen de beste is, waarbij de kwaliteit van de voorspelling van (toekomstig) individueel gedrag als voornaamste vergelijkingscriterium wordt gehanteerd. Sommige auteurs vinden dat modellen met continue heterogeniteit het beter doen, terwijl anderen aantonen dat modellen met discrete heterogeniteit even goed in staat zijn om individueel gedrag te voorspellen.

In dit artikel vinden we segmenten met behulp van clusteranalyse technieken (Wedel en Kamakura, 2000). De specifieke techniek die wij gebruiken is het latente klasse model, ook wel aangeduid als mixture model. In tegenstelling tot andere clusteranalyse methoden betreft het hier een op een statistisch model gebaseerde clustertechniek. Vandaar dat ook de term “model-based clustering” wordt gebruikt (McLachlan en Peel, 2000). Om precies te zijn: we veronderstellen dat de geobserveerde data tot stand zijn gekomen via een bepaald proces dat kan worden gemodelleerd. Het simpelste voorbeeld van een dergelijk proces is dat de populatie bestaat uit S segmenten met verschillende preferenties. De indeling in segmenten leidt tot verschillen in de kans om, bijvoorbeeld, een bepaald product te bezitten.

Doel van dit artikel is inzicht te geven in het gebruik van latente klasse analyse voor marktsegmentatie. Hierbij gaan we vooral in op een veelbelovende nieuwe variant, het multi-niveau latente klasse model. Na de inleiding vergelijken we methoden voor clusteranalyse en latente klasse analyse. Daarna bespreken we standaard latente klasse analyse en enkele voor marktonderzoek relevante varianten. Vervolgens besteden we aandacht aan latente klasse modellen voor multi-niveau data. We illustreren het multi-niveau latente klasse model met een toepassing waarin consumenten worden geclusterd op basis van hun portfolio aan financiële producten en landen worden geclusterd op basis van de verdeling van consumenten over de verschillende segmenten. De marketingimplicaties van de bevindingen van deze analyse worden vervolgens besproken en het artikel wordt afgesloten met een korte discussie over bredere toepassingen van het multi-niveau latent klasse model voor marktonderzoek.

2. VERGELIJKING TUSSEN VERSCHILLENDE METHODEN VOOR CLUSTERANALYSE

Latente klasse analyse, clusteranalyse op basis van een statistisch model, verschilt sterk van andere clustermethoden, zoals hiërarchische en niet-hiërarchische clustertechnieken (Wedel en Kamakura, 2000). Bij hiërarchische clustertechnieken wordt eerst een tabel met paarsgewijze afstanden (similariteiten) berekend tussen respondenten en vervolgens worden clusters gevormd door, via een bepaald algoritme, de respondenten (en clusters) die op elkaar lijken samen te voegen. Hoewel deze methoden een aantal ad hoc elementen bevatten (welke afstandsmaat, welke procedure van samenvoegen), werken ze over het algemeen goed bij kleine steekproeven.

Latente klasse analyse maakt geen gebruik van een dergelijke similariteitsmatrix en is evenals niet-hiërarchische clustertechnieken zoals K-means juist meer geschikt voor segmentatie op basis van grotere steekproeven. K-means clustering en latente klasse analyse zijn sterk verwante technieken. Hoewel de eerste niet is gebaseerd op een expliciet populatiemodel, is het wel degelijk mogelijk om de impliciete aannames te vertalen in een zeer restrictief model-based clustering model (Magidson en Vermunt, 2001). Een belangrijk verschil is dat K-means een harde partitionering geeft terwijl latent klasse analyse leidt tot een fuzzy partitioning. Dat wil zeggen, bij clusteranalyse moet er worden bepaald tot welk segment elk individu behoort, terwijl bij latente klasse analyse voor elk individu een segmentspecifieke lidmaatschapskans wordt bepaald. Ambroise en Govaert (2000) tonen echter aan dat het heel gemakkelijk is om een variant van latente klasse analyse te construeren waarin de mate van fuzzyness kan worden gemanipuleerd. Als de fuzzyness parameter op 0 wordt gezet, verkrijgt men

een harde partitioning. De fuzzyness parameter kan echter niet alleen gebruikt worden ter verkrijging van minder fuzzy of zelfs een harde partitioning, maar ook van juist een meer fuzzy partitioning.

Ondanks de overeenkomsten tussen K-means clusteranalyse en latente klasse analyse, merken we op dat model gebaseerde clustering als belangrijke voordeel heeft dat men een statistisch model gemakkelijk kan aanpassen aan het type data dat men wil gebruiken. Zo kan men in plaats van een standaard latente klasse model een wat ingewikkelder model specificeren dat, bijvoorbeeld, geschikt is voor de analyse van data, die zijn verkregen via een keuze experiment (choice-based conjoint studies). Ook zijn allerlei verfijningen mogelijk zoals clustering in meerdere dimensies, modellen voor longitudinale data en de in dit artikel gepresenteerde modellen voor multi-niveau data. Bovendien kan er relatief gemakkelijk worden omgegaan met variabelen van verschillende meetniveaus. Nadeel van een statistisch model is dat modelaannames noodzakelijk zijn, zoals, bijvoorbeeld, de aanname van locale onafhankelijkheid. Over het algemeen kunnen deze echter worden getoetst en deels worden losgelaten. Bovendien maken ook andere methoden aannames, al zijn die vaak impliciet en niet toetsbaar.

3. LATENTE KLASSE ANALYSE

Er zijn veel varianten van latente klasse analyse en mixture modellen. Hieronder zullen we een aantal daarvan kort bespreken. We beginnen met het standaard ongerestricteerde model voor dichotome of polytome responsvariabelen - de gemeten variabelen in de latente klasse analyse. Daarna laten we zien hoe het model wordt aangepast voor de analyse van antwoorden verkregen via een keuze experiment. Tot slot bespreken we kort modellen voor continue responsvariabelen.

Startpunt is dat we voor iedere persoon de waarden kennen op een set responsvariabelen. De respons van individu i op variabele t duiden we aan met y_{it} en het aantal variabelen als T . De volledige responsvector voor individu i wordt aangeduid als \mathbf{y}_i . Een bepaalde latente klasse (segment) zal worden aangeduid als s en het totaal aantal klassen als S , waarbij $1 \leq s \leq S$. Het basis idee van latente klasse modellen is dat de kansverdeling van $\mathbf{y}_i - P(\mathbf{y}_i)$ - een gewogen soms is van S klasse specifieke kansverdelingen - $P(\mathbf{y}_i|s)$ (McLachlan and Peel, 2000; Vermunt en Magidson, 2002). Weging is gebaseerd op de proportie individuen in elke latente klasse. Dat wil zeggen dat:

$$P(\mathbf{y}_i) = \sum_{s=1}^S P(s)P(\mathbf{y}_i | s) \quad (1)$$

In het standaard latent klasse model wordt dit basis idee gecombineerd met de zogenaamde aanname van locale onafhankelijkheid (Goodman, 1974, Magidson en Vermunt, 2004). The T geobserveerde variabelen worden verondersteld onafhankelijk van elkaar te zijn binnen klassen, hetgeen als volgt wordt weergegeven:

$$P(\mathbf{y}_i | s) = \prod_{t=1}^T P(y_{it} | s) \quad (2)$$

De geschatte klasse-specifieke responskansen $P(y_{it}|s)$, voor de T indicatoren, geven aan hoe klassen van elkaar verschillen en kunnen worden gebruikt om de latente klassen te

labelen. Wanneer we de twee basisvergelijkingen (1) en (2) combineren krijgen we het volgende model voor $P(\mathbf{y}_i)$:

$$P(\mathbf{y}_i) = \sum_{s=1}^S P(s) \prod_{t=1}^T P(y_{it} | s) \quad (3)$$

De onbekende coëfficiënten in het model zijn de modelkansen $P(s)$ en $P(y_{it}|s)$. Deze worden over het algemeen geschat via de maximum likelihood methode. Bij modelselectie maakt men meestal gebruik van likelihood-ratio toetsen en informatiecriteria zoals AIC, BIC en CAIC. Zoals bij elke vorm van clusteranalyse betreft het belangrijkste modelselectie vraagstuk het aantal klassen of clusters dat men nodig heeft om de data in voldoende mate te beschrijven.

Wanneer latente klasse analyse gebruikt wordt als een clusteranalyse techniek zal men niet enkel de onbekende modelkansen willen schatten maar tevens personen willen toewijzen aan klassen. De kans dat persoon i , gegeven haar/zijn waardes op de responsvariabelen, behoort tot klasse s , $P(s|y_i)$, verkrijgen we met behulp van de Bayes formule. De meest gebruikte classificatieregels is modale toewijzing. Dit is toewijzen aan het segment waarvoor $P(s|y_i)$ het grootste is. Het zal duidelijk zijn dat in een goede clusteroplossing, een oplossing waarin men clusters goed van elkaar kan onderscheiden, de grootste kans voor de meeste respondenten dicht bij 1 ligt.

Latente klasse analyse is voor veel marketing doeleinden gebruikt. Bijvoorbeeld, het zogenaamde latente klasse keuzemodel wordt gebruikt voor de analyse van data verkregen via keuze experimenten. In keuze experimenten heeft elk van de responscategorieën (elk van de producten waar men uit kiest) bepaalde kenmerken (een bepaalde prijs, formaat, kwaliteit, kleur, etc.). Verondersteld wordt dat de kans dat een respondent een bepaald product kiest afhangt van de utiliteit van deze kenmerken: naarmate de utiliteit van een product in vergelijking tot die van de andere producten in de keuzeset hoger is, is de kans dat men dat product kiest groter. Deze utiliteiten kunnen worden geschat met behulp van een zogenaamd discrete-keuzemodel of conditional logit model dat is ontwikkeld door Nobelprijswinnaar McFadden (1974). Typisch voor de latente klasse analyse variant van dit type model is dat verondersteld wordt dat de populatie heterogeen is wat betreft de utiliteiten van productkenmerken: iedere klasse of segment heeft haar eigen set utiliteiten (Kamakura, Wedel en Agrawal, 1994). Een latente klasse choice-based conjoint model wordt verkregen door restricties op te leggen aan de responskansen $P(y_{it}|s)$, met behulp van een conditional logit model:

$$P(y_{it} = m | s) = \frac{\exp(\sum_{j=1}^J \beta_{js} z_{mjt})}{\sum_{m'=1}^{M_t} \exp(\sum_{j=1}^J \beta_{js} z_{m'tj})} \quad (4)$$

Hier is m een bepaalde responscategorie, j een bepaald productkenmerk, z_{mjt} de waarde van product m in keuzeset t voor kenmerk j en β_{js} de utiliteit van kenmerk j voor latente klasse s .

Behalve met discrete responsvariabelen kan latente klasse analyse ook worden gebruikt met continue variabelen, en zelfs met combinaties van discrete en continue responsvariabelen. In het geval van een set continue variabelen zal de kans $P(\mathbf{y}_i|s)$ in vergelijking (1) de vorm hebben van een multivariaat normale verdeling (McLachlan en Peel, 2000; Vermunt en Magidson, 2002, 2004). In het meest algemene model zal dit een volledig

ongerestricteerde verdeling betreffen, wat inhoudt dat gemiddelden, varianties en covarianties vrij geschat worden en ongelijk zijn tussen klassen. Een voorbeeld van een meer gerestricteerde variant is een model waarin de gemiddelden verschillen tussen klassen, maar de klassen dezelfde variantie-covariantie matrix hebben. Dit is een specificatie die ook wordt gebruikt in lineaire discriminant analyse. Een andere mogelijke restrictie is om, zoals in het standaard latente klasse model, te veronderstellen dat de variabelen onafhankelijk zijn binnen klassen, hetgeen bij continue indicatoren inhoudt dat hun covarianties op nul worden gezet. Een zeer gerestricteerde covariantie structuur die zeer lijkt op wat impliciet wordt verondersteld bij K-means clusteranalyse wordt verkregen door de covarianties op nul te zetten en de varianties gelijk te maken tussen klassen en tussen indicatoren (Magidson en Vermunt, 2001, 2004).

Andere voorbeelden van relevante uitbreidingen model zijn modellen met meerdere latente variabelen en modellen waarin klasse lidmaatschap wordt voorspeld met behulp van covariaten (zie Magidson en Vermunt, 2004 voor een overzicht). Hieronder bespreken we een relatief recente uitbreiding van het latente klasse model, die veelbelovend is voor marketingtoepassingen, namelijk het multi-niveau latente klasse model.

4. LATENTE KLASSE MODELLEN VOOR MULTI-NIVEAU DATA

Multi-niveau latente klasse modellen worden gebruikt in situaties waarin het onjuist is om te veronderstellen dat waarnemingen onafhankelijk van elkaar zijn. Voorbeelden zijn data verkregen via clustersteekproeven, data verkregen via huishoudpanels, en allerlei vormen van geneste data, zoals van consumenten behorende tot bepaalde regio's, postcodes of filialen van een supermarktketen.

Er zijn op zijn minst drie manieren om dit type afhankelijke waarnemingen te behandelen binnen de context van latente klasse modellen. Dat zijn:

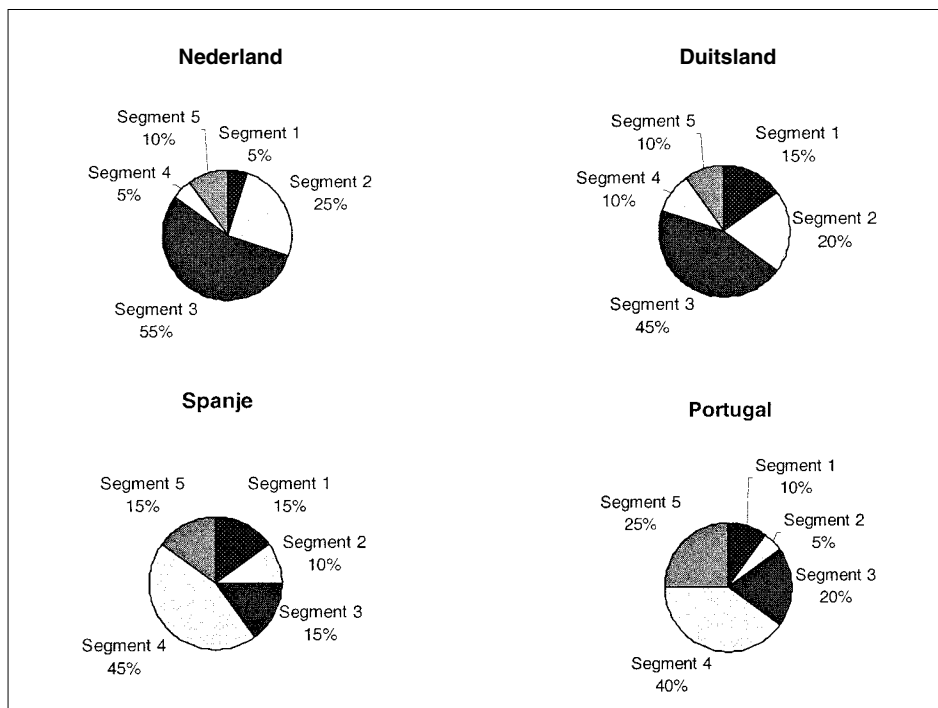
1. het gebruik van pseudo maximum likelihood schattingen met zogenaamde "survey design corrected" standaardfouten,
2. het gebruik van normaal verdeelde random effecten,
3. het gebruik van niet-parametrische random effecten, hetgeen neerkomt op het clusteren van hogere niveau eenheden, ofwel de multi-niveau variant van het latente klasse model dat we hier zullen gebruiken.

De eerste wijze van aanpak werd voorgesteld door Wedel, Ter Hofstede en Steenkamp (1998) en Vermunt (2001). Afhankelijkheid tussen waarnemingen wordt daarin als een statistisch probleem gezien en dat probleem wordt opgelost door zogenaamde "survey corrected standaardfouten" voor de modelparameters te berekenen. Deze zullen over het algemeen wat groter zijn dan de niet gecorrigeerde standaardfouten.

De tweede aanpak is sterk verwant aan hoe een normale multi-niveau regressieanalyse wordt uitgevoerd. Daarin wordt verondersteld dat de intercept (de constante) en mogelijk ook één of meerdere regressiecoëfficiënten variëren tussen groepen (Snijders en Bosker, 1999). Dit wordt gemodelleerd door aan te nemen dat deze parameters uit een multivariate normale verdeling komen. Vandaar dat in deze context ook wel de term random effect of random coëfficiënt wordt gebruikt. Vermunt (2003) stelde voor om hetzelfde principe toe te passen in de context van het latente klasse model. Dit kan door de omvang van de klassen te laten variëren tussen groepen. Daarbij wordt aangenomen dat (na een logistische transformatie) de klassenproporties normaal verdeeld zijn tussen groepen. In feite wordt een random-effects multinomiaal logistisch regres-

siemodel (Hedeker, 2003) geschat voor de klassenproporties.

In dit artikel kiezen we voor de derde benadering, die overigens sterk verwant is aan de tweede. Heterogeniteit in de klassenproporties tussen groepen wordt niet gemodelleerd als een continu verschijnsel, maar als een discreet fenomeen. Er wordt aangenomen dat groepen geclusterd kunnen worden in een beperkt aantal clusters (latente klassen), op basis van het vóórkomen van de verschillende latente klassen van respondenten in elke groep (Vermunt, 2003). Voordeel van deze aanpak is dat er inzicht wordt gegeven in de verbanden op verschillende niveaus. Als er, bijvoorbeeld, individuen in verschillende landen zijn geïnterviewd, dan levert het multi-niveau latente klasse model inzicht in de gelijkenis tussen individuen onderling en tussen de verschillende landen in de dataset. Dit soort informatie kan zeer relevant zijn voor marketing. Niet alleen wordt duidelijk welke klantsegmenten er per land zijn, maar ook worden landen op relevante kenmerken geclusterd. Stel dat de situatie in Figuur 1 opgaat met betrekking tot een segmentatie op productbezit. Er zijn vijf consumentsegmenten en in elk van de vijf segmenten vinden we consumenten met een bepaald productbezit patroon. Het voorkomen van deze segmenten, met bepaald productbezit patronen, kan per land verschillen. In Nederland vinden we, bijvoorbeeld, dat slechts 5% van de consumenten het productbezit patroon heeft dat segment 1 typeert. Verder is in dit land 25% van de consumenten ingedeeld in segment 2 en maar liefst 55% in segment 3. In Duitsland is een soortgelijke segmentatiestructuur gevonden. In een multi-niveau latente klasse analyse worden Nederland en Duitsland daarom waarschijnlijk in hetzelfde landencluster ingedeeld. In Spanje en Portugal is er echter sprake van een sterk afwijkende verdeling over segmenten: er zijn veel consumenten met productbezit zoals in segment 4. Deze landen zullen



Figuur 1. Fictief voorbeeld van segmentstructuren in verschillende landen.

daarom waarschijnlijk zijn in te delen in een ander landencluster dan Nederland en Duitsland, maar wel bij elkaar in hetzelfde landsegment. Een bedrijf kan op basis van deze clustering selecteren welke landen de juiste waarden hebben op relevante kenmerken voor, bijvoorbeeld, de introductie van een nieuw product.

Hoewel de meer technische aspecten van dit model – zoals het schatten van model parameters via maximum likelihood – niet eenvoudig zijn, is het conceptueel wel een relatief eenvoudig model. De volledige responsvector voor case i uit groep j is nu aangeduid als \mathbf{y}_{ij} . Het enige dat verder verandert ten opzichte van het basismodel beschreven in vergelijking (1) is dat de klassengrootten – $P(s)$ – nu afhangen van het groepssegment, r , waartoe groep j behoort. Voor de meer wiskundige lezer betekent dat:

$$P(\mathbf{y}_{ij} | r) = \sum_{s=1}^S P(s | r) P(\mathbf{y}_{ij} | s) \quad (5)$$

Verdere details over hoe dit model kan worden geschat en uitgebreid zijn te vinden in Vermunt (2003). Het multi-niveau latente klasse model is geïmplementeerd in het softwarepakket Latent GOLD 4.0 (Vermunt en Magidson, 2005), een pakket dat is toegespitst op de toepassing van een breed scala aan latente klasse modellen en dat draait onder MS-Windows.

5. EEN TOEPASSING BIJ HET SIMULTAAN SEGMENTEREN VAN LANDEN EN CONSUMENTEN

De toepassing die we willen gebruiken om de multi-niveau extensie van het latente klasse model te illustreren betreft een onderzoek naar het bezit van financiële producten in 15 Europese landen. De dataset die we gebruiken betreft de Eurobarometer 56.0, verzameld tussen 22 augustus en 27 september 2001 door een consortium van marktonderzoekbureaus in opdracht van de Europese Commissie, Directoraat-generaal Media en Communicatie, Opinie Enquêtes (Christensen, 2001). De Eurobarometer enquête omvat de populatie (15 jaar en ouder) van de 15 EU lidstaten in 2001. Er zijn 17 steekproef zones: Duitsland is verdeeld in Oost en West, en het Verenigd Koninkrijk in Groot-Brittannië en Noord-Ierland. Met de term “land” duiden we deze steekproef zones aan.

De steekproefomvang is ongeveer 1000 per land, met uitzondering van Luxemburg (600) en Noord-Ierland (300). De totale steekproefomvang is 16,200. We gebruiken gewichten uit de Eurobarometer database zodat elke nationale steekproef representatief is met betrekking tot demografische variabelen en er wordt gecorrigeerd voor de verschillende steekproef ratio's tussen landen (zie Tabel 1).

De informatie over het bezit van financiële producten leiden we af uit vragen 25 tot met 28 van de Eurobarometer 56.0 survey (Christensen, 2001). Het betreft het bezit van een lopende rekening, een spaarrekening, een creditkaart, een andere bankpas, een chequeboek, rood staan, een hypotheek, en andere leningen. Het gaat om het bezit van individuele consumenten. Tabel 1 geeft beschrijvende informatie met betrekking tot het bezit van deze 8 producten in de 17 landen.

De theoretische motivatie van de hier gepresenteerde analyse wordt gedetailleerd beschreven in Bijmolt, Paas en Vermunt (2004). Hoewel we het voorbeeld hier gebruiken om het multi-niveau latente klasse model te illustreren, is het van belang om de

Tabel 1. Beschrijvende Statistieken voor de Internationale Steekproef.

Land	Bezit van Financiële Producten (Steekproefproporties)									
	Steekproef- omvang	Gemiddeld gewicht	Lopende rekening	Spaar- rekening	Credit- kaart	Bankpas	Cheque- boek	Rood staan	Hypo- theek	Lening
Oostenrijk	1093	.34	71.5	82.3	33.7	61.0	21.6	41.4	17.7	21.8
België	1031	.43	85.2	85.9	39.2	74.0	34.2	33.3	25.9	21.1
Denemarken	1001	.22	78.5	63.2	48.2	60.8	33.9	55.2	51.8	36.3
Finland	1023	.21	87.5	50.2	31.5	84.7	0.7	16.0	22.0	26.5
Frankrijk	1002	2.42	87.8	69.8	57.7	31.0	87.9	50.6	18.6	26.6
Oost Duitsland	1024	.67	91.8	76.1	22.5	81.2	41.5	35.7	13.0	23.0
West Duitsland	1023	2.87	89.5	84.2	29.9	78.0	40.9	39.8	16.8	16.5
Groot-Brittannië	1041	2.37	75.2	77.1	52.3	58.7	76.4	29.3	37.1	20.1
Griekenland	1001	.45	11.0	79.7	18.7	25.9	6.3	3.4	14.6	11.6
Ierland	1002	.15	51.4	71.7	32.3	40.3	45.1	16.2	25.7	26.6
Italië	998	2.52	65.6	19.4	36.3	51.3	62.7	10.0	12.3	12.8
Luxemburg	609	.03	84.7	81.8	65.0	69.6	49.6	50.9	29.7	30.0
Nederland	1047	.65	89.5	82.5	37.2	94.3	26.6	63.6	33.6	14.9
Noord Ierland	305	.22	62.3	59.7	42.3	41.3	62.3	21.3	35.7	14.1
Portugal	1000	.42	70.0	44.2	33.0	33.0	60.8	2.5	13.0	8.0
Spanje	1000	1.70	61.6	67.2	52.1	33.2	17.4	8.2	19.4	17.2
Zweden	1000	.37	76.1	77.5	57.0	59.4	19.7	19.0	34.7	25.8

twee elementen van de theoretische motivatie te benadrukken. Ten eerste past de studie in de traditie van onderzoek naar acquisitiepatronen (Kamakura, Ramaswami en Srivastava, 1991; Paas, 2003; Paas en Molenaar, 2005). De vraag is hierbij of producten in een bepaalde volgorde worden aangeschaft, ofwel of er consumentsegmenten bestaan die verschillen in productbezit en die kunnen worden beschouwd als personen in verschillende levensfasen. Het tweede element in deze studie is het segmenteren van landen (Steenkamp en Ter Hofstede, 2002). In plaats van landen te segmenteren op basis van macro gegevens op landniveau gebruiken we hier micro data, ofwel de informatie over het productbezit van consumenten binnen landen. Om precies te zijn, we gaan na of er landclusters te vinden zijn die vergelijkbaar zijn wat betreft de mate van vóórkomen van de gevonden consumentclusters.

De eerste stap in de analyse is het vaststellen van het aantal consument en landsegmenten dat noodzakelijk is om de gegevens in voldoende mate te beschrijven. Dat doen we met hulp van CAIC (Consistent Akaike Information Criterion). We hebben het model geschat met 1 t/m 15 consumentsegmenten en 1 t/m 8 landsegmenten. Het model met 14 clusters van consumenten en 7 van landen heeft de laagste CAIC waarde en is daarom het beste model.

Tabel 2. Model Resultaten: Landsegmenten.

Landsegment	Relatieve Omvang	A Posteriori Kansen voor Landsegment Lidmaatschap $\{P(r y_j)\}^*$	
		Land	Kans
1	.256	België, Oost Duitsland, West Duitsland, Nederland	1.00
		Luxemburg	.81
2	.260	Oostenrijk, Denemarken, Finland, Zweden	1.00
		Luxemburg	.19
3	.175	Groot-Brittannië, Ierland, Noord Ierland	1.00
4	.119	Italië, Portugal	1.00
5	.064	Spanje	1.00
6	.064	Griekenland	1.00
7	.064	Frankrijk	1.00

* Alle niet gerapporteerde a posteriori kansen zijn kleiner dan 0.01.

Tabel 2 laat a posteriori klasse lidmaatschapkansen zien voor de 17 landen in onze studie. De classificatie van landen in segmenten valt sterk samen met de kaart van Europa. Het eerste cluster bevat de lage landen (België and Nederland) en Duitsland. Het tweede cluster bevat de Scandinavische landen en Oostenrijk. Luxemburg blijkt zowel kenmerken van cluster 1 als cluster 2 te hebben. Groot-Brittannië, Ierland, en Noord-Ierland vormen cluster 3. In deze drie landen lijken de segmentstructuren sterk op elkaar. In tegenstelling tot de andere delen van Europa bestaat Zuid-Europa uit meerdere kleine segmenten: Italië en Portugal vormen samen cluster 4 en Spanje, Griekenland en Frankrijk vormen ieder een eigen cluster. In elk van deze landen komt een unieke segmentstructuur voor.

Het bovenste deel van Tabel 3 geeft informatie over het productbezit binnen elk van de 14 consumentsegmenten. Om de interpretatie te vereenvoudigen zijn de clusters zodanig geordend dat de geaggregeerde penetratie over alle producten oploopt. In segment 1 bezitten consumenten de minste producten en consumenten in segment 14 de meeste. Opvallend is de zeer lage penetratiegraad in de eerste drie segmenten van zowel een

Tabel 3. Model Resultaten: Consumentsegmenten.

	Consumentsegmenten:													
	1	2	3	4	5	6	7	8	9	10	11	12	13	14
Financiële producten:														
Lopende Rekening	.05	.05	.33	1.00	1.00	.92	.98	1.00	1.00	.88	1.00	.99	.98	1.00
Spaarrekening	.39	.85	.61	.14	.65	.61	.91	.91	.15	.71	.78	.73	.93	.86
Credietkaart	.00	.09	.00	.29	.84	.23	.16	.62	.76	.81	.96	.49	.51	.83
Bankpas	.00	.19	.00	.62	.32	.80	.85	.53	.90	.69	.00	.99	.95	.87
Chequeboek	.09	.01	.86	.79	.13	.02	.29	1.00	.95	.30	.98	1.00	.50	1.00
Rood staan	.00	.03	.31	.02	.04	.42	.19	.11	.22	.29	.65	.68	.68	.60
Hypotheek	.01	.11	.25	.01	.14	.06	.00	.03	.29	.60	.22	.06	.55	.89
Lening	.02	.08	.21	.01	.14	.31	.00	.02	.27	.47	.33	.33	.30	.39
Landsegmenten:														
	Relatieve Omvang van Consumentsegmenten $\{P_{(s r)}\}$													
1	.07	.03	.01	.00	.00	.17	.39	.03	.00	.00	.00	.08	.22	.01
2	.05	.17	.04	.00	.06	.26	.11	.01	.00	.18	.00	.01	.11	.01
3	.15	.11	.09	.03	.01	.00	.04	.26	.00	.02	.00	.07	.00	.23
4	.37	.00	.06	.29	.02	.00	.00	.01	.26	.00	.00	.00	.00	.00
5	.16	.27	.01	.01	.34	.00	.02	.03	.01	.15	.01	.00	.00	.01
6	.13	.79	.01	.00	.00	.00	.01	.02	.00	.04	.00	.00	.00	.00
7	.06	.00	.11	.07	.00	.02	.02	.14	.00	.00	.41	.14	.00	.04

basis betaalmiddel (lopende rekening) als van meer geavanceerde betaalmiddelen (creditkaart of andere bankpas). Segment 1 heeft een lage kans voor alle producten, terwijl segment 2 enkel een vrij hoge penetratiegraad heeft voor spaarrekening en segment 3 voor spaarrekening en chequeboek. Consumenten in segmenten 4 tot en met 9 hebben penetratiegraden van bijna 1 voor lopende rekening en enkele andere betaalmiddelen, waarbij het type betaalmiddel verschilt tussen deze segmenten. Bijvoorbeeld, creditkaart bezit is heel hoog in segment 5 terwijl chequeboek bezit juist meer voorkomt in segmenten 8 and 9. Segmenten 10 tot en met 14 kenmerken zich door een hoge kans op het bezitten van minstens één kredietproduct (rood staan, hypotheek of andere lening). Segmenten 13 and 14 bevatten de meest intensieve gebruikers met relatief hoge penetratiegraden voor alle acht financiële producten, waarbij vooral de hoge waarde voor hypotheek in segment 14 opvalt.

De resultaten in het onderste gedeelte van Tabel 3 koppelen de landsegmenten aan de consumentsegmenten. Het betreft informatie over de relatieve omvang van de 14 consumentclusters binnen ieder van de 7 landclusters. Veertien consumentsegmenten lijkt veel, maar veel segmenten bevatten maar 1% of zelfs maar 0.1% van de consumenten in de meeste landsegmenten. Dit betekent dat het werkelijke aantal consumentsegmenten in een bepaald landsegment steeds beduidend kleiner is dan 14. Bijna alle consumentsegmenten, vooral degene met hoge productpenetraties (3 t/m 14), zijn groot (zeg groter dan 10%) in maar één of twee landsegmenten. Anderzijds zijn consumenten segmenten 1 and 2, met een lage penetratie voor alle producten, vrij groot in vier landsegmenten.

Voor elk landsegment zien we een paar veel voorkomende consumentsegmenten, een aantal met lage penetratiegraden (consumentsegmenten 1 t/m 3), een aantal met midden tot hoge graden maar lage graden voor kredietproducten (consumentsegmenten 4 t/m 9) en een aantal met hoge graden voor de meeste producten (segmenten 10 t/m 14). Landsegment 2 (Oostenrijk en de Scandinavische landen) bevat bijvoorbeeld consumenten in segmenten 2, 6, 7, 10 en 13, die allemaal een lage penetratiegraad hebben voor chequeboek. De zeer diverse set van consumentsegmenten die in landsegment 3 (Ierland, Noord-Ierland, Groot-Brittannië) veel voorkomen gaat van segmenten met zeer lage kansen voor alle producten (consumentsegment 1) tot en met segmenten met extreem hoge kansen voor alle producten (consumentsegment 14). Consumentsegment 2, met lage penetratiegraden voor alle producten behalve voor de spaarrekening, is extreem dominant in Griekenland (landsegment 6). Frankrijk (landsegment 7) bevat vrijwel alleen consumentsegmenten 2, 8, 11, and 12, allen met een hoge kans op bezit van een chequeboek.

6. MARKETING IMPLICATIES VAN DE RESULTATEN

Onze empirische resultaten zijn zeer relevant voor internationaal opererende financiële instellingen. We zien dat, tegen verwachtingen in (Berger, Dai, Ongena en Smith, 2003; Ganesh, 1998), er in Europa nog steeds sprake blijkt te zijn van grote verschillen tussen landen en consumenten. Er zijn wel een aantal landen in hetzelfde landsegment in te delen, bijvoorbeeld, Duitsland, Nederland, Luxemburg en België, maar dit is wellicht een gevolg van culturele overeenkomsten van voor het EU-tijdperk. Voor financiële dienstverleners suggereren deze bevindingen dat Europa nog niet als een geheel kan worden benaderd. Verder is ook duidelijk geworden welke groepen landen moge-

lijk hetzelfde kunnen worden benaderd, bijvoorbeeld, Duitsland, Nederland, Luxemburg en België, en welke specifieke landen niet, Frankrijk, Griekenland en Spanje. Ook voor het targetten van een enkel product zijn onze bevindingen zeer relevant. Bijvoorbeeld, een creditcard bedrijf kan de bevindingen die zijn gerapporteerd in dit artikel gebruiken om landsegmenten te selecteren voor de introductie van hun product. Vervolgens kan het bedrijf ook bepalen welke segmenten binnen elk land voor benadering in aanmerking komen. Als het bedrijf een nieuwe creditcard introduceert, waarmee men bestaande gebruikers van creditcards hoopt aan te spreken, dan kan de financiële dienstverlener consumentensegment 5 in landensegment 5 (Spanje) benaderen. Andere segmenten met veel bezitters van creditcards zijn consumentensegment 11, dat veel voorkomt in landensegment 7 (Frankrijk) en consumentensegment 14, dat veel is te vinden in landensegment 3 (Groot Brittannië, Ierland en Noord-Ierland).

7. DISCUSSIE

In dit artikel hebben we aandacht besteed aan een voor marktonderzoek zeer interessante uitbreiding van het latente klasse model. Deze uitbreiding maakt het mogelijk om een simultane clusteranalyse uit te voeren van eenheden op een lager en op een hoger niveau. Het model werd geïllustreerd met behulp van een voorbeeld waarin consumenten werden geclusterd op basis van het bezit van financiële producten en landen op basis van de omvang van de verschillende consumentensegmenten.

Inhoudelijk gezien is dit een goed gekozen voorbeeld omdat het gemakkelijk te volgen is en de marketingrelevantie evident is. Aan de andere kant zijn er wellicht nog betere voorbeelden denkbaar. In ons voorbeeld was het aantal hogere niveau eenheden (landen) vrij klein terwijl multi-niveau analyse nog meer voordelen biedt in situaties waarin het aantal groepen groot is (50 of meer). Ook is in ons voorbeeld het aantal personen per groep groot, meestal ongeveer 1000 consumenten per land, terwijl tussen de 5 à 30 personen per groep de meer standaard situatie is bij een multiniveau analyse. Er zijn waarschijnlijk verschillende marktonderzoektoepassingen die nog beter passen bij de meer standaard situatie waarin veel meer groepen onderzocht worden die ieder bestaan uit veel minder onderzochte consumenten.

LITERATUUR

- Ambrose, C, en G. Govaert, 2000. 'Clustering by maximizing a fuzzy classification maximum likelihood criterion'. In: W. Jansen and J.G. Bethlehem (eds.), *Proceedings in Computational Statistics 2000*. Heidelberg: Physica-Verlag, pp. 187-192.
- Berger, A.N., Q. Dai, S. Ongena, en D.C. Smith, 2003. To what extent will the banking industry be globalized? A study of bank nationality and reach in 20 European nations. *Journal of Banking & Finance*, 27, 383-415.
- Bijmolt, T.H.A., Paas, L.J., en J.K. Vermunt, 2004. Country and consumer segmentation: Multi-level latent class analysis of financial product ownership. *International Journal of Research in Marketing*, 21, 323-340.
- Christensen, T., 2001. Eurobarometer56.0: Information and communication technologies, financial services, and cultural activities, August-September 2001. Brussels: European Opinion Research Group.

- Ganesh, J., 1998. Converging trends within the European Union: Insights from an analysis of diffusion patterns. *Journal of International Marketing*, 6, 32-48.
- Goodman, L.A., 1974. The analysis of systems of qualitative variables when some of the variables are unobservable: Part I - A modified latent structure approach. *American Journal of Sociology*, 79, 1179-1259.
- Hedeker, D., 2003. A mixed-effects multinomial logistic regression model. *Statistics in Medicine*, 22, 1433-1446.
- Kamakura, W.A., S.N. Ramaswami en R.K. Srivastava, 1991. Applying latent trait analysis in the evaluation of prospects for cross-selling of financial services. *International Journal of Research in Marketing*, 8, 329-349.
- Kamakura, W.A., Wedel, M. en J. Agrawal, 1994. Concomitant variable latent class models for the external analysis of choice data. *International Journal of Research in Marketing*, 11, 451-464.
- Magidson, J. en J.K. Vermunt, 2002. Latent class modeling as a probabilistic extension of K-means clustering. *Quirk's Marketing Research Review*, March 2002, 20 & 77-80.
- Magidson, J. en J.K. Vermunt, 2004. 'Latent class analysis'. In: D. Kaplan (ed.), *The Sage Handbook of Quantitative Methodology for the Social Sciences*. Thousand Oakes: Sage Publications, pp. 175-198.
- McFadden, D., 1974. 'Conditional logit analysis of qualitative choice behavior'. In: I. Zarembka (Ed.), *Frontiers in econometrics*. New York: Academic Press, pp. 105-142.
- McLachlan, G.J., en D. Peel, 2000. *Finite Mixture Models*. New York: John Wiley & Sons, Inc.
- Paas, L.J. (2003). 'De toepassing van het Mekkenschaalmodel voor acquisitiepatroonanalyse in de financiële dienstverlening'. In: A.E. Bronner, P. Dekker, J.C. Hoekstra, E. de Leeuw, Th.B.C. Poiesz, K. de Ruyter en A. Smidts (Eds.), *Jaarboek Marktonderzoeksassociatie, Haarlem: Vrieseborch*, pp. 121-137.
- Paas, L.J. en I.W. Molenaar, 2005. Analysis of acquisition patterns: A theoretical and empirical evaluation of alternative methods. *International Journal of Research in Marketing*, 22, 87-100.
- Snijders, T.A.B. en R.J. Bosker, 1999. *Multilevel Analysis*. London: Sage Publications.
- Steenkamp, J.-B.E.M., en F. Ter Hofstede, 2002. International market segmentation: Issues and perspectives. *International Journal of Research in Marketing*, 19, 185-213.
- Vermunt, J.K., 2002. Comments on "Latent class analysis of complex sample survey data". *Journal of the American Statistical Association*, 97, 736-737.
- Vermunt, J.K., 2003. Multilevel latent class models. *Sociological Methodology*, 33, 213-239.
- Vermunt, J.K. en J. Magidson, 2002. 'Latent class cluster analysis'. In: J. Hagenaars en A. McCutcheon (eds.), *Applied latent class models*. Cambridge: Cambridge University Press, pp. 89-106.
- Vermunt, J.K. en J. Magidson, 2005. *Latent GOLD User's Manual*. Boston: Statistical Innovations Inc.
- Wedel, M. en W.A. Kamakura, 2000. *Market segmentation: Conceptual and methodological foundations (second edition)*. Dordrecht: Kluwer.
- Wedel, M., F. Ter Hofstede en J.-B.E.M. Steenkamp, 1998. Mixture model analysis of complex samples. *Journal of Classification*, 15, 225-244.

11. *Imputeren in plaats van amputeren: De voordelen van meervoudige imputatie*

J.A. VOERMAN

SAMENVATTING

Wat doe je als onderzoeker als de te verklaren variabele uit je model voor een groot deel niet ingevuld blijkt te zijn in de enquête? De meest radicale oplossing is het weglaten van de respondenten die de enquête niet volledig hebben ingevuld (*listwise deletion*). Probleem daarbij is dat het aantal bruikbare cases drastisch kan verminderen. Zo ook in het hieronder beschreven onderzoek naar exportsucces van Europese Midden- en Kleinbedrijven. Veel respondenten weigerden de exportverkoppen in te vullen, de afhankelijke variabele in dit geval.

In dit artikel wordt ingegaan op de mogelijkheid om missende waarden *meervoudig* te imputeren op basis van Bayesiaanse technieken en op de voordelen (en nadelen) van deze methode ten opzichte van andere imputatietechnieken. Dit wordt geïllustreerd met behulp van een studie naar het effect van de attitude van de entrepreneur en het informatiegedrag op de exportintensiteit van MKB-bedrijven. Het blijkt dat meervoudig imputeren tot betere conclusies kan leiden ten aanzien van de significantie, richting en grootte van de te schatten parameters.

1. INLEIDING

Bij de uitvoering van kwantitatief onderzoek middels enquêtes is non-respons vaak een probleem. Niet alleen *survey non-respons*, maar ook *item non-response* kan leiden tot een suboptimale dataset. Het niet willen participeren in het onderzoek (*survey non-response*) is op te lossen door vooraf goed na te denken over het onderzoeksontwerp en bijvoorbeeld gebruik te maken van *incentives* (e.g. Malhotra, 2004). Bij het niet invullen van één of meerdere vragen in de enquête (*item non-response*) staat de onderzoeker voor de vraag hoe met niet-volledige vragenlijsten om te gaan. Immers, missende waarden resulteren in minder efficiënte schattingen en het niet kunnen gebruiken van analysemethoden die uitgaan van complete data (e.g., Rubin, 1987). Er zijn meerdere mogelijkheden voor het hanteren van incomplete datasets, van het simpel weglaten van incomplete enquêtes (*listwise deletion*) tot het toepassen van statistische imputatietechnieken om de missende waarden toch een ‘gezicht’ te geven en zo de dataset volledig te maken. In dit artikel wordt gekozen voor het meervoudig imputeren van missende waarden via Bayesiaanse technieken, gebruik makend van het programma NORM (Schafer, 1999). De dataset die wordt geïmputeerd is afkomstig van de Interstratosⁱ studie naar exportsucces van Europese MKB-bedrijven. De missende waarden in deze dataset worden drie keer geïmputeerd, om zo het effect van exportmarktorientatie en de attitude van de entrepreneur op exportsucces op een verantwoorde manier te kunnen schatten.

In paragraaf 2 wordt verder ingaan op meervoudige imputatie van missende waarden versus het enkelvoudig imputeren. Tevens komt het programma NORM aan bod. In paragraaf 3 wordt de dataset besproken waarop de meervoudige imputatie wordt toegepast. Vervolgens wordt het conceptuele model besproken (paragraaf 4), evenals de operationalisatie van dit model met behulp van de geïmputeerde variabelen in de dataset (paragraaf 5). In paragraaf 6 en 7 worden vervolgens de resultaten van de drie geschatte modellen en de conclusies behandeld.

2. MEERVOUDIGE IMPUTATIE VAN MISSENDE WAARDEN

2.1. Enkelvoudige en meervoudige imputatie

De meest simpele manier om item non-respons in enquêtes op te lossen is alleen de rijen in de dataset te gebruiken met volledige informatie; *listwise deletion*. Nadeel is dat dit vaak resulteert in een drastische vermindering van het aantal cases. In het in deze studie gebruikte Interstratos onderzoek bijvoorbeeld, hebben slechts 250 van de 1122 respondenten informatie gegeven over zowel export- en totale verkopen, de afhankelijke variabelen. Nog geen kwart van het oorspronkelijke aantal respondenten is dus bruikbaar bij *listwise deletion*. In plaats van het weggooien van cases, is er de mogelijkheid de missende waarden enkelvoudig dan wel meervoudig te imputeren. Een populaire enkelvoudige imputatiemethode is het vervangen van de missende waarden met de gemiddelde waarde van de geobserveerde waarden (Rubin, 1987). Deze methode is simpel toe te passen, maar kent ook nadelen. De variantie wordt namelijk onderschat en er bestaat de kans dat een mogelijke non-respons fout een vertekend beeld geeft van de relaties die er bestaan tussen de variabelen in de dataset. Een andere manier is het gebruiken van de Hotdeck-procedure waarbij voor iedere non-responder een vergelijkbare respondent wordt gezocht, gebaseerd op variabelen die voor beiden geobserveerd zijn (Rubin, 1987). Helaas onderschat ook deze methode de variabiliteit in de dataset door ervan uit te gaan dat elke missende waarde met zekerheid geïmputeerd wordt.

In tegenstelling tot enkelvoudige imputatie is het bij meervoudige imputatie (MI) wel mogelijk de aanwezige informatie te gebruiken om de onzekerheid omtrent de te imputeren waarden weer te geven, door elke missende waarde te vervangen door twee of meer geïmputeerde waarden (Rubin, 1987). Daarnaast kunnen valide conclusies worden getrokken uit het combineren van de analyseresultaten, als de meervoudige imputaties voortkomen uit herhaaldelijke aselechte trekkingen. Immers, de imputatie leidt tot meerdere uitwisselbare complete datasets waarop alle mogelijke analyses kunnen worden gedaan die uitgaan van complete data. Een ander voordeel is dat MI de imputatiefase scheidt van de analysefase, wat met name voordelig is als de variabelen die gebruikt worden voor het schatten van de missende waarden niet dezelfde zijn als de variabelen die gebruikt worden om het uiteindelijke model te schatten. Zo is het mogelijk in de imputatiefase variabelen op te nemen, die later niet worden meegenomen in de analysefase, maar wel van belang kunnen zijn voor het schatten van de missende waarden (Schafer, 2003). MI is gebaseerd op de Bayesiaanse theorie. In de oorspronkelijke dataset zitten de zogenoemde incomplete Y-variabelen waar missende waarden aanwezig zijn en X-variabelen die volledig geobserveerd zijn. De *posterior* verdeling van de missende waarden (Y_{mis}) wordt gesimuleerd onder een expliciet Bayesiaans

model. De informatie (covariantiematrix, gemiddelden en varianties) in de volledig geobserveerde variabelen (de X -variabelen) en de geobserveerde waarden in de te imputeren variabelen (Y_{obs}) wordt gebruikt om Y_{mis} m keer te schatten ($Y_{\text{mis}}|X, Y_{\text{obs}}$). Nadelen zijn de imputatie- en analysetijd die nodig is en de opslagruimte voor meerdere datasets.

In meerdere studies wordt MI vergeleken met andere imputatiemethoden. Zo bewijzen Rubin en Schenker (1986) dat MI superieur is ten opzichte van enkelvoudige imputatie. Deze conclusie wordt ook getrokken door Noh, Kwak en Han (2004) als zij MI vergelijken met (1) het niet meenemen van onvolledige cases en (2) het imputeren van het gemiddelde voor de missende waarden. Olinsky, Chen en Harlow (2003) vergelijken MI met *expectation maximization (EM)*, *full information maximum likelihood (FIML)*, imputatie van het gemiddelde en regressie imputatie in een situatie waarbij het uiteindelijke model wordt geschat middels *structural equations modeling* (een methode die ik ook toepas in deze studie). MI komt uit de test als een “excellente” schatter, mede gezien het feit dat MI een complete dataset genereert (in tegenstelling tot FIML, een directe methode voor het schatten van een model met missende waarden). Gezien het voorgaande wordt gekozen voor MI als methode voor het omgaan met missende waarden in de Interstratos dataset.

2.2. NORM – de methode

Gebaseerd op bovenstaande theorie heeft Schafer (1997) het pakket NORM ontwikkeld voor het meervoudig imputeren onder de aanname van multivariate normaliteit. NORM bestaat uit twee fasen, de EM-fase en de DA-fase (*data augmentation* fase). In de EM-fase wordt met behulp van het EM-algoritme een efficiënte schatting van de gemiddelden, varianties en covarianties gemaakt op basis van alle cases (ook de niet complete) in de dataset. Vervolgens bestaat de DA-fase uit het simuleren van aselechte waarden van de parameters en de missende waarden uit hun *posterior* verdeling in twee stappen, de I-stap en de P-stap. In de I-stap worden de missende data geïmputeerd door ze uit hun conditionele verdeling gegeven de geobserveerde waarden en veronderstelde waarden van de parameters te trekken. Daarna worden in de P-stap nieuwe waarden voor de parameters en missende waarden gesimuleerd uit een Bayesiaanse *posterior* verdeling gegeven de geobserveerde data en de meest recent geïmputeerde waarden voor de missende data. De EM-fase levert de startwaarden voor de DA-fase, waarin datavermeerdering plaatsvindt op iteratieve wijze, middels het opslaan van een aantal I-stappen. Bovendien geeft de EM-fase een indicatie van het aantal iteraties (k) dat nodig is om tot convergentie te komen. Hoe meer informatie mist in de dataset, hoe meer iteraties nodig zijn voor convergentie. In de DA-fase wordt dan voor n complete datasets het aantal iteraties op $n*k$ gezet, waarbij elke k -de iteratie de geïmputeerde waarden worden opgeslagen. Op deze manier wordt de missende data meerdere malen correct geïmputeerd resulterend in een n aantal volledige datasets. Voor het kunnen toepassen van NORM moet wel vastgesteld worden of de variabelen normaal verdeeld zijn. Ook veronderstelt NORM dat de missende waarden *missing at random (MAR)* zijn (Schafer 1997).

3. CONCEPTUEEL MODEL

In het empirische gedeelte van dit artikel probeer ik de relatie tussen informatieverza-

meling als onderdeel van exportmarktoriëntatie en exportsucces van MKB-bedrijven vast te stellen, waarbij ook de risicohouding van de entrepreneur wordt meegenomen.

3.1. Exportmarktoriëntatie en informatieverzameling

De eerste stap tot een marktgerichte organisatie bestaat uit het genereren van markt-informatie (Kohli en Jaworski 1990; Narver en Slater 1990; Jaworski en Kohli 1993). Ook in een internationale context blijkt dit een cruciale stap (Cadogan en Diamantopoulos 1995). Exportmarkt-informatie kan worden verkregen door een verscheidenheid aan bronnen te raadplegen. Zo maken Johanson en Vahlne (1977) een onderscheid tussen objectieve informatie, afkomstig uit secundaire of primaire bronnen, en 'ervarings-informatie', persoonlijk verworven door direct contact met de markt of de consument. Souchon en Diamantopoulos (1996, 1999) classificeren de mogelijke informatiebronnen als *export market research* (formele, systematische en objectieve informatieverzameling), *export assistance* (overheidssteun en -advies), en *export market intelligence* (informele informatieverzameling, verkregen tijdens het dagelijkse zakendoen). Zij stellen vast dat MKB-bedrijven *export market research* vaak te duur vinden en daardoor voornamelijk vertrouwen op *export market intelligence*.

Uit een overzicht van veertig cross-sectionele exportsuccesstudies blijkt duidelijk dat exportprestatie mede bepaald wordt door informatiegedrag, waarbij het verzamelen en gebruiken van informatie een positief effect heeft (Voerman, 2004). Bovendien verbetert het gebruik van het meer formele *export market research* en het verkrijgen van het informele *export market intelligence* exportsucces, terwijl de resultaten aangaande *export assistance* tegenstrijdig zijn. Een verklaring hiervoor is dat gedurende het proces van internationalisatie, exportinformatie van overheidsbronnen te algemeen is om van nut te zijn. Gezien bovenstaande kan de volgende hypothese worden geformuleerd:

H1: De hoeveelheid exportmarkt-informatie die een MKB-onderneming verzamelt, heeft een positief effect op de exportprestatie.

3.2. Risicohouding, exportmarkt-informatie en exportsucces

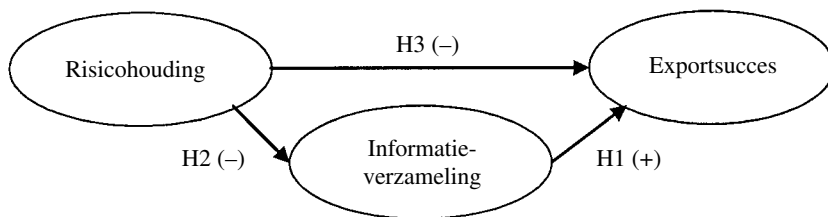
Interessant bij MKB-studies is de speciale plaats die de eigenaar (of entrepreneur) inneemt. Zo vindt Voerman (2004) dat in MKB-bedrijven weinig tot geen gebruik wordt gemaakt van informatiebronnen voor het verbeteren of opzetten van de exportactiviteiten, hoewel het verzamelen van informatie theoretisch gezien misschien een juiste keuze is. In deze studie zegt slechts de helft van de ondervraagde ondernemers buitenlandse informatiebronnen te hebben geconsulteerd, terwijl een kwart zelfs aangeeft in het geheel geen binnenlandse informatiebronnen te consulteren ten behoeve van de exportactiviteiten. De vraag is of de persoonlijkheid van de entrepreneur hierbij misschien een rol speelt. Met name de risicohouding blijkt een karakteriserende eigenschap van een MKB-ondernemer te zijn (Ahmed, 1985; Kuratko en Hodgetts, 2001). Kohli en Jaworski (1990) stellen ook duidelijk dat senior managementfactoren marktgerichtheid kunnen bevorderen, dan wel belemmeren: "kenmerken van de top, zoals opleiding of houding tegenover verandering of risico, zijn van groot belang". In een latere studie stellen Jaworski en Kohli (1993) dan ook de hypothese op dat een risico-avers topmanagement ertoe leidt dat medewerkers minder marktkennis genereren uit angst voor een mislukking. Uit de resultaten van Avlonitis en Gounaris (1999) blijkt

ook dat de mate van risicoaversie negatief gerelateerd is aan de ontwikkeling van marktgerichtheid. In een internationale context toetsen Zaheer en Zaheer (1997) de invloed van onzekerheidsvermijding op informatiegedrag en vinden inderdaad dat bedrijven uit onzekerheidsmijdende landen minder informatie verzamelen. Daarnaast blijkt uit de exportliteratuur dat een risicomijdend management een lagere exportprestatie behaalt (zie bijvoorbeeld Holzmüller en Stöttinger 1996). De volgende hypothesen resulteren:

H2: In MKB-bedrijven met een risicomijdende ondernemer wordt minder exportmarkt-informatie verzameld.

H3: In MKB-bedrijven met een risicomijdende ondernemer is de exportprestatie lager.

In figuur 1 staat het in de analysefase te schatten model, gebaseerd op de hierboven beschreven hypothesen.



Figuur 1. Het conceptuele model.

4. DATASET, IMPUTATIE EN OPERATIONALISERING

4.1. Dataset

Het empirische gedeelte van dit artikel is gebaseerd op een dataset die verzameld is door de internationale onderzoeksgroep Interstratos. Deze onderzoeksgroep heeft als doel de internationalisatie van industriële MKB-bedrijven in Europa in kaart te brengen. In dit geval wordt een MKB-bedrijf gedefinieerd als een onafhankelijk bedrijf met minder dan 500 werknemers. Met behulp van een schriftelijke enquête zijn in 1995 gegevens verzameld van bedrijven afkomstig uit zeven Europese landen (Oostenrijk, België, Nederland, Zwitserland, Noorwegen, Zweden en Finland) en vijf industriële sectorenⁱⁱ. De steekproef bestaat dan uit 1122 industriële, Europese MKB-bedrijven. De in deze studie gebruikte variabelen zijn slechts een klein onderdeel van de totale beschikbare verzameling variabelen. Met de totale verzameling van variabelen is door Voerman (2004) een integraal model van exportsucces geschat. Voor deze exercitie, waarbij de nadruk ligt op het meervoudig imputeren van missende waarden, is een onderdeel van het totale model geschat.

4.2. Imputatie middels NORM

Gezien het grote aantal missende waarden in de afhankelijke variabele 'exportverkoop' (872 van de 1122) en het feit dat ook andere variabelen missende waarden bevat-

ten, zou de oorspronkelijke dataset van 1122 respondenten drastisch slinken bij het toepassen van *listwise deletion*. Om toch de informatie in alle 1122 cases te gebruiken wordt de dataset drie keer geïmputeerd met behulp van het programma NORM. Vervolgens wordt het model zoals weergegeven in figuur 1 drie keer geschat, waarna de parameters in de eerste dataset worden afgezet tegen die in de tweede en derde dataset met als doel de juistheid van de significantie, richting en grootte te bepalen. Door de dataset drie keer te completeren en de drie geschatte modellen te vergelijken is de onzekerheid omtrent de te imputeren waarden terug te vinden in de verschillen tussen de parameters. Teneinde de geschatte modellen met elkaar te kunnen vergelijken wordt de modelstructuur zoals gebruikt voor de eerste dataset precies zo opgelegd aan de data in de tweede en derde dataset.

Om NORM te kunnen toepassen moet eerst gekeken worden of de missende waarden MAR zijn. Hoewel MAR zonder externe informatie niet 100% vastgesteld kan worden, is het volgens Schafer (1997) niet zo zeer van belang te bewijzen dat de responskans volkomen ongerelateerd is aan de missende data, maar dat deze relatie kan worden verklaard door de wél geobserveerde data. Met andere woorden, door het meenemen van achtergrondvariabelen die mogelijk een relatie hebben met het patroon van missende waarden, komt men dicht bij MAR. De relatie tussen de achtergrondvariabelen en de incompleetheid van de Y-variabelen genereert extra informatie die leidt tot realistischer schattingen van de missende waarden. In de Interstratos dataset blijkt dat de respons met name afhangt van de grootte van de MKB-onderneming (gemeten in aantal werknemers), de sector en het land van herkomst. Door deze achtergrondvariabelen mee te nemen bij het simuleren van de missende waarden (evenals de overige variabelen in de oorspronkelijke dataset), wordt MAR aannemelijker geacht. Dit is in overeenstemming met de conclusie van Demirtas (2004) dat het gebruik van supplementaire variabelen die mogelijke oorzaken zijn van of correleren met het patroon van missende waarden of van de missende waarden zelf tot betere resultaten kan leiden in vergelijking met simpelere imputatiemethoden. Ook Collins, Schafer en Kam (2001) pleiten voor een inclusieve strategie (het opnemen van talrijke extra variabelen) versus een restrictieve strategie (het opnemen van weinig of geen extra variabelen). Hierdoor is MAR aannemelijker te maken dan zonder het opnemen van deze variabelen.

Ook de multivariate normaliteit wordt gecontroleerd, waarbij transformaties worden toegepast op variabelen met scheve verdelingen. Na de EM-stap moet voor de DA-stap het aantal iteraties gekozen worden. We stellen dit aantal gelijk aan drie keer het aantal iteraties dat benodigd is voor convergentie in de EM-stap ($k = 600$). In ons geval wordt dus na elke 600^e iteratie een dataset geïmputeerd, resulterend in drie complete datasets. In tabel 1 zijn de gemiddelde waarden weergegeven van de variabelen waaruit de afhankelijke variabele 'exportratio' is opgebouwd.

De cijfers in tabel 1 laten duidelijk zien dat hoe meer informatie mist in een variabele, hoe groter de afwijking is ten opzichte van de oorspronkelijke waarde van het gemiddelde. Immers, 'totale verkopen' vertoont veel minder non-respons dan 'exportverkopen'. Daarbij is het logisch dat de gemiddelde exportverkopen in de drie volledige datasets lager zijn dan in de oorspronkelijke dataset, aangezien met name de kleinere bedrijven item non-respons vertonen op deze variabele. Hierdoor zijn de gemiddelde waarden van 'exportverkopen' realistischer. Daarbij is het niet de bedoeling van

Tabel 1. Gemiddelde waarden 'totale verkopen' en 'exportverkopen' in oorspronkelijke en geïmputeerde datasets.

Variabele	oorspronkelijke dataset	1 ^e geïmputeerde dataset	2 ^{de} geïmputeerde dataset	3 ^{de} geïmputeerde dataset
Totale verkopen	3861.69	3819.16	3925.47	3794.87
Exportverkopen	2431.76	1507.76	1978.99	1551.85

NORM de gemiddelden van de in te vullen variabelen zo precies mogelijk weer te geven, maar de covariantiematrix van de oorspronkelijke variabelen zo goed mogelijk weer te geven. Met de complete datasets wordt het model drie keer geschat om zo de parameters te kunnen vergelijken en een uitspraak te kunnen doen over de onzekerheid rondom de missende waarden en de uitkomsten van de hypothesen.

4.3. Operationalisering van de variabelen

De variabelen die in deze studie gebruikt worden zijn gerelateerd aan exportsucces, exportinformatie en risicohouding. Exportsucces wordt gemeten middels de exportratio, dat wil zeggen het gedeelte van de totale verkopen dat verkregen wordt door exportverkopen. Dit is een gangbare maatstaf van exportsucces en heeft als voordeel dat de relatieve grootte van de onderneming er al uit gefilterd is, dit in tegenstelling tot bijvoorbeeld exportverkopen (Voerman, 2004). De mate waarin informatie wordt verzameld is formatief en wordt gemeten middels het totale aantal informatiebronnen dat een MKB-bedrijf consulteert voor de internationale activiteiten ingedeeld naar *export market research* (4 bronnen), *export assistance* (3 bronnen) en *export market intelligence* (5 bronnen)ⁱⁱⁱ. Hierbij is gevraagd of de respondent een specifieke informatiebron op de thuismarkt dan wel in het buitenland heeft geraadpleegd in het afgelopen jaar. Er zijn Likert-stellingen opgesteld om de risicohouding van de entrepreneur te meten. Na een factor- en betrouwbaarheidsanalyse blijkt dit reflectieve concept tweedimensionaal te zijn, dat wil zeggen de 'risicohouding met betrekking tot veranderingen' wordt gemeten met twee stellingen ('changes in a business should be avoided at all costs' en 'a firm should not leave the region where it is established') en de 'risicohouding met betrekking tot planning' wordt ook gemeten met twee stellingen ('firms should only introduce proven office procedures and production techniques' en 'managers should plan rather than follow their intuition'). In tabel 2 zijn de variabelen en de manier waarop ze totstandkomen, vermeld.

4.4. Methodologie

Voor het schatten van de directe en indirecte effecten tussen de constructen en variabelen zoals gespecificeerd in de hypothesen wordt *structural equations modeling* (SEM) toegepast, gebruik makend van het programma LISREL (Jöreskog & Sörbom, 1998). Bij SEM moeten normaliter twee modellen worden gespecificeerd: Het structurele model en de meetmodellen (Hair *et al*, 1998; Leeflang *et al*, 2000). Hierbij hanteer ik de twee-staps aanpak van Anderson & Gerbing (1988), waarbij eerst de meetmodellen worden geschat. Dat wil zeggen, de kwaliteit van de latente constructen wordt vastge-

Tabel 2. De constructen / indices en operationalisatie.

Variabele	Samenstelling
Aantal <i>export market research</i> bronnen op binnenlandse markt (Res_{bi})	Sommatie van 4 mogelijke bronnen
Aantal <i>export assistance</i> bronnen op binnenlandse markt (Ass_{bi})	Sommatie van 3 mogelijke bronnen
Aantal <i>export intelligence</i> bronnen op de binnenlandse markt (Int_{bi})	Sommatie van 5 mogelijke bronnen
Aantal <i>export market research</i> bronnen op buitenlandse markt (Res_{bu})	Sommatie van 4 mogelijke bronnen
Aantal <i>export assistance</i> bronnen op buitenlandse markt (Ass_{bu})	Sommatie van 3 mogelijke bronnen
Aantal <i>export intelligence</i> bronnen op de buitenlandse markt (Int_{bu})	Sommatie van 5 mogelijke bronnen
Export ratio	Exportverkopten / totale verkopen
Risicohouding met betrekking tot veranderingen (RV)	Gemiddelde van 2 5-punts Likertschalen
Risicohouding met betrekking tot planning (RP)	Gemiddelde van 2 5-punts Likertschalen

steld (in dit geval de twee reflectieve risicohouding-variabelen) door deze in een confirmatieve factoranalyse te beoordelen op interne betrouwbaarheid volgens de criteria die genoemd worden in Bagozzi en Yi (1988) en Hair *et al.* (1998). Aangezien de informatie-indices niet reflectief maar formatief zijn, worden deze in de eerste stap op kwaliteit getoetst middels de richtlijnen van Diamantopoulos en Winklhofer (2001). Vervolgens wordt in de tweede stap het structurele of conceptuele model geschat met behulp van een MIMIC ('Multiple Indicator and Multiple Causes') specificatie, waarbij de parameters in de meetmodellen uit de eerste stap worden vastgezet om zodoende te kunnen focussen op de relaties in het conceptuele model (MacCallum & Browne, 1993; Bollen, 1994).

Voorgaande werkwijze wordt toegepast op elk van de drie complete datasets. De eerste geïmputeerde dataset wordt gebruikt om de meetmodellen en het structurele model te schatten. Vervolgens worden deze structuren opgelegd op de tweede en derde geïmputeerde dataset. Met de resultaten van deze laatste twee analyses wordt gekeken of de oplossing uit de eerste dataset robuust is: zijn de parameters (significantie, teken en grootte) stabiel over de drie datasets, of worden de geïmputeerde waarden gekenmerkt door een hoge mate van onzekerheid (veroorzaakt door een hoog percentage geïmputeerde missende waarden) met als gevolg dat de drie modeloplossingen afwijken van elkaar?

5. RESULTATEN

In deze paragraaf worden achtereenvolgens de resultaten van de meetmodellen, de resultaten van het schatten van het structurele model, de gevolgen voor de hypothesen en de verschillen dan wel overeenkomsten van de drie geïmputeerde datasets besproken.

5.1. De meetmodellen

Bij alle drie datasets zijn de *goodness of fit* maatstaven van de confirmatieve factoranalyse op 'risicohouding ten aanzien van veranderingen' en 'risicohouding ten aanzien van planning'^{iv} in overeenstemming met de gestelde criteria, evenals de interne betrouwbaarheid, de convergente validiteit en de discriminante validiteit (Bagozzi en Yi 1988, Hair *et al.* 1998).

Ook de informatie-indices voldoen aan de eisen die Diamantopoulos en Winklhofer (2001) hebben gesteld ten aanzien van de inhoud- en indicatorspecificatie, de maximale multicollineariteit en de externe validiteit. Zo lijkt de index qua definitie veel op de al eerder toegepaste schalen die de hoeveelheid verzamelde informatie meten of de uitgebreidheid van het verzamelen van informatie, zoals in Marandu (1995) of in Yeoh (2000).

5.2. Schatten van het conceptuele model

In de tweede stap wordt het structurele model geschat op de drie volledige datasets, zodat de stabiliteit van de parameters kan worden geëvalueerd. Zoals eerder gezegd gebruik ik hiervoor een MIMIC specificatie in het programma LISREL. De resultaten van deze exercities staan weergegeven in tabel 3. Significante parameters zijn vet gemaakt ($\alpha = .05$) of cursief ($\alpha = .10$), terwijl opvallende verschillen tussen de drie datasets zijn gemarkeerd.

De verklarende kracht van de drie modellen is goed, aangezien de χ^2 niet significant is, de CFI en GFI boven de .90 liggen en de RMSEA onder de .05. Bij het interpreteren van de resultaten is het interessant (1) of de hypothesen zijn aanvaard dan wel verworpen en (2) of het gebruik van drie complete datasets tot andere conclusies leidt aangaande de hypothesen dan slechts één dataset. Als er inderdaad verschillen zijn, komt dat doordat er onzekerheid is omtrent de geïmputeerde waarde die tot uitdrukking komt in verschillen tussen de parameters. Dat zou ook betekenen dat bij enkelvoudige imputatie, zelfs als deze gebaseerd is op een model zoals MI, de uitkomsten kunnen leiden tot foutieve conclusies omtrent verbanden tussen variabelen. Hiertoe worden de resultaten van de eerste dataset^v vergeleken met die van de tweede en derde dataset.

5.3. Toetsing van de hypothesen

De drie getoetste hypothesen worden deels door de data ondersteund^{vi}. Ten aanzien van hypothese 1 (het effect van informatiecollectie op exportsucces) kan gesteld worden dat met name het consulteren van buitenlandse informatiebronnen een significant positief effect heeft op de exportratio, waarbij het effect van informele of *intelligence* informatie het sterkst is. Deze resultaten zijn stabiel over de drie datasets heen. Hoewel het verband tussen binnenlandse *assistance* informatie en exportratio in de eerste dataset niet significant is, blijkt deze in dataset 2 en 3 wel positief significant te zijn en zelfs vergelijkbaar met het effect van buitenlandse *assistance* bronnen.

De tweede hypothese stelt dat een risicoaverse entrepreneur minder exportinformatie verzamelt dan een risiconemende entrepreneur. Er blijkt duidelijk verschil te zijn tus-

Tabel 3. Resultaten LISREL op de drie geïmputeerde datasets.

Fitmaatstaven			
Totale model	$\chi^2 = 30.31$ (p=.09, df = 21)	$\chi^2 = 32.29$ (p=.06, df = 21)	$\chi^2 = 28.11$ (p=.14, df = 21)
	RMSEA = .02 (1.00)	RMSEA = .022 (1.00)	RMSEA = .018 (1.00)
	GFI = 1.00	GFI = .99	GFI = 1.00
	CFI = 1.00	CFI = .99	CFI = 1.00
Grootte en significantie van de parameters			
Op 'het verzamelen van <i>export intelligence</i> informatie op de binnenlandse markt' van			
RV	-.15 (-2.54)	-.10 (-1.78)	-.06 (-1.05)
RP	.09 (1.47)	.10 (1.80)	.04 (.78)
Op 'het verzamelen van <i>export market research</i> informatie op de binnenlandse markt' van			
RV	-.16 (-2.65)	-.16 (-2.66)	-.11 (-2.05)
RP	.09 (1.52)	.14 (2.39)	.12 (2.09)
Op 'het verzamelen van <i>export assistance</i> informatie op de binnenlandse markt' van			
RV	-.10 (-1.67)	-.07 (-1.29)	-.07 (-1.25)
RP	.14 (2.22)	.10 (1.81)	.10 (1.68)
Op 'het verzamelen van <i>export intelligence</i> informatie op de buitenlandse markt' van			
RV	-.17 (-2.96)	-.14 (-2.38)	-.20 (-3.68)
RP	-.06 (-1.03)	.03 (.44)	-.03 (-.45)
Op 'het verzamelen van <i>export market research</i> informatie op de buitenlandse markt' van			
RV	-0.9 (-1.59)	-.09 (-1.61)	-.10 (-1.80)
RP	-0.2 (-.30)	-.04 (-.64)	-.02 (-.28)
Op 'het verzamelen van <i>export assistance</i> informatie op de buitenlandse markt' van			
RV	-.10 (-1.67)	-.19 (3.26)	-.12 (-2.19)
RP	-.02 (.29)	.09 (1.50)	.01 (.14)
Op 'exportratio' van			
RV	-.10 (-1.69)	-.16 (-2.73)	-.27 (-4.74)
RP	-.08 (-1.32)	.01 (.19)	.17 (2.91)
Int _{bi}	.02 (.58)	.00 (-.17)	.01 (.29)
Res _{bi}	-.02 (-.73)	-.02 (-.52)	-.01 (-.58)
Ass _{bi}	.02 (.65)	.06 (1.98)	.07 (2.37)
Int _{bu}	.21 (7.17)	.27 (9.45)	.22 (7.28)
Res _{bu}	.12 (4.11)	.09 (3.12)	.09 (2.87)
Ass _{bu}	.12 (4.22)	.09 (2.93)	.11 (3.58)

sen de effecten van de twee attitudecomponenten. De veranderingscomponent heeft een veel significantere (en negatieve) link met informatiegedrag dan de planningscomponent. Uit de drie datasets komt uniform naar voren dat een entrepreneur die negatiever staat tegenover veranderingen zowel minder binnenlandse *export market research* als buitenlandse *export intelligence* informatie verzamelt dan een meer risiconemende ondernemer. Twee van de drie datasets leveren ook significant bewijs voor een lager gebruik van *export assistance* bronnen. Dit wordt in mindere mate ondersteund in de

eerste analyse ($\alpha \leq 10$). Het is moeilijk iets te zeggen over het effect van de houding ten aanzien van veranderingen op het verzamelen van binnenlandse informele informatie. In alle datasets is het teken van de parameter negatief, maar de significantie loopt uiteen van significant (dataset 1), net niet significant (dataset 2) tot zeker niet significant (dataset 3). Er is dus geen duidelijk bewijs voor een negatief effect, maar wel een indicatie. In tegenstelling tot de hypothese is de risicohouding ten opzichte van planning juist positief gelieerd aan informatieverzameling, zij het maar aan twee van de zes informatiev variabelen. Des te meer een entrepreneur zekerheid wil door te plannen, des te meer binnenlandse *export assistance* informatie hij verzamelt (deze uitspraak kent voor datasets 2 en 3 een betrouwbaarheid van slechts 90%). Dataset 2 en 3 indiceren ook duidelijk dat deze entrepreneurs vaker gebruik maken van binnenlandse *export research* informatie; in de eerste dataset is deze parameter niet significant. De laatste hypothese veronderstelt een negatief verband tussen risicoaversie en export-succes. Alle analyses duiden op een significant negatief effect van 'risicohouding ten aanzien van veranderingen' op exportratio (weliswaar in de eerste dataset op 90%), waarmee de hypothese deels aanvaard wordt.

5.4. Verschillen en overeenkomsten tussen de drie geïmputeerde datasets

Doordat het conceptuele model drie maal is geschat, kan met meer zekerheid iets worden gezegd over de gevonden verbanden. Zo is het duidelijk dat de positieve relatie tussen buitenlandse informatie en exportratio stabiel is over de drie sets, waardoor er sterk bewijs is voor dit verband. Dat geldt ook voor de effecten van RV op informatieonderdelen Res_{bi} en Int_{bu} . Andere verbanden, die niet significant schijnen bij analyse van alleen de eerste dataset (zoals bij enkelvoudige imputatie), worden in een ander licht gezet met de significante resultaten uit de tweede en derde dataset. Voorbeelden zijn de relaties tussen RP en Int_{bi} , tussen RV en Ass_{bu} , tussen RV en ER en tussen Ass_{bi} en ER. Als de eerste dataset de maatstaf zou zijn voor het accepteren of verwerpen van de hypothesen zouden deze verbanden als niet significant en dus niet belangrijk worden aangemerkt. Zo zou in de conclusies niet terugkomen dat het voordelig zou zijn voor ondernemers om binnenlandse *export assistance* bronnen te consulteren, terwijl dit toch een positief effect heeft op exportsucces gezien de resultaten van de andere twee analyses. De omgekeerde situatie is ook mogelijk: niet-significante resultaten in de tweede en derde dataset kunnen vragen oproepen over een relatie die in de eerste dataset wel significant is. Een voorbeeld is het vermeende negatieve effect van RV op Int_{bi} in de eerste dataset, dat in de tweede en derde dataset respectievelijk net niet, en helemaal niet significant is. Dit is nog duidelijker als de resultaten uit de derde dataset worden vergeleken met die uit de eerste of tweede dataset. Er wordt een significante, positieve relatie gevonden tussen RP en ER, terwijl deze in de andere datasets helemaal niet terug te vinden is. Ook het teken van de parameter verschilt tussen de datasets. In zo'n geval is er dus te weinig bewijs voor een relatie, gezien de resultaten uit de drie geïmputeerde datasets gezamenlijk.

6. CONCLUSIES

In dit artikel is een conceptueel model empirisch getoetst met drie complete datasets, waarbij de missende waarden in deze datasets drie keer zijn geïmputeerd door het MI

programma NORM. De conclusies van deze studie zijn tweeledig. Enerzijds zeggen de resultaten iets over het waarheidsgehalte van de opgestelde hypothesen. Anderzijds is een conclusie te trekken aangaande de meerwaarde van MI ten opzichte van enkelvoudige imputatie bij het op juiste manier interpreteren van de parameterschattingen.

6.1. Conclusies ten aanzien van het conceptuele model

Met deze analyses is aangetoond dat er een positieve relatie bestaat tussen het verzamelen van met name buitenlandse informatie en exportsucces. Gegeven het feit dat MKB-bedrijven slechts weinig buitenlandse informatiebronnen consulteren, is er ruimte voor deze internationale bedrijven hun exportsucces te verbeteren door meer op zoek te gaan naar informatie in het buitenland, ondanks de hogere kosten. Een voorzichtige conclusie is ook dat MKB-bedrijven profiteren van de exporthulp van de overheid. Zo bestaat in Nederland bijvoorbeeld de Export Voorlichtingdienst (EVD) waar bedrijven informatie kunnen verkrijgen over velerlei landen en onderwerpen (www.evd.nl). Ook is inzicht verkregen in het effect van de risicohouding van de entrepreneur op informatiegedrag en exportsucces. Een manager die een aversie heeft tegen veranderingen is geneigd minder informatie te verzamelen en is daarmee minder marktgeoriënteerd. Deze aversie heeft dus, naast een significante directe negatieve invloed, ook een indirect negatieve invloed op exportintensiteit. Aan de andere kant geven de resultaten aan dat managers die planning verkiezen boven het risicovollere intuïtieve beslissen juist meer informatiebronnen consulteren, wat daarmee een indirect positief effect heeft op exportsucces. Het is dus interessant voor het MKB, maar ook voor de overheid en banken, zich bewust te zijn van de persoonlijke eigenschappen van de entrepreneur en het effect van deze op het te verwachten exportsucces.

6.2. Conclusies ten aanzien van meervoudige imputatie

De keuze voor MI in plaats van een enkelvoudige imputatie van missende waarden blijkt een juiste te zijn geweest. Het gebruik van de drie geïmputeerde datasets om tot één resultaat te komen, levert inderdaad extra informatie over de stabiliteit van en onzekerheid rondom parameters. Zo kan gesteld worden dat meer bewijs is voor een bepaald verband tussen twee variabele als deze significant is en hetzelfde teken heeft in elke dataset. Ondanks de incomplete informatie in de oorspronkelijke dataset kunnen door imputatie van missende waarden toch met analysemethoden worden gebruikt die gebaseerd zijn op complete data. Deze studie levert bewijs dat MI een krachtige methode is om onzekerheid omtrent missende waarden mee te nemen in de datavermeerdering. Aangezien bij enkelvoudige imputatie slechts één dataset wordt gebruikt om een theoretisch model te schatten, kan de onderzoeker slechts deze resultaten als een zo goed mogelijke afspiegeling van de werkelijkheid beschouwen. Zoals de resultaten hebben aangetoond, kan dit leiden tot een foutieve evaluatie van de parameters. Immers, significante parameters in één dataset kunnen worden tegengesproken in vergelijkbare geïmputeerde datasets. Daartegenover staat dat als parameters consequent in elke geïmputeerde dataset naar voren komen, dit een sterker bewijs is voor een werkelijk empirische relatie. In deze studie is een relatief simpel model gebruikt ter illustratie. Bij het schatten van het integrale exportsuccesmodel met een veel groter aantal variabelen en relaties, zijn de verschillen tussen de drie datasets nog opvallender (Voerman, 2004). Onderzoekers die zich in de situatie bevinden dat er eigenlijk te veel mis-

sende waarden zijn om een krachtige analyse uit te voeren en overwegen de missende data te imputeren, moeten dus altijd gebruik maken van meervoudige imputatie en niet van enkelvoudige imputatie.

In de literatuur is de belangstelling voor de MI methode groeiend. Daarnaast zijn, voortbouwend op de MI methode, al meerdere extensies of gerelateerde technieken ontwikkeld die toepasbaar zijn als niet alleen missende waarden het probleem zijn. Voorbeelden zijn het gebruik van MI bij het samenvoegen van afzonderlijke vragenlijsten of steekproeven (Brown 2002), de *multiple edit/multiple imputation* extensie van Ghosh-Dastidar en Schafer (2003), of een op *bootstrapping* gebaseerde imputatietechniek in combinatie met het factor model (Kamakura en Wedel 2003).

Het belangrijkste doel van dit artikel, onderzoeken of het gebruik van MI tot een andere interpretatie van parameters en dus tot andere conclusies leidt, is behaald. Doordat MI de onzekerheid rondom een te imputeren variabele in ogenschouw neemt, kunnen resultaten uit meerdere geïmputeerde datasets vergeleken worden. Bij het gebruik van enkelvoudige imputatie in plaats van MI was ik tot andere conclusies gekomen ten aanzien van het geschatte model. Deze conclusies kunnen negatief uitwerken in de aanbevelingen naar het MKB, de overheid en bijvoorbeeld banken in hun benadering van internationaliserende entrepreneurs en MKB-bedrijven. De benodigde extra inspanningen zijn dan, zowel vanuit theoretisch als empirisch oogpunt, zeker de moeite waard.

NOTEN

- i De internationale onderzoeksgroep INTERSTRATOS (INTERNationalization of STRATEGic orientation of small- and medium-sized enterprises) bestaat uit J. Hanns Pichler, Christian Lettmayer, Erwin Fröhlich en Peter Voithofer (Oostenrijk), Rik Donckels, Ria Aerts en Jan Degadt (België), Antti Haahti en Petri Ahokangas (Finland), Yvonne Prince, Peter Zwart en Liane J.A. Voerman (Nederland), Per-Anders Havnes en Arild Saether (Noorwegen), Håkan Boter en Carin Holmquist (Zweden), Margrit Habersaat en Hans Jobst Pleitner (Zwitserland).
- ii NACE 43 & 45 textiel en kleding, NACE 34 elektronica, NACE 41 & 42 voedingsmiddelen, dranken en tabak, NACE 46 houten producten, inclusief meubels, NACE 31 & 32 metalen producten en werktuigbouwkunde.
- iii De Interstratos enquête vraagt de respondent aan te geven of er de afgelopen 12 maanden een bepaalde informatiebron is geconsulteerd. Dit gaat om twaalf mogelijke informatiebronnen die voor deze studie zijn gecategoriseerd op basis van de indeling van Souchon en Diamantopoulos (1996).
- iv Hoewel deze maatstaven bij de tweede risicoattitude niet consequent boven de gestelde afkappunten uitkomen, liggen deze daar slechts net onder. Daarom is besloten dit construct toch mee te nemen, mede gezien het doel van deze exercitie: het evalueren van MI.
- v. Welke van de drie geïmputeerde datasets wordt gebruikt voor het vaststellen van het structurele model is arbitrair. Immers, elke geïmputeerde dataset is uitwisselbaar.
- vi. Alleen de significante resultaten worden besproken.

LITERATUUR

Ahmed, S. U. 1985, "nAch, Risk-Taking Propensity, Locus of Control and Entrepreneurship", *Personality & Individual Differences*, vol. 6, no. 6, pp. 781-782.

- Anderson, J. C. & Gerbing, D. W. 1988, "Structural Equation Modeling in Practice: A Review and Recommended Two-Step Approach", *Psychological Bulletin*, vol. 103, no. 3, pp. 411-423.
- Avlonitis, G. J. & Gounaris, S. P. 1999, "Marketing Orientation and its Determinants: An Empirical Analysis", *European Journal of Marketing*, vol. 33, no. 11/12, pp. 1003-1036.
- Bagozzi, R. P. & Yi, Y. 1988, "On the Evaluation of Structural Equation Models", *Academy of Marketing Science*, vol. 16, no. 1, pp. 74-94.
- Bollen, K. A. 1994, *Structural Equations with Latent Variables*, 6 edn, Wiley, New York.
- Brown, J. G. 2002, "Using a Multiple Imputation Technique to Merge Data Sets", *Applied Economics Letters*, vol. 9, no. 5, pp. 311-314.
- Cadogan, J. W. & Diamantopoulos, A. 1995, "Narver and Slater, Kohli and Jaworski and the Market Orientation Construct: Integration and Internationalization", *Journal of Strategic Marketing*, vol. 3, pp. 41-60.
- Collins, L. M., Schafer, J. L. & Kam, C. 2001, "A Comparison of Inclusive and Restrictive Strategies in Modern Missing Data Procedures", *Psychological Methods*, vol. 6, no. 4, pp. 330-351.
- Demirtas, H. 2004, "Simulation Driven Inferences for Multiply Imputed Longitudinal Datasets", *Statistica Neerlandica*, vol. 58, no. 4, pp. 466-482.
- Diamantopoulos, A. & Winklhofer, H. M. 2001, "Index Construction with Formative Indicators: An Alternative to Scale Developments", *Journal of the Academy of Management Science*, vol. 38, no. 2, pp. 269-277.
- Ghosh-Dastidar, B. & Schafer, J. L. 2003, "Multiple Edit/Multiple Imputation for Multivariate Continuous Data", *Journal of the American Statistical Association*, vol. 98, no. 464, pp. 807-817.
- Hair, J. F. jr., Anderson, R. E., Tatham, R. L., & Black, W. C. 1998, *Multivariate Data Analysis*, 5 edn, Prentice-Hall International, London.
- Jaworski, B. J. & Kohli, A. K. 1993, "Market Orientation: Antecedents and Consequences", *Journal of Marketing*, vol. 57, no. 3, pp. 53-70.
- Johanson, J. & Vahlne, J. E. 1977, "The internationalization process of the firm – A model of knowledge development and increasing foreign market commitments", *Journal of International Business Studies*, vol. 8, no. 1, pp. 23-32.
- Jöreskog, K. G. & Sörbom, D. 1999, *LISREL 8: Structural Equation Modeling with the SIMPLIS Command Language*, 4 edn, Scientific Software International, Chicago.
- Kamakura, W. A. & Wedel, M. 2003, "List Augmentation with Model Based Multiple Imputation: A Case Study using a Mixed-Outcome Factor Model", *Statistica Neerlandica*, vol. 57, no. 1, pp. 46-57.
- Kohli, A. K. & Jaworski, B. J. 1990, "Market Orientation: The Construct, Research Propositions, and Managerial Implications", *Journal of Marketing*, vol. 54, no. April, pp. 1-18.
- Kuratko, D. F. & Hodgetts, R. M. 2001, *Entrepreneurship: A Contemporary Approach*, 5 edn, The Dryden Press, Fort Worth.
- Leeflang, P. S. H., Wittink, D. R., Wedel, M., & Naert, P. A. 2000, *Building Models for Marketing Decisions* Kluwer Academic Publishers, Dordrecht.
- MacCallum, R. C. & Browne, M. W. 1993, "The Use of Causal Indicators in Covariance Structure Models: Some Practical Issues", *Psychological Bulletin*, vol. 114, no. 3, pp. 533-541.
- Marandu, E. E. 1995, "Impact of Export Promotion on Export Performance: A Tanzanian Study", *Journal of Global Marketing*, vol. 9, no. 1/2, pp. 9-39.
- Narver, J. C. & Slater, S. F. 1990, "The Effect of a Market Orientation on Business Profitability", *Journal of Marketing*, vol. 54, October, pp. 20-35.
- Nijssen, E. J. & Douglas, S. P. 1999, "Gathering and Using Information for the Selection of Trading Partners", *European Journal of Marketing*, vol. 33, no. 1/2, pp. 143-162.
- Noh, H., Kwak, M. & Han, I. 2004, "Improving the Prediction Performance of Customer Behavior Through Multiple Imputation", *Intelligent Data Analysis*, vol. 8, no. 6, pp. 563-577.

- Olinsky, A., Chen, S. & Harlow, L. 2003, "The Comparative Efficacy of Imputation Methods for Missing Data in Structural Equations Modeling", *European Journal of Operational Research*, vol. 151, no. 1, pp. 53-79.
- Rubin, D. B. 1987, *Multiple Imputations for Non-Response in Surveys*, Wiley, New York.
- Rubin, D. B. & Schenker, N. 1986, "Multiple Imputation for Interval Estimation From Simple Random Samples With Ignorable Responsiveness", *Journal of the American Statistical Association*, vol. 81, no. 394, pp. 366-364.
- Schafer, J. L. 1997, *Analysis of Incomplete Multivariate Data*, Chapman & Hall, London.
- Schafer, J. L. 2003, "Multiple Imputation in Multivariate Problems When the Imputation and Analysis Model Differ", *Statistica Neerlandica*, vol. 57, no. 1, pp. 19-35.
- Souchon, A. L. & Diamantopoulos, A. 1996, "A Conceptual Framework of Export Marketing Information Use: Key Issues and Research Propositions", *Journal of International Marketing*, vol. 4, no. 3, pp. 49-71.
- Souchon, A. L. & Diamantopoulos, A. 1999, "Export Information Acquisition Modes: Measure Development and Validation", *International Marketing Review*, vol. 16, no. 2, pp. 143-168.
- Voerman, J. A. 2004, *The Export Performance of European SMEs*, Labyrinth Publications, Alblasterdam.
- Yeoh, P. L. 2000, "Information Acquisition Activities: A study of Global Start-Up Exporting Companies", *Journal of International Marketing*, vol. 8, no. 3

12. Integrated Campaign Evaluation: Vragen naar mediagebruik op basis van fusie als alternatief voor recall

A. VAN MEURS

SAMENVATTING

Voor respondenten in ad-hoc onderzoek is het vaak lastig om in te schatten hoe vaak ze met een reclamecampagne in contact zijn geweest, bijvoorbeeld hoe vaak ze deze op televisie hebben gezien. Er kunnen vraagtekens worden gezet bij de validiteit en betrouwbaarheid van een dergelijke frequency recall vraag. Voor televisiecampagnes biedt Integrated Campaign Evaluation (ICE) een schatting van het aantal malen dat men met de campagne in contact is geweest door middel van fusie van het effectonderzoek met het Kijkonderzoek, als alternatief voor de uitkomsten van de frequency recall vraag. Het geschat aantal contacten met de televisiecampagne in de steekproef van het effectonderzoek biedt een alternatief voor frequency recall vragen zoals “Hoe vaak heeft u deze campagne op televisie gezien”.

Integrated Campaign Evaluation is nu naast televisie ook ontwikkeld voor radio, print en buitenreclame. In dit hoofdstuk wordt het gebruik van ICE voor zowel televisie als voor de overige media gepresenteerd. Hierbij worden onder meer de verklarende modellen toegelicht en worden voorbeelden gegeven hoe deze resultaten kunnen worden toegepast in een effectstudie uit de praktijk.

1. INLEIDING: TWIJFEL OVER FREQUENCY RECALL

Bij campagnetracking en andere effectonderzoeken lijken vragen naar het aantal malen dat respondenten in contact zijn geweest met deze campagne onvermijdelijk. Bij het evalueren van televisiecampagnes ligt het voor de hand om bijvoorbeeld een vraag als “Hoe vaak heeft u deze commercial op televisie gezien” op te nemen in de vragenlijst. Zodra een dergelijke frequency recall vraag in de vragenlijst is opgenomen, bij online onderzoek bijvoorbeeld met foto's of een filmpje van de commercial, zullen de meeste respondenten een antwoord proberen te geven. Maar in hoeverre is de informatie die dit oplevert bruikbaar? Kunnen respondenten zich werkelijk herinneren hoe vaak ze een campagne op televisie hebben gezien? Hoe vaak zullen ze het aantal contacten onderschatten omdat ze zich niet realiseren hoe vaak ze een campagne hebben gezien? En hoe vaak zullen ze bij multimediacampagnes het aantal televisiecontacten overschatten omdat ze ook contacten met tijdschriftreclame of buitenreclame meetellen?

Er is een aantal redenen om de validiteit en betrouwbaarheid van dergelijke frequency recall vragen te wantrouwen. Ten eerste is het zelfs voor getrainde en goed gemotiveer-

de profpersonen lastig om het werkelijke aantal confrontaties met commercials bij te kunnen houden. Verder is het bij campagnes waar de uitingen van de verschillende media goed op elkaar aansluiten (zoals de identieke bioscoop- en televisiecommercials voor Bacardi in ons voorbeeld) goed voorstelbaar dat respondenten de bron van de reclameherkenning door elkaar halen. Een dominant medium als televisie kan hierbij makkelijk worden overschat. Het horen van een reclame op de radio kan de beelden van de televisiecommercial oproepen. Deze zogenaamde visual transfer (Prance, 1995) kan zo sterk zijn dat men achteraf wellicht vergeet de commercial ook op de radio gehoord te hebben. Daarnaast is het niet voor alle media even goed mogelijk om respondenten te vragen naar de mate waarin zij met de betreffende campagne in dat medium zijn geconfronteerd: wie kan immers goed bijhouden hoeveel buitenreclameposters dagelijks voorbijflitsen? Tot slot zijn consumenten niet altijd bewust van de effecten die reclame op ons heeft, reclame kan immers ook werken via low involvement of pre-attentieve verwerking (Heath, 2000; Franzen, 1998). We kunnen van respondenten niet verwachten dat ze het aantal contacten met dergelijke low involvement processing kunnen bijhouden. Bovendien kan de mate waarin de werking van de campagne door middel van low involvement processing plaatsvindt per medium verschillen, waardoor een vraagstelling die geen rekening houdt met low involvement processing niet geschikt is om campagne effecten te vergelijken tussen verschillende mediatypes.

Als we het effect van de inzet van de verschillende mediamiddelen willen onderzoeken is het ook geen oplossing deze in te schatten op basis van de bruto of netto bestedingen per medium. Nog afgezien van de verschillen in effectiviteit van de uitingen voor elk van de media die voor de campagne zijn ingezet wordt bij een dergelijke benadering voorbijgegaan aan de individuele verschillen in confrontatie tussen respondenten: niet iedereen kijkt bijvoorbeeld evenveel televisie.

2. INTEGRATED CAMPAIGN EVALUATION

Om toch per medium apart voor individuele respondenten een inschatting te kunnen maken van de mate van confrontatie met een campagne is door Intomart GfK de Integrated Campaign Evaluation (ICE) ontwikkeld. Bij deze methode worden aan respondenten algemene vragen naar de tijdsbesteding van verschillende mediatypen gevraagd (zogenaamde media imperatives). Door middel van analyses op mediabereiksonderzoek wordt voor elk medium een voor die campagne specifiek model ontwikkeld waarbij wordt vastgesteld welke combinatie van media imperatives en eventuele achtergrondkenmerken het beste een voorspelling kan geven van het aantal contacten met de campagne van dat medium. Vervolgens wordt elk mediumpspecifieke model gerepliceerd in de steekproef van de campagnetracking, zodat kan worden geschat wat de bijdrage in de campagne effectiviteit is geweest van elk van de campagnes.

Voor de voorbeeld campagne van Bacardi zijn apart modellen opgesteld voor televisie, buitenreclame, bioscoop, print en radio. Als voorbeeld wordt in de volgende paragraaf getoond hoe het ICE model voor televisie van deze campagne is geschat. Ondanks de in de inleiding genoemde methodologische problemen zijn betrouwbare en valide schattingen van het aantal contacten met een campagne, bijvoorkeur voor ieder

medium apart van groot belang bij de analyse van de effecten. Integrated Campaign Evaluation biedt een alternatief voor de onbetrouwbare antwoorden die respondenten geven in enquêtes. Door middel van voorspellende modellen gebaseerd op campagne specifieke achtergrondkenmerken en vragen naar algemeen mediagebruik kan het aantal contacten met de campagne worden ingeschat. ICE biedt hiermee apart voor elk mediumtype een alternatief voor frequency recall vragen in elke vorm van effectonderzoek.

3. DE BACARDI CAMPAGNE

De toegevoegde waarde van ICE in een campagne evaluatie zal worden gedemonstreerd aan de hand van de resultaten van de campagne voor Bacardi witte rum zoals die liep in Nederland van januari tot met maart 2005. Met meer dan 1.000 GRP's lag de nadruk van deze campagne qua mediabestedingen op televisie. De televisiecampagne bestond uit twee commercials van 20 en 30 seconden met elk een verschillende verhaallijn ("strong arm" en "lemon squeeze"). De televisiecampagne werd ondersteund door de inzet van buitenreclame en reclame in bioscopen, tijdschriften, radio en horecagelegenheden gedurende dezelfde periode. In 2004 liep voor Bacardi een vergelijkbare campagne met de nadruk op televisie in de maanden april, mei, oktober en november en bioscoop in de zomermaanden.

Gedurende de maanden januari tot met maart 2005 werden wekelijks 100 respondenten van 18-39 jaar (de communicatiedoelgroep) on-line ondervraagd door Intomart GfK. In deze tracking werd allereerst gevraagd naar de spontane merk- en reclamebekendheid (brand en advertising awareness). Vervolgens werd gevraagd naar de geholpen merk- en reclamebekendheid, campagneherkenning per medium en een serie imago statements over het merk. In de vragenlijst werd de campagne getoond door middel van het afspele van een filmpje en het geluid en het tonen van de uitingen voor tijdschriften en buitenreclame. De vragenlijst werd afgesloten met een vraag naar het gebruik van het merk.

In figuur 1 zijn een aantal campagne effecten gedurende de campagneperiode weergegeven. Bacardi heeft een hoge top of mind positie en een grote merkbekendheid. Het product heeft zichzelf met eerdere campagnes een dermate sterke positie verworven, dat gedurende de campagne van januari tot en met maart 2005 de bekendheid van het merk niet verder toeneemt. De spontane merkbekendheid lijkt zelfs enige tijd te dalen, maar herstelt zich gedurende de campagne. Door een voorafgaande en vergelijkbare campagne in 2004 ving de campagne zelf in 2005 aan met een relatief hoge aanvangswaarde in campagneherkenning. Dat de campagne desondanks niet ongezien is verlopen blijkt uit het feit dat gedurende de 13 weken van de campagne de campagneherkenning toeneemt.

Aan het begin van de campagne daalt de spontane reclamebekendheid, mogelijk door campagnes voor andere merken. Gedurende de campagne herstelt de spontane reclamebekendheid zich grotendeels.

We zullen in deze beschouwing niet verder ingaan op de specifieke effecten van deze campagne voor het merk Bacardi. Campagne effecten zijn moeilijk over merken gene-

Vragenlijst campagnetracking

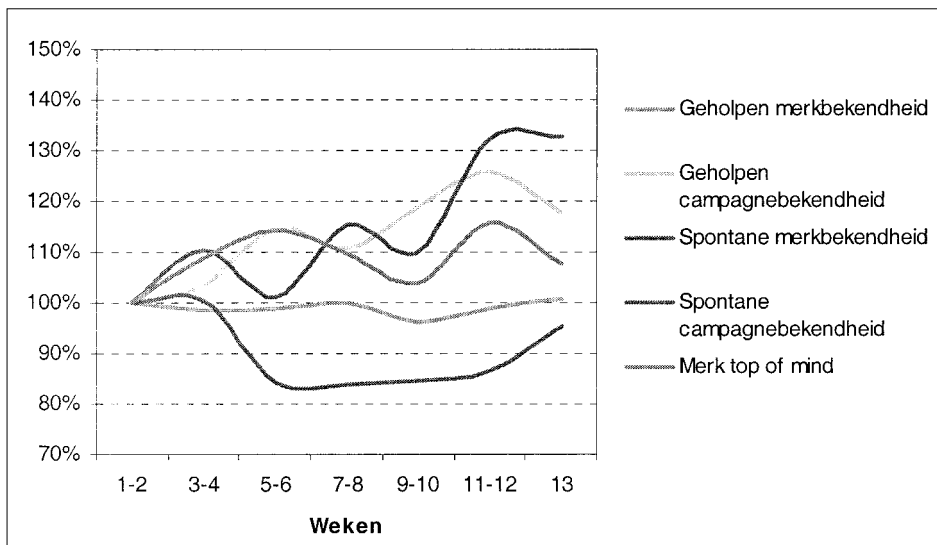
Merk- en campagnebekendheid

- Top of Mind: Als u denkt aan <product>, welk merk schiet u dan als eerste te binnen?
- Spontane merkbekendheid: Welke merken <product> kent u nog meer, al is het alleen maar van naam?
- Spontane reclamebekendheid: Als u denkt aan reclame voor <product>, van welke merken, kunt u zich herinneren recentelijk reclame te hebben gezien of gehoord, in welke vorm dan ook?
- Geholpen merkbekendheid: Welke van de volgende merken <product> kent u, ook al is het alleen maar van naam?
- Geholpen reclamebekendheid: Voor welke merken alcoholische dranken kunt u zich herinneren recentelijk reclame te hebben gezien of gehoord, in welke vorm dan ook?

Campagneherkenning

- Waar, in welke media, heeft u reclame voor Bacardi gezien of gehoord??
 - televisie
 - bioscoop
 - buitenreclame
 - tijdschriften/kranten
 - horeca gelegenheid
 - radio
 - winkel
- Image statements over het merk (random selectie van 25 respondenten)
- Herkenning televisiecommercial
- Herkenning radiocommercial
- Herkenning buitenreclameposter
- Herkenning tijdschriftuiting

Gebruik van het merk



Figuur 1. Effecten tijdens de campagne (januari-maart 2005, index week 1-2 = 100).

raliseerbaar omdat deze te afhankelijk zijn van de creatieve executie en kracht van de uiting, de boodschap, de positie van het merk en campagnes van concurrerende producten. De vraag of de campagne effect heeft gehad en zo ja, welke effecten dat zijn geweest zullen we in dit hoofdstuk dan ook niet beantwoorden. Maar stel dat de campagne het gewenste effect heeft gehad, aan welk medium of welke combinatie van media is dit effect dan toe te schrijven?

4. GEBRUIK EN BEPERKINGEN VAN DE FREQUENCY RECALL VAN DE BACARDI CAMPAGNE PER MEDIUMTYPE

Als we onderscheid willen maken bij de campagne effecten tussen de verschillende ingezette mediumtypes kunnen bijvoorbeeld de uitkomsten van de campagneherkenning gedurende de campagne periode worden geanalyseerd. In tabel 1 staan de resultaten van een dergelijke frequency recall vraag voor de campagne van Bacardi gemiddeld over de 13 weken van de campagne evaluatie. In figuur 2 worden deze percentages door de tijd weergegeven.

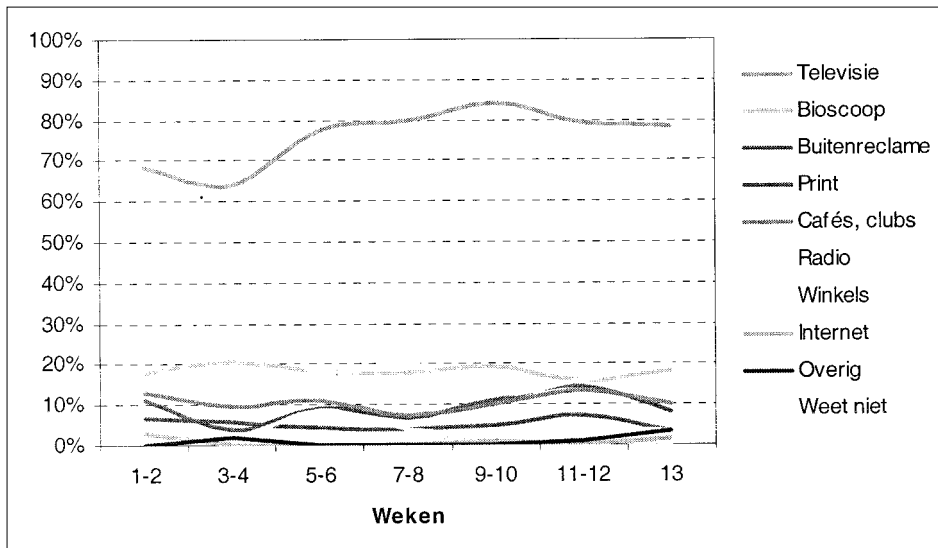
Tabel 1. Campagneherkenning "Waar, in welke media, heeft u reclame voor Bacardi (Witte Rum) gezien of gehoord?" (gemiddelde week 1-13 2005).

Televisie	76%
Bioscoop	19%
Buitenreclame	5%
Print	10%
Bars, clubs	11%
Radio	2%
Winkels	5%
Overig	1%
Internet	1%
Weet niet	20%

Uit tabel 1 blijkt dat consumenten zeggen de campagne primair van televisie te kennen. Bioscoopreclame scoort opvallend hoog als tweede met gemiddeld 19%. Uit figuur 2 blijkt dat gedurende de campagneperiode het percentage respondenten dat aangeeft de campagne niet te kennen geleidelijk afneemt. Televisie wordt gedurende de campagne steeds vaker genoemd als medium waar men de campagne heeft gezien. De campagne heeft duidelijk een bereik opgebouwd gedurende de looptijd van drie maanden.

Zou het eventuele succes van de campagne nu voornamelijk het gevolg zijn van de televisiecampagne? Als we afgaan op de percentages in tabel 1 zou je zeggen van wel. Televisie wordt immers door een overgrote meerderheid van de respondenten genoemd als het medium waar men de campagne heeft gezien. Maar in hoeverre zijn de antwoorden op de vraag in welk medium men de vraag heeft gezien betrouwbaar?

Analyse in het Kijkonderzoek¹ over de periode januari tot en met maart 2005 van het gecombineerde bereik van de beide commercials uit de Bacardi campagne wijst uit dat na deze drie maanden 93,5% van de personen van 13 jaar en ouder waren bereikt (één



Figuur 2. Campagneherkenning “Waar, in welke media, heeft u reclame voor Bacardi (Witte Rum) gezien of gehoord?” (januari-maart 2005).

contact of meer), met een gemiddelde contactfrequentie van 11,9. Dit bereik is wat hoger dan het maximum van 84% met de frequency recall vraag in week 9 en 10 van de campagne. Verdere analyse in het Kijkonderzoek leert dat 82,4% 3 of meer contacten met de campagne hebben gehad. Zou men zich pas na drie contacten een campagne herinneren, of zijn er andere redenen waarom de uitkomsten van de frequency recall vraag verschillen met die van het Kijkonderzoek?

ICE biedt een alternatief voor het gebruik van de frequency recall vraag. In de paragrafen 6 en verder wordt de werking van de ICE modellen toegelicht aan de hand van de Bacardi case en zal de toepasbaarheid van de ICE variabelen worden vergeleken met de frequency recall vragen. Allereerst zal in de volgende paragraaf worden ingaan op het principe en gebruik van fusie, een techniek waaraan de ICE modellen ten grondslag liggen.

5. FUSIE IN DE PRAKTIJK

Het fuseren van onderzoeksbestanden is het toekennen van gegevens van respondenten uit een donorbestand aan vergelijkbare respondenten in een ander “recipiënt” onderzoek door deze te matchen op gemeenschappelijke kenmerken (de koppelvragen) uit beide bestanden. In Nederland is het gebruik van fusie vrij zeldzaam, Bronner (1989) spreekt in dit verband zelfs over een fusie-fobie. Ook in het buitenland wordt fusie moeizaam geaccepteerd en er is slechts een beperkt aantal gepubliceerde succesvolle fusieprojecten (O’Brien, Harris en Baker, 1997).

In deze paragraaf zullen kort een aantal verschillende fusietechnieken worden bespro-

ken, wordt ingegaan op de voorwaarden en beperkingen voor het gebruik van deze techniek en worden een aantal toepassingen genoemd van het gebruik van fusie in media onderzoek in binnen- en buitenland.

Fusie kan op verschillende manieren worden uitgevoerd (voor een overzicht: zie Bock en Jephcott, 1998). Traditioneel worden bij fusie individuele observaties uit twee steekproeven in een één op één relatie paarsgewijs aan elkaar gematched, aan elke case uit het recipientbestand wordt één case uit het donor bestand toegekend (nearest neighbour methode). Een Nederlands voorbeeld hiervan is de gastprocedure waarbij kijkgedrag van gasten aan panelleden van het Kijkonderzoek wordt toegekend (Kalfs en van Meurs, 2003). Een minder transparante methode is predictive modelling (Weiss en Indurkha, 1998) waarbij een recipiënt niet één donor heeft, maar waarbij de te fuseren kenmerken vanuit het donorbestand in een statistisch model worden geschat uit de gemeenschappelijke kenmerken. Dit verklarende model wordt vervolgens gerepliceerd met dezelfde gemeenschappelijke kenmerken in de recipiënt database om in dit bestand het te fuseren kenmerk te schatten. Deze fusie techniek wordt ook bij ICE toegepast: de kijkfrequentie van panelleden uit het Kijkonderzoek wordt op basis van gemeenschappelijke kenmerken aan vergelijkbare respondenten uit het trackingsonderzoek toegekend. Predictive modeling is een techniek die met name wordt gebruikt in database marketing, data mining, direct marketing en onderzoek ten behoeve van financiële dienstverlening.

Een ander onderscheid dat gemaakt kan worden tussen verschillende fusietechnieken is het aantal gemeenschappelijke kenmerken waarop wordt gefuseerd (de "lijm" tussen de samen te voegen bestanden) en de wijze waarop deze kenmerken worden gekozen. Een fusiemethode met slechts één gemeenschappelijke kenmerk (bijvoorbeeld de postcode) betreft de ook in Nederland met name voor brievenbusreclame en direct marketing toegepaste geomarketing, waarbij gegevens per postcodegebied worden samengevoegd. Meestal wordt echter op meer dan één kenmerk gefuseerd. Voor een succesvolle fusie dienen de gemeenschappelijke kenmerken uit het donorbestand het te fuseren kenmerk zo volledig mogelijk te verklaren.

Bij fusies kan tevens onderscheid worden gemaakt tussen generieke en specifieke fusies. Bij een generieke fusie wordt vaak uit praktische overwegingen voor een groot aantal te fuseren gegevens dezelfde algemene set gemeenschappelijke kenmerken gekozen zoals bijvoorbeeld een selectie van de meest relevant geachte achtergrondkenmerken. Maar hoe algemener de fusie, des te minder waarschijnlijk dat deze voor een specifiek doel een optimale oplossing geeft. Bij een specifieke fusie of 'fusion on the fly' (McDonald, 2004) waarbij de gemeenschappelijke kenmerken speciaal gekozen zijn voor één te doneren kenmerk zal deze samenhang over het algemeen groter zijn dan bij generieke fusies.

Een speciale vorm van fusie is imputatie waarbij missende waarden binnen één steekproef worden geschat (Rubin, 1987), waarbij gegevens van donoren worden toegekend aan recipienten binnen dezelfde steekproef. Deze techniek wordt in Nederland in grote onderzoeksprojecten van met name de rijksoverheid gebruikt. Door het toepassen van imputatie is het ook mogelijk om ondanks beperkingen in de vragenlijstduur voor respondenten ontbrekende gegevens te schatten. Van 1999 t/m 2001 kon het aantal

titels in de SummoScanner worden vergroot door de vragenlijst te splitsen en het bereik van ontbrekende titels met imputatie te schatten (Van der Noort en Tchaoussoglou, 1999).

Een beperking van alle soorten fusie is dat de sterkte van het verband van het te fuseren kenmerk tussen het donor- en recipiëntbestand kenmerk altijd wordt onderschat (regressie naar het gemiddelde). Een bij het toepassen van fusie bij mediaonderzoek optredend probleem is verder dat bij bijvoorbeeld het fuseren van kijkgegevens uit het Kijkonderzoek of van leesgedrag uit het NOM het niveau als standaard of "industry currency" alleen met speciale technieken na fusie kan worden behouden.

Elms (1997) noemt succesvolle fusieprojecten met mediaonderzoek in vijf Europese landen (België, Frankrijk, Italië, Spanje en Groot-Brittannië) waarbij productgebruik aan media onderzoek wordt gekoppeld. Het Britse voorbeeld betreft de fusie tussen het BARB Kijkonderzoek en het TGI doelgroepenonderzoek waarbij sinds 1991 productgebruik aan kijkmeterdata wordt gekoppeld (Roberts, 1996). Sinds de oprichting van het TN Sofres MediaSpan single-source kijkmeterpanel is deze fusie echter deels achterhaald. In het Amerikaanse MARS/NTI project wordt de Nielsen Television Index gefuseerd met het MARS OTC/DTC farmaceutische onderzoek (McDonald, 2004).

Ten behoeve van een zo optimaal mogelijke kwaliteit van de fusie wordt bij Integrated Campaign Evaluation voor elke campagne een specifieke set koppelvragen geschat (zie paragraaf 8) waarbij naast algemene achtergrondkenmerken zoveel mogelijk gegevens over het mediagedrag worden gebruikt. Op deze wijze wordt een optimale samenhang verkregen tussen het te fuseren kenmerk (hier het aantal in het Kijkonderzoek geregistreerde contacten met een reclamecampagne) met de gemeenschappelijke kenmerken.

6. ICE IN DE PRAKTIJK: TELEVISIE EN DE ANDERE MEDIA

Een alternatief voor de frequency recall vraag is niet om direct naar de herkenning van of het aantal contacten met een campagne te vragen, maar om indirect te vragen naar het contact met de media waarin deze campagne heeft gelopen. Maar zelfs voor hele breed ingezette campagnes hoeft bijvoorbeeld de vraag "Hoe vaak kijkt u televisie?" niet direct inzicht te bieden in het aantal contacten met de campagne op televisie. Om er achter te komen met welke vragen het aantal contacten met de Bacardi campagne het beste kan worden voorspeld is hiervoor door middel van regressie een verklarend model opgesteld.

Voor een steekproef van 1.922 respondenten van 13 jaar en ouder uit het Nederlandse Kijkonderzoek is het aantal contacten met de campagne voor Bacardi vastgesteld. Om dit kenmerk te kunnen fuseren in de steekproef van het Bacardi trackingsonderzoek dienen gemeenschappelijke kenmerken te worden gekozen uit het Kijkonderzoek. Hiertoe zijn allereerst de in de jaarlijks herhaalde Basisondervraging van het Kijkonderzoek beschikbare achtergrondkenmerken en onder meer de interesse in een aantal verschillende programmagenres geselecteerd. Verder is in het Kijkonderzoek over januari t/m maart 2005 de kijkfrequentie vastgesteld gedurende 5 verschillende tijdvakken en voor de 9 best bekeken televisiezenders. Door middel van stepwise multiple regression is

geschat in welke mate en met welk van de beschikbare kenmerken het contact met de campagne van Bacardi kan worden geschat. Normaliter gebeurt dit alleen voor het aantal contacten met de campagne, maar het bereik en de contactfrequentie van de campagne voor Bacardi is dermate hoog dat een model ter verklaring van het aantal contacten met de campagne vooral individuele verschillen tussen een hoge en een zeer hoge contactfrequentie (ongeveer 5% van de steekproef had 30 of meer contacten met de campagne) lijkt te verklaren. Daarom is ook een tweede model ontwikkeld waarbij de afhankelijke variabele is gedefinieerd of men wel of niet 3 of meer contacten heeft gehad met de campagne. Een uitgangspunt voor deze variabele is de stelling dat de optimale contactfrequentie van een campagne drie contacten is binnen de aankoopcyclus en dat het verdere aantal contacten vanaf die drempelwaarde relatief minder effect heeft (Krugman, 1972; Naples, 1979)².

In tabel 2 worden het contactfrequentiemodel en het 3 contactenmodel voor het contacten met de campagne van Bacardi verklaard met respectievelijk 5 en 6 onafhankelijke variabelen. De beta coëfficiënten geven de mate aan waarmee de betreffende variabelen bijdragen aan de verklaring van de afhankelijke variabelen. De verklaarde variantie (R^2) van het contactfrequentiemodel (CF) en het 3 contactenmodel (R3+) is respectievelijk 77% en 25%³.

Tabel 2. Regressie model (beta coëfficiënten) ter verklaring van het contact met Bacardi (januari-februari 2005, $n=1.922$).

	CF	R3+
Kijktijd RTL4	0,402	0,208
Kijktijd SBS6	0,350	0,154
Kijktijd Yorin	0,087	0,093
Kijktijd Net5	0,266	0,092
Kijktijd Veronica	0,233	0,105
Interesse in speelfilms		0,158
R^2	77,0%	24,5%

CF: Contactfrequentie met campagne Bacardi

R3+: Bereik 3 contacten of meer met campagne Bacardi

Bron: Kijkonderzoek SKO/Intomart GfK

Alle significanties beta < 0,05

In de twee regressiemodellen in tabel 2 bleken getoetste achtergrondkenmerken als leeftijd, geslacht, regio en opleidingsgraad geen bijdrage te leveren aan de verklaring van het contact met de campagne van Bacardi. Bij de fusie van het contact met de Bacardi campagne worden als gemeenschappelijke kenmerken de in de tabel opgenomen media imperatives gebruikt: algemene inschattingen van de mate van het gebruik van televisie per zender en genre (speelfilms).

De ervaring met de toepassing van ICE voor televisiecampagnes heeft geleerd dat de mate waarin het contact met de campagne kan worden voorspeld op basis van de gebruikte achtergrondkenmerken en media imperatives sterk verschilt. Het contact met een campagne die alleen op specifieke zenders, gedurende speciale tijdvakken of alleen rond bepaalde programmagenres is uitgezonden kan beter worden ingeschat dan meer

generieke campagnes. Het aantal contacten met Bacardi blijkt goed verklaarbaar te zijn. De kans op drie of meer contacten is moeilijker te verklaren, mede ook door het zeer grote percentage van respondenten met drie of meer contacten (82%).

Om in de steekproef van de wekelijkse campagnetracking van Bacardi een schatting te maken van het aantal contacten met deze campagne op televisie en de kans dat men de campagne drie maal of meer heeft gezien zijn de in de regressiemodellen gevonden gemeenschappelijke kenmerken opgenomen in de vragenlijst van de campagnetracking van Bacardi. Hiertoe zijn vragen gesteld over het kijkgedrag per zender, tijdvak en per programmagenre. Daarnaast zijn ook media imperatives gevraagd voor andere media die ook in de campagne van Bacardi waren ingezet:

- leesfrequentie van kranten en tijdschriften;
- luisterfrequentie van radiozenders;
- inschatting van de mate waarin men buitenreclame ziet;
- bezoekfrequentie bioscoop.

Deze media imperative variabelen zijn begin april 2005 in een aanvullende vragenlijst voorgelegd aan alle 1.304 respondenten die in de periode januari tot en met maart 2005 in de campagnetracking van Bacardi hebben geparticipeerd. Van 1.055 (81%) konden op deze wijze de media imperatives die benodigd zijn voor ICE worden verzameld.

In de volgende tabel wordt een vergelijking gemaakt tussen de op media imperatives gebaseerde ICE variabelen en de frequency recall vraag “Waar, in welke media, heeft u reclame voor Bacardi (Witte Rum) gezien of gehoord?” uit de vragenlijst van de campagnetracking.

Tabel 3. Correlaties tussen de frequency recall vraag uit de campagnetracking en de ICE variabelen.

Televisie	0,067 *
Bioscoop	0,205 **
Buitenreclame	-
Print	-
Radio	0,064 *

** : sig < 0,01 * : sig < 0,05 - : sig > 0,05

De samenhang tussen de uitkomsten van de frequency recall vraag uit de campagnetracking en de op basis van media imperatives geschatte ICE variabelen is opmerkelijk laag. Bij bioscoopreclame is er nog enige samenhang tussen het herinneren de commercial voor Bacardi in de bioscoop te hebben gezien en de uitkomsten van de media imperatives vraag “Hoe vaak gaat u gemiddeld naar de bioscoop?” (correlatie 0,205), maar bij de andere media is er niet of nauwelijks sprake van enige samenhang.

Dit gebrek aan samenhang tussen de herinnering de campagne van Bacardi in een bepaald medium te hebben gezien en een algemene vraag naar de mate van contact met dat medium lijkt erop te duiden dat de frequency recall vraag inderdaad onbetrouwbare antwoorden geeft. In de volgende paragraaf zal worden nagegaan welke informatie

beter bruikbaar is bij de interpretatie en verklaring van campagne effecten: de frequency recall vragen of de ICE-variabelen.

7. VERKLARING VAN CAMPAGNE EFFECTEN: DE TOEGEVOEGDE WAARDE VAN ICE

In welke mate leveren de op basis van media imperatives geschatte ICE variabelen een bijdrage aan de verklaring van de effecten van de campagne van Bacardi? Om op deze vraag een antwoord te geven is (opnieuw met regressie) geschat in welke mate diverse campagne-effecten van Bacardi verklaard kunnen worden.

Omdat een campagne verschillende effecten kan bewerkstelligen is effect hierbij door middel van de volgende 10 variabelen geoperationaliseerd:

- Drink Drinken van Bacardi
- SMB Spontane merkbekendheid Bacardi
- GMG Geholpen merkbekendheid Bacardi
- SRB Spontane reclamebekendheid Bacardi⁴
- Stijlvol Bacardi is een stijlvol merk
- Trots Bacardi is een merk waarvan je trots (blij) zou zijn het te bestellen aan de bar
- Gezien Bacardi is een goede drank om mee gezien te worden
- Jong Bacardi is een drank voor jongeren
- Vriend Bacardi is een drank waar je met vrienden van geniet
- Populair Bacardi wordt steeds populairder

De 8 imagostatements (5 punt schaal helemaal niet van toepassing t/m zeer van toepassing) zijn alleen beschikbaar voor een random selectie van 25% van de steekproef. In tabel 4 staat de steekproefgrootte uit het trackingsonderzoek vermeld voor elk van de bovenstaande 10 effectvariabelen.

In tabel 4 worden de R^2 en bèta's weergegeven van een verklarend model voor elk van de 10 effectvariabelen. In de modellen zijn deze verklaard door middel van de volgende variabelen:

- achtergrondkenmerken leeftijd, geslacht en schoolopleiding, de week van het interview en het drinken van Bacardi (de bèta's voor deze variabelen zijn niet in de tabel opgenomen);
- de media waarin men de campagne denkt te hebben gezien (zie tabel 1 voor een volledig overzicht van deze variabelen);
- de op media imperatives gebaseerde ICE variabelen voor televisie, radio, print en bioscoop.

De verklaarde variantie van de 10 modellen in tabel 4 is laag. Alleen de spontane reclamebekendheid kan redelijk goed worden verklaard ($R^2 = 24,7\%$). Waarom een respondent Bacardi drinkt, zich het merk en/of de campagne herinnert en welk imago men aan Bacardi toekent is maar in geringe mate te verklaren door de verklarende variabelen

Tabel 4. Gebruik frequency recall vragen en ICE variabelen per medium ter verklaring effectvariabelen campagne Bacardi (beta coëfficiënten regressie).

	Drink	SMB	GMG	SRB	Stijlvol	Trots	Gezien	Jong	Vriend	Popu- lair
n	1.181	1.304	261	256	245	248	240	258	253	221
R ²	10,8%	11,6%	5,8%	24,7%	7,3%	6,3%	6,9%	13,1%	17,2%	12,8%
Reclame										
gezien: TV	0,178	0,128	0,225	0,452				0,186	0,139	0,148
Reclame gezien in horeca					0,146				0,157	
ICE TV: 3+	0,108		0,071							0,174
ICE print	0,096									
ICE cinema			0,095		0,198	0,196	0,155			
ICE radio	0,138	0,067								

De modellen zijn gecontroleerd voor leeftijd, geslacht, opleiding, de week van het interview en het gebruik van Bacardi. De bèta's voor deze variabelen zijn niet in de tabel opgenomen.

Alle significanties beta < 0,05

(het drinken van Bacardi, de drie achtergrondkenmerken, het moment van ondervraging gedurende de campagne en het contact met de reclamecampagne van Bacardi).

In verband met de lage verklaarde variantie kunnen alleen met enige voorzichtigheid globale conclusies worden ontleend aan de 10 verklarende modellen uit tabel 4. We zullen ons in de volgende paragraaf daarom beperken tot het beantwoorden van de vraag in welke mate de ICE variabelen bruikbaar zijn in de 10 campagne-effectmodellen in vergelijking met de frequency recall vragen.

8. UITKOMSTEN VAN HET EXPERIMENT

De doelstelling van deze paper was aan de hand van de campagne van Bacardi aan te tonen in hoeverre de op media imperatives gebaseerde ICE variabelen een bijdrage leveren aan het verklaren van diverse effecten van de campagne voor dit merk. Geconcludeerd kan worden dat ICE hierin inderdaad een bijdrage levert. In 8 van de 10 modellen worden de campagne effectvariabelen mede verklaard door de ICE variabelen.

De ICE variabelen blijken geen volledig alternatief te bieden voor de frequency recall vragen (de mate waarin men zich kan herinneren reclame voor Bacardi te hebben gezien in de diverse media). In 8 van de 10 modellen blijken deze frequency recall vragen een rol te spelen, namelijk in hoeverre men zich kan herinneren reclame voor Bacardi te hebben gezien op televisie (in 7 modellen) en in horecagelegenheden (in twee modellen). Het sterkste effect hierbij is de relatie tussen de spontane reclamebekendheid van Bacardi en de mate waarin men zich kan herinneren een televisie-commercial voor dit merk op televisie gezien te hebben. Deze relatie is deels te beredeneren uit de opbouw van de vragenlijst⁵. Maar ook de zeer hoge contactfrequentie op televisie van de campagne voor Bacardi zal zeker hebben bijgedragen aan de betrouwbaarheid van de herinnering van de respondent of men de campagne heeft

gezien op dit specifieke medium. Voor andere media en andere campagnes zal dit zeker niet altijd het geval zijn. Maar zelfs voor deze zeer zware televisiecampagne blijkt de herinnering van de steekproef haar beperkingen te hebben. Het drinken van Bacardi, de geholpen merkbekendheid en de mate waarin men denkt dat Bacardi steeds populairder wordt blijkt niet alleen samen te hangen met de mate waarin men zich kan herinneren de commercial op televisie gezien te hebben (de frequency recall vragen), maar ook met de mate waarin men over het algemeen naar de zenders zegt te kijken waarop deze commercial is uitgezonden (de ICE-variabele).

Een opvallende uitkomst is dat de frequency recall vraag voor televisie sterker samenhangt met het drinken van Bacardi dan de ICE variabele. Een mogelijke verklaring hiervoor is dat gebruikers van Bacardi het aantal contacten met de televisiecampagne in de frequency recall vraag eerder overschatten dan niet-gebruikers van Bacardi.

In vier van de tien modellen blijkt de bezoekfrequentie van bioscoop één van de variabelen te zijn waarmee de campagne effectvariabelen kunnen worden verklaard, en in drie modellen blijkt de ICE variabele "Bereik 3 contacten of meer met campagne Bacardi" een rol te spelen.

Ook de mate waarin aangeeft de bioscoop te bezoeken, de tijdschriften te lezen en de radiozenders te beluisteren waarin Bacardi heeft geadverteerd gedurende de campagneperiode blijkt samen te hangen met het drinken van het merk, het spontaan herinneren van het merk, de reclame hiervoor en vijf imagostatements van Bacardi. Bij het verklaren van de drinkfrequentie van Bacardi en de spontane merkbekendheid speelt tevens de ICE radiovariabele een rol.

Voor het contact met reclame voor Bacardi in horecagelegenheden zou in het vervolg eveneens een ICE vraag kunnen worden gesteld. In plaats van de respondent te vragen in hoeverre men zich kan herinneren reclame voor Bacardi te hebben gezien op gezien in horecagelegenheden zou kunnen worden gevraagd naar de bezoekfrequentie van horecagelegenheden.

9. CONCLUSIE

De ICE variabelen voor televisie, print, bioscoop en radio blijken een goede bijdrage te leveren aan het verklaren van diverse aspecten van de effectiviteit van de campagne van Bacardi. Gezien de lage verklaarde variantie bieden de ICE variabelen een wellicht bescheiden maar desondanks praktische aanvulling op het bestaande onderzoeksinstrumentarium.

Het blijkt dat we de herinnering van respondenten in welke media ze een campagne hebben gezien niet altijd kunnen vertrouwen. Een dergelijke vraag kan haar waarde hebben, maar het verdient aanbeveling om vragenlijsten van effectiviteitsonderzoek aan te vullen met een paar algemene vragen naar het gebruik van de in de betreffende campagne ingezette media. Deze vragen leveren ondersteund met analyses op gegevens uit standaard mediabereiksonderzoek een bruikbare bijdrage aan het verklaren van de effectiviteit van reclamecampagnes.

NOTEN

1. Het Kijkonderzoek wordt in opdracht van Stichting Kijkonderzoek uitgevoerd door Intomart GfK en TV Times.
2. De discussie omtrent de effectieve contactfrequentie voert uiteraard verder dan alleen het aantal contacten binnen de aankoopfrequentie (zie bijvoorbeeld Dun en Bremer, 1997; Faasse & Neijens, 2003). In het kader van dit experiment beperken we ons hier echter tot het meest gebruikte alternatief voor één of meer contacten met een campagne, namelijk drie of meer contacten.
3. De R^2 was nog te verhogen door meer verklarende variabelen in het model op te nemen (respectievelijk maximaal 79% en 27%), maar gekozen is voor deze twee "zuinigere" modellen.
4. De Spontane Advertising Awareness was voor de campagne van Bacardi zo hoog dat deze variabele te weinig variantie had om verder te kunnen analyseren.
5. Dit sterke verband is mede een gevolg van de routing van de vragenlijst: wie spontaan of geholpen geen advertising awareness heeft wordt niet gevraagd in welke media men zich kan herinneren reclame voor Bacardi gezien te hebben: deze groep kan zich immers geen reclame voor Bacardi herinneren.

LITERATUUR

- Bock, T. & Jephcott, J. (1998) The Application and Validation of Data Fusion. *International Journal of Market Research*, 40 (3).
- Boon, A.K. den. (2003) Bereik van media en reclame. In: A.K. den Boon & P.C. Neijens (Red.), *Media & Reclame*, 4e druk (pp 199-232). Groningen: Wolters-Noordhoff.
- Bronner, A.E. (1989) Einde fusie-fobie in Nederland. In: *Jaarboek van de Nederlandse Vereniging van Marktonderzoekers 1988-1989*, p 9-18. Haarlem: Uitgeverij de Vrieseborch.
- Dun, W. van & Bremer, D. (1997). *Wear-in en Wear-out*. Amsterdam: SWOCC (nr 6).
- Elms, S. (1997) Linking sales tracking data to television viewing. *Admap*, 32 (4).
- Faasse, J. & Neijens, P.C. (2003). Mediastrategie. In: A.K. den Boon & P.C. Neijens (Red.), *Media & Reclame*, 4e druk (pp 233-252). Groningen: Wolters-Noordhoff.
- Franzen, G. (1998) *Merken & Reclame*. Deventer: Kluwer Bedrijfsinformatie.
- Heath, R. (2000) Low Involvement processing – a new model of brands and advertising. *International Journal of Advertising*, 19 (3).
- Kalfs, N. & Van Meurs, A. (2003). Is anybody home? Out-of-home viewing and guest viewing. In: *ESOMAR, TV Audience Measurement, LA, June 2003*.
- Krugman, H.E. (1972). Why three exposures may be enough. *Journal of Advertising Research*, 12 (6), pp 11-14.
- McDonald, C. (2004) What is fusion? *Admap*, 451, pp.12-13.
- Naples, M.J. (1979). *Effective frequency: the relationship between frequency and advertising effectiveness*. New York: Association of National Advertisers, Inc.
- Van der Noort, W. & Tchaousoglou, C. (1999). Divide and unite: Splitting the SummoScanner and data fusion. In: *Worldwide Readership Research Symposium, Florence 1999*, p 325-331.
- O'Brien, J.; Harris, P. & Baker, K. (1997) Data fusion: An appraisal and experimental evaluation. *International Journal of Market Research*, 39 (1).
- Prance, S. (1995) The mind's eye – Can television advertising be visually transferred by the medium of radio? In: *ESOMAR, Radio Research Symposium, July 1995*.
- Roberts, A. (1996) What do we know about advertising's short-term effects? *Admap*, 31 (2).
- Rubin, D.B. (1987). *Multiple imputation for nonresponse in surveys*. Brisbane: John Wiley & Sons.
- Weiss, S.M. and Indurkha, N. (1998) *Predictive Data Mining: A Practical Guide*. San Francisco, California: Morgan Kaufmann Publishers Inc.

13. *Burgersegmenten en vertrouwen in de regering*

M.C.M. ARNOLD

SAMENVATTING

In het publieke debat over de politieke situatie in Nederland worden de termen ‘burgerrevolte’ en ‘vertrouwenscrisis’ niet geschuwd. Dit doet vermoeden dat er veel aan de hand is in ons eens zo rustige en vredige land. Studies geïnitieerd na diverse politiek turbulente perioden in de afgelopen jaren hebben dit beeld echter genuanceerd. Er blijkt geen sprake van almaar toenemend wantrouwen, wel van afname van politieke onverschilligheid. Ook in vergelijking met ons omringende landen lijkt er zich in Nederland geen sterk afwijkend patroon van ontwikkelingen voor te doen. Wel zijn er groepen in de samenleving die meer vatbaar zijn voor mobilisatie van onvrede. Zorgwekkend is dat de overheid deze groepen nog maar moeizaam bereikt en dat het vertrouwen in de regering relatief laag is, aldus de Commissie Toekomst Overheidscommunicatie. Op welke wijze komt het vertrouwen in de regering in deze groepen tot stand? In hoeverre spelen beleidstevredenheid en politiek cynisme een rol, of laat vertrouwen zich vooral verklaren vanuit de toestand van de economie? Analyses op de belevingsmonitor laten duidelijke verschillen tussen burgersegmenten zien in de wijze waarop vertrouwen in de regering tot stand komt.

1. INLEIDING

Op 27 augustus 2001 heeft de Commissie Toekomst Overheidscommunicatie (Commissie Wallage) haar eindrapport aan Minister-President Kok aangeboden. Het rapport kreeg de titel “in dienst van de democratie” om het belang van communicatie in een moderne democratie in veranderende maatschappelijke omstandigheden te onderstrepen. Burgers zijn mondiger, meer individualistisch en hebben meer keuzevrijheid voor het eigen handelen. Zij hebben een minder vaste binding met medeburgers en instituties en beoordelen het overheidshandelen in toenemende mate ad-hoc. De media spelen daarbij een belangrijke rol door als voorhoede van een kritische bevolking en van belangen indringend en veelvuldig om verantwoording te vragen. Informeren is in deze sterk veranderde context onvoldoende, aldus de commissie. Het afleggen van verantwoording zal steeds belangrijker worden om het vertrouwen van burgers te winnen of te behouden. De Commissie merkt daarbij op dat “de burger” niet bestaat. Er kunnen verschillende groepen in de samenleving worden onderscheiden die in normen en waarden in de relatie tot de overheid leiden tot zeer uiteenlopende verwachtingen als het – vanuit de burger gezien – om communicatie met de overheid gaat. Bovendien kan een en dezelfde burger tot de overheid in geheel verschillende betrekkingen staan: men is kiezer, coproducent, klant en onderdaan tegelijk (Wallage 2001, P. 8). De commissie

stelt dat het vanuit democratisch oogpunt aanbeveling verdient om het perspectief van de overheid te verbreden, door actief contact te zoeken met alle burgers. Slaagt de overheid hier niet in, dan dreigen grote groepen burgers zich van de overheid af te wenden. Er zijn groepen in de samenleving die dat reeds doen (Wallage 2001, P. 30). Het rapport van de Commissie Wallage sluit af met zeventien adviezen die door de Voorlichtingsraad¹ zijn vertaald in een actieprogramma.

Een van de projecten binnen dat programma is de belevingsmonitor. De belevingsmonitor heeft tot doel om de regering inzicht te geven in de kloof tussen overheid en samenleving. Welke onderwerpen houden burgers bezig? Op welke punten verschilt de burgeragenda van de agenda van de regering? Wat vindt men van verschillende oplossingsrichtingen en hoe waardeert de bevolking beleid en communicatie? Om te kunnen werken aan het herstel van contact en om op termijn vertrouwen te herstellen is het antwoord op deze vragen van belang. In navolging van de Commissie Wallage worden in de Belevingsmonitor zes burgersegmenten onderscheiden.

In dit artikel wordt verslag gedaan van analyses naar het vertrouwen in de regering die op data van de belevingsmonitor werden uitgevoerd. Inzicht in de wijze waarop vertrouwen tot stand komt en hoe groepen in de samenleving daarin verschillen is relevant om de dynamiek in de publieke opinie beter te kunnen duiden. In de volgende paragraaf wordt beschreven welke variabelen hiervoor binnen de monitor beschikbaar zijn. Vervolgens wordt het theoretische model toegelicht en empirisch ingevuld voor het algemeen publiek. In de laatste paragraaf wordt beschreven welke burgersegmenten in de belevingsmonitor worden onderscheiden en in hoeverre zij verschillen in de wijze waarop het vertrouwen in de regering tot stand komt. Het artikel wordt afgesloten met enkele conclusies.

2. DE BELEVINGSMONITOR

In februari 2003 is de eerste meting van de belevingsmonitor gehouden. Het instrument bestond aanvankelijk uit maandelijkse kwantitatieve metingen. Kwartaalmetingen met telkens dezelfde beleidsonderwerpen, in de tussenliggende maanden werden clusters van verwante thema's meer in de breedte onderzocht. Na een tussentijdse evaluatie is de meetfrequentie teruggebracht naar eens per kwartaal en werd kwalitatief onderzoek toegevoegd om opvallende resultaten uit de kwantitatieve metingen te verdiepen en te duiden. Tevens is een media-analyse toegevoegd om meer inzicht te krijgen in de relatie tussen de agenda van burgers – wat voor burgers belangrijke onderwerpen zijn - en de media-agenda – hoeveel er in kranten geschreven wordt over beleidsonderwerpen. Voor de analyses in het kader van dit artikel werden twee kwantitatieve kwartaalmetingen samengevoegd: december 2004 en maart 2005. Hierdoor ontstond een bestand met voldoende omvang om uitsplitsingen naar burgersegmenten mogelijk te maken (n=2128, representatief voor 18 jaar en ouder). Het veldwerk werd uitgevoerd door bureau Marketresponse in samenwerking met the Smart Agent Company (voor meer informatie zie: www.regering.nl dossier belevingsmonitor).

De kwantitatieve vragenlijst van de belevingsmonitor bestaat uit verschillende blokken.

¹ Overleg waarin alle departementale directeuren voorlichting zitting hebben.

Het eerste blok bevat vragen naar vertrouwen in een aantal instellingen en onderwerpen, waaronder de Nederlandse regering en de Nederlandse economie. Vervolgens komen algemene beleidstevredenheid, tevredenheid met de regering, de houding ten aanzien van het politieke bedrijf en politici en betrokkenheid bij politiek aan bod. Een groot deel is gereserveerd voor vragen over *beleidsthema's*. Er zijn 12 beleidsthema's in de monitor opgenomen. Per beleidsthema zijn telkens vijf specifieke beleidsaspecten nader uitgevraagd. Het laatste deel van de vragenlijst bevat achtergrondvragen waaronder waardenoriëntaties. Die waardenoriëntaties liggen ten grondslag aan het BSR segmentatiemodel dat verderop wordt beschreven.

Het theoretisch model dat ten grondslag ligt aan de analyses (zie volgende paragraaf) bestaat uit negen variabelen. Vier variabelen in het model zijn gebaseerd op een enkelvoudige vraagstelling, de overige vijf zijn samengesteld op basis van meerdere schaalvragen. Hieronder volgt een overzicht van de wijze waarop die variabelen in de belevingsmonitor worden gemeten en hoe de schalen werden geconstrueerd.

Vertrouwen in de Nederlandse economie en regering (10-puntsschaal: 1 geen enkel vertrouwen – 10 zeer veel vertrouwen):

Hoeveel vertrouwen heeft u op dit moment in de volgende onderwerpen en instellingen?

- De Nederlandse economie?
- De Nederlandse regering?

Tevredenheid regering (5 puntsschaal: 1 zeer ontevreden – 5 zeer tevreden)

Hoe tevreden of ontevreden bent u over:

- De mate waarin de regering bij haar plannen rekening houdt met de belangen van burgers?
- De mate waarin de regering weet wat er in de samenleving leeft?
- De slagvaardigheid waarmee de regering te werk gaat
- Het behalen van de beoogde resultaten door de regering
- De mate waarin burgers bij de uitvoering van regeringsbeleid betrokken worden
- De informatieverstrekking door de regering over plannen en maatregelen

Cronbach's alpha: .88

Volgen van beleidsthema's in de media (3 puntsschaal: 1 niet, 2 soms, 3 vaak)

Kunt u steeds aangeven in welke mate u het aspect *bewust* volgt via radio, TV, op Internet, in kranten of anderszins?

Schaalconstructie: 5 aspecten per beleidsthema zijn samengevoegd tot totaalmaat voor per thema (somscore gedeeld door 5). De 12 geconstrueerde issuematen zijn samengevoegd tot een totaalmaat (somscore gedeeld door 12). Van de onderstaande issues zijn 5 aspecten bevraagd:

- Verkeer en vervoer
- Onderwijs en kennis
- Sociale zekerheid
- Economie en werkgelegenheid
- Financiën en koopkracht
- Criminaliteit
- Veiligheid

- Integratie & immigratie
- Relatie burger-overheid
- Zorg en welzijn
- Europa en internationaal beleid
- Natuur en milieu

Tevredenheid beleid (totaalmaat over 12 beleidsthema's gemeten op een 5 puntsschaal: zeer ontevreden – zeer tevreden)

De onderwerpen in dit hoofdstuk hadden alle betrekking op < zie *bovenstaande 12 issues* >. Kunt u tot slot van dit hoofdstuk aangeven hoe tevreden of ontevreden u bent over het beleid van de regering op dit onderwerp? Schaalconstructie: Somscore gedeeld door 12.

Politieke betrokkenheid (5 puntsschaal: 1 paste helemaal niet – 5 past heel erg)

Hieronder volgt een aantal uitspraken die te maken hebben met uw betrokkenheid bij de politiek. Kunt u aangeven in welke mate de uitspraak bij u past?

- Ik zorg ervoor dat ik goed op de hoogte blijf van het regeringsbeleid
- Ik praat met vrienden en kennissen vaak over politieke onderwerpen
- Ik ben zeer geïnteresseerd in politieke onderwerpen

Cronbach's alpha: .88

Politiek cynisme (5 puntsschaal: 1 zeer oneens – 5 zeer eens)

- Mensen zoals ik hebben geen enkele invloed op wat de regering doet (gehercodeerd)
- Ik denk niet dat kamerleden en ministers veel geven om wat mensen als ik denken (gehercodeerd)
- De meeste politici zijn bekwame mensen die weten wat ze doen
- Kamerleden letten te veel op het belang van enkele machtige groepen in plaats van op het algemeen belang (gehercodeerd)
- Tegen beter weten in beloven politici meer dan ze waar kunnen maken (gehercodeerd)
- Soms lijkt de politiek zo ingewikkeld dat mensen zoals ik niet echt begrijpen wat er speelt (gehercodeerd)

Cronbach's alpha: .76

Tabel 1. Gemiddelde scores op de (samengestelde) variabelen voor het algemeen publiek.

	gem.	SD	n
Vertrouwen in de regering (10 puntsschaal)	4,53	1,97	1994
Vertrouwen in de economie (10 puntsschaal)	5,34	1,60	1982
Tevredenheid regering (5 puntsschaal)	2,28	0,68	2003
Politiek cynisme (5 puntsschaal)	2,33	0,65	2002
Tevredenheid beleid (5 puntsschaal)	2,62	0,65	2010
Politieke betrokkenheid (5 puntsschaal)	3,19	0,92	2000
Sociaal vertrouwen (5 puntsschaal)	2,44	0,96	1988
Altruïsme (5 puntsschaal)	2,93	0,93	1962
Volgen beleidsthema's in media (3 puntsschaal)	2,22	0,38	2015

NB Alle variabelen zijn zo gecodeerd dat een hoge score positief geïnterpreteerd dient te worden. Een hoge score op politiek cynisme is bijvoorbeeld positief.

Altruïsme (5 puntsschaal: 1 paste helemaal niet – 5 past heel erg)

Kunt u aangeven in welke mate de uitspraak bij u past?

Ik voel me pas goed als ik me kan inzetten voor een gemeenschappelijk doel

Sociaal vertrouwen (5 puntsschaal: 1 zeer oneens – 5 zeer eens)

Kunt u aangeven in hoeverre u het met onderstaande uitspraak eens bent?

- Tegenwoordig weet je echt niet meer op wie of wat je nog kunt vertrouwen (geherco-deerd)

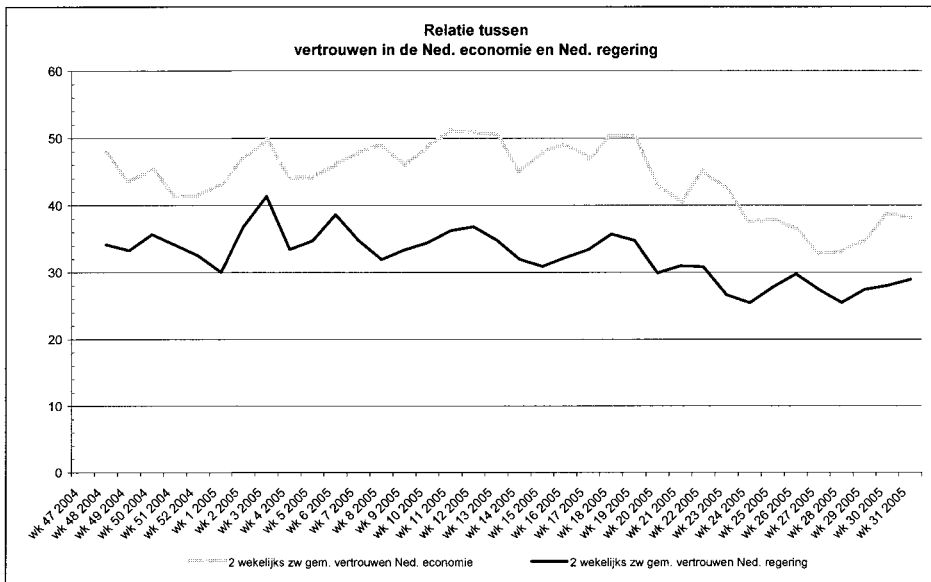
3. BESCHRIJVING THEORETISCH MODEL

De afhankelijke variabele in het model is het vertrouwen in de regering. Over het belang van vertrouwen en over de specifieke definitie als het gaat om de overheid en/of de regering bestaat veel literatuur. We laten een uitgebreide beschouwing daarvan achterwege. Bij het belang van vertrouwen willen we wel de volgende opmerking maken. Hoewel er erg veel te doen is over het vertrouwen in de regering is het de vraag in hoeverre het een noodzakelijke voorwaarde is voor het functioneren en overleven van een politiek systeem. Daarover bestaan twee tegengestelde visies. Aan de ene kant staan zij die vertrouwen in de overheid en interpersoonlijk vertrouwen als noodzakelijk beschouwen en deze twee dan ook met elkaar verbinden. Aan de andere kant staan zij die wantrouwen in de overheid zien als een uiting van een gezonde democratische attitude. Dit veronderstelt dat vertrouwen in de regering problematisch is als het zeer laag is, maar ook als het zeer hoog is. Volgens sommigen is 50% vertrouwen in de regering een mooi streefcijfer.

In het model wordt van de overige acht variabelen verondersteld dat ze een deel van het vertrouwen in de regering kunnen verklaren, hetzij direct, hetzij indirect door de tussenkomst van één of meerdere andere variabele(n). Het analysemodel is gebaseerd op veronderstellingen over de richting van de causaliteit. Als men echter de literatuur over vertrouwen in de overheid en de regering in beschouwing neemt bestaat over de richting van die causaliteit niet altijd evenveel eenduidigheid. Daarom zijn keuzes gemaakt die aansluiten bij bepaalde visies hierover. Elke keuze die daarbij gemaakt wordt kan dan ook ter discussie worden gesteld. Zonder naar volledigheid te streven beperken we ons tot verwijzingen die aansluiten bij het model dat voor dit artikel werd geconstrueerd.

Laten we beginnen met de relatie tussen het vertrouwen in de Nederlandse economie en het vertrouwen in de regering. Deze relatie is uitgebreid bestudeerd binnen de politieke economie. Frey en Schneider beschrijven in hun model de samenhang van politiek en economie als een kringloop. De ene helft van de kringloop heeft betrekking op de invloed van overheidsbeleidsinvesteringen op de economie (policy function). De andere helft van de kringloop heeft betrekking op de relatie tussen de economische toestand en de populariteit van de regering en de beoordeling van het politieke bedrijf (evaluation function) (van Winden, 1987: 65-66). De 'government overload' hypothese is hieraan sterk gerelateerd. Deze stelt dat het voor de overheid nagenoeg onmogelijk is om naar wens te presteren omdat aan de verwachtingen ten aanzien van de overheid niet kan worden voldaan. Dit vormt het uitgangspunt in het model. Vertrouwen in de

economie wordt opgenomen als variabele die slechts verklaard wordt uit de achtergrondkenmerken “altruïsme” en “sociaal vertrouwen”. Op basis van resultaten van eerdere studies wordt een vrij sterke relatie verwacht tussen vertrouwen in de economie, tevredenheid en vertrouwen in de regering. Deze veronderstelling is mede gebaseerd op de onderstaande grafiek die de samenhang tussen beide maten van vertrouwen goed zichtbaar maakt. De gegevens zijn afkomstig uit het continue effectonderzoek naar Postbus 51 campagnes (trackingonderzoek).



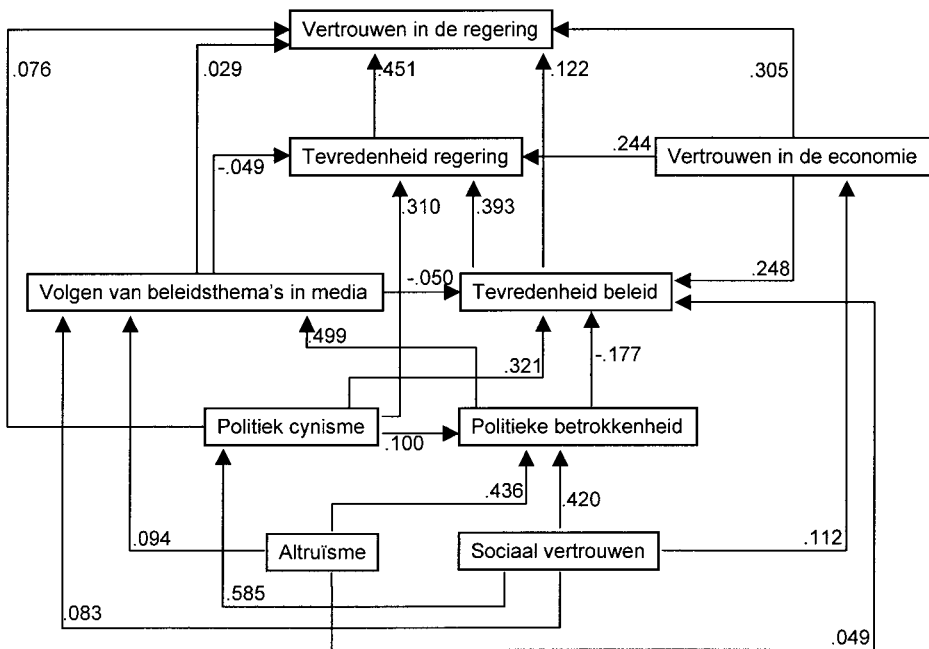
Figuur 1. De gegevens in de grafiek zijn afkomstig uit het trackingonderzoek naar Postbus51 campagnes. De punten in de grafiek zijn samengesteld op basis van het twee wekelijks zwevend gemiddelde (n circa 200 per week) ($r=.61$).

Tevredenheid met de regering kan worden opgevat als de uitkomst van een deels rationele en deels emotionele evaluatie. Aan die evaluatie ligt kennis ten grondslag die voor een belangrijk deel uit de media wordt verkregen. We veronderstellen dan ook een relatie tussen de mate waarin beleidsthema's bewust worden gevolgd in de media en de tevredenheid met de regering. Daarnaast wordt verondersteld dat eigen observaties en ervaring de tevredenheid met beleid en daarmee de tevredenheid met de regering beïnvloeden. Informatie uit de media over succes en falen en steun en kritiek op beleid blijkt eveneens in bepaalde omstandigheden de tevredenheid met beleid te bepalen. Op basis van de media analyse veronderstellen we tevens een relatie tussen het volgen van de media en opvattingen over politici en het politieke bedrijf, in het model samengenomen onder de noemer politiek cynisme. Verondersteld mag worden dat politiek cynisme van invloed is op de politieke betrokkenheid. Betrokkenheid bij politiek is niet erg bevredigend als men zeer cynisch is over de openheid en oprechtheid van politici.

De rol van twee variabelen is nog niet volledig besproken. Altruïsme en sociaal vertrouwen. Deze variabelen zijn opgenomen als exogene variabelen (worden niet beïn-

vloed door andere variabelen in het model). Beide variabelen zijn indicatoren voor sociaal kapitaal. Sociaal kapitaal kan worden gezien als een veralgemenisering van vertrouwen tussen mensen: van specifiek ‘particularized’ vertrouwen in de eigen groep via algemeen ‘generalized’ vertrouwen in onbekende medeburgers gaat men over tot vertrouwen in hun anonieme voortbrengselen en verbanden (Dekker, 2001). Verwacht wordt dat beide variabelen een bijdrage leveren aan het model, hetzij direct of indirect vooral via politieke betrokkenheid en politiek cynisme.

Tot zover de theoretische beschouwing. Hieronder volgt een grafische weergave van het padmodel. Alleen significante relaties zijn aangegeven met een pijl en bijbehorende regressiecoëfficiënt. Niet significante relaties zijn buiten beschouwing gelaten omdat deze de inzichtelijkheid van de presentatie teveel geweld zou aandoen.



Figuur 2. Met het bovenstaande model wordt 59% van de variantie in regeringsvertrouwen verklaard.

Het vertrouwen in de regering wordt hoofdzakelijk bepaald door de tevredenheid met de regering. De tevredenheid met de regering wordt respectievelijk bepaald door tevredenheid met beleid, politiek cynisme, vertrouwen in de economie en in negatieve zin door het volgen van beleidsthema's in de media. Naarmate de berichtgeving over beleidsonderwerpen beter wordt gevolgd is men minder tevreden met de regering. Ook is men dan minder tevreden met beleid. Opvallend is dat het volgen van beleidsthema's in de media wel rechtstreeks een positief effect heeft op het vertrouwen in de regering. Dit rechtstreekse positieve effect compenseert de negatieve indirecte effecten echter niet (dit verschijnsel is een gevolg van vrij sterke maar net niet problematische samen-

hang tussen de onafhankelijk variabelen). Het vertrouwen in de economie bepaalt vervolgens het grootste deel van het vertrouwen in de regering. Het totaal van directe en indirecte effecten (via tevredenheid beleid en via tevredenheid regering) is maar net iets minder sterk dan het rechtstreekse effect van de tevredenheid met de regering (0,445 versus 0,451). Het lage vertrouwen in de economie heeft negatieve consequenties voor oordelen over beleid en de regering. Op de derde plaats levert tevredenheid met beleid een bijdrage aan het vertrouwen in de regering. De tevredenheid met beleid is groter naarmate men minder politiek cynisch is en lager onder politiek betrokkenen. Politieke betrokkenheid is onder andere geoperationaliseerd met de vraag hoe vaak men met vrienden en kennissen praat over politieke onderwerpen. Op basis van de gevonden negatieve relatie met beleidstevredenheid mag worden verondersteld dat die gesprekken hoofdzakelijk een negatief karakter hebben. Politiek cynisme heeft slechts een zwak rechtstreeks effect op vertrouwen in de regering, het indirecte effect is echter aanzienlijk. Politieke betrokkenheid levert uitsluitend indirect een negatieve bijdrage aan regeringsvertrouwen. Dit kan mogelijk worden verklaard doordat politiek betrokkenen verwachtingen hebben waaraan de regering niet voldoet of niet kan voldoen. Politieke betrokkenheid leidt tevens tot het meer bewust volgen van beleidsthema's in de media. Opvallend is dat de veronderstelde relatie tussen het bewust volgen van de media en politiek cynisme zich niet voor blijkt te doen. De variabelen die sociaal kapitaal representeren leveren indirect een bijdrage aan regeringsvertrouwen. Sociaal vertrouwen hangt in positieve zin vrij sterk samen met politiek cynisme, politieke betrokkenheid, vertrouwen in de economie en het volgen van beleidsthema's in de media. De neiging tot altruïstisch gedrag bevordert politieke betrokkenheid en de tevredenheid met beleid.

Al met al geeft het model goed inzicht in directe en indirecte effecten op regeringsvertrouwen. De op basis van theoretische inzichten veronderstelde relaties blijken empirisch te kunnen worden onderbouwd, met uitzondering van de relatie tussen het volgen van beleidsthema's in de media en politiek cynisme. Politiek cynisch zijn die groepen in de samenleving die sowieso weinig sociaal vertrouwen hebben. In de volgende paragraaf wordt het model ingevuld voor de verschillende BSR segmenten.

4. VERKLARING VAN VERTROUWEN IN DE REGERING VOOR ZES BURGERSEGMENTEN

In de vragenlijst is een blok opgenomen met waardenoriëntaties en leefstijlkenmerken. Binnen het domein 'overheidsbeleving' wordt op basis van dit vraagblok onderscheid gemaakt tussen zes groepen burgers die zich kenmerken door overeenkomstige waardenpatronen (BSR model). De correspondentieanalyse die ten grondslag ligt aan deze zes clusters bestaat uit een tweetal dimensies, een sociologische dimensie en een psychologische dimensie. Alvorens in te gaan op de analyse van het padmodel wordt eerst het segmentatiemodel samengevat. Deze samenvatting is gebaseerd op een uitgebreide beschrijving van de segmenten waarover door Marketresponse een afzonderlijke rapportage is opgesteld (burgersegmentatie overheidsbeleving, Marketresponse: 2003).

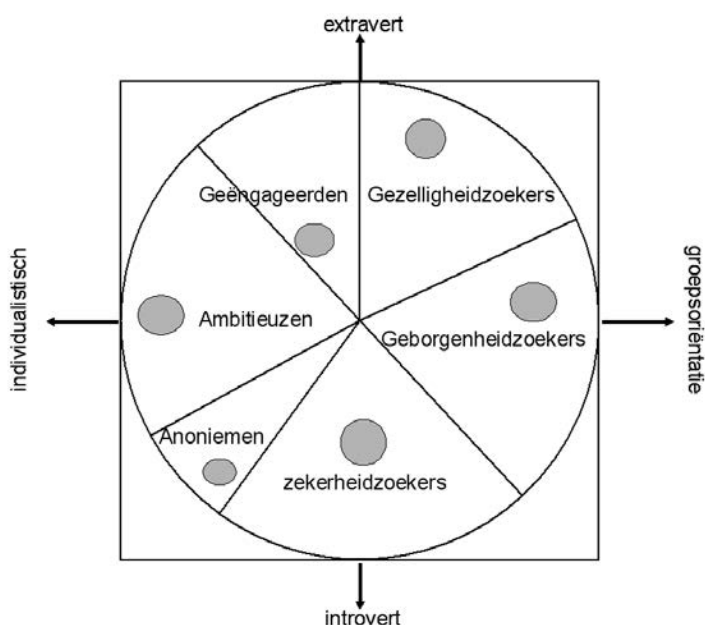
De sociologische dimensie van het BSR model

De sociologische dimensie laat een tegenstelling zien tussen ego-oriëntatie en een groepsoriëntatie. De ego-oriëntatie wordt met name bepaald door een zakelijke, carriè-

regerelateerde oriëntatie. In houdingen ten opzichte van de overheid komt deze zakelijke instelling tot uiting, bijvoorbeeld in uitspraken als: ‘de overheid moet veel efficiënter werken’ en ‘de overheid moet zich vooral bezig houden met een solide economische basis’. Het is met name dit uiteinde van de dimensie waar betrokkenheid bij de landelijke politiek groot is. De groepsoriëntatie binnen het model wordt in belangrijke mate bepaald door karaktergerelateerde kenmerken als ‘verlegen’, ‘zachtaardig’ en ‘behulpzaam’. De gewenste rol van de overheid komt tot uiting in uitspraken als ‘de overheid moet ervoor zorgen dat bepaalde groepen burgers niet buiten de boot vallen’ en ‘overheidsbeleid moet vooral gericht zijn op het vergroten van onderlinge betrokkenheid, hulpvaardigheid, solidariteit en onderling respect’. De betrokkenheid bij (landelijke) politiek is aanzienlijk lager dan onder meer ego-georiënteerden.

De psychologische dimensie van het BSR model

De psychologische dimensie laat een tegenstelling zien tussen een extraverte en introverte oriëntatie. De extraverte oriëntatie kan getypeerd worden door karakterkenmerken als enthousiast, spontaan, vlot, vrolijk en hartstochtelijk. Interesse in of betrokkenheid bij landelijke en lokale politiek is hier gering. De rol die de overheid wordt toegekend is primair een ‘stimulerende’ rol. De introverte oriëntatie binnen het model wordt vooral getypeerd door kenmerken die wijzen op isolement: niet geschikt voor het gezinsleven, geïsoleerd gezin en werkloos. Dit isolement komt ook naar voren in de houding ten opzichte van de overheid: ‘de overheid moet zich niet bemoeien met privé-zaken van burgers’ en ‘de overheid moet de privacy van de burgers meer respecteren’. De betrokkenheid bij en interesse in politiek is overigens gemiddeld. In dit assenstelsel kunnen de volgende groepen worden geplaatst:



Figuur 3. De burgersegmenten.

Geëngageerden (9% van de Nederlandse samenleving)

Geëngageerden zijn het meest altruïstisch ingesteld en het minst politiek cynisch. Het sociaal vertrouwen is het hoogst binnen deze groep. Het segment bestaat uit iets meer vrouwen (56%) dan mannen (44%) en relatief veel alleenstaanden (22%). De gemiddelde leeftijd is 46 jaar. De meesten hebben een hoge (beroeps-) opleiding en een hoog inkomen.

Ambitieuzen (18% van de Nederlandse samenleving)

Het vertrouwen in de economie en de politieke betrokkenheid zijn onder ambitieuzen het hoogst.

De gemiddelde leeftijd in deze groep ligt met 42 jaar wat lager dan gemiddeld. Ze volgen bewust de beleidsthema's in de media. Het cluster bevat veel jonge tweepersoonshuishoudens of huishoudens met jonge kinderen. Het gemiddelde opleidingsniveau in dit cluster is hoog, men werkt overwegend fulltime en het inkomen ligt boven modaal. 11% is lid van een beroepsorganisatie.

Anoniemen (14% van de Nederlandse samenleving)

Anoniemen hebben het minste vertrouwen in zowel de regering als de economie. Zij zijn het minst geneigd tot altruïstisch gedrag, volgen beleidsthema's nauwelijks in de media en zijn het minst tevreden over het regeringsbeleid. Tweederde van de personen in dit clusters is van het mannelijke geslacht en bijna driekwart is alleenstaand (72%). De opleiding van deze mensen is gemiddeld, het inkomen echter merendeels beneden modaal. Dit laatste is mede een gevolg van het feit dat dit cluster relatief de meeste werklozen, arbeidsongeschikten en invaliden, bijstandstrekkers en studenten / scholieren bevat.

Zekerheidzoekers (20% van de Nederlandse samenleving)

Zekerheidzoekers zijn een vrij gemiddeld segment. Alleen op tevredenheid met beleid scoren zij het laagst. Een belangrijk deel van deze groep, circa tweederde, is 45 jaar of ouder en van het mannelijke geslacht; in die zin vormt dit cluster min of meer de mannelijke tegenhanger van de geborgenheidszoekers. Een belangrijk verschil echter is de meer introverte oriëntatie van deze groep. Dit cluster bestaat voor een relatief groot deel uit oudere tweepersoonshuishoudens ('empty-nesters'). Het opleidingsniveau binnen dit cluster is overwegend laag, het inkomen modaal. Tevens bevinden zich in deze groep de meeste gepensioneerden.

Geborgenheidszoekers (19% van de Nederlandse samenleving)

Geborgenheidszoekers hebben het meeste vertrouwen in de regering. Tevens zijn zij het meest tevreden met de regering en met het beleid. Het cluster bestaat hoofdzakelijk uit vrouwen (89%) en de gemiddelde leeftijd is 45 jaar. Het opleidingsniveau is gemiddeld, het inkomen grotendeels modaal. Een groot deel van dit cluster is moeder / huisvrouw, veelal gecombineerd met parttime werk en/of vrijwilligerswerk.

Gezelligheidszoekers (20% van de Nederlandse samenleving)

De gezelligheidszoekers zijn het minst politiek betrokken en hebben een laag sociaal vertrouwen. Het politiek cynisme is groot en de tevredenheid met de regering het laagst. Gezelligheidszoekers bestaan in belangrijke mate uit relatief jonge vrouwen. Het opleidingsniveau is gemiddeld en het inkomen modaal; bijna de helft van deze mensen

heeft een middelbare beroepsopleiding gevolgd. Een belangrijk deel van dit cluster bestaat uit gezinnen met kinderen en jonge alleenstaanden en tweepersoonshuishouders aan. Men werkt over het algemeen: parttime en / of fulltime.

Tabel 2. Gemiddelde scores op de (samengestelde) variabelen per BSR segment.

	Geënga- geer- den N= 170	Ambi- tieuzen n= 352	Anonie- men n= 280	Zeker- heid- zoekers N= 404	Gebor- genheid- zoekers n= 383	Gezellig- heid- zoekers n= 404
Vertrouwen in de regering (10 pnts) *	4,36	4,71	4,22	4,51	4,88	4,33
Vertrouwen in de economie (10 pnts) *	5,57	5,71	5,12	5,20	5,30	5,25
Tevredenheid regering (5 pnts) *	2,26	2,37	2,23	2,26	2,39	2,19
Politiek cynisme (5 pnts) *	2,55	2,54	2,28	2,20	2,38	2,17
Tevredenheid beleid (5 pnts) *	2,59	2,57	2,60	2,57	2,77	2,61
Politiek betrokkenheid (5 pnts) *	3,33	3,58	3,04	3,04	3,17	3,03
Sociaal vertrouwen (5 pnts) *	2,79	2,67	2,35	2,34	2,42	2,29
Altruïsme (5 pnts) *	3,14	3,14	2,73	2,86	2,94	2,85
Volgen beleidsthema's in media (3 pnts) *	2,24	2,28	2,13	2,19	2,23	2,24

Per variabele is het burgersegment met de hoogste score **vet** weergegeven, de laagste score *cursief*.

* voor alle variabelen werden significante verschillen tussen de BSR segmenten gevonden ($p < 0,01$)

Alle schalen zijn gehercoreerd zodat een hoge score positief geïnterpreteerd dient te worden.

Voor elk van bovenstaande segmenten is het eerder gepresenteerde model doorgerekend. Op basis van bovenstaande beschrijvingen mag verondersteld worden dat een aantal segmenten zal afwijken van het model voor het algemeen publiek zoals gepresenteerd in de vorige paragraaf. Om de resultaten voor de verschillende segmenten gemakkelijker te kunnen vergelijken is ervoor gekozen om ze in een tabel naast elkaar weer te geven.

4.1. Saillante resultaten per segment

In tegenstelling tot de verwachting hebben *Geëngageerden* opvallend weinig vertrouwen in de regering. Dit komt vooral door het sterke effect van het vertrouwen in de economie. Meer vertrouwen in de economie leidt onder alle segmenten tot meer vertrouwen in de regering, bij dit segment is dat effect echter veel sterker. Het effect van tevredenheid regering op het vertrouwen is in vergelijking tot de andere segmenten juist relatief laag maar in absolute zin nog altijd vrij sterk. Tevredenheid met de regering wordt vooral verklaard door tevredenheid met beleid. Tevredenheid met beleid is eveneens direct een voorspeller van vertrouwen in de regering. De tevredenheid met beleid kan vooral worden verklaard vanuit politiek cynisme. Weinig cynisme leidt tot een hogere beleidstevredenheid. Politiek cynisme wordt op zijn beurt voor een belangrijk deel verklaard vanuit sociaal vertrouwen. In vergelijking met de andere segmenten is die relatie bij geëngageerden het sterkst. Politieke betrokkenheid wordt bepaald door de geneigdheid tot altruïstisch gedrag en door politiek cynisme.

Ambitieuzen hebben relatief veel vertrouwen in de regering hetgeen samengaat met relatief veel vertrouwen in de economie. Het vertrouwen in de regering laat zich vooral voorspellen vanuit tevredenheid met de regering dat op zijn beurt vooral wordt bepaald

Tabel 3. Resultaten padanalyses voor de zes BSR burgersegmenten.

	Geënga- geer- den	Ambi- tieuzen	Anonie- niemen	Zeker- heid- zoekers	Gebor- genheid- zoekers	Gezellig- heid- zoekers
Vertrouwen in de regering verklaard						
tevredenheid regering	0,376	0,481	0,485	0,467	0,499	0,387
vertrouwen in de economie	0,431	0,258	0,296	0,365	0,222	0,335
politiek cynisme		0,086		0,124		0,126
tevredenheid beleid	0,167	0,127	0,115		0,146	0,099
Volgen beleidsthema's in media				0,076		
sociaal vertrouwen					0,103	
Adjusted R Square	0,64	0,60	0,56	0,62	0,59	0,56
Vertrouwen economie verklaard						
sociaal vertrouwen						0,212
Tevredenheid regering verklaard						
vertrouwen in de economie	0,190	0,225	0,351	0,300	0,198	0,216
politiek cynisme	0,288	0,281	0,336	0,211	0,305	0,374
tevredenheid beleid	0,516	0,436	0,266	0,348	0,479	0,336
Volgen beleidsthema's in media			-0,116	-0,171		
politiek betrokkenheid			-0,154	0,135		
altruïsme				0,090		
Volgen van beleidsthema's in media verklaard						
politiek cynisme		-0,110				0,115
politiek betrokkenheid	0,540	0,525	0,414	0,560	0,536	0,482
altruïsme		0,147	0,186			0,099
sociaal vertrouwen		-0,108				-0,182
Tevredenheid beleid verklaard						
vertrouwen in de economie	0,217	0,195	0,267	0,330	0,207	0,273
politiek cynisme	0,411	0,332	0,336	0,275	0,340	0,290
volgen beleidsthema's in media		-0,174			-0,163	
politiek betrokkenheid	-0,281		-0,159	-0,357		-0,125
altruïsme				0,195		
Politiek cynisme verklaard						
sociaal vertrouwen	0,634	0,492	0,587	0,596	0,590	0,593
Politiek betrokkenheid verklaard						
altruïsme	0,489	0,392	0,378	0,517	0,402	0,357
politiek cynisme	0,182				0,156	
sociaal vertrouwen		0,132	-0,119	0,102		

De cijfers in deze tabel zijn gestandaardiseerde beta coëfficiënten ($p < .001$) afkomstig uit padanalyse.

door tevredenheid met beleid. Opmerkelijk is dat de ambitieuzen die cynischer zijn over de politiek, beleidsthema's meer bewust volgen in de media. Bij gezelligheidszoekers leidt minder cynisme tot meer bewust volgen van de media, bij andere segmenten werd geen relatie gevonden met als gevolg dat ook in het model voor het algemeen publiek deze relatie niet significant bleek te zijn. Politieke betrokkenheid wordt bepaald door de neiging tot altruïstisch gedrag en door sociaal vertrouwen.

Onder de *Anoniemen* is het vertrouwen in de regering en de tevredenheid met beleid het laagst. Tevredenheid met beleid is het meest bepalend voor vertrouwen in de regering. Vertrouwen in de economie is ook sterk bepaald voor het vertrouwen in de regering. Zowel direct als indirect via beleidstevredenheid en tevredenheid met de regering. Minder politiek cynisme leidt tot meer tevredenheid met beleid en tevredenheid met de regering. Daarmee levert het indirect een bijdrage aan het verklaren van vertrouwen in de regering. Opmerkelijk is dat meer sociaal vertrouwen leidt tot afnemende politieke betrokkenheid. Dit kan een indicatie zijn van het feit dat het contact met dit segment slecht is. Sociaal vertrouwen leidt wel tot minder politiek cynisme.

Het vertrouwen in de regering is onder *zekerheidszoekers* gemiddeld. Ook voor dit segment zijn tevredenheid met de regering en vertrouwen in de economie sterk bepaalde factoren. Opvallend is dat het volgen van beleidsthema's een directe maar wel vrij zwakke positieve relatie heeft met vertrouwen in de regering. Bij geen van de andere segmenten werd dit positieve verband gevonden. De tevredenheid met beleid wordt hoofdzakelijk bepaald door vertrouwen in de economie. In vergelijking met de andere segmenten is die relatie sterk. Politieke betrokkenheid levert een opvallend sterke negatieve bijdrage aan de tevredenheid met beleid. Politieke betrokkenheid wordt binnen dit segment relatief sterk bepaald door de neiging tot altruïstisch gedrag en door sociaal vertrouwen.

Onder *geborgenheidszoekers* wordt het hoge vertrouwen in de regering zowel direct als indirect (via tevredenheid regering en tevredenheid beleid) zeer sterk bepaald door het vertrouwen in de economie. Opvallend is dat sociaal vertrouwen een rechtstreeks positief effect heeft op vertrouwen in de regering. Dit effect is bij geen van de andere clusters aangetroffen. Het volgen van beleidsthema's in de media heeft een lagere tevredenheid met beleid tot gevolg en dat effect is sterker dan bij de overige segmenten.

Onder de *gezelligheidszoekers* is het vertrouwen in de regering relatief laag. Het directe effect van tevredenheid met de regering is daarvoor het meest bepalend. Als echter naar het totaal van directe en indirecte effecten van vertrouwen in de economie wordt gekeken dan blijkt dit de belangrijkste voorspeller van regeringsvertrouwen. De tevredenheid met beleid levert in vergelijking met de andere segmenten weinig bij aan het vertrouwen in de regering. Opvallend is ook het rechtstreekse en vrij sterke effect van politiek cynisme op regeringsvertrouwen.

5. CONCLUSIES

Het onderwerp 'vertrouwen in de regering' kan zich in veel aandacht verheugen. Diverse politieke ontwikkelingen hebben dit onderwerp ongekend actueel gemaakt. De ana-

lyses in dit artikel beogen een bijdragen te leveren aan de inzichten in de totstandkoming van regeringsvertrouwen. Daartoe is met behulp van een theoretische basis en een padanalyse een verklarend model voor vertrouwen in de regering samengesteld. Dit model beschrijft de relaties tussen de onafhankelijke variabele (vertrouwen in de regering) en acht afhankelijke variabelen en tussen de afhankelijke variabelen onderling. De veronderstelde relaties bleken in grote lijnen overeen te komen met de analyseresultaten. Zowel vertrouwen in de economie, de tevredenheid met beleid en met de regering blijken belangrijke voorspellers van regeringsvertrouwen. Het totale effect van vertrouwen in de economie bleek een net iets minder goede voorspeller dan tevredenheid met de regering. Dit model vormde de opmaat om te komen tot specifieke modellen voor zes verschillende BSR burgersegmenten. De analyses op deze burgersegmenten hebben opmerkelijke verschillen aan het daglicht gebracht. Een positieve ontwikkeling in de economische situatie - met meer vertrouwen in de economie tot gevolg - heeft bijvoorbeeld voor alle segmenten een positieve uitwerking op regeringsvertrouwen. Dit effect is echter sterker onder gezelligheidszoekers en geëngageerden. De tevredenheid met beleid en met de regering liften mee op positieve ontwikkelingen in economisch vertrouwen. Ontwikkelingen in politiek cynisme hebben vooral gevolgen voor het regeringsvertrouwen van ambitieuzen, zekerheidszoekers en gezelligheidszoekers. Indirect heeft sociaal kapitaal via politieke betrokkenheid een relatie met regeringsvertrouwen. Sociaal kapitaal hangt in positieve zin samen met politiek cynisme. Daarmee leveren sociaal kapitaal, politieke betrokkenheid en politiek cynisme een belangrijke bijdrage aan het model. Per burgersegment verschillen de onderlinge relaties echter vrij sterk. Bijna elke ontwikkeling in één van de opgenomen variabelen zal anders doorwerken binnen de verschillende burgersegmenten. Nadere analyses zijn van belang om hier meer inzicht in te verkrijgen.

LITERATUUR

- Commissie toekomst overheidscommunicatie. (2001): In dienst van de democratie. SDU: Den Haag
- Dekker, P. (2001): Vertrouwen in de overheid. Een verkenning van actuele literatuur en enquêtegegevens. Katholieke Universiteit Brabant.
- Erikson, R., Wlezien, C. (1996), 'Temporal horizons and presidential election Forecasts.' In *American Politics Quarterly* 24: 492-505.
- Instituut voor de overheid (2002): Identity vs Performance: An overview of theories explaining trust in government. Katholieke Universiteit Leuven: Leuven.
- Kleinnijenhuis, J., Ridder, J.A. de, Hoof, A.M.J. van, Oegema, D. (2004) Media-analyse, Belevingsmonitor 2004 meting 3, 2004. UvA i.s.m. VU
- Kleinnijenhuis, J., Ridder, J.A. de, Hoof, A.M.J. van, Oegema, D. (2004) Media-analyse, Belevingsmonitor 2004 meting 4, 2004. UvA i.s.m. VU
- Marketresponse (2003): burgersegmentatie overheidsbeleving. Amersfoort.
- Winden, VAN, F., Bruno Frey (1987): De economische theorie van de politieke besluitvorming', 58-87 in De Beus, J., Vuijsje, F. (red.), *Politieke economie*, Utrecht: Het spectrum BV, 319 p.

Medewerkers

M.C.M. Arnold studeerde Sociale wetenschappen aan de Universiteit van Nijmegen. Hij werkt als onderzoeksadviseur bij de Dienst Publiek en Communicatie waar hij ondermeer projectleider is van de Belevingsmonitor.

T.H.A. Bijmolt is Hoogleraar Marketing-onderzoek bij het Departement Marketing van de Economische Faculteit, Rijksuniversiteit Groningen. Hij promoveerde cum laude in 1996 bij dit departement en werkte tot 2004 bij de Universiteit van Tilburg. Tammo Bijmolt is voorzitter van de NIMA/MOA-examencommissie Marketing-onderzoek en Informatiemanagement (MIM). Hij heeft gedoceerd aan verschillende doelgroepen als doctoraal studenten, promovendi en postdoctoraal managers over onderwerpen als CRM, marktonderzoek en segmentatie. Hij heeft een groot aantal “research-based consultancy” projecten uitgevoerd voor onder andere: Bavaria, Brova, Carlson Marketing Group, GfK en Unilever. Onderzoeksmethodologie, CRM en loyaliteitsprogramma’s vormen een belangrijk deel van zijn onderzoeksportefeuille. Tammo Bijmolt heeft over deze en vergelijkbare onderzoeksprojecten een groot aantal lezingen gegeven bij (inter-)nationale congressen en publiceerde hierover in internationale toptijdschriften zoals: *Journal of Marketing Research*, *Journal of Consumer Research*, *International Journal of Research in Marketing* en *Journal of Classification*; en daarnaast regelmatig in Nederlandse vak-tijdschriften zoals het *Jaarboek van de MOA* en *Tijdschrift voor Marketing*. E-mail adres: t.h.a.bijmolt@eco.rug.nl

J. Bleijerveld studeerde International Management aan de Faculteit der Economische Wetenschappen en Bedrijfskunde van de Universiteit Maastricht. Als onderdeel van deze studie heeft hij vijf maanden doorgebracht in Århus Denemarken, waar hij studeerde aan de Århus School of Business. In zijn economische studie heeft hij zich vooral geconcentreerd op marketing en organisatie. Momenteel is Jeroen bezig met de afronding van zijn tweede studie, het European master programme on Society, Science and Technology aan dezelfde universiteit. Zijn afstudeeronderzoek richt zich op de innovatieprocessen in de zorg en de rol daarin van met name de patiënt, de zorgverzekeraar, het bedrijfsleven en de kenniscentra. Tevens is hij sinds september 2005 werkzaam als promovendus bij het marketing departement van de faculteit der Economische Wetenschappen en Bedrijfskunde van de Universiteit Maastricht, waar hij zich richt op de internationale marketing van universiteiten.

A.E. Bronner studeerde politicologie met als specialisatie methoden en technieken van onderzoek. Na zijn studie is hij in het marktonderzoek gaan werken bij Veldkamp. De banden met de universiteit bleven en hij promoveerde op een proefschrift over beslisdgedrag van mensen. Momenteel is hij hoogleraar Communicatiewetenschap aan de Universiteit van Amsterdam en onderzoeksadviseur bij Veldkamp en TNS NIPO te Amsterdam. Hij publiceerde in vele binnen- en buitenlandse tijdschriften. Won de Marijn Veraartprijs en drie keer de Hans du Chatinierprijs voor het meest innovatieve mediaonderzoek, in 2004 betraf de prijs het project “de familie Polder” over de implicaties voor reclame van de toenemende democratisering in gezinnen.

G. van Bruggen is hoogleraar marketing aan RSM Erasmus University te Rotterdam. Zijn huidige onderzoek heeft betrekking op de toepassing en effecten van informatie technologie in marketing, informatie delen in marketing kanalen en het gebruik van virtual stock markets voor marktvoorspellingen. Zijn publicaties verschenen in wetenschappelijke tijdschriften zoals *International Journal of Research in Marketing*, *Journal of Marketing*, *Journal of Marketing Research*, *Management Science* en *Marketing Science*.

B.G.C. Dellaert is hoogleraar bij de afdeling Marketing aan de Universiteit Maastricht. Zijn onderzoeksinteressen betreffen consumenten beslisgedrag, consument-product interactie, customisatie en personalisatie, retailing en toerisme. Zijn werk verscheen in wetenschappelijke tijdschriften zoals de *International Journal of Research in Marketing*, de *Journal of Interactive Marketing*, de *Journal of Marketing Research* en *Marketing Letters*.

T. Kuijlen studeerde Economische Psychologie en methoden en technieken. Hij is naast zijn werk bij het bedrijfsleven aan de universitaire wereld verbonden gebleven. Na zijn vertrek bij Postbank waar hij hoofd was van de afdeling statistische methoden en technieken, heeft hij samen met drie partners zijn eigen bedrijf, Cmotions, opgericht. Hij is gepromoveerd op het onderwerp "The Scenario Approach". Deze unieke manier van computergestuurd interviewen "leert" welke vragen relevant zijn om een maximum aan "actionable" informatie te krijgen. Als bijzonder Hoogleraar aan de Katholieke Universiteit van Brabant doceert Ton op het gebied van CRM (Marketing Engineering en Databasemarketing). Hij heeft een aantal op dit vakgebied gerichte publicaties op zijn naam staan. Momenteel geeft hij aan de KUB leiding aan een groot project ter verdere ontwikkeling van het Klantcarriere-concept. Dit project is gericht op het identificeren van de parameters die bepalend zijn voor de groei in "share of wallet", switch gedrag, retentie and defectie en de ontwikkeling van de life cycle van consumenten.

P.S.H. Leeftang bezet sinds september 2005 de Frank M. Bass leerstoel in marketing aan de Rijksuniversiteit Groningen. Vanaf 1976 tot september 2005 was hij hoogleraar marketing aan dezelfde universiteit. Hij studeerde econometrie aan de Erasmus Universiteit (1964-1969) waar hij in 1974 promoveerde. Naast tal van boeken (o.m. *Marketing*, 2003, *Building Models for Marketing Decisions*, 2000) publiceerde hij o.m. in de *Journal of Marketing*, *Journal of Marketing Research* (JMR), *Marketing Science*, *Management Science*, *International Journal of Research in Marketing* (IJRM) en *Quantitative Marketing and Economics*. Hij is lid van de editorial boards van o.m. JMR, IJRM en Marketing Science. Leeftang is tevens verbonden aan de Johann Wolfgang Goethe Universiteit te Frankfurt am Main, het European Institute for Advanced Studies in Management (Brussel) en in 1999 is hij benoemd tot lid van de Koninklijke Academie van Wetenschappen.

J. Leenheer is universitair docent aan de Vrije Universiteit Amsterdam, afdeling Marketing. In 2004 promoveerde zij aan de Universiteit van Tilburg met het proefschrift: "The Adoption and Effectiveness of Loyalty Programs in Retailing". Haar onderzoeksinteresses gaan uit naar loyaliteit, CRM, retailstrategie en keuzemodellen. (jleenheer@feweb.vu.nl).

J.G.A.M. Lemmink (Jos) is hoogleraar Marketing en Marktonderzoek, Faculteit der Economische Wetenschappen en Bedrijfskunde, Universiteit Maastricht, en directeur van het Maastricht Academic Center for research in Services (MAXX). Hij studeerde bedrijfseconomie aan de Rijksuniversiteit Groningen en promoveerde in Maastricht. Voorts was hij visiting professor aan de University of Southern Queensland, Australië. Voor zijn academische loopbaan was hij marktonderzoeker bij de hoofddirectie van PPT Post en Telecom in Den Haag. Zijn publicaties gaan vooral over marketing van dienstverlening en merkstrategie. Hij publiceerde veelvuldig in wetenschappelijke tijdschriften (o.a. *Journal of Marketing*, *International Journal of Research in Marketing*, *Journal of Retailing*) en publiceerde 7 boeken over de genoemde onderwerpen. Hij was van 2000-2005 editor-in-chief van het *International Journal of Service Industry Management*.

A. van Meurs is sinds 1990 werkzaam bij Intomart GfK, waar hij thans als research manager verantwoordelijk is voor de methodologie en ontwikkeling van media-onderzoek, zoals bijvoorbeeld het Kijkonderzoek, het on-line Waarderingspanel, het Bereiksonderzoek buitenreclame en het Ster Prognosesysteem. Sinds januari 2001 is hij tevens parttime Universitair Hoofddocent aan de Afdeling Communicatiewetenschap en The Amsterdam School of Communications Research ASCoR van de Universiteit van Amsterdam.

Hij studeerde methoden en technieken van politicologisch onderzoek (1989) en communicatiewetenschap (1989, cum laude), beide aan de Universiteit van Amsterdam. Als freelance onderzoeker werkte hij gedurende korte tijd voor onder andere de Weekbladpers, de Universiteit van Amsterdam en de NOS afdeling Kijk- en LuisterOnderzoek. In september 1999 promoveerde hij aan de Universiteit van Amsterdam met het proefschrift "Switching during commercial breaks". Won drie maal de Hans du Chatinier Prijs van de MWG voor het beste, meest innovatieve en meest toepasbare media onderzoek: in 1995 samen met P. Sijtsma met het onderzoek "Aandacht voor Televisie", in 1997 samen met J. Ligthart met het SPOT-onderzoek naar schakelgedrag van televisiekijkers tijdens reclameblokken, en in 2000 samen met C. Tchaoussoglou, M. van der Kooi en S.G. de Boer met het Bereiksonderzoek buitenreclame. Hij publiceerde eerder in wetenschappelijke- en vakbladen over onder andere de evaluatie van meetinstrumenten, segmentatie van gegevens uit de SUMMO Scanner, het bereiksonderzoek buitenreclame 2000, aandacht voor televisie, het voorspellen van kijkdichtheid, verkiezingscampagnes op televisie en schreef in diverse publicaties over schakelgedrag, de methodologie van het Kijkonderzoek en de effectiviteit van buitenreclame.

L.J. Paas is als Universitair Docent verbonden aan het departement Marketing (Faculteit der Economische Wetenschappen en Bedrijfskunde) van de Vrije Universiteit in Amsterdam. Veel van zijn onderzoek gaat over de afname van producten en diensten door consumenten in de financiële sector. Andere onderzoeksinteresses betreffen consumentengedrag in de kunstsector, onderzoek en methodologie betreffende meetschalen in marketing en ontwikkelingen in de marketingliteratuur. Zijn werk is o.a. verschenen in *Journal of Economic Psychology*, *International Journal of Research in Marketing* en *The Service Industries Journal*. E-mail adres: lpaas@feweb.vu.nl

A. Ronteltap studeerde in 2001 af als voedingskundig ingenieur met de specialisaties voedingsleer en marktkunde. In 2003 begon zij als promovendus bij de Marktkunde en Consumentengedrag Groep van Wageningen Universiteit, begeleid door Prof. J.C.M. van Trijp, Prof. L.J. Frewer en dr. R.J. Renes. Haar project is getiteld "Zicht op de toekomst: de ontwikkeling van nutrigenomics met experts en consumenten" en valt binnen het NWO programma "De maatschappelijke component van genomics-onderzoek". Dit project beoogt met behulp van expert kennis, literatuur- en consumentenonderzoek een beter inzicht te krijgen in toekomstige consumentenreactie op een technologische innovatie op het gebied van voeding en gezondheid, namelijk *nutrigenomics*: het begrijpen van de invloed van voeding op menselijke genen. Door middel van semi-structureerde interviews met experts op het gebied van nutrigenomics zijn de belangrijkste determinanten die de implementatie van deze technologie beïnvloeden geïdentificeerd. Deze determinanten zijn gevisualiseerd in korte filmpjes, waarmee de invloed van elke factor afzonderlijk op hun reactie op de toekomstbeelden kan worden gemeten. Op deze manier krijgen wij meer inzicht in de toekomstige consumentenreactie op nutrigenomics.

S. Stremersch is hoogleraar Marketing aan de Erasmus Universiteit Rotterdam (Nederland) en visiting associate professor aan Emory University (Atlanta, VS). Hij is in 2001 cum laude promoveerd aan de Universiteit van Tilburg. In zijn onderzoek bestudeert hij de marketing van innovaties en technologie, hetgeen werd gepubliceerd in o.a. *International Journal of Research in Marketing*, *Journal of Marketing*, *Journal of Marketing Research* en *Marketing Science*. Zijn werk werd bekroond met prijzen zoals, onder andere, de J.C. Ruigrokprijs (2005), de Erasmus Onderzoeksprijs (2004), de Harold H. Maynard voor beste artikel in *Journal of Marketing* (2002). Hij is lid van de editorial board van *Journal of Marketing*. (stremersch@few.eur.nl).

I.A.L. Stoop is hoofd van de afdeling Informatievoorziening en Automatisering van het Sociaal en Cultureel Planbureau. Ze houdt zich onder andere bezig met het aanbesteden van dataverzameling, datakwaliteit en surveymethodologie, en is onlangs gepromoveerd op een proefschrift over nonrespons. Ze is lid van het Central Coordinating Team van het European Social Survey, van het International Statistical Institute (ISI) en het European Advisory Committee on Statistical Information in the Economic and Social Spheres (CEIES). i.stoop@scp.nl

H.C.M. van Trijp is Hoogleraar Marktkunde en Consumentengedrag aan Wageningen Universiteit, en één dag per week verbonden aan het Unilever Health Institute als Senior Scientist Consumer Behaviour for Functional Foods. In 1986 studeerde hij af in Humane Voeding aan de Universiteit van Wageningen, en hij promoveerde in 1995 bij Prof. MTG Meulenberg en Prof. JEBM Steenkamp op het proefschrift: "Variety seeking in product choice behaviour: theory with applications in the food domain".

De academische interesse van Hans van Trijp richt zich in algemene zin op marketing strategie en consumentengedrag in relatie tot de food and agribusiness. Er ligt thans een bijzonder accent op de impact van nieuwe technologie op marketing en consumentengedrag. Professor Van Trijp heeft ruime ervaring in de voedingsmiddelenindustrie en publiceerde in diverse tijdschriften zowel op het terrein van voeding, voedsel en marketing. Zijn publicaties verschenen onder andere in *Journal of Marketing Research*, *International Journal of Research in Marketing and Marketing Letters*, *European Review of Agricultural Economics*, *Appetite*, *Food Quality & Preference* and *Journal of Sensory Studies*. Hij heeft zitting in de Redacties van *International Journal of Research in Marketing* en *Food Quality and Preference*.

K. de Valck is universitair docent aan HEC School of Management – Paris. Zij is op 28 april 2005 gepromoveerd aan de RSM Erasmus University met een onderzoek naar de rol van virtuele communities in marketing. Naast dit onderwerp is zij geïnteresseerd in online consumentengedrag, interpersoonlijke beïnvloeding en het gebruik van het Internet als onderzoeksinstrument. Ze heeft haar onderzoek gepresenteerd op wetenschappelijke conferenties zoals *European Marketing Academy Conference* en de conferentie van de *Association of Consumer Research*.

P.C. Verhoef is hoogleraar Customer Based Marketing aan de Rijksuniversiteit Groningen. Hij is in 2001 gepromoveerd aan de Erasmus Universiteit Rotterdam en daar tot 2005 werkzaam geweest als universitair docent. In zijn onderzoek specialiseert hij zich vooral op customer management. In dit onderzoek werkt hij vaak intensief samen met bedrijven. Zijn onderzoek is of wordt gepubliceerd in internationale tijdschriften, zoals *Journal of Marketing*, *Journal of Marketing Research*, *Marketing Science* en de *Journal of Retailing*. Zijn werk werd o.a. beloond met de Donald R. Lehmann Award (2004) voor de best dissertation-based paper in *Journal of Marketing* en *Journal of Marketing Research*. Hij is lid van de editorial boards van de *Journal of Marketing* en de *Journal of Service Research*. (p.c.verhoef@rug.nl)

J.K. Vermunt is als Hoogleraar verbonden aan het departement Methoden en Technieken van Onderzoek (Faculteit Sociale Wetenschappen) van de Universiteit van Tilburg. Zijn onderzoek-sinteresses betreffen methodologische onderwerpen bij sociaal en gedragswetenschappelijk onderzoek. Hij ontwikkelde samen met Jay Magidson de Latent GOLD and Latent GOLD Choice programma's voor latent klasse analyse. Zijn werk verscheen in onder andere in *International Journal of Research in Marketing*, *Canadian Journal of Marketing Research*, *Quirk's Marketing Research Review*, *Sociological Methodology*, *Applied Psychological Measurement* en *Psychometrika*. E-mail adres: j.k.vermunt@uvt.nl

J.A. Voerman is universitair docent bij de vakgroep Marketing van de Economische Faculteit, Rijksuniversiteit Groningen. Zij promoveerde in 2004 op een internationaal, longitudinaal MKB-onderzoek naar exportsucces en exportgedrag. Haar huidige onderzoeksinteresses liggen op het gebied van exportmarktorientatie en retailmarketing. Zij heeft gedoceerd over onderwerpen als basis marketing, marktonderzoek, internationale marketing en retailmarketing aan propedeuse- en doctoraal studenten en managers. (j.a.voerman@rug.nl)

S. Wendel werkt momenteel aan haar proefschrift bij de afdeling Marketing aan de Universiteit Maastricht. Haar onderzoeksinteresses gaan uit naar consumentengedrag, diensten marketing en communicatie. Zij is vooral geïnteresseerd in het nader onderzoeken van de communicatie tussen fabrikant en consument.

B. Wierenga is hoogleraar marketing aan RSM Erasmus University te Rotterdam. Zijn onderzoek heeft vooral betrekking op het beslissingsgedrag van marketeers en op de vraag hoe deze beslissingen kunnen worden ondersteund met zogenaamde Marketing Management Support Systemen (MMSS). Hij is geïnteresseerd in alles wat zich afspeelt in de interface tussen marketing en informatietechnologie. Hij geeft, onder andere, onderwijs op het gebied van Customer Relationship Management (CRM) en online marketing. Zijn publicaties verschenen in wetenschappelijke tijdschriften zoals *International Journal of Research in Marketing*, *Journal of Marketing*, *Journal of Marketing Research*, *Management Science* en *Marketing Science*. Berend Wierenga maakte van 1984 tot en met 2000 deel uit van de redactie van dit Jaarboek.

J. van der Zouwen studeerde sociologie. Promoveerde in 1970 op een proefschrift op het gebied van de organisatiesociologie. Was van 1969 tot 2001 hoofd van de afdeling Methoden en Technieken van de VU. Was voorzitter van diverse organisaties op het terrein van het sociaal (methodologisch) onderzoek (w.o. SISWO, NOSMO). Voornaamste onderzoeksinteresses: (1) de verbetering van de kwaliteit van data die worden verzameld in interviews met behulp van gestandaardiseerde vragenlijsten; (2) de toepassing van begrippen en methoden van de cybernetica bij het onderzoek van sociale processen. Sinds 2004 emeritus hoogleraar aan de Faculteit der Sociale Wetenschappen van de VU. E-mail: j.van.der.zouwen@fsw.vu.nl

Summary

1.

J. Bleijerveld and J. Lemmink: Consumer reactions on product placements in TV-programs.

The placement of products in television programs is a relatively new and increasingly used marketing instrument. This article presents the results of a study on the influence of the product placement type, program type and product type on the effect of product placements on consumer attention, attitude and buying intention. In addition the research addressed possible influences of gender, TV watching frequency and pre-existing program and product preferences on this effect. The research made use of conjoint analysis and presented respondents with a survey including product placement descriptions. The results indicate that a “product show” type product placement in a “soap” program and involving a FMCG is most successful at creating buying intentions among consumers. This combination of placement and program type is also most successful at attracting a consumer’s attention and creating a positive attitude. It remains unclear however which type of product placed would attract most attention and create positive consumer attitudes. The results further show a positive effect of watching frequency on consumers’ attitudes towards a product placement. Moreover, a consumer will develop more attention, more positive attitudes and stronger buying intentions for a placed product, when its “gender identifications” correspond to his or her gender. Finally, pre-existing preferences for product or program are proven to have a positive influence on a product placement’s effect on attention, attitude and buying intention. On the basis of these results managers can better predict the effect of a specific product placement. However due to the experimental nature of this research, future research is needed to further validate these findings.

2.

K. de Valck, G.H. van Bruggen and B. Wierenga: Virtual communities.

Consumers increasingly interact with other consumers through the Internet. Electronically based discussion forums, bulletin boards, list servers, chat rooms, newsgroups, and web logs provide consumers worldwide with the ability to share their knowledge, experiences, and opinions. The popularity of electronic consumer exchanges is reflected in the vast number of virtual communities that specifically focus on consumption-related interests. These virtual communities of consumption represent substantial networks of consumer knowledge and companionship that affect consumer behavior. In this article, we address several new insights with respect to the functioning of virtual communities that will enable community managers, marketers, and market researchers to strategically use the opportunities that are oftentimes left unexploited. We discuss various community platforms and point out how the organizational structure may reinforce their potential as an online reference group. Furthermore, we give a classification of virtual community member types based on their participation behavior that allows marketers to locate interesting target groups and that gives them insight in how to address them. Finally, we show how market researchers can use netnographic research to gain insight not only in the functioning of the virtual community as an information source and reference group, but also in the processes that underlie the buying decisions of the community members.

3.

J. Leenheer: Loyalty programs.

Loyalty programs can be an appropriate instrument for retailers to enhance customer loyalty and knowledge. Especially in sectors in which customers make purchases frequently and retailers can hardly distinguish themselves with their product assortment, many customer-oriented retailers have adopted loyalty programs. However, retailers seem to underestimate the technical challenges of loyalty program adoption. Therefore, the analysis of loyalty card data is often disap-

pointing. An analysis of the supermarket sector reveals that through loyalty programs firms are able to enhance behavioral loyalty (share-of-wallet). However, the effectiveness is easily overestimated if the self-selection of members or two-sided causality is ignored. Still the net revenues turn out to increase because of the loyalty program. The effectiveness of loyalty programs can be improved with an optimal program design. It appears that rewards that give incentives to make future purchases (such as store vouchers) can lower costs substantially.

4.

J. van der Zouwen: The interviewer, help of hindrance?

The explanation of behavior requires the use of subjective concepts like opinions and attitudes. For the measurement of these non-directly observable concepts, strict standardization of the process of question answering is necessary. If interviewers have to be used to ensure a proper process of data collection, these interviewers may also hamper the standardization. Is the interviewer an indispensable *help* for researcher and respondent, or a difficult *obstacle* for obtaining unbiased and comparable answers?

To answer this question, transcripts of survey interviews have been analyzed; especially the methods interviewers use to repair inadequate answers of respondents. The outcome of this analysis, and insights gained from cybernetics, make clear that it is ineffective to strongly standardize repair behavior of the interviewers.

5.

A.E. Bronner and T. Kuijlen: The living or digital interviewer: a comparison between CASI, CATI and CAPI.

Three recent developments had a strong influence upon data collection in the Netherlands. Increasing non-response, rising interviewers costs and the growing idea within market research agencies that respondents should be treated as 'customers' have prompted market research organisations to look for alternatives. Access panels combined with CASI are gaining popularity. This implies that in longitudinal studies two types of changes will have effect:

(a) a sampling change (from ad hoc/fresh samples to access panels) and (b) a data collection change (from CATI/CAPI with a living interviewer to CASI with a digital interviewer). In this contribution we concentrate upon the effects of the method of the data collection and keep modality (sample characteristics) constant. It is shown that the presence or absence of an interviewer can have a strong influence upon the answers to certain questions. This implies that researchers should be careful if comparing data gathered by different methods.

6.

I.A.L. Stoop: Cross-national surveys.

Cross-national surveys aim explicitly at comparing countries. They differ from national surveys because attention has to be paid to translation into different languages, the use of different sampling frames, possibly different research organisations, and different fieldwork and survey traditions. In aiming for equivalence the question comes up if and to what extent countries are comparable, how to define comparability and whether a survey can be designed and implemented in different countries in a comparable way. Without knowledge of context and survey traditions in the participating countries it will be difficult to design surveys that result in comparable data and to draw relevant inferences from the data. More than national surveys, cross-national surveys require extensive documentation. This chapter gives a short overview of the problems inherent to cross-national. The European Social Survey is introduced as a study in which optimal comparability in all relevant areas is aimed for.

7.

P.S.H. Leeflang: Benchmarks

Managers benefit from having benchmarks relative to the competition. Empirical generalizations, derived from meta-analyses of market response estimates, provide one basis for benchmarks. For

example, extent research includes average price, advertising and promotion elasticities, and decompositions of sales effects resulting from temporary price cuts. In this article we discuss these and other generalizations. Limitations and future research areas are also discussed.

8.

A. Ronteltap, S. Wendel, H.C.M. van Trijp and B.G.C. Dellaert: Marketing products that match the genetic structure of the individual.

Recent developments in the life sciences, particularly those concerning the human genes, give companies more and more opportunities to develop products and services that match the individual physical needs of the consumer. Delivering these products and services poses new challenges for marketing in a number of areas.

In this paper, the authors describe these challenges and develop and test hypotheses for consumer acceptance of products and services based on recent insights from the life sciences. The results show, among other findings, that it is important to analyze these services as integral systems, and not only evaluate their separate characteristics.

9.

P.C. Verhoef and S. Stremersch: Publication Productivity of Dutch Marketing Scientists.

In this study we provide an overview of the productivity of Dutch marketing scientists. Our results show, that Dutch marketing scientists have become very productive since the mid '80s of the past century. This resulted in a number 2 position in the most recent years. We also show that the most productive universities are Erasmus University Rotterdam, University of Groningen and Tilburg University. Based on our data collection, we also discuss a list of the most productive individual scientists during this scientific marketing evolution. The three top universities also employ the most productive individual scientists. Finally, we discuss the research focus of each of the universities. Our results show a divergence in focus between universities. Only two universities, Erasmus University Rotterdam and Tilburg University, have a rather broad research focus, while the other universities usually focus on a specific number of topics, such as service marketing, quantitative marketing or customer relationship management.

10.

J.K. Vermunt, T.H.A. Bijmolt and L.J. Paas: Latent class analysis.

In this paper we describe latent class analysis, a model-based clustering technique. Special attention is given to a recently introduced extension of latent class analysis for the analysis of multi-level data, such as data in which respondents are nested within postal codes, clients nested within branch offices, or consumers nested within countries. After an introduction in which we describe latent class analysis and compare it with other clustering methods, we describe in more detail the multilevel latent class model. We illustrate the method with an application in which consumers are clustered on the basis of their financial product portfolios and countries are clustered based on the distribution of consumers over segments. We discuss the marketing implications of our findings, as well as the broader applicability of the multilevel latent class model for marketing research.

11.

L. Voerman: Imputation instead of Amputation.

What do you do as a researcher, when a large part of your respondents left blank the most important variable in your survey? The most radical solution is to delete all respondents that did not fill out the survey completely (list wise deletion). A problem with this is the often-drastic drop in the number of useful cases. This is exactly what happens in the export performance research, used in this study. Many respondents refused to fill in the export sales, the dependent variable.

In this paper, I focus on the possibility to *multiply* impute missing values using Bayesian techniques, including the advantages (and disadvantages) of this method compared to other imputation techniques. This multiple imputation is illustrated with a study on the effect of the attitude of the

entrepreneur and the information behavior on the export performance of SMEs. It appears that multiple imputation leads to better conclusions regarding the significance, the direction, and the size of the parameters to be estimated.

12.

A. van Meurs: Integrated Campaign Evaluation.

Respondents in media effect studies such as campaign tracking studies have difficulty estimating how often they have been in contact with a campaign, for instance how many times they have seen a commercial on television. There are reasons to question the validity and reliability of the response to such a frequency recall question. For television campaigns Integrated Campaign Evaluation (ICE) offers an estimated number of contacts with the campaign based on a fusion of ad-hoc research with people meter data as an alternative to the frequency recall question.

The estimated number of contacts with the television campaign in the sample of the media effect study offers an alternative to the questionable data obtained from questionnaire items like "How many times did you see this ad on television?"

In addition to television Integrated Campaign Evaluation is now also developed for radio, print and outdoor. In this paper the use of ICE for television and these additional media types will be presented, providing a detailed discussion of the predictive models used (multivariate regression), their explained variance and examples how these results were implemented in a real life effect study.

13.

M.C.M. Arnold: Segmentation and the level of trust in government.

In the public debate on the political situation in the Netherlands people express themselves in terms of "civil revolt" and "crisis in public trust". This makes one think that there are big problems in the Netherlands. Recent studies tend to give a more balanced image of the actual situation. There is no trend in rising levels of distrust, but there are indications that political indifference is increasing. Also compared to neighboring countries there are now different patterns in these developments. On the other hand there are indications that larger groups in Dutch society are subject to mobilization of distrust and dissatisfaction. It's alarming that the government tends to lose contact with these groups and that the level of trust in government within these groups is low. How is trust formed in these groups? To what extent play political cynicism and low satisfaction with government policy a role in this lack of trust. Is the economical situation a factor that plays a role? The analysis in this article shows that there are different groups in the Dutch society in the way trust is formed.