

# 1. *Het consumentengedrag na Abraham*

D. SIKKEL en E.A. KEEHNEN

## SAMENVATTING

Ook bij het ouder worden blijven mensen interessante consumenten, maar de achterliggende motieven bij het consumentengedrag veranderen voor een deel. Er is een groot aantal thema's dat van invloed is op de motivatie van oudere consumenten. Deze zijn te ordenen in de tijd. De oerinvloed is die van de *evolutie*. Mensen worden oud omdat dit hun nakomelingen grotere overlevingskansen biedt. Ouderen zijn daar, bijvoorbeeld in hun gedrag naar hun kleinkinderen, op geprogrammeerd. De variatie in genetische eigenschappen bij de *geboorte* is even groot als bij jongeren. Die aspecten van het consumentengedrag die afhangen van genetisch bepaalde persoonlijkheidstrekken veranderen dan ook niet bij het ouder worden. De gebeurtenissen in de *formatieve periode* zijn bepalend voor generatie-effecten, maar erg veel duidelijkheid omtrent deze effecten bestaat er niet. In het *volwassen leven* kunnen mensen psychologisch groeien. Sommigen doen dat, anderen niet. In het *heden* gaat het verwerken van informatie steeds slechter, maar veel ouderen compenseren dit door hun ervaring. Het dichterbij het *levenseinde* staan heeft zijn gevolgen voor sociale voorkeuren en waarden. Al deze thema's hebben hun consequenties voor consumentengedrag en marketing.

## 1. INLEIDING

Dat de vergrijzing grote consequenties heeft voor de marktverhoudingen binnen West-Europa is al lang geen nieuws meer. Het aantal vijftigplussers zal de komende 25 jaar met 2 miljoen toenemen en in 2025 uitkomen op ruim 7 miljoen. Anno 2005 is hun inkomenspositie zeer goed; of dit zo zal blijven hangt enigszins af van politieke keuzen die worden gemaakt, maar omdat ze een groot deel van de stemmende bevolking uitmaken lijkt het niet waarschijnlijk dat die positie dramatisch zal verslechteren. Vergelijken met jongeren is de vermogenspositie van ouderen voortreffelijk. Waar jongeren voor hun opleiding en als starters op de huizenmarkt zich aanzienlijk in de schulden moeten steken, hebben ouderen dat in veel mindere mate moeten doen. De schulden zijn dan ook in de regel afbetaald en de eigen huizen vrij van hypotheek. Vijftigplussers hebben alle ruimte om van het leven te genieten. Gezondheidsproblemen die er zwaar inhakken treden gemiddeld pas op na het 75<sup>e</sup> jaar. Kortom, het gaat goed met de vijftigplussers en ze vormen dan ook in vele markten een aantrekkelijke doelgroep.

Hoewel men het vergrijzingsprobleem al decennia lang heeft kunnen zien aankomen is de kennis over het marketen van producten aan vijftigplussers zeker in Nederland nog onderontwikkeld. In de Verenigde Staten verschijnt al 15 jaar lang het ene na het andere boek over *mature marketing*, en in wat mindere mate is er ook in de ons omringende

landen al sprake van de opkomst van dit nieuwe vak. De studie van veroudering in het algemeen is begonnen in de jaren zestig. Deze richtte zich vooral op de medische en psychologische kant van het ouder worden. Boeken met de titel "Aging and communication" richten zich dan ook voornamelijk op de communicatie tussen patiënt en therapeut. Ook voor marketing zijn dit soort boeken leerzaam, maar de lezer moet zelf de vertaalslag maken. Beoefenaars van het marketingvak baseren zich vaak op onderzoeken die rechtstreeks de product- en communicatiebehoeften van de consument meten, zonder zich druk te maken om achterliggende wetmatigheden en mechanismen. In dit artikel volgen we de omgekeerde weg. Het is een overzicht van wetenschappelijke invalshoeken, die de pretentie van algemeenheid hebben, maar niet in detail gericht zijn op marketingproblemen. Veelal geven ze inzicht in het waarom van bepaalde gedragingen. Het belangrijkste probleem van deze aanpak is dat het aantal thema's zo groot is dat het moeilijk is om aan te geven wat in welke situatie het meest relevant is. In het AgeWise basisonderzoek<sup>1</sup>, een verkennend onderzoek naar gedragsveranderingen na het vijftigste levensjaar, zijn we toch veel uitkomsten tegengekomen die in overeenstemming met algemene wetmatigheden zijn. Deze uitkomsten zullen we dan ook gebruiken om de theorie te illustreren. Als ordenend principe hanteren we het moment waarop oorzaken van wetmatigheden in het gedrag ontstaan. We beginnen bij de oertijd en eindigen in de toekomst.

## 2. DE OERTIJD

De mens is het enige zoogdier waarvan de vrouwelijke exemplaren tot ver na hun vruchtbare periode doorleven. De vraag hoe dit komt heeft veel wetenschappers in de afgelopen decennia bezig gehouden, maar langzamerhand zijn er voldoende onderzoekresultaten bekend die het mogelijk maken om een plausible theorie op te stellen. Vijf miljoen jaar geleden, toen de menselijke evolutielijn zich afsplitste van de chimpansee, kon onze gemeenschappelijke voorouder, de homo australopithecus, net als de chimpansee nu, niet ouder worden dan vijftig jaar. Uit paleontologisch onderzoek bleek dat 2,5 miljoen jaar geleden de homo erectus al maximaal 75 jaar kon worden. Toen zo'n 50.000 jaar geleden de homo sapiens werd geboren was de maximale leeftijd 100 jaar (O'Connell e.a., 1999). Het aantal mensen dat deze leeftijd in het verleden daadwerkelijk bereikte was natuurlijk gering, maar ze waren er wel. Dit blijkt onder meer uit Romeinse grafschriften (Parkin, 2003). De natuur geeft deze extra tijd om te leven natuurlijk niet cadeau. Voor ouderen die alleen maar schaars voedsel consumeren en niets bijdragen is in de evolutie geen plaats. Met andere woorden: dit kan alleen maar wanneer ouderen bijdragen aan de overlevingskansen van hun genen. Er zijn twee onderzoekslijnen die deze hypothese bevestigen. In de eerste plaats blijkt dat bij sommige primitieve volkeren de vrouwen in hun vruchtbare leeftijd vrijwel permanent zwanger zijn en in periodiek slecht in staat zijn om voor hun kinderen te zorgen. Deze taak wordt overgenomen door de grootmoeder van moeder's kant, wanneer deze tenminste nog in leven is. De kans dat kinderen zonder grootmoeder overleven is aanmer-

1 Het AgeWise basisonderzoek is uitgevoerd door AgeWise in opdracht van wijlen het DMSA. Hierin is voor een aantal markten op basis van open vragen onderzocht in welke opzichten vijftigplussers nieuw consumentengedrag gaan vertonen. Het is uitvoerig beschreven in Sikkell en Keehnen (2004).

kelijk geringer (Hawkes e.a., 1997). Het tweede type onderzoek is het bestuderen van bevolkingsregisters uit het verleden, toen er nog geen anti-conceptie middelen bestonden. Ook hier bleek dat de aanwezigheid van grootouders, en vooral de grootmoeder van moeder's kant de overlevingskansen van de kleinkinderen verhoogden (Beise en Volland, 2002; Lahdenperä, 2004). Met andere woorden: mensen met een genetische aanleg om goed voor kleinkinderen te zorgen hebben een relatief grote kans om hun genen door te geven. Daarom hebben we nu grootouders met een instinctieve drang om goed voor kleinkinderen te zorgen. De marketing-technisch relevante vraag is: "hoe dan?"

Een belangrijk onderscheid tussen mensen en andere dieren is het gebruik dat ze kunnen maken van hun hersenen. De hypothese ligt dan ook voor de hand dat hier de sleutel ligt voor feit dat mensen in relatie tot hun vruchtbare periode langer leven. Tot voor de uitvinding van het schrift, zo'n 5000 jaar geleden, was alle aanwezige kennis opgeslagen in de hoofden van mensen. De ouderen beschikten over de meeste kennis. Bij olifanten, die een ouderdom kunnen bereiken die vergelijkbaar is met die van mensen, zien we dat het onthouden van met name geografische kennis van cruciaal belang is voor het overleven van de groep. Maar ook in sociaal opzicht zijn oudere olifanten leidend, bijvoorbeeld bij het vaststellen wie vrienden en vijanden van een kudde zijn (McComb e.a. 2001). Bij veel primitieve volkeren werd kennis overgedragen via ouders; met name oudere mannen waren leider en bewaker van rituelen terwijl grootmoeders de rol van een soort reservemoeder of aanvulmoeder hadden (Simmons, 1945). De instinctieve gevoelens voor een kleinkind kunnen we in het AgeWise basisonderzoek terughoren als

*Ik verbaas me over de sterke band die ik voel.*

*Ik kijk anders tegen het leven aan sinds mijn kleinzoon er is. Vroeger dacht ik: een mens gaat dood, wat maakt dat uit.*

De drang tot het doorgeven van iets wat in het leven van de grootouders als zeer bijzonder werd ervaren klinkt door in

*Ik wil mijn kleinkind Frankrijk laten zien, want voor mij is dat het mooiste land ter wereld*

In het onderzoek bleek dat het nut dat kleinkinderen later van cadeaus kunnen hebben voor grootouders een belangrijk keuzecriterium is. Deze uitkomst is volledig in overeenstemming met het motief, dat de kleinkinderen, de dragers van genen van de grootouders, goede kansen moeten krijgen om te overleven. Dit is geen sentimenteel mechanisme. De schaarse commercials over grootouders en kleinkinderen druipen echter van het sentiment. Hoewel oma zeker een warm gevoel kan krijgen bij een blij gezicht van een dankbare peuter, is deze benadering wel erg eenzijdig.

### 3. DE GEBOORTE

Voor de geboorte liggen de menselijke genetische eigenschappen al vast. Deze bepalen

het uiterlijk, vatbaarheid voor verschillende ziekten, maar ook voor een belangrijk deel de persoonlijkheid. Over de definitie van persoonlijkheid lopen de meningen van psychologen uiteen. Deze weinig vruchtbare academische discussie zullen we hier uit de weg gaan. Een voor het consumentengedrag wel vruchtbare benadering is het karakter van mensen beschrijven in termen van alledaagse woorden die men uit het woordenboek kan halen. Omdat we onze meningen over anderen vaak op bloemrijke wijze verwoorden zijn dat er veel. Gelukkig bestaat er factoranalyse om hierin ordening te brengen. Uit vele onderzoeken blijkt dat er vijf dimensies zijn, die vanwege hun universele toepasbaarheid ook wel de Big Five worden genoemd (Digman, 1996; McCrae en Costa, 2003). Het zijn:

- Neuroticisme: is iemand emotioneel onbekommerd of lijdt hij sterk onder emotionele problemen
- Extravertheid: treedt iemand gemakkelijk op de voorgrond of kijkt hij meer de kat uit de boom
- Vriendelijkheid: de mate waarin we mensgericht, attent of meer zakelijk en direct zijn
- Gewetensvolheid: de mate waarin iemand geneigd is zijn verplichtingen na te komen
- Openheid: de mate waarin iemand open staat voor nieuwe ervaringen of zich voor nieuwe dingen afsluit

Het interessante aan deze dimensies is dat de posities van mensen hierop, zeker na het dertigste levensjaar, vrijwel niet veranderen. Iemand die op 30-jarige leeftijd stressbestendig is, is dat op 60-jarige leeftijd ook. Een jonge nurks wordt na verloop van tijd vrijwel zeker een oude nurks. Iemand die graag experimenteert op zijn dertigste zal dat op zijn zestigste nog steeds graag doen. McCrae en Costa geven goede argumenten dat deze karaktertrekken ofwel aangeboren zijn ofwel op zeer jonge leeftijd vrijwel onuitwisbaar worden ontwikkeld, al zijn ze in de adolescentieperiode nog niet in hun uiteindelijke vorm aanwezig. Het gevolg van het constant zijn van de scores op de Big Five is dat de variatie aan karakters onder jonge mensen even groot is als de variatie aan karakters bij oude mensen. Op grond hiervan kan worden verwacht dat veranderingen in aankopen van merken en producten, voor zover die het gevolg zijn van de drang naar variatie, bij ouderen even sterk voorkomen als bij jongeren. Dit blijkt ook in Sikkels en Keehnen (2004) waarin omstandig wordt beschreven dat de consument ook na het vijftigste levensjaar tot nieuwe aankopen overgaat. De oudere consument is dus niet alleen interessant vanwege zijn financiële draagkracht, maar ook vanwege het feit dat er, net als bij jongeren, een aanzienlijk segment is dat op nieuwe producten uit is.

#### 4. DE FORMATIEVE PERIODE

Een verschijnsel dat regelmatig bij het verklaren van het consumentengedrag van baby boomers en ouderen wordt genoemd is het optreden van generatie- of cohort-effecten. Dit onderwerp is in Nederland gepopulariseerd door de Utrechtse socioloog Henk Becker (1992, 1997). Hij definieert generaties ruwweg als groepen mensen die op basis van hun geboortjaar in hun formatieve periode veel gemeenschappelijke gebeurtenissen hebben meegemaakt. Zijn veel gebruikte indeling in generaties is

– *De vooroorlogse generatie*, geboren tussen 1910 en 1930. Ze zijn geboren en opge-

- groeit in een sterk verzuilde maatschappij. Omdat ze de crisistijd en de tweede wereldoorlog hebben meegemaakt, weten ze uit de eerste hand wat armoede en dreiging is.
- *De stille generatie* (1930 – 1945). De periode waarin zij volwassen werden was de opbouwperiode na de tweede wereldoorlog. In materiële zin is dit een zeer succesvolle generatie, want zij hebben een enorme welvaartsstijging gegenereerd. Hun naam ontleent ze aan het feit dat ze noeste werkers waren, die verder weinig van zich lieten horen, iets wat we tegenwoordig met ‘niet lullen maar poetsen’ aangeven. Velen van hen gaan nu met pensioen.
  - *De protestgeneratie* (1946 – 1954). Dit zijn de baby boomers voor wie materiële welvaart alleen niet meer genoeg was. Zij bevochten in de jaren zestig en zeventig democratisering, doorbreking van taboes en een subcultuur voor jongeren. Sommigen van hen zijn reeds met vervroegd pensioen, de meesten nemen nog deel aan het arbeidsproces.

De twee jongere generaties noemt Becker de verloren generatie (1955-1970) en de pragmatische generatie (na 1970). Zij behoren voorlopig nog niet tot de vijftigplussers. Het is uitermate moeilijk vast te stellen in welke mate generatie-effecten mede bepalend zijn voor het consumentengedrag. Het feit dat mensen in verschillende leeftijdsgroepen zich anders gedragen is als argument onvoldoende omdat er verschillende alternatieve verklaringen zijn. Het feit dat mensen een bepaalde leeftijd hebben is op zichzelf al een verklaring. Het tijdsbeeld, met alle technologische en culturele ontwikkelingen, heeft voor verschillende leeftijdsgroepen verschillende consequenties. Daardoor wordt het moeilijk om verschillen in gedrag tussen leeftijdsgroepen met zekerheid toe te schrijven aan verschillen in generatie. Toch zijn er wel aanwijzingen dat generatie-effecten bestaan en dat het generatiebegrip met vrucht in de marketingpraktijk kan worden gehanteerd. We geven drie voorbeelden.

1. *Consumenten rekenen zichzelf tot een generatie.* Uit onderzoek (Diepstraten e.a., 1998) blijkt dat juist voor de ouderen geldt dat ze vinden dat ze tot een duidelijke generatie behoren, zie tabel 1. Dat geldt het meest voor de vooroorlogse generatie: 91% vindt zichzelf tot een specifieke generatie behoren. Referentie in advertenties naar generatiekenmerken zal een wij-gevoel oproepen dat bepaalde producten aantrekkelijker maakt. Dit effect kan zich zelfs voordoen wanneer er geen enkel objectief verband tussen generaties en consumentengedrag bestaat.

*Tabel 1. Respondenten die zich tot een generatie rekenen; percentages.*

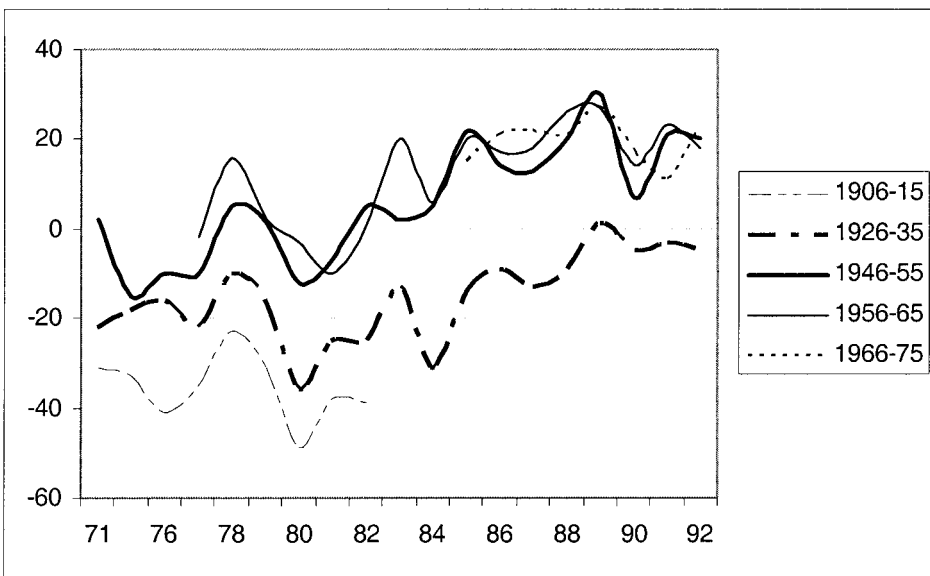
	rekent zich tot een generatie
vooroorlogs (1910-1929)	91
stil (1930-1945)	86
protest (1946-1954)	75
verloren (1955-1970)	51
pragmatisch (na 1970)	65
<b>totaal</b>	<b>70</b>

Bron: Diepstraten e.a. (1998)

2. *Herinneringen aan vroeger komen terug bij het ouder worden.* Er zijn methoden van psychologisch onderzoek waarbij men steekproeven van herinneringen kan nemen (Draaisma, 2001; Rubin, 2000). Voor jongeren hebben de meeste herinneringen betrekking op het recente verleden; het aantal herinneringen aan het verdere verleden neemt exponentieel af naarmate de betreffende gebeurtenis zich langer geleden voordeed. Tussen het veertigste en het vijftigste levensjaar gaat deze wet niet meer op. De meeste herinneringen betreffen nog steeds het recente verleden, maar het aantal herinneringen aan de tijd tussen het 10<sup>e</sup> en het 25<sup>e</sup> levensjaar is veel groter dan aan de periode dat men ouder is. De formatieve periode komt terug, en dit blijft zo na het vijftigste levensjaar. Voor dit verschijnsel zijn vele verklaringen gegeven. Eén van de verklaringen is dat ouderen in evolutionaire zin hun bestaan danken aan het feit dat ze veel kennis uit het verleden hebben (zie paragraaf 2). Omdat het recente verleden ook bekend is bij jongeren hebben ouderen daar niet veel voordeel van; voor het verdere verleden, dat jongeren niet uit de eerste hand kunnen kennen, hebben ze dat voordeel wel. De consequentie van dit mechanisme zou zijn dat generatie-effecten bij de protestgeneratie pas sinds een korte tijd een rol spelen. Hun herinneringen aan vroeger zijn immers pas kort geleden weer op de voorgrond getreden.
3. *Concrete aanwijzingen uit de crisisjaren.* In Elder (1974) wordt een longitudinaal onderzoek beschreven van mensen die in 1920 en 1921 in de Verenigde Staten zijn geboren en de crisisjaren vanaf 1928 als kind hebben meegemaakt. Voor sommigen betekende dit bittere armoede, maar bij gezinnen waarvan de kostwinner zijn baan niet verloor vielen de gevolgen betrekkelijk mee. Deze generatie van 1920/1921 had het voordeel dat op het moment dat zij de arbeidsmarkt betraden de economie weer aantrok. Maar wat betekende de crisis later voor degenen die als kind zwaar waren getroffen? Elder laat zien dat voor de mensen uit arbeidersklasse de problemen uit de kindertijd inderdaad beperkte negatieve gevolgen hadden in termen van gezondheid en carrière. Voor kinderen uit de middenklasse bleken de gevolgen echter uiterst verrassend: de financiële terugval van het gezin in de crisisperiode had een zeer heilzame werking. Op latere leeftijd was hun gezondheid beter en hun inkomen hoger dan bij degenen die zo'n terugval niet hadden meegemaakt. De problemen in de jeugd hadden de weerbaarheid versterkt. Dit voorbeeld, dat weliswaar in een andere tijd en op een andere plaats speelt, laat zien dat generatie-effecten wel bestaan maar zeer situatie-specifiek kunnen zijn. Het ligt dan ook niet voor de hand dat ze in gemakkelijke statistische generalisaties kunnen worden gevangen.

In Inglehart (1977) wordt een theorie afgeleid dat zich in de steeds welvarender wordende westerse samenlevingen een "stille revolutie" voltrekt. De politieke agenda van verschillende generaties zou steeds meer verschuiven van materiële naar postmateriële behoeften. De materiële behoeften zijn economische welvaart en veiligheid, postmateriële behoeften zijn democratisering en inspraak. In Abramson en Inglehart (1995) wordt deze trend uitvoerig onderbouwd aan de hand van een lange tijdreeks die voor een groot aantal landen is gemeten. In figuur 2 is deze tijdreeks voor Nederland weergegeven voor een beperkt aantal geboortecohorten (de weggelaten cohorten voldoen aan dezelfde wetmatigheden, maar maken de figuur onoverzichtelijk). De figuur laat goed zien hoe het belang van postmateriële behoeften varieert met het cohort en met de kalendertijd. In perioden van economisch optimisme, zoals in 1978, worden de postmateriële behoeften sterker, gaat het economisch minder, zoals in 1980, dan worden de

materiële behoeften sterker. De cohorten blijven daarbij op ongeveer dezelfde afstand van elkaar. Slechts bij de jongste cohorten is deze afstand afwezig. Van den Broek (1996) toonde aan dat in Nederland het verschil tussen de generaties gepaard gaat met een verschil in opleidingsniveau. Waren de generaties even goed opgeleid geweest, dan waren ze ook even (post)materialistisch.



Figuur 1. Postmaterialisme index in Nederland in verschillende jaren voor verschillende geboortecohorten. Bron: Abramson en Inglehart (1995).

Over de consequenties voor het consumentengedrag wordt vooral in de Verenigde Staten zeer opgewonden gedaan. Een zeer bekend en lekker leesbaar boek is Smith en Clurman (1997), waarin zonder enige vorm van serieuze onderbouwing generational marketing als het vak van de toekomst wordt aangeprezen. Ook in het AgeWise basisonderzoek vinden we aanwijzingen voor de bruikbaarheid van generaties als invalshoek, getuige citaten als

*“Ik kom uit een echt arbeidersmilieu waar het belangrijk was nuttig te zijn. Tegen mijn pensionering uit dit jaar kijk ik wel op want hoe ik mij daarna nuttig moet maken, weet ik nog niet.”*

Man, 62 jaar

*“Door de oorlog moesten wij heel creatief omgaan met de beperkte middelen die wij hadden. Vaak herken ik die creativiteit in dingen die ik nu doe zoals zelf dingen verbouwen of aanpassen in huis.”*

Man, 71 jaar

Deze referenties aan het verleden zijn duidelijk, maar een andere respondent verklaarde op de vraag waarom hij minder deed aan doe-het-zelven:

“Geen concrete aanleiding, wel een gevoel van: na 40 jaar werken hoef ik niet zo nodig meer.”

Man, 63 jaar.

Kortom, generatie-effecten moeten in concrete situaties zorgvuldig op hun consequenties worden onderzocht. Veelgehoorde beweringen als “de baby boomers zijn veel mondiger” zijn in sommige opzichten misschien juist, maar daarmee nog niet noodzakelijk een generatie-effect en bovendien veel te algemeen om bruikbaar te zijn. Een uitspraak als “er is een zeer mondig segment onder de baby boomers” komt waarschijnlijk dichter bij de waarheid.

## 5. HET VOLWASSEN LEVEN

Eén van de meest intrigerende vragen rondom ouder worden is of mensen in psychologische zin ook blijven groeien, of met de jaren ook de wijsheid komt. Deze vraag is alleen al lastig, omdat de definitie van wat we onder groei verstaan zeer moeilijk te geven is. Eén ding is duidelijk: in termen van de karaktertrekken die behoren tot de Big Five groeien mensen in ieder geval niet, want die blijven constant. Toch beweren grote psychologen uit de vorige eeuw als Jung (1933), Maslov (1954) en Erikson (1982) dat de mens tot op hoge leeftijd groeicapaciteit heeft. Voor Jung had dit te maken met het in harmonie leven met het onbewuste, voor Maslow met de wijze waarop de mens wordt gemotiveerd en voor Erikson met een complete ontwikkeling. Het zijn theorieën die zijn ingegeven door directe waarneming in de psychiatrische praktijk, niet door grootschalig empirisch statistisch onderzoek bij gezonde mensen. Zulk onderzoek is inmiddels wel gedaan. Een uitgebreide onderzoekslijn is die van Gisela Labouvie-Vief, die groei operationaliseert als een vorm van toenemende complexiteit in de organisatie van psychologische concepten. Een voorbeeld is het zelf-concept. In Labouvie-Vief e.a (1995) werden de volgende niveaus onderscheiden.

0. *Concreet*. Er worden simpele en concrete termen gebruikt: “ik ben een ingenieur”, “ik ben aardig”, “ik ben lang”
1. *Interpersoonlijk, in termen van ik*. Er worden eenvoudige uitspraken gedaan die de relatie met de groep weergeven: “ik heb veel vrienden”, “ik houd van mijn gezin”.
2. *Interpersoonlijk, in termen van relatie*. De persoon beschrijft zijn relatie met een sociale groep: “ik ben een goede moeder”, “ik heb in mijn leven weinig succes gehad”.
3. *Contextueel, in termen van ontwikkeling*. De persoon toont zich bewust van zijn ontwikkeling in de tijd: “ik ben zangeres en actrice en wil deze talenten creatiever gaan gebruiken”, “ik ben mezelf opnieuw aan het ontdekken”.
4. *Dynamisch, intersubjectief*. De persoon beschrijft zichzelf op een complex psychologisch niveau en toont zich bewust van achterliggende motivaties en interacties: “mijn ouders zijn nu van mij afhankelijk en ik heb moeite om mijn ‘volwassen’ zelf te blijven en niet te worden teruggezogen naar mijn ‘jongere’ zelf als ik bij ze ben”.

In tegenstelling tot de Big Five zijn er psychologisch hoog geschoolde codeurs nodig om de beschrijvingen van respondenten van zichzelf aan de niveaus toe te wijzen. Empirisch bleek er inderdaad sprake van groei bij volwassenen te zijn, al verliet de



groei minder hard dan bij kinderen en adolescenten. De gemiddelde niveau's waren niet buitengewoon hoog. Het hoogste gemiddelde werd gehaald tussen het 46<sup>e</sup> en het 59<sup>e</sup> jaar, namelijk 1,87, dus nog onder niveau 2. Kennelijk is het bereiken van het maximale niveau 4 wel erg uitzonderlijk. Na het 60<sup>e</sup> jaar zet een daling in. Bij 70-plussers zakt dit niveau gemiddeld terug naar 1,34. Interessant was dat, alleen bij ouderen, er een forse correlatie bleek te zijn met de persoonlijkheidskenmerken extraversie en flexibiliteit. Hoewel deze kenmerken zelf niet veranderen, stellen ze kennelijk iemand in staat om door zijn levenservaring psychologisch te groeien.

Dit soort bevindingen suggereert dat de manier waarop mensen in het leven staan, wat ze hebben meegemaakt en hoe ze het hebben verwerkt belangrijke marketingconsequenties kan hebben. In het AgeWise basisonderzoek vinden we grote verschillen, bijvoorbeeld hoe vijftigplussers met hun uiterlijk omgaan of welke verzekeringen ze belangrijk vinden. Deze verschillen zijn in belangrijke mate het gevolg van hoe volwassen (= volgroeid!) ze tegen veroudering aankijken en ook van de expertise die ze op bepaalde productvelden hebben ontwikkeld. Psychologische groei gecombineerd met praktische ervaring is daarmee een interessant segmentatiecriterium. Een toepassing is de vakantiemarkt. In het AgeWise basisonderzoek werden twee segmenten gevonden van consumenten die hun toegenomen vrije tijd benutten om meer te gaan reizen. Enkele kenmerken van de segmenten zijn

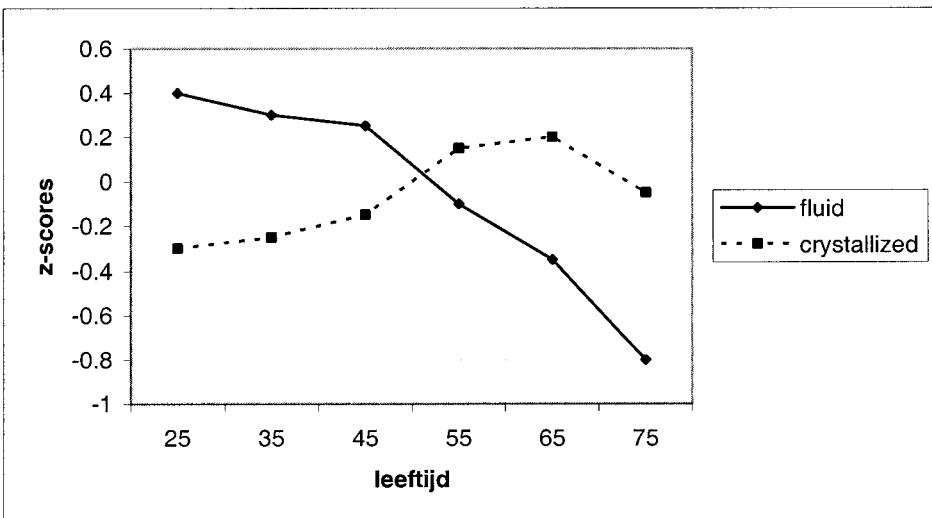
*Segment 1: hoog niveau.* Staat open voor alle bestemmingen; onbekend is leuk; heeft zelfvertrouwen, is geneigd internet te benutten; veel interesse in cultuur van de bestemming; georganiseerde reis moet functie hebben dat je dan dingen ziet/doet die je anders niet te zien/doen zult krijgen

*Segment 2: laag niveau.* Bestemmingen niet te exotisch, maar moeten wel interessant klinken; geen zelfvertrouwen; niet geneigd of in staat internet te benutten; wil combinatie van veiligheid en toch sterke verhalen aan vrienden kunnen vertellen; weinig interesse in cultuur van de bestemming; georganiseerde reis moet functie hebben van risicoreductie; spannende elementen moeten met waarborgen omgeven zijn

Een opmerkelijk aspect aan deze indeling is verder dat er nauwelijks socio-demografische verschillen tussen de segmenten bestaan.

## 6. NU

Vijftigplussers voelen zich vaak nog jong van lichaam en geest. Wanneer het gaat om meekunnen in de maatschappij, fysiek en geestelijk functioneren, is dat begrijpelijk maar niet helemaal conform de feiten. De meest zichtbare vorm van aftakeling is dat vrijwel elke vijftigplusser een soort prothese draagt, namelijk een (lees)bril. Veel organen functioneren op jonge leeftijd met zware overcapaciteit. Op zestigjarige leeftijd functioneren deze organen nog steeds goed, maar de overcapaciteit is een stuk teruggegaan. Maar het verschil tussen zware overcapaciteit en gewoon goed functioneren wordt natuurlijk door niemand ervaren.



Figuur 2. Gemiddelde z-scores van een aantal tests waarin fluid en crystallized cognition wordt gemeten. Bron: Salthouse (1999).

Het mentale functioneren verandert ook. De cognitieve eigenschappen kunnen we in twee onderdelen splitsen: *crystallized cognition*, de algemene kennis en het vocabulaire, en *fluid cognition*, het vermogen tot opnemen van informatie en redeneren. Figuur 3 laat zien dat crystallized cognition tot op hoge leeftijd verbetert, terwijl fluid cognition vanaf 25-jarige leeftijd alleen maar verslechtert. Concreet betekent dit dat ouderen beter zijn in herkennen, jongeren beter zijn in verwerken. Ouderen functioneren goed op basis van ervaring, jongeren op basis van het tempo waarmee ze nieuwe informatie aankunnen. In veel situaties kunnen ouderen hun lager tempo goed met ervaring compenseren. Voor het consumentengedrag is de consequentie dat de verhouding herkenning/nieuwe informatie bij ouderen anders moet liggen dan bij jongeren. Dit geldt zowel voor productontwerp als bij commerciële communicatie. Maar het betekent ook dat ouderen moeilijker hun weg kunnen vinden in een supermarkt die net opnieuw is ingedeeld.

## 7. DE TOEKOMST

Het is in de sociale wetenschappen hoogst ongebruikelijk om een verband te leggen van gebeurtenissen die later in de tijd plaatsvinden naar gebeurtenissen die eerder plaatsvinden. Toch blijken bij het (consumenten)gedrag van ouderen zulke verbanden te bestaan. Ze worden onder meer beschreven in Carstensen e.a. (1999). Ze zijn gebaseerd op een hypothese die socio-emotionele selectie theorie heet.

Uitgangspunt bij de socio-emotionele selectie theorie is dat interactie tussen mensen en sociale binding essentieel is om te kunnen overleven en bovendien vaak met emoties gepaard gaat. Deze emoties hoeven niet altijd positief te zijn. Contact tussen mensen kan ook leiden tot confronterende situaties, moeilijke gesprekken en zelfs ruzies. Emo-

tionele risico's die met contacten worden genomen worden geaccepteerd, omdat het alternatief is dat men een eenling blijft. In de jeugd- en adolescentiefase is het hebben van veel vrienden ook een investering in de toekomst. Je doet mensenkennis op en houdt er mogelijk een paar mensen aan over die vrienden voor het leven blijven. Wanneer de toekomst echter van beperkte duur is, heeft het minder zin om te investeren in een groot aantal nieuwe contacten. Ouderen nemen daarom minder het risico van onprettige contacten. Ze zijn geneigd zich beperken tot een wereld van bekenden, van wie ze weten dat de contacten ermee plezierig zijn. Dat dit te maken heeft met een beperkte tijdshorizon bleek uit twee voorbeelden waarbij dezelfde effecten zijn geconstateerd. In de eerste plaats betrof het jonge mensen die seropositief waren. Op het moment dat dit bij hen werd geconstateerd werd de subjectieve levensverwachting opeens een stuk korter. Hun gedrag ten aanzien van contacten met anderen ging op dat van ouderen lijken. Het tweede voorbeeld speelde zich af in Hong Kong rond de overdracht van Engeland aan China. Een jaar voor de overdracht werden dezelfde verschillen tussen ouderen en jongeren geconstateerd als elders in de wereld. Twee maanden voor de overdracht werd deze gebeurtenis door de inwoners van Hong Kong gepercipieerd als een soort naderend einde, en ging ook bij hen het gedrag van de jongeren op dat van de ouderen lijken. Een jaar later waren de verschillen tussen jong en oud echter weer als voorheen.

In het algemeen is het, al dan niet bewuste, besef dat het grootste deel van het leven voorbij is een reden om positieve emoties op te zoeken. Dit blijkt aan de hand van een veelheid van experimenten. Wanneer ouderen verhaaltjes aan kinderen vertellen hebben ze een meer positieve toonzetting dan jongeren. Van geschreven teksten zijn ouderen meer geneigd de positieve passages te onthouden dan jongeren. Afbeeldingen die positieve emoties uitstralen blijven bij ouderen relatief beter hangen dan bij jongeren (Charles e.a., 2003). Dit geldt ook voor advertenties. Met name advertenties die saamhorigheid uitstralen worden door ouderen relatief goed onthouden (Fung en Carstensen, 2003).

## 8. CONCLUSIE

Wie in één zin een samenvatting wil van "hoe" oudere consumenten zijn had dit artikel beter niet kunnen lezen. Elke one-liner is fout. Een groot aantal onvergelijkbare motieven zorgt tezamen voor de verschillende gedragstendensen van ouderen. Ze variëren van genetische programma's uit de oertijd tot de naderende dood. En dan hebben we het nog niet eens gehad over situationele kenmerken als empty nester zijn, gepensioneerd zijn en het bezit van een aanzienlijk vermogen. Over de vraag welke tendensen relevant zijn voor de consument in een concrete situatie valt op voorhand weinig algemeen te zeggen. Gelukkig is het wel goed onderzoekbaar. Theoretisch kader is er te over. De marktonderzoekwereld kan daar dankbaar gebruik van maken, vooropgesteld dat de bedrijven de oudere consument überhaupt zien staan. Dit lijkt echter een kwestie van tijd. De vergrijzing zal zelfs de meest gerontofobe marketing manager er toe dwingen rekening te houden met klanten die Abraham hebben gezien.

## LITERATUUR

- Abramson , P.R. en R. Inglehart, 1995, Value Change in Global Perspective. The University of Michigan Press, Ann Arbor.
- Becker, H.A., 1992, Generaties en hun kansen. Meulenhoff, Amsterdam.
- Becker, H.A., 1997, De toekomst van de verloren generatie. Meulenhoff, Amsterdam.
- Beise, J en E. Voland, 2002, A multilevel event history analysis of the effects of grandmothers on child mortality in a historical German population Krummhörn, Ostfriesland, 1720-1874. Demographic Research 7, blz. 470-498.
- Carstensen, L.L., D. Isaacowitz en S.T. Charles, 1999, Taking Time Seriously. A theory of Socioemotional Selectivity. American Psychologist no 54, blz. 165-181.
- Charles, S. T., M. Mather en L.L. Carstensen, 2003 Aging and Emotional Memory: The Forgettable Nature of Negative Images for Older Adults. Journal of experimental psychology 132 p.310.
- Fung, H.H. en L.L. Carstensen, 2003, Sending Memorable Messages to the Old: Age Differences in Preferences and Memory for Advertisements. Journal of personality and social psychology vol.85 (2003) nr.1 p.163-177.
- Diepstraten, I., P. Ester en H. Vinken, 1998, Mijn Generatie. Syntax Publishers.
- Digman, J.M., 1996, The curious history of the five-factor model. In J.S. Wiggins (ed.), The five factor model of personality, blz. 1-20. The Guilford Press, New York.
- Draaisma, D., 2001, Waarom het leven sneller gaat als je ouder wordt. Historische uitgeverij, Groningen.
- Erikson, E. en J.M. Erikson, 1982, The Life Cycle Completed, Extended Version. Norton & Company, New York.
- Hawkes, K, J. O'Connel en N. Blurton Jones, 1997, Hazda's women's time allocation, offspring provisioning en the evolution of long postmenopausal life spans. Current Anthropology, 38, blz. 551-577.
- Inglehart, R, 1977, The silent revolution. Changing values and political styles among western publics. Princeton University Press.
- Jung, C.G., 1933, Modern man in search of a soul. Harcourt, New York.
- Labouvie-Vief, G., L.M. Chiodo, L.A. Goguen, M. Diehl, en L. Orwoll, 1995, Representations of self across the life span. Psychology and Aging 10, blz. 404-415.
- Lahdenperä, M., V Lummaa, S. Helle, M. Tremblay, A.F. Russell, 2004, Fitness benefits of prolonged post-reproductive lifespan in women. Nature 428 blz. 178-181.
- Maslov, A., 1954, Motivation and Personality. Harper, New York.
- McComb, K., C. Moss, S. M. Durant, L. Baker, S. Sayialel, 2001, Matriarchs As Repositories of Social Knowledge in African Elephants. Science 292, blz. 491-494.
- McCrae, R.M. en P.T. Costa, 2003, Personality in adulthood. A five factor theory perspective. Second edition. The Guilford Press, New York.
- O'Connel, J.F., K. Hawkes and N.G. Blurton Jones, 1999, Grandmothering and the evolution of homo erectus. Journal of Human Evolution, 36, blz. 461-485.
- Parkin, T. G., 2003, Old age in the Roman world. A cultural and social history. John Hopkins University Press.
- Rubin, D.C., 2000, Autobiographical Memory and Aging. In Park, D. en N. Schwartz (eds.), Cognitive Aging. A Primer. Psychology press, Philadelphia.
- Salthouse, T.A., 1999, Theories of cognition. In V.L. Bengtson en K.L.Schaie (eds.). Handbook of Theories of Aging. Springer, New York.
- Sikkel, D. en Keehnen, E.A., 2004, Ervaren maar veranderlijk: het consumentengedrag van de vijftigplusser. Kluwer, Amsterdam.
- Simmons, L.W., 1945, The Role of the Aged in Primitive Society. Yale University Press, London.
- Smith, J.W. en A. Clurman, 1977, Rocking the ages. The Yankelovich report on generational marketing. HarperCollins, New York.

## 2. *Vergrijzen of verzilveren?* *Senioren, merken en communicatie*

L. GERRITSEN

### SAMENVATTING

Senioren genieten van het leven. Hoe men geniet en de behoeften en doelen die men heeft, verschillen per senior. Senioren kunnen dus niet worden gezien als één homogene groep. Senioren kunnen worden gesegmenteerd op basis van de mate waarin men behoefte heeft aan controle in het leven en de sociale oriëntatie van een senior. Elk van de vier segmenten heeft andere behoeften op het gebied van merk- en productgebruik. De voorkeur voor een product verschilt per segment, evenals de voorkeur voor A- of B-merken. Tevens worden er do's and don'ts gegeven voor aansprekende communicatie voor senioren.

### 1. INLEIDING

#### *1.1. Aanleiding tot onderzoek*

Nederland telt momenteel 4,8 miljoen 50-plussers. In 2020 is dit aantal gegroeid tot zo'n 6,7 miljoen. Tussen nu en 2020 zullen de bestedingen van mensen beneden de vijftig jaar met 1% toenemen, terwijl de aankopen van 50-plussers zullen verdubbelen. Toch is 90% van alle reclamecampagnes gericht op consumenten onder de vijftig. De nadruk ligt op jong, functioneel en mooi, terwijl het product of de dienst net zo interessant is voor de 50-plusser.

In commercieel opzicht is er nog relatief weinig over de seniorendoelgroep bekend. Hoewel zij inmiddels voorzichtig als marketingdoelgroep onderkend worden, worden zij nog niet als communicatiedoelgroep benaderd.

Het is de vraag of senioren op dezelfde manier op merken en producten reageren als andere doelgroepen en of zij op dezelfde manier benaderd dienen te worden. Veel communicatie en beeldvorming rondom senioren bevestigt nog het vooroordeel dat senioren inactief, hulpbehoevend, eenzaam en klagerig zijn, terwijl senioren tegenwoordig gezonder en gelukkiger zijn en langer een kwalitatief goed leven leiden.

De definitie van een senior verschilt. Mensen van 50 worden als senior bestempeld, maar ook van 55. In dit onderzoek zijn senioren 55 jaar en ouder.

Senioren kunnen niet worden gezien als één homogene groep. MARE en Ogilvy hebben onderzoek gedaan naar behoeften en drijfveren van senioren, hun merkgebruik en merkbeleving. Dit heeft geresulteerd in een segmentatie van deze groep.

Er zijn tevens aanknopingspunten verkregen voor de manier waarop met senioren over merken kan worden gecommuniceerd.

### *1.2. Doelstellingen en methode*

De doelstellingen van het onderzoek waren geformuleerd als het verkrijgen van inzicht in:

- Welke behoeften hebben senioren
- Wat beweegt/drijft senioren, welke levenshouding heeft men
- Welke doelen hebben zij
- Welke merken, producten en diensten spelen een rol in hun leven
- Wat is de attitude van senioren ten aanzien van merken en hun communicatie
- Wat zijn aansprekende merken en proposities voor senioren
- Hoe kunnen adverteerders met senioren over merken communiceren

Er zijn diverse methoden gehanteerd om de onderzoeksvragen te kunnen beantwoorden. Daartoe is van deskresearch, kwantitatief en kwalitatief onderzoek gebruik gemaakt.

#### *1.2.1. Kwantitatief onderzoek*

BrandZ™ is een methodologie van Millward Brown waarmee de relatie die consumenten met merken hebben, wordt gekwantificeerd. De sterke en zwakke punten van merken worden onderscheiden, gebaseerd op een analyse van bekendheid en loyaliteit van consumenten met de gemeten merken. Aan de hand van data uit BrandZ™ zijn mogelijke verschillen in de relatie met merken tussen senioren en jongere leeftijdsgroepen onderzocht. Hiervoor zijn data gebruikt van 17 metingen in 1999 en 2001 in Nederland in 13 verschillende productcategorieën. De resultaten zijn gebaseerd op een steekproef van 3578 respondenten, waarvan 1466 respondenten in de leeftijdsgroep van 55 tot 70 jaar vallen.

#### *1.2.2. Kwalitatief onderzoek*

Om behoeften en drijfveren in kaart te brengen was kwalitatief onderzoek de geëigende methode. Kwalitatief onderzoek is hypothese-genererend van aard. Het levert indicaties op van richtingen waarin antwoorden moeten worden gezocht. Door middel van open vragen kan inzicht worden verkregen in dieperliggende attitudes, motieven, drijfveren en emoties. Kwalitatief onderzoek gaat de diepte in en geeft antwoord op de vraag waarom.

#### *Respondenten*

De kwalitatieve bevindingen zijn gebaseerd op diverse onderzoeken in 2003 waaraan in totaal 140 senioren hebben meegewerkt. Hun leeftijd varieerde van 56 tot 85 jaar en zij waren afkomstig uit verschillende delen van Nederland. 40% van de participerende senioren kan omschreven worden als actief (bij de tijd, midden in de maatschappij), 40% als semi-actief (actief, maar met beperkingen door bijvoorbeeld gezondheid) en 20% van de senioren kan omschreven worden als in-actief (immobiel, geen doelen en ambities, weinig participatie in de maatschappij).

Het was de bedoeling een evenredige spreiding in actieve, semi-actieve en in-actieve senioren te realiseren. De in-actieve respondenten bleken echter nauwelijks bereid of in staat in onderzoek te participeren.

### *Methodieken*

In de kwalitatieve onderzoeken is gebruik gemaakt van verschillende methodieken.

- *In-home interviews*

Om leefgewoontes, product- en merkgebruik en attitudes met betrekking tot merken in kaart te brengen zijn in-home interviews zeer inzichtgevend. Bij mensen thuis kan bekeken worden welke merken, producten en diensten een rol spelen in hun leven, terwijl tegelijkertijd de leefomgeving bestudeerd wordt.

- *Schriftelijke vragenlijst*

Om basale doelen, behoeften en drijfveren te achterhalen die moeilijker mondeling beantwoord kunnen worden, is een schriftelijke vragenlijst verstuurd.

- *Focusgroep met experts: consumer shoes group*

Het is voor marketing-, product-, brand- en/of communicatiemanagers moeilijk om zich in de seniorendoelgroep in te leven. Een manier om daaraan tegemoet te komen is een zogenaamde expert-meeting, ofwel consumer-shoes group. Bovengenoemde managers hebben deelgenomen aan een focusgroep. Zij moesten zich elk aan de hand van een echt profiel van een senior inleven in diens leven. Tijdens de focusgroep speelden zij die senioren, die meekeken en naderhand commentaar gaven.

- *Lagerhuisdebat*

De attitude met betrekking tot merken en merkcommunicatie is onderzocht aan de hand van een ‘Lagerhuisdebat’ met als onderwerp merken en hun communicatie. Door de interactie van de groep stimuleerden de deelnemers elkaar tot het uitdiepen van de eigen en andermans mening.

- *Diepte-interviews*

Om de levenshouding te achterhalen is gebruik gemaakt van diepte-interviews.

- *Focusgroepen*

Om meer algemene doelen, behoeften en drijfveren te achterhalen zijn focus-groepen gehouden.

## 2. BELANGRIJKSTE BEVINDINGEN

De diverse onderzoeken hebben verschillende, maar aanvullende inzichten opgeleverd. Het deskresearch heeft met name voor cijfermateriaal en referentiekaders gezorgd, terwijl het kwantitatieve BrandZ™ inzichten heeft gegeven over merkgebruik en merkvoorkeuren. Meer ‘zachte’ informatie over doelen, ambities, drijfveren of angsten is uit de kwalitatieve onderzoeken gekomen. Aangezien de informatie zoals gezegd aanvullend is, zijn de inzichten op geaggregeerd niveau weergegeven.

### *2.1. Senioren en hun levenshouding*

#### *Een senior is geen bejaarde*

Senioren krijgen vaak te maken met clichés en vooroordelen in de maatschappij: senioren zouden langzaam, klagerig, ouderwets en gierig zijn en niet openstaan voor nieuwe dingen. Een groot deel van de senioren identificeert zich niet met dit beeld, wat naar hun mening de omschrijving van een *bejaarde* is. Een bejaarde is in hun ogen iemand die niet meer midden in de maatschappij staat, veel thuis zit en in hun ogen ‘afgeschreven’ is. De meeste senioren voelen zich niet bejaard en willen niet als een bejaarde gezien, noch behandeld worden.

Men voelt zich meer een *senior*: iemand die volop in de maatschappij staat, actief en geïnteresseerd is. Men wordt door een hogere leeftijd liever gezien als iemand met levenservaring, kennis en wijsheid.

Toen men jonger was zag men ouderen ook als bejaarden. Nu men zelf ouder is voelt men zich echter niet bejaard. Iemands referentiekader en daarmee ook de beoordeling van wat oud is, verandert naarmate men ouder wordt. Men voelt zich 10-15 jaar jonger en wil en kan ook nog hetzelfde als iemand die zoveel jaar jonger is.

### *Een senior geniet*

Senioren zijn over het algemeen tevreden. Ondanks de beperkingen die men kan hebben in gezondheid of inkomen, is men relatief gelukkig en geniet men. Het motto van veel senioren is genieten van elke dag die men nog heeft en deze optimaal te gebruiken.

### *Ook een senior maakt zich zorgen*

Maar er zijn ook zorgen. Senioren maken zich voornamelijk druk over maatschappelijke problemen, zoals criminaliteit. Men voelt zich vaak onveilig op straat, waardoor men zich beperkt voelt. Het verharden van de maatschappij en het vervagen van normen en waarden ervaart men als beangstigend.

Verder heeft men soms het gevoel geen overzicht meer te hebben. Er komen veel dingen op senioren af waar men nog weinig van afweet. Soms gaan maatschappelijke en technologische ontwikkelingen voor senioren te snel om bij te houden. Sommige senioren zijn door deze snelle ontwikkelingen bang om de grip op de maatschappij te verliezen en proberen derhalve ontwikkelingen zoals internet en mobiele telefonie bij te houden.

### *Senioren en hun activiteiten*

Een senior heeft een drukke agenda. Men is volop bezig. De redenen hiervoor en hoe men dit invult lopen echter uiteen.

Men vult de agenda voornamelijk om zoveel mogelijk te kunnen genieten. Men wil veel doen. Voor de senioren die net stoppen met werken betekent dit ook tijd inhalen. Was men voorheen niet in de gelegenheid om allerlei dingen te doen, nu heeft men hier wel de tijd voor. Er wordt meer tijd besteed aan familie, hobby's ontstaan en men reist en gaat uit.

Voor anderen bestaat vrije tijd uit sociale (aan)gelegenheden. Men wil graag onder de mensen zijn en zoekt de gezelligheid op, omdat men anders 'maar alleen of met zijn tweetjes thuis zit.' Men is vaak lid van een vereniging.

Daarnaast wordt de vrije tijd ingevuld uit schuldgevoel. Er is veel vrije tijd en men wil iets terug doen voor de maatschappij. Dit uit zich veelal in vrijwilligerswerk.

Ook vult men de vrije tijd in uit angst voor eenzaamheid of uit angst de grip op de maatschappij te verliezen. Men zoekt om die reden sociaal contact om binding te houden met die samenleving. Ook is men bang af te takelen en zoekt een hobby of een (thuis)studie om dit tegen te gaan. Om nog mee te tellen en van betekenis te zijn probeert men zich zowel binnenshuis als buitenshuis nuttig te maken.

## *2.2. Segmentatie*

### *2.2.1. Factoren*

Er zijn factoren die een rol spelen in het leven van senioren en die bepalend kunnen



zijn voor de mate van activiteit, drijfveren, ambities, behoeften en merk- en productgebruik van senioren. Voor elke factor is bestudeerd of senioren op basis van deze factor gesegmenteerd kunnen worden.

#### *Gezondheid*

De gezondheid van een senior bepaalt en beïnvloedt weliswaar mogelijkheden, activiteiten en behoeften van een senior, maar gezondheid verklaart niet *alle* behoeften van een senior en is om die reden geen goede segmentatiefactor.

#### *Werk en beschikbare tijd*

Het leven van een senior wordt sterk bepaald door werk. Of men werkzaam is of niet bepaalt hoeveel (vrije) tijd men heeft. Ambities, drijfveren en activiteiten van werkende en niet-werkende senioren zijn verschillend. Aangezien een groot deel van de senioren veel vrije tijd heeft, maar zij deze tijd verschillend indelen, is tijd geen verklarende factor voor de behoeften van senioren en om senioren naar te kunnen segmenteren.

#### *Inkomen*

Wanneer men ophoudt met werken treden er op financieel gebied veranderingen op. Het vaste inkomen verdwijnt. Elke senior gaat hier echter anders mee om en het *willen* genieten is onafhankelijk van het inkomen, alleen de concrete invulling daarvan is afhankelijk van het te besteden budget. Geld is een middel, maar verklaart geen doelen en behoeften waarop senioren kunnen worden gesegmenteerd.

#### *Leeftijd*

Iemands leeftijd is mede bepalend voor hoe men in het leven staat. Op hogere leeftijd is er meer kans op ouderdomsverschijnselen en hoe ouder men is, hoe korter de tijd die nog rest. Dat besef speelt een rol bij de aankoop van met name duurzame producten, maar leeftijd verklaart de *behoefte* niet.

#### *Generatie*

De tijdgeest van de periode waarin men opgroeit, kan iemands normen en waarden bepalen. Becker (1992) deelt senioren in naar generaties. De generatie waartoe de senior behoort, heeft mede bepaald wat men belangrijk vindt en welke behoeften men heeft. Binnen één generatie senioren zijn echter grote verschillen in doelen en behoeften waar te nemen. De factor generatie is derhalve onvoldoende bruikbaar om senioren naar te segmenteren.

#### *Levenshouding*

De keuze van activiteiten, doelen en behoeften van senioren hangen ook samen met een positieve of negatieve levenshouding. Levenshouding is een belangrijke onderliggende factor in de keuze van activiteiten en bepaalt doelen en behoeften.

### 2.2.2. Onderliggende dimensies

Aan genoemde factoren ligt een tweetal dimensies ten grondslag.

#### *Psychologische oriëntatie*

In verschillende beschrijvingen van senioren komen structuur en grip op de maatschappij steeds terug. De onderliggende dimensie is *persoonlijke of psychologische oriënta-*

*tie*, de behoefte aan zekerheid en controle. Deze behoefte bepaalt hoe men omgaat met de overgang van werken naar niet-werken, de behoeften die men heeft in het leven en de behoefte die men heeft met betrekking tot een merk of een product.

Een deel van de senioren leeft bij de dag en ziet wat de dag hen brengt. Zij willen genieten en een prettig leven hebben. Deze senioren hebben behoefte aan avontuur en hebben niet veel controle nodig. 'Pluk de dag' is hun motto.

Andere senioren hebben meer behoefte aan controle en aan zekerheid. Ze houden vast aan gewoonten en routines en hebben structuur nodig. Zo proberen zij grip te houden op de maatschappij.

Deze twee uitersten, namelijk *avontuur* en *zekerheid*, zijn de twee polen van deze dimensie.

### *Sociale oriëntatie*

Er zijn ook verschillen in de mate waarin senioren op zichzelf of juist naar buiten gericht zijn. Terwijl de ene senior een 'en nu ik'-houding heeft, is de andere senior gericht op gezelligheid, deze te delen met anderen of het verzorgen van anderen.

De eerstgenoemde senioren kunnen *autonomen* worden genoemd. Ze zijn voornamelijk bezig het voor zichzelf zo aangenaam mogelijk te maken. De sociale contacten zijn veelvuldig, maar zijn primair bedoeld om het eigen genot te vergroten.

De andere pool wordt gevormd door de *socialen*. Deze senioren zijn vooral op zoek naar gezelligheid, houden van het verenigingsleven en betekenen graag wat voor anderen. De sociale contacten zijn niet primair gericht op het eigen genot, maar vooral op het genot van anderen. Men geniet samen en door het anderen naar de zin te maken. De onderliggende dimensie is *sociale oriëntatie*.

### 2.2.3. Segmenten

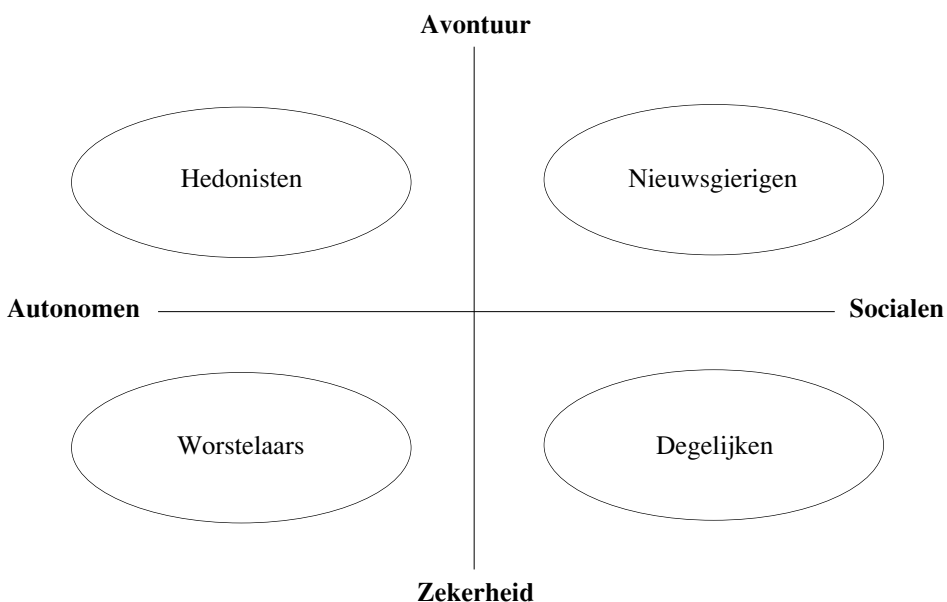
Een segmentatie op basis van deze dimensies geeft vier segmenten: hedonisten (avontuurlijke autonomen), nieuwsgierigen (avontuurlijke socialen), degelijken (zekerheidszoekende socialen) en worstelaars (zekerheidszoekende autonomen).

### *Hedonisten*

Belangrijke levensdoelen van deze groep senioren zijn genieten van het leven en vrijheid. Niet-werken is een bevrijding. Niets hoeft meer, alles mag. Hedonisten hebben een positieve kijk op het leven en maken het zo plezierig mogelijk. Ze zijn hierbij voornamelijk op zichzelf gericht. Hoewel het eigen welzijn voorop staat, gaat dat niet ten koste van een ander. 'Nu is het mijn beurt' is hun mentaliteit. Ze zijn met een open blik op zoek naar een nieuwe invulling van het niet-werkende leven.

De vrije tijd wordt besteed aan activiteiten waar *zij* plezier in hebben, die ze misschien altijd al hadden willen doen, maar waar ze eerder niet aan toe kwamen. De activiteiten zijn op henzelf (en de eventuele partner) gericht en worden gedaan om *hun* leven te verrijken. Zo wordt gezelligheid gezocht bij familie en vrienden, omdat *zij* hiervan genieten; er wordt avontuur gezocht in reizen en hobby's, omdat *zij* dat altijd al hadden willen doen.

Deze groep senioren doet haar best gezond en fit te blijven, aangezien ze van het leven willen genieten en niet bejaard willen worden.



Figuur 1. Segmentatie senioren.

#### *Nieuwsgierigen*

Het levensdoel van deze groep senioren is zoveel mogelijk genieten en samen zijn met anderen. Voor deze groep senioren is de zin van het leven doen wat je leuk vindt en daar zoveel mogelijk mensen van laten meegenieten. Hun oriëntatie is naar buiten gericht. Ze zijn nieuwsgierig en staan open voor nieuwe dingen en mensen. Ze genieten ervan nieuwe dingen bij te leren en mee te maken en genieten volop. Dit genot bestaat uit het met en voor anderen iets doen. Genieten is ook een ander laten genieten. Omdat een ander geholpen wordt of plezier heeft, heeft men dat zelf ook.

Ze hebben een actief leven. Door de diversiteit aan activiteiten hebben ze een grote sociale omgeving. Hun overvolle agenda staat vol afspraken met vrienden, familie, clubjes en verenigingen. Ze houden van mensen om zich heen, met mensen zijn en voor mensen zorgen. Met behulp van deze mensen en een druk bezette agenda houden ze contact met en grip op de maatschappij en proberen ze sociale isolatie tegen te gaan.

#### *Degelijken*

Het levensdoel van deze groep senioren is het zoveel mogelijk behouden van de dingen die men heeft en kan. Het zijn de oudere, nuchtere Hollanders. Motto's van deze groep zijn: 'Doe maar gewoon, dan doe je al gek genoeg' en 'Wie spaart die heeft wat.' Ze hebben een Calvinistische inslag en leven een simpel en eenvoudig bestaan. Luxe is niet nodig. Ze zijn met weinig tevreden. Ze zijn dit leven gewend en houden zich hieraan vast. Ze genieten echter wel. In hun vrije tijd zijn ze actief en genieten ze van sociale contacten. Familie en vrienden spelen hierbij een belangrijke rol.

Ze hebben er moeite mee wanneer ze door beperkingen af moeten wijken van gewoontes en daardoor zekerheden moeten loslaten. Ze houden namelijk sterk vast aan gewoontes. Ze kunnen derhalve moeilijk omgaan met ingrijpende gebeurtenissen. Ze zoeken zekerheid en die zekerheid valt bij een ingrijpende gebeurtenis weg.

### *Worstelaars*

Belangrijke levensdoelen van deze groep zijn controle over het leven en grip op de maatschappij houden en van nut en betekenis zijn. Deze groep wordt gedreven door de angst om controle te verliezen over hun eigen bestaan.

De achteruitgang van lichamelijke capaciteiten is een voortdurende bron van zorg en frustratie en blokkeert het genieten van de extra tijd. In de vrije tijd zijn deze senioren bezig van nut te zijn voor de maatschappij door vrijwilligerswerk te doen, mensen te verzorgen of een thuisstudie of hobby te beginnen om niet afgetakeld te raken. Ze voelen zich beter door deze activiteiten.

Pensionering wordt beleefd als verlies van maatschappelijke status. Ze voelen zich betekenisloos en zijn bezig om deze status terug te vinden in raden, besturen en vrijwilligerswerk. Ze genieten van het aanzien dat men krijgt van deze activiteiten. Ze zijn erg gericht op hun directe omgeving, hun familie en goede vrienden en genieten van hun gezelschap.

Ze kunnen slecht omgaan met achteruitgang in gezondheid. Ze zijn hierdoor bang dat ze niet meer mee kunnen doen en niet meer meetellen. Ze zijn gericht op zichzelf en wat het ouder worden voor hen betekent. Ze hebben medelijden met zichzelf.

Ze hebben moeite om te gaan met ingrijpende gebeurtenissen. Ze zoeken immers zekerheid. Veranderingen kunnen voor isolement zorgen.

### *2.3. Product- en merkgebruik door senioren*

Product- en merkgebruik door senioren wordt door een aantal factoren beïnvloed. De factoren die worden besproken zijn niet uitputtend, noch bepalen zij uitsluitend het productgebruik van *senioren*. De manier waarop de factoren van invloed zijn, kan echter verschillen voor senioren en voor andere leeftijdsgroepen.

#### *Factoren die van invloed zijn op de aankoop en het gebruik van producten en merken*

Naast factoren die de aankoop en het gebruik van producten en merken door senioren beïnvloeden, is ook het segment waartoe een senior gerekend kan worden van invloed op het gebruik van producten en merken.

#### *Merk*

Merken zijn voor jong en oud belangrijk. Welke merken aansprekend zijn is echter voor jong en oud verschillend. Over het algemeen zijn senioren geneigd vast te houden aan oude, vertrouwde en gerenommeerde merken. Dit geldt meer voor de 'worstelaars' en de 'degelijken' dan voor de andere segmenten. Oude, vertrouwde, gerenommeerde merken zijn over het algemeen bekende, vaak Nederlandse merken die men veelal al jarenlang gebruikt. Men heeft vertrouwen in deze merken en weet wat men aan deze merken heeft. Dergelijke merken hebben zich immers bewezen. Men is over het algemeen wars van allerlei beloften van nieuwe merken. Op voorhand is men in eerste instantie sceptisch en zet men vraagtekens bij de geloofwaardigheid van sommige claims, zeker als deze ingewikkeld zijn, vaag of niet inzichtelijk. Vooral de 'worstelaars' en de 'degelijken' staan hier kritisch tegenover. Zij houden sterk vast aan het merk dat zij gewend zijn. Het merk geeft hen zekerheid. Wanneer men goede ervaringen heeft met een merk, is men geneigd dit te blijven gebruiken.

### *Kwaliteit*

Kwaliteit is een factor die senioren belangrijk vinden bij het gebruik van een product. Kwaliteit kan vele vormen aannemen, zoals bijvoorbeeld de *versheid*, *gezondheid* en *puurheid* van levensmiddelen. Senioren zijn erg bezig met hun gezondheid en letten om die reden goed op wat zij eten. Zij gebruiken zo veel mogelijk verse en gezonde producten. Met behulp van gezonde maaltijden probeert men de benodigde vitamines binnen te krijgen en zo de gezondheid op pijl te houden. Men is dan ook minder geïnteresseerd in producten met toevoegingen. Doordat men veelal zelf de maaltijden met zoveel mogelijk verse producten bereidt, is men van mening dat men dergelijke producten niet nodig heeft.

*Smaak* is tevens een belangrijk kwaliteitscriterium bij de aankoop van levensmiddelen, welke staat voor de kwaliteit van een product. Senioren consumeren graag een product met een optimale smaak waarvan men kan genieten.

Vaak wordt een duur product als een goed product gezien. Als het duur is zal het wel kwaliteit hebben. A-merken zijn vaak duurder en worden dus als kwalitatief hoogwaardige merken gezien. Kwaliteit is belangrijk bij duurzame producten zoals een TV, een wasmachine of andere wit- of bruingoed producten, waarbij het aankoopbedrag hoog is en men veelal een lange tijd met het product wil doen. Men ontleent bij dergelijke producten de kwaliteit dus vaak aan het merk of de ervaring die men heeft met een ander product van datzelfde merk.

### *Prijs*

Bij de aankoop van een product waarbij de duurzaamheid of de kwaliteit minder belangrijk is, is prijs een belangrijke factor. Men is geneigd een goedkoper product te kiezen indien concessies kunnen worden gedaan aan de kwaliteit. Oudere senioren kiezen hier vaak voor. Men is zich ervan bewust dat men wellicht niet meer lang zal leven en heeft er om die reden minder behoefte aan om een product te kopen dat nog jarenlang meegaat. Men kiest dan eerder voor een goedkopere variant. De levensduur van een product wordt minder belangrijk naarmate men echt oud wordt. Prijs is vooral belangrijk voor 'worstelaars' en 'degelijken'. 'Worstelaars' vinden producten snel overbodig en willen niet teveel uitgeven. Men is erg kritisch. 'Degelijken' zijn Calvinistisch en willen om die reden niet veel geld uitgeven. Dit is 'zonde'. Dit betekent niet dat men kwaliteit onbelangrijk vindt, maar men is erg kritisch op de prijs/kwaliteit verhouding en heeft liever een product van mindere kwaliteit als dit scheelt in de prijs. Deze groep let erg op aanbiedingen en doet veel moeite om het product zo goedkoop mogelijk aan te schaffen. Aangezien voor 'hedonisten' en 'nieuwsgierigen' genieten centraal staat, is kwaliteit voor deze twee segmenten belangrijker. Dit geldt nog meer voor 'hedonisten' dan voor 'nieuwsgierigen', aangezien hun genot voornamelijk voor zichzelf is. Men wil voor zichzelf 'het beste van het beste'. De 'nieuwsgierige' haalt zijn genot uit de omgang met anderen, waarbij het samenzijn meer genot geeft dan het product zelf.

### *Aanbiedingen*

Senioren letten over het algemeen goed op aanbiedingen, maar voornamelijk 'degelijken' vergelijken prijzen en kopen het product daar waar deze het goedkoopst wordt aangeboden. Men gaat daartoe folders en huis-aan-huis bladen met aanbiedingen na. Men heeft en neemt de tijd om deze folders goed uit te pluizen. Dit geldt voor diverse productcategorieën. 'Degelijken' doen ook moeite om producten goedkoper aan te

schaffen door deze in verschillende winkels te kopen. Aangezien een aantal senioren geen groot budget te besteden heeft, zijn goedkopere supermarkten zoals de Lidl populair.

Men gebruikt aanbiedingen ook om nieuwe producten te proberen. Dit is voornamelijk het geval bij levensmiddelen. Op deze manier kan men tegen een goede prijs kennis maken met een nieuw product en deze afwegen tegen het vertrouwde product. Het risico van een miskoop is ook laag, omdat de aanschafprijs laag is.

Een nadeel van aanbiedingen is dat deze vaak grote hoeveelheden bevatten, terwijl men veelal kleine hoeveelheden consumeert omdat men hooguit een tweepersoonshuishouding voert.

### *Gebruiksvriendelijkheid*

Het is belangrijk voor een senior dat een product gebruiksvriendelijk is. Verpakkingen, gebruiksaanwijzingen en de uitvoering van bepaalde apparaten beïnvloeden de gebruiksvriendelijkheid.

Hoe ouder men wordt, hoe meer moeite men met *verpakkingen* krijgt. Sommige verpakkingen zijn te ingewikkeld omdat het niet duidelijk is hoe en waar ze geopend moeten worden. Vaak zijn letters of aanwijzingen op verpakkingen te klein. Veel senioren hebben slechte ogen. Soms is men niet in staat om verpakkingen te openen omdat er kracht of behendigheid voor nodig is.

Bij wit- en bruingoed is de *gebruiksaanwijzing* vaak een struikelblok, hetgeen een goed en optimaal gebruik van het betreffende product in de weg staat. Gebruiksaanwijzingen zijn vaak in een vreemde taal. Oudere senioren beheersen anderen talen dan Nederlands veelal niet goed genoeg en hebben daarom behoefte aan een Nederlandstalige gebruiksaanwijzing. Dit is een extra reden om voor een Nederlands merk te kiezen. Ook in de gebruiksaanwijzing zijn de lettertjes veelal te klein.

Senioren hebben daarnaast vaak moeite om apparaten optimaal te gebruiken of te bedienen. Doordat men de gebruiksaanwijzing niet goed kan lezen, weet men niet welke functies een product allemaal heeft. Men gebruikt daardoor meestal maar een beperkt deel van het product. Men kiest daarom ook vaak voor een simpel model zonder veel 'poespas'.

Doordat een senior meestal minder goede ogen heeft, heeft men vaak ook moeite met de grootte van de knopjes op producten en de functies die op de knopjes worden aangegeven. Vooral op moderne en compacte apparaten zijn de knopjes vaak klein. Men kan dan bijvoorbeeld slecht onderscheid maken tussen de functies op onder andere afstandsbedieningen, video's of mobiele telefoons.

Omdat senioren niet met geavanceerde techniek opgegroeid zijn, moeten zij zich dat zelf eigen maken. Of men dat al dan niet doet is mede afhankelijk van de angst voor technologie. Die angst is meer aanwezig bij 'worstelaars' en 'degelijken', die liever bij het vertrouwde blijven. 'Hedonisten' en 'nieuwsgierigen' zijn nieuwsgierig ingesteld en zijn bereid met nieuwe technologieën kennis te maken.

### *Gemak*

Gemaksproducten hebben niet altijd de interesse van senioren. De niet-werkende senioren hebben veel vrije tijd en hebben dan minder voordeel van gemak. Sommige gemaksproducten, zoals bijvoorbeeld *kant-en-klare-(magnetron)maaltijden*, stellen senioren en dan voornamelijk 'hedonisten' en 'nieuwsgierigen', echter in staat meer tijd over te houden om te genieten. Een situatie waarin een kant en klare maaltijd ook

aantrekkelijk is, is wanneer een senior alleen komt te staan. Een ander voordeel van kant-en-klaar-maaltijden zijn de kleine porties waarin deze verkocht worden, hetgeen aansluit op het kleine huishouden. Een groot deel van de senioren geeft er echter de voorkeur aan de maaltijden zelf klaar te maken.

Een ander gemakproduct dat voornamelijk ‘hedonisten’ en ‘nieuwsgierigen’ gebruiken, is de *computer* om te e-mailen en te internetten. Zo boekt men losse vliegtickets, zoekt men naar een theatervoorstelling of mailt men met de (klein)kinderen.

Daarnaast gebruiken zij de *magnetron* om bijvoorbeeld ingevroren porties te ontdoien.

#### *Invloed van de sociale omgeving op (het proberen van) nieuwe producten en merken*

Ook in de sociale omgeving van senioren zijn er factoren die de aankoop van producten en merken kunnen beïnvloeden.

- Senioren worden beïnvloed door de *kinderen* (indien men die heeft). Uit deskresearch is gebleken dat senioren voordeel dienen te zien in een nieuw product, wil men hiervoor openstaan. De kinderen van senioren komen eerder in aanraking met nieuwe producten dan de senioren zelf. Wanneer de kinderen tevreden zijn over een product, zal men deze de ouders aanraden en de voordelen ervan uitleggen. Voornamelijk ‘hedonisten’ en ‘nieuwsgierigen’ zijn geneigd een nieuw product te proberen. Indien het product bevalt of meerwaarde heeft boven het vertrouwde product, zal men het opnieuw gebruiken. ‘Degelijken’ en ‘worstelaars’ zijn moeilijker te overtuigen van de meerwaarde van een product. Zij hebben vaak behoefte aan ‘harder’ bewijsmateriaal.
- Senioren brengen veel tijd door met hun *kleinkinderen* (indien men die heeft). Door kleinkinderen komen senioren ook met nieuwe producten in aanraking. De kleinkinderen gebruiken andere producten dan de senioren en nemen deze mee of vragen de grootouders deze voor hen te kopen. Daarnaast willen senioren kleinkinderen vaak verwennen en koopt men producten die de kleinkinderen graag eten of gebruiken. Hierdoor komen senioren bijvoorbeeld in aanraking met speelgoed, dvd’s, nieuwe smaken chips of andere snacks, koeken zoals Liga, kinderverzorgingsproducten en drankjes in éénpersoonspakjes.
- Indien *vrienden* een product en diens voordelen aan een senior voorleggen, zijn ‘hedonisten’ en ‘nieuwsgierigen’ geneigd dit nieuwe of andere product te proberen. Een senior zal een product echter alleen gaan gebruiken, wanneer hij of zij er zelf ook voordeel van heeft.
- Veel senioren raadplegen de *Consumentengids* om na te gaan wat de beste aankoop is, bijvoorbeeld wat betreft de prijs/kwaliteit verhouding en welke voordelen producten hebben. Door de Consumentengids krijgt men vertrouwen in de aankoop. ‘Degelijken’ en ‘worstelaars’ zijn vaak onzeker over een aankoop en testen halen onzekerheid weg.
- Om dezelfde reden laten voornamelijk deze twee segmenten zich beïnvloeden door *consumentenprogramma’s* waarin producten worden vergeleken en de voor- en nadelen op een rijtje worden gezet. ‘Degelijken’ en ‘worstelaars’ hebben vaak weinig vertrouwen in nieuwe producten en geloven de beloften niet zonder meer. Consumentenprogramma’s kunnen ‘bewijs’ leveren van de voordelen van een (nieuw) product en het gebruik van een product uitleggen.

*Productkeuze en productbehoefte per segment*

- *Hedonisten* willen genieten. Geld wordt aangewend om kwaliteit en plezier in te kopen. Merken of producten die het hen makkelijker of plezieriger maken, waardoor zij nog meer kunnen genieten, zijn interessant. Een product of dienst moet genot verschaffen. A-merken zijn belangrijk, omdat deze staan voor kwaliteit. Ze zoeken het beste van het beste, omdat kwaliteit hen de gelegenheid geeft optimaal te genieten. Prijs is meestal van ondergeschikt belang.

Ze ontdekken en leren graag nieuwe dingen en staan open voor nieuwe merken en producten. Ze zijn niet bang voor technologie en hebben weinig moeite met apparaten, gebruiksaanwijzingen of verpakkingen.

Gemak wordt ingekocht indien het product een meerwaarde heeft ten opzichte van andere producten en indien men niet inlevert op kwaliteit en smaak. Gemak geeft deze doelgroep de mogelijkheid tijd te besteden aan andere dingen die weer genot opleveren.

- *Nieuwsgierigen* proberen te genieten. Ze zijn nieuwsgierig en proberen bij te blijven. Ze zoeken naar producten, diensten en merken waardoor zij samen met anderen kunnen genieten. Omdat samenzijn de behoefte aan genieten vervult, is deze groep minder gericht op merken. Daardoor behoren ook B-merken tot de evoked set, mits deze van voldoende kwaliteit zijn om optimaal te kunnen genieten. Wel hebben zij bij gerenommeerde merken meer vertrouwen in de kwaliteit. Wanneer zij kiezen voor een gerenommeerd merk, zullen zij kiezen voor het merk met de beste prijs/kwaliteit verhouding.

Zij staan open voor nieuwe producten en diensten en zijn hier ook nieuwsgierig naar, maar zijn daar iets gereserveerder en behoudender in dan 'hedonisten'. Dit segment heeft een sterkere behoefte aan advies en raad van anderen dan 'hedonisten'. Zij proberen nieuwe producten en diensten uit indien dit op aanraden is van influentials uit de sociale omgeving of wanneer men een duidelijke meerwaarde in het nieuwe product of dienst ziet. Soms is deze groep ook nieuwsgierig naar een nieuw product of dienst, wanneer deze in de samenleving al veelvuldig wordt gebruikt (zoals internet).

- *Degelijken* genieten van de dingen die ze hebben en kunnen, maar met een Calvinistische inslag. Ze doen dit graag samen met anderen. De prijs/kwaliteit verhouding is belangrijk. Ze willen niet veel betalen. Ze zijn gevoelig voor aanbiedingen en doen moeite om een product zo goedkoop mogelijk aan te schaffen. Ze zoeken naar producten en diensten die zekerheid verschaffen.

'Degelijken' staan minder open voor nieuwe producten. Deze groep gebruikt al jaren dezelfde producten en merken. Deze zijn vertrouwd. Men weet wat men aan deze producten heeft en men is hieraan gewend. Men is zeer kritisch. De mening van vrienden en familie, de Consumentengids en consumentenprogramma's kan een bijdrage leveren aan het overtuigen van deze doelgroep. Het is moeilijk om deze doelgroep te overtuigen van de meerwaarde van een product of een dienst, maar wanneer het lukt heeft men wel een trouwe klant.

- *Worstelaars* zijn op zoek naar zekerheid en controle. Ze zoeken ook zekerheid in producten, dus ze willen van de producten op aan kunnen. Zij hechten grote waarde aan de kwaliteit van een product en kwaliteit verbinden zij aan merken. Voornamelijk oude, vertrouwde merken geven hen de zekerheid een goed product in handen te



hebben. Wel is de prijs/kwaliteit verhouding van belang. Men is nauwelijks bereid veel voor een product te betalen en men vindt veel zaken al snel te duur, tenzij men overtuigd is van het nut en de kwaliteit van een product. In dat geval is men wel bereid daar geld aan uit te geven.

Deze groep staat nauwelijks open voor nieuwe merken, producten of diensten, die veelal overdreven, overbodig of onzinnig worden gevonden. Een product of dienst moet zijn nut bewezen hebben alvorens men er gebruik van maakt.

#### *2.4. TV en reclame bij senioren*

Moderne commercials gaan vaak te snel voor senioren, met name de commercials waarbij korte en snel opeenvolgende beelden worden gebruikt. Men mist vaak waar de commercial over gaat en begrijpt daarom de boodschap niet. Verhaallijnen waarbij aan het begin niet duidelijk is welk product of dienst wordt aangeprezen, worden veelal niet begrepen. Men is zó bezig de verhaallijn te begrijpen, dat men de plot mist en vervolgens de boodschap van de commercial niet meekrijgt of zelfs het merk of het product waar het over gaat mist.

Vooraf oudere senioren hebben moeite met een tag-on. Tegen de tijd dat de tag-on verschijnt, is men de commercial, waar de tag-on bij hoort, al vergeten.

Verder heeft men moeite met commercials in het Engels, waarbij geen ondertiteling is gegeven. Men heeft vaak moeite Engels te begrijpen en kan om die reden de commercial niet verstaan. Hierdoor komt de boodschap van deze commercial niet over en voelt men zich niet aangesproken.

Om bovenstaande redenen geeft de senior de voorkeur aan een commercial met een simpele verhaallijn. Daarnaast wordt men graag vermaakt en worden commercials met humor derhalve sterk gewaardeerd.

Senioren hebben geen voorkeur voor commercials waarin ouderen voorkomen. Men kan zich ook herkennen in jongere mensen, aangezien men zich over het algemeen niet oud voelt en zich identificeert met mensen die 10-15 jaar jonger zijn.

Senioren prefereren commercials waarin senioren op een positieve manier worden voorgesteld, dus niet zijn afgebeeld naar de clichés en vooroordelen die over hen in de maatschappij heersen. Indien deze vooroordelen gebaseerd zijn op humor, duidelijk overdreven zijn en met een knipoog bekeken moeten worden, worden dergelijke commercials wel gewaardeerd. Een goed voorbeeld hiervan is de commercial van Lays, waarin twee senioren vechten om een zak chips en de oma het kunstgebit van de opa steelt.

Aansprekende communicatie voor senioren:

- communiceert een single minded propositie
- communiceert een duidelijke benefit
- communiceert een geloofwaardige benefit
- laat de afzender in een vroeg stadium bekend worden
- heeft een rechte verhaallijn
- gebruikt geen tag-on's
- toont herkenbare, geloofwaardige en relevante situaties
- beeldt een senior positief en waardig af
- gebruikt humor

- heeft geen hele korte, flitsende beelden
- heeft geen te snel opeenvolgende beelden
- is ondertiteld indien de taal niet Nederlands is.

### 3. CONCLUSIES

Senioren hebben veel last van vooroordelen vanuit de maatschappij. Een groot deel van de senioren identificeert zich niet met het beeld dat over hen bestaat. Senioren worden door jongere mensen echter allemaal over één kam geschoren en als bejaarden gezien. Men voelt zich echter jonger dan men is. Men is over het algemeen tevreden met wie men is en hoe men leeft. Men heeft als belangrijkste doel niet te snel bejaard te worden en ondertussen te genieten van wat men kan en heeft.

Daardoor heeft een senior het druk. De mate waarin men het druk heeft, waarmee men het druk heeft en de behoeften en doelen die men heeft, verschillen echter. Senioren kunnen dus niet worden gezien als één homogene groep. Senioren kunnen worden gesegmenteerd naar psychologische oriëntatie (de mate waarin men structuur en zekerheid nodig heeft) en sociale oriëntatie (is men op zichzelf of op anderen gericht). Zo zijn vier segmenten te identificeren:

#### *Hedonisten*

Genieten staat voorop en is op henzelf gericht. 'Nu is het mijn beurt' is de mentaliteit. Activiteiten moeten *hun* leven verrijken. Ze staan positief in het leven en kunnen goed omgaan met veranderingen. Ze staan open voor nieuwe producten, diensten en merken. Merken of producten die het hen makkelijker maken en waardoor ze meer kunnen genieten zijn interessant. A-merken en kwaliteit zijn belangrijk om optimaal te kunnen genieten.

#### *Nieuwsgierigen*

Deze groep wil samen met anderen genieten. Ze houden van mensen om zich heen, met mensen zijn en voor mensen zorgen. Zo houden ze contact met en grip op de maatschappij. Ze zijn van nature nieuwsgierig en proberen bij te blijven. Ze staan open voor nieuwe producten en diensten, maar doen dit wel op aanraden van influentials of een betrouwbare bron. Om te kunnen genieten met anderen is de kwaliteit van een product belangrijker dan het merk. B-merken horen ook tot de evoked set, mits deze van voldoende kwaliteit zijn om te kunnen genieten.

#### *Degelijken*

Deze groep geniet van de dingen die men heeft en kan, maar heeft een Calvinistische inslag. Ze leven een eenvoudig bestaan. Ze zijn dit gewend en houden hieraan vast. In hun vrije tijd zijn ze actief en genieten ze van sociale contacten. Ze gebruiken al jaren dezelfde producten en merken. Ze zijn kritisch en moeilijk te overtuigen. Nieuwe producten en diensten vinden hooguit een ingang via betrouwbare bronnen als kinderen, kleinkinderen of vrienden, een Consumentengids of consumentenprogramma. De prijs/kwaliteit verhouding van een product is erg belangrijk. Daarnaast houdt men vast aan oude, vertrouwde merken.

### *Worstelaars*

Belangrijke levensdoelen van deze groep zijn controle houden over het leven en de maatschappij en van nut en betekenis zijn. Deze groep wordt gedreven door de angst om controle te verliezen over hun bestaan.

Een deel van de worstelaars is in de vrije tijd bezig van nut te zijn door vrijwilligerswerk te doen, mensen te verzorgen of een thuisstudie of hobby te beginnen om niet afgetakeld te raken. Een ander deel wil nog van betekenis zijn en probeert dat door status terug te vinden in raden, besturen en vrijwilligerswerk.

Ze staan minder open voor nieuwe producten en zijn behoudend. Ze zoeken zekerheid in vertrouwde producten, diensten en A-merken, aangezien ze hiermee verzekerd zijn van een bepaalde kwaliteit. Een product of dienst moet zijn nut bewezen hebben, alvorens ze er gebruik maken.

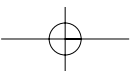
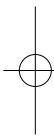
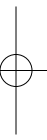
### *Do's en don'ts voor producten en communicatie voor senioren*

- Een strategie die eerlijk, helder en duidelijk de voordelen van een product communiceert, is aan te bevelen. Senioren hebben behoefte aan een duidelijke, simpele en geloofwaardige 'reason to believe'. Deze reason to believe moet bovendien nog niet eerder door andere producten of merken geclaimd zijn, zodat de meerwaarde van dit specifieke product of merk duidelijk is. Senioren verschillen hierin dus niet veel van andere doelgroepen, maar behoeven wel een andere aanpak met betrekking tot communicatie.
- De waardering en het begrip voor tv-commercials vallen samen met een heldere en duidelijke verhaallijn. Er dient een duidelijke link te zijn tussen het product en de verhaallijn. Voor een goed begrip dient de afzender aan het begin van de commercial bekend te zijn. Daarnaast moet het product of dienst dat wordt aangeboden relevant zijn. Senioren geven de voorkeur aan een single minded propositie. Voor een goede waardering is humor belangrijk. Het gebruik van tag-ons wordt afgeraden, evenals korte en snel opeenvolgende beelden. Engelstalige communicatie moet vertaald of ondertiteld worden.  
Over het algemeen heeft men geen waardering voor commercials waar senioren naar de vooroordelen en de clichés die in de maatschappij leven, worden afgebeeld. Wanneer het echter duidelijk is dat dit overdreven, met humor en een knipoog is gebracht, is er wel waardering.
- De hoeveelheid waarin producten verpakt zijn, zijn voor senioren veelal te groot, aangezien men een klein huishouden heeft. Men heeft een voorkeur voor portieverpakkingen.
- Veel producten en apparatuur worden door senioren niet optimaal gebruikt, omdat men niet alle functies kent. Dit is te wijten aan onduidelijke gebruiksaanwijzingen of niet-gebruiksvriendelijke bedieningen.
- Gebruiksaanwijzingen zijn vaak onduidelijk. Een gebruiksaanwijzing dient in het Nederlands te zijn geschreven met een groot lettertype en dient helder te zijn in de uitleg.

- Knoppen op apparatuur en afstandsbedieningen zijn vaak te klein, staan te dicht op elkaar en de functies zijn niet duidelijk weergegeven. Senioren hebben grote knoppen met een duidelijk opschrift nodig.
- Senioren geven de voorkeur aan duidelijke en handige verpakkingen. Daarnaast wenst men grote letters op verpakkingen met duidelijke verwijzingen naar de opening of het gebruik en moeten verpakkingen met weinig kracht te openen zijn.
- Op basis van het voorgaande lijkt het moeilijk een nieuw merk of product te introduceren dat ook voor senioren aantrekkelijk is. Aangezien men merktrouw is en veelal eerst van het nut van een nieuw product overtuigd moet zijn, lijkt men niet zondermeer open te staan voor iets nieuws. Duidelijk is echter dat de bereidheid een nieuw product te proberen mede afhankelijk is van het segment waartoe men behoort en het advies van derden. En als men deze nieuwe klant eenmaal aan zich gebonden heeft, heeft men ook een trouwe klant.

#### LITERATUUR

- 50+ Plus Marketing. Marketing, Communication and Selling to the over 50s Generations.* (2002)
- Beale, C. (1998). Over-55s turned off by "clever" TV advertising. *Campaign*. 20-11-1998. p. 8.
- Becker, H.A. (1992). *Generaties en hun kansen*.  
BrandZ. WWP Group USA.
- Datamonitor, (2003). *Targeting seniors effectively*.
- Doormans, C. (2003). Oud geld. Uw koper is niet jong en blits. *Automotive*. 10-4-2003, nr 7 p. 18.
- Ed Nijpels zet ouderen aan het werk. *Midi 4*, 2003. p. 10-13.
- Geron, (2003) *Tijdschrift over ouder worden en maatschappij*. Juni 2003.
- Engelen, M.W.H. (2003). Basis. Van grijze muis tot grijze panter. *Tijdsbesteding van senioren in Zuid-Holland*. 31-3-2003, nr 1, p 8.
- Parmar, A. (2003). Knowledge of mature market pays reward. *Marketing News*, 28-4-2003. p. 5.
- SCP (2001). *Het nieuwe consumenten*.
- SCP (2001). *Rapportage Ouderen*.
- SCP (2003). *Rapportage Sport*.
- Senioren in Nederland*, onderzoeksrapport MARE, September 2003.
- Senmar Consultancy (2000). *Handboek seniorenmarketing*.
- Succes 50+beurs schuilt in aftersales aan senioren. *Adformatie*. 7-10-1999.
- Tempest, S., Barnatt, C. & Coupland, C. (2003). Nieuwe Strategie voor senioren. *Management Review*. Nr 89-2003.
- Ward, D. (1984). *Over 50s: How to reach out to an ignored consumer*. Ogilvy & Mather. New York.



### *3. Aandacht voor aandacht*

L. HULSEBOS, D. SILLE en J.L.C.M. WOLTMAN ELPERS

#### SAMENVATTING

Technologische en methodologische doorbraken hebben reclameonderzoek voorgoed veranderd. Het is nu mogelijk om met directe metingen als eye-tracking het gedrag van consumenten te meten, in plaats van naar hun gedrag te vragen (met gebruikmaking van het imperfecte geheugen van de respondent). Tevens gebeurt dit in een natuurlijkere situatie – zonder gedwongen confrontatie met de reclame-uiting – zodat ook het reclamebereik kan worden bepaald. Dit vult de ontbrekende schakel van het huidige mediaonderzoek in; men kan nu ook rekening houden met de communicatieve kracht van een individuele uiting. Voor pretesten betekent dit dat men vooraf de communicatieve kracht van een reclame-uiting kan bepalen en – door middel van vaak kleine aanpassingen aan de executie – aanzienlijk kan verbeteren. Voor posttesten en tracking-onderzoek betekent dit dat men op ieder moment de bijdrage van een individuele uiting aan de communicatiedoelstellingen vast kan stellen. Dit betekent in de praktijk een efficiënter gebruik van het beschikbare mediabudget.

#### 1. INLEIDING

Het aantal reclameboodschappen waar we iedere dag mee geconfronteerd worden neemt nog steeds toe. Voor de afzender is het van belang dat reclame opgemerkt wordt en in voldoende mate aandacht krijgt, zodat de boodschap begrepen kan worden.

In de loop der jaren is veel geschreven over de werking van reclame. Aanvankelijk zag men aandacht als de eerste fase in het proces van reclamewerking. Dit werd zelfs geformaliseerd in de traditionele reclamewerkingsmodellen zoals bijvoorbeeld AIDA (Attention, Interest, Desire, Action). Aandacht werd in de praktijk vaak gelijk gesteld aan het bereik van een tijdschrift of programma. We zien dit ook terug in de gebruikte methoden en technieken voor het onderzoeken van reclameboodschappen waar de aandacht voor een uiting wordt geforceerd, zonder dit daadwerkelijk te meten.

De laatste jaren is de theorievorming rondom de rol die (visuele) aandacht in het proces van reclamewerking heeft echter sterk veranderd. Aandacht is niet slechts de eerste fase in het proces van reclamewerking. Steeds duidelijker blijkt dat visuele aandacht het hele proces van reclamewerking stuurt en richting geeft. Onderzoek van Humphreys in het Verenigd Koninkrijk, Iti en Koch en anderen in de VS, en Pieters en Wedel in Nederland en de VS laten zien dat (1) de eerste en tweede functie van aandacht cruciaal zijn in reclamewerking en in de waarneming van andere complexe stimuli, dat (2) deze functies kunnen worden vastgesteld met nieuwe technieken van oogbewegingsregi-

stratie, dat (3) deze processen en effecten niet met andere methodieken, zoals geheugenmetingen of vragenlijsten kunnen worden vastgesteld en dat (4) de inzichten uit oogbewegingmetingen direct gebruikt kunnen worden om de effectiviteit van media- en reclameplanning aan te pakken.

In dit artikel geven we een overzicht van de ontwikkelingen in het meten van visuele aandacht. Daartoe staan we in sectie 2 eerst stil bij de wijze waarop mensen waarnemen en hoe kijkgedrag werkt. Vervolgens zullen we de wijze waarop aandacht (oftewel reclamecontact) gemeten kan worden bespreken in sectie 3. Daarna wordt in sectie 4 ingegaan op de nieuwste ontwikkelingen op het gebied van eye-tracking en de inzichten die dit tot dusver heeft opgeleverd. In sectie 5 besluiten we met conclusies.

## 2. WAARNEMEN EN KIJKGEDRAG

Door onderzoek zijn we steeds meer te weten gekomen over hoe mensen visueel waarnemen. Een belangrijke conclusie luidt dat ‘zien’ niet bestaat uit het creëren van een representatie van de werkelijkheid in de hersenen, maar eerder uit de wetenschap dat de werkelijkheid om ons heen op elk moment toegankelijk is door daar aandacht aan te schenken.<sup>1</sup> We nemen zogezegd geen ‘foto’ van onze omgeving die we naar willekeur in ons geheugen kunnen bekijken. We gebruiken de werkelijkheid meer als een soort ‘buitengeheugen’, een encyclopedie die we kunnen gebruiken door er aandacht aan te schenken. Dit heeft verregaande gevolgen voor de manier waarop men reclame waarneemt en verwerkt.

Hiervoor dacht men altijd dat het oog fungeerde als een ‘camera’: het oog registreert een beeld dat men in het geheugen kan scannen op informatie. Deze theorie houdt niet langer stand. We weten nu dat het oog maar een beperkt stukje van de omgeving werkelijk kan waarnemen, namelijk door daar op te fixeren. De rest van de omgeving blijft een vaag gebied, totdat men daarop fixeert. De aandacht van het oog kan wel actief worden getrokken door objecten in de omgeving, met name beweging (verandering). Dit is ook de reden dat bijvoorbeeld waarschuwingslichten meestal knipperen.

Vaak wordt dus niet alles binnen een reclame-uiting gezien. Het is daarom erg belangrijk de uiting zo te creëren, dat de belangrijkste delen de aandacht van het oog trekken zodat daar op wordt gefixeerd. Binnen een printadvertentie betekent dit vaak de visual, (een deel van) de tekst en het merk (de afzender). Pas wanneer het oog op al deze onderdelen heeft gefocust, kan er overdracht van de boodschap plaatsvinden.

Hetzelfde geldt ook voor een tv-commercial. Het is – juist binnen een bewegend medium waar veel onderdelen om aandacht vragen – belangrijk dat de juiste onderdelen de aandacht van de consument krijgen.

De wijze waarop mensen waarnemen is een belangrijk gegeven dat in de dagelijkse praktijk van communicatie niet of nauwelijks wordt gebruikt. Het aantal mensen dat

<sup>1</sup> J. Kevin O’Regan, Laboratoire de Psychologie Expérimentale, Centre National de Recherche Scientifique, Paris, France.

televisie kijkt is belangrijk, maar het aantal mensen dat naar een bepaalde commercial kijkt en ook nog de juiste dingen ziet bepaalt het effect.

Een ander probleem is dat consumenten steeds meer keuzemogelijkheden hebben. Het actief mijden van reclame is een normaal verschijnsel geworden. We zappen massaal commercials weg en pagina's die niet de aandacht weten te behouden worden gewoon doorgebladerd. Uit verschillende onderzoeken (Adformatie september 2002 en 2003) blijkt dat meer dan 80 van de 100 mensen afhaken bij het zien van reclame. Dat dit niet alleen een Nederlands verschijnsel is blijkt uit meerdere artikelen (Woltman Elpers, 2003).

Het zap- en bladergedrag van consumenten kan voor de adverteerder dan ook grote gevolgen hebben. Het betekent namelijk dat het geïnvesteerde bedrag in mediaplacement niet het gewenste effect zal hebben. Belangrijk is daarom te bekijken wat het effect van zappen is op reclame- en merkeffecten.

### 3. HET METEN VAN RECLAME-CONTACT

Uit de voorgaande sectie blijkt dat aandacht krijgen voor een uiting niet vanzelfsprekend is. Het is niet voldoende dat de kijker zijn of haar ogen open heeft op het moment dat de reclameboodschap wordt getoond. Hier is meer voor nodig. In dit kader wordt vaak gesproken over 'aandacht'; wij denken echter dat 'reclamecontact' een beter begrip is om aan te duiden wat noodzakelijk is voor het kunnen verwerken van de boodschap. Oftewel: hoe komen we van "Opportunity to See" naar "Certainty to See".

Om de stap van mediumbereik naar reclamebereik te kunnen maken, is het belangrijk om vast te stellen op welke wijze reclamecontact gemeten zou moeten worden. Sinds het begin van de 20e eeuw wordt onderzoek naar reclamecontact uitgevoerd (Nixon, 1924 en Poffenberger, 1925). Dit gebeurt op twee fundamenteel verschillende manieren, welke in de volgende paragrafen verder worden uitgelegd:

1. Indirecte metingen op basis van geheugen, zoals advertising recognition
2. Directe metingen zoals eye-tracking

Ondanks de aantoonbare grote problemen met de validiteit van ad recognition (tijdschrift doornemen, met herkenningsvragen achteraf, zie ook paragraaf 3.1) wordt deze methode nog steeds gebruikt. Dit artikel laat zien hoe directe metingen, gebaseerd op eye-tracking (het registreren van gedrag tijdens het doornemen van een tijdschrift, zie paragraaf 3.2) tot nauwkeurige, nieuwe inzichten over gerealiseerd reclamebereik leiden die voordelig zijn voor de marktonderzoeksbureaus die ze hanteren, en voor de adverteerders en reclamebureaus die er gebruik van maken.

Communicatie, oftewel reclamebereik is een functie van mediabereik en kennisname van de individuele uiting oftewel:

*Mediabereik x reclamecontact = gerealiseerd reclamebereik.*



Doel van beide methoden (direct en indirect) is vast te stellen of consumenten voldoende contact hebben gehad met een bepaalde reclame-uiting om tot verwerking van de boodschap te kunnen komen.

Beide methoden kunnen echter niet vaststellen of een reclame-uiting effectief is (geweest). Immers, reclame-effect is het resultaat van herhaalde aanbieding van de reclame-uiting, en van reclame-uitingen van de concurrentie, in steeds wisselende contexten, die gecombineerd worden met de bestaande kennis van consumenten.

Dergelijke geaccumuleerde reclame-effecten:

- zijn moeilijk te herleiden tot een individuele reclameboodschap;
- vinden vaak vertraagd plaats (reclame heeft veelal niet een rechte effect);
- vinden vaak onder lage betrokkenheid (low involvement) plaats, wat herhaalde reclamecontacten vereist.

Om deze redenen wordt en werd reclamecontact gemeten. Bovendien heeft het meten van reclamecontact een belangrijk voordeel. Directe metingen van reclamecontact maken het mogelijk om voor elke reclame-uiting afzonderlijk vast te stellen of hij voldoet aan de minimale voorwaarden om het beoogde effect te kunnen hebben. Een valide meting maakt het daarom mogelijk om (a) de juiste reclame-uitingen te selecteren (pretesting), (b) de inzet van verschillende reclame-uitingen binnen eenzelfde media type vast te stellen (intra-mediaplanning), (c) reclame-uitingen tussen verschillende mediumtypen onderling te vergelijken (inter-mediaplanning), (d) te bepalen in hoeverre een reclame-uiting nog kan bijdragen aan het realiseren van de communicatie doelstellingen (post testing en wear-out).

### 3.1. Ad recognition

Ad recognition werd geïntroduceerd rond 1915 door Daniel Starch voor dagblad- en tijdschriftreclame. In vrijwel ongewijzigde vorm wordt het nu nog steeds toegepast voor print en in afgeleide vorm voor televisiereclame. Aan mensen wordt gevraagd zich te *herinneren* of ze contact hebben gehad met een reclame-uiting, op basis van confrontatie met de oorspronkelijke advertentie (vandaar ad recognition). Hierbij wordt verondersteld dat mensen weten waar ze naar gekeken hebben, en dit op de juiste manier kunnen rapporteren. Deze methode wordt tegenwoordig steeds vaker via Internet uitgevoerd, waardoor de antwoorden en de setting nog minder controleerbaar zijn.

Het belangrijkste probleem met ad recognition is de validiteit, dat wil zeggen wat het eigenlijk meet. Bij ad recognition is dat *“de uitkomst van alle voorafgaande reclamecontacten met deze advertentie en andere advertenties van dit merk, advertenties van vergelijkbare merken, de associaties van consumenten met het merk en de advertenties van dit en andere merken, en het idee dat consumenten hebben van wat de onderzoeker wil weten.”* Met andere woorden, ad recognition meet juist niet het reclamecontact, en zelfs alles behalve het reclamecontact. Dit komt door verschillende processen:

- Geheugenproblemen

Het menselijk geheugen is heel plastisch, en niet een “archiefkast” waarin we dingen opslaan onder een duidelijk adres, en dat later weer precies goed ophalen van dat

adres. Mensen denken dingen gezien te hebben die er niet zijn, en ze zien dingen niet die er wel zijn. Als je een goede vriend ziet die zijn snor heeft afgeschoren, vraag je hem of hij een andere bril op heeft. We zien dat er iets is veranderd aan de vriend, maar niet wat. Met andere woorden, het geheugen maakt het niet altijd mogelijk goed te rapporteren.

– Motivatieproblemen

Mensen willen niet altijd het juiste ‘geheugen-antwoord’ geven, ook al zou dat wel kunnen. Mensen hebben de automatische neiging zich beter voor te stellen dan ze zijn, en dat hebben ze vaak zelf niet door. We denken dat we vaker naar het nieuws kijken, en minder vaak naar een soap opera dan we feitelijk doen. We lezen de Volkskrant beter dan Privé, denken we, als we intellectualistisch zijn belangrijk vinden. Mensen overschatten het aantal keren dat ze naar de bioscoop gaan, sporten om dezelfde redenen, omdat het ons goed doet voelen over onszelf.

Deze processen werken onmiddellijk en voortdurend. Zelfs als we ons ervan bewust zijn, kunnen we ons er moeilijk of niet tegen wapenen.

Het gevolg is dat de verkregen antwoorden op ad recognition vragen niet valide zijn. Op de Universiteit van Tilburg is – in samenwerking met Verify International – groot-schalig onderzoek gedaan naar dit fenomeen (Draisma 2000). Het Starch onderzoek werd uitgevoerd (respondenten nemen een tijdschrift door, waarna achteraf de advertenties nogmaals worden vertoond en er wordt gevraagd of men deze tijdens het lezen heeft gezien. Tegelijkertijd werd ook eye-tracking onderzoek (zie paragraaf 3.2) onder dezelfde consumenten gedaan.

Wat bleek:

- hoe vaker mensen naar de illustratie van een advertentie kijken, hoe vaker ze rapporteren (ad recognition) dat ze de tekst van de advertentie hebben gelezen.
- hoe bekender een merk is, hoe meer mensen rapporteren dat ze het merk hebben gezien, onafhankelijk van de mate waarin ze het merk daadwerkelijk hebben bekeken.

Deze bevindingen zijn zeer systematisch, over vele advertenties en consumenten en ze zijn in gelijksoortige vorm ook door anderen gevonden (bijvoorbeeld Finn, 1988). Ze betekenen bijvoorbeeld dat de prestatie van advertenties van bekende merken systematisch wordt overschat door ad recognition onderzoek. Ook wordt de mate waarin we denken dat mensen tekst binnen advertenties lezen systematisch overschat. Bovendien zijn ad recognition maten traag: ze pikken niet de onmiddellijke effecten van advertenties op, maar de vertraagde effecten van allerlei uitingen uit het verleden.

Dit alles betekent dat ad recognition scores wel gemakkelijk te verzamelen, te rapporteren en te begrijpen zijn (het zijn namelijk eenvoudigweg percentages), maar dat ze maar heel weinig, en vaak zelfs het verkeerde zeggen over wat het feitelijke reclame-contact van een individuele uiting is.

### 3.2. Directe metingen

Bij directe metingen zoals eye-tracking (in sectie 4 wordt eye-tracking uitvoerig besproken) wordt direct vastgesteld of en in welke mate mensen kennis nemen van de inhoud van de advertentie. Deze metingen zijn niet gebaseerd op het geheugen van respondenten. Alhoewel men er altijd van overtuigd was dat deze directe methoden meer betrouwbaar, nauwkeurig en valide waren dan indirecte, was grootschalige toepassing van eye-tracking tot voor kort niet mogelijk.

Wat nodig is, is een meting van reclamecontact die direct en valide is (het meet datgene wat je wil weten, en niets anders), onmiddellijk (en niet vertraagd) en nauwkeurig (detail informatie over het proces). Bovendien dient een dergelijke meting op industriële schaal toegepast te kunnen worden, tegen lage kosten. Het is momenteel de enige methode om onweerlegbaar reclamecontact vast te stellen, en hiermee de kans te bereiken dat een reclame-uiting kan bijdragen aan reclame-effect.

## 4. EYE-TRACKING: DE STAND VAN ZAKEN

In de vorige subparagraaf hebben we uiteengezet waarom eye-tracking de meest geschikte methode is om aandacht voor reclame te meten. Echter, grootschalige toepassing van eye-tracking was tot voor kort niet mogelijk. Eye-trackers waren vaak ingewikkelde apparaten die voor respondenten heel onnatuurlijk aandeden. Bovendien waren de voorbereidingen, metingen en evaluatie zo arbeidsintensief, dat hooguit bij een paar deelnemers per dag metingen konden worden gedaan.

In principe waren er voorheen twee soorten eye-trackers, te weten:

- Brillen en/of helmen die op het hoofd konden worden geplaatst. Afgezien van de problemen voor de respondent heeft dit grote gevolgen voor de nauwkeurigheid. De hoofdhuid is flexibel en elke hoofdbeweging betekent opnieuw kalibreren.
- Apparaten waaraan de respondent letterlijk gefixeerd wordt. Dit kan enerzijds door het hoofd in een soort klem te zetten en anderzijds door de respondent in een soort mondstuk te laten bijten.

In de jaren '90 veranderde deze situatie totaal. Door verschillende technologische en methodologische doorbraken kan reclamecontact middels eye-tracking nu betrouwbaar, grootschalig en relatief goedkoop worden uitgevoerd. In figuur 1 is een voorbeeld van de nieuwe meetsituatie weergegeven.

Een opsomming van de belangrijkste ontwikkelingen:

1. De nauwkeurigheid van de metingen is behoorlijk verbeterd. Door gebruik te maken van onzichtbare infrarood camera's die niet aan de respondent zijn bevestigd, is de respondent tijdens een test vrij in zijn bewegingen. Desondanks is het toch mogelijk om met een zeer grote nauwkeurigheid te volgen waar de respondent naar kijkt.
2. Men onderzocht in het verleden voornamelijk of mensen kort of lang naar iets kijken. Feitelijk betekent de duur van het kijken weinig. Het gaat erom wat mensen zien en of dat voldoende is om reclamecontact te realiseren. Hiervoor zijn indicatoren nodig die in benchmark kunnen worden weergegeven. Vergelijking met directe en indirecte concurrenten wordt hierdoor vergemakkelijkt.



*Figuur 1. Nieuwe meetsituatie eye-tracking Verify Nederland.*

3. Eye-tracking wordt vaak gebruikt als een kwalitatieve methode om het kijkpatroon te volgende gedurende een geforceerde blootstelling. Dit is niet logisch, omdat op deze manier geen rekening wordt gehouden met het ontwijken (bijv. doorbladeren en/of wegzappen) van reclame. Tegenwoordig kan het natuurlijke mediagedrag van consumenten worden gemeten.
4. Tegenwoordig kan eye-tracking worden uitgevoerd op bewegende televisiebeelden. Dit was voorheen niet mogelijk. Het probleem daarbij is namelijk dat voor iedere seconde van een commercial 25 frames worden gebruikt. Voor het afstemmen van “waar kijken mensen naar en op welk moment zien ze iets?”, is bijzonder veel computercapaciteit nodig. De stand van de techniek was enkele jaren geleden nog onvoldoende om dit uit te voeren.
5. Voor de betrouwbaarheid van de resultaten is een grote steekproef nodig. Vanaf minimaal 60 respondenten beginnen resultaten betrouwbaar worden. Resultaten gebaseerd op minder dan dit aantal moeten altijd met de nodige voorzichtigheid worden behandeld. Dankzij recente technologische doorbraken is het mogelijk om vele tientallen of zelfs honderden respondenten per dag te meten.
6. Veel onderzoek houdt geen rekening met de dagelijkse praktijk van communicatie, waarin korte doorlooptijden en snelle beslissingen centraal staan. Doordat de doorlooptijden voor eye-tracking tegenwoordig sterk verkort zijn is het mogelijk om elke communicatie-uiting (print, televisie) te (pre-)testen, voorafgaand aan en tijdens mediaplacings. Dit maakt het mogelijk om de informatie te gebruiken in het proces van communicatie en mediaplanning waarin vaak snel gehandeld moet worden en in principe altijd te weinig tijd is.

#### 4.1. Nieuwe inzichten

De wetenschap gebruikt eye-tracking al decennia lang om nieuwe inzichten in het communicatieproces te verkrijgen.

In onderzoek (zie Pieters et al, 1999) zien we dat bijvoorbeeld de verschillen tussen advertenties in hun effect veel groter zijn dan de verschillen tussen consumenten. Dat betekent niet dat individuele verschillen niet belangrijk zijn, maar wel dat de meeste winst in reclame-effect te behalen is door de prestaties van advertenties nauwkeurig vast te stellen, en dat een nadruk op het verbeteren van de advertenties zelf meteen winst oplevert.

In een recent artikel in *Journal of Marketing* (Pieters en Wedel 2004), dat gebruik maakt van gedetailleerde gegevens die door Verify International zijn verzameld, wordt nogmaals aangetoond dat eye-tracking nieuwe inzichten kan geven in reclamecontact die op geen enkele andere manier kunnen worden vastgesteld. En dat deze nieuwe inzichten radicaal ingaan tegen gevestigde ideeën over reclamewerking. Zo blijkt bijvoorbeeld dat de illustratie, tekst en het beeld in advertenties andere functies hebben in het reclameproces dan tot nu toe werd gedacht. Aandacht voor het merk speelt een centrale rol, de illustratie trekt veelal de meeste aandacht, maar doet dat onafhankelijk van de grootte (ook kleine illustraties trekken relatief veel aandacht), en de omvang van de tekst blijkt een grotere rol te spelen dan tot nu toe werd gedacht.

Ander onderzoek met eye-tracking toont het grote effect van originele reclame op het **trekken**, en vooral **houden** van aandacht bij herhaald reclamecontact aan (Pieters et al. 2002). Dit onderzoek toont onder meer aan dat originele reclames, in overeenstemming met wat reclamemakers steeds hebben gedacht, maar wat niet werd erkend, veelal juist meer aandacht vestigen op de merkvermelding. Op deze manier maakt eye-tracking het mogelijk om vast te stellen welke advertenties de grootste bijdrage leveren aan gerealiseerd reclamebereik.

Ook het proefschrift van Woltman Elpers (2003) verschaft inzicht in a) hoe verwerking van tv-reclamemomenten door consumenten invloed heeft op de waardering van de *gehele* commercial en b) hoe moment-tot-moment verwerking van tv-reclames het kijkgedrag van de consumenten *tijdens* de tv-reclame beïnvloedt. Enkele conclusies van het onderzoek van Woltman Elpers zijn:

- Het affectieve eendoordeel over de commercial wordt sterk beïnvloed door de *maximale* en *laatste* affectieve beoordeling en de *snelheid* in de affectieve reacties;
- Reclame met een hoog amusementsgehalte behoudt de aandacht van de consument langer dan reclame met een hoog informatiegehalte;
- Reclame met *tegelijktijd* een hoog amusementsgehalte en een hoog informatiegehalte verhoogt de kans op zapgedrag;
- Wanneer het merk (product, naam of logo) in beeld wordt gebracht, stijgt de kans dat consumenten geen aandacht meer aan de commercial besteden. Dit heeft verschillende oorzaken, die door oogbewegingregistratie kunnen worden verklaard.

## 5. CONCLUSIE

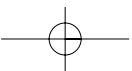
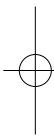
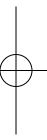
De rol van aandacht in reclame- en mediaplanning wordt steeds belangrijker. Het bepaalt namelijk of een uiting in staat is reclamecontact te realiseren. Dit kan weer gebruikt worden voor mediaplanning. Het gebrek aan een objectieve maat voor de creatieve kracht van een individuele uiting vormt een belangrijk probleem voor de moderne mediaplanning. Het weerhoudt mediaplanning ervan mediadruk in relatie te kunnen brengen tot het uiteindelijk communicatie-effect.

Tot voor kort werd reclamecontact gemeten door gebruik te maken van het geheugen. Alhoewel algemeen werd aangenomen dat dit geen valide methode is werd deze maat gebruikt bij gebrek aan beter. De nieuwe ontwikkelingen op het gebied van eye-tracking hebben hier verandering in gebracht.

Eye-tracking levert niet alleen relevante informatie op voor het reclamebureau (dat in staat wordt gesteld om door de ogen van de consument naar het eigen werk te kijken), maar ook voor mediaplanning. Deze informatie maakt het mogelijk om reclame actief te begeleiden en relevante maatregelen te treffen indien de situatie hiertoe aanleiding mocht geven.

## LITERATUUR

- Draisma, K. (2000), "Out of Sight, Out of Mind?" Doctoraalscriptie, departement Marketing, Universiteit van Tilburg.
- Finn, A. (1988), "Print Ad Recognition Readership Scores: An Information Processing Perspective," *Journal of Marketing Research*, 25, 168-177.
- Nixon, H. K. (1924), "Attention and Interest in Advertising," *Archives of Psychology*, 72, 5-67.
- Pieters, R.G.M., E. Rosbergen en M. Wedel (1999), "Visual Attention to Repeated Print Advertising: A Test of Scanpath Theory," *Journal of Marketing Research*, 36 (November), 424-438.
- Pieters, R.G.M., E. Rosbergen, M. Wedel and L. Warlop (2002), "Breaking Through the Clutter: Benefits of Advertisement Originality and Familiarity on Brand Attention and Memory," *Management Science*, 48 (6), 765-781.
- Pieters, R.G.M., E. Rosbergen, M. Wedel (2004), "Attention Capture and Transfer in Advertising: Brand, Pictorial, and text Size Effects," *Journal of Marketing*, April, forthcoming.
- Poffenberger, A.T. (1925), *Psychology in Advertising*. Chicago, Ill: A.W. Shaw Company.
- Rayner, K. (1998), "Eye movements in reading and information processing: Twenty years of research." *Psychological Bulletin*, 1998, 124, 372-422
- Rossiter & Percy (1998) "Advertising Communications & Promotion Management, 1998", The McGraw-Hill Companies Inc.
- Woltman Elpers J.L.C.M., M. Wedel, and R.G.M. Pieters 'Why do consumers stop viewing television commercials? Two experiments on the influence of moment-to-moment entertainment and information value'; *Journal of Marketing Research*, November 2003.
- Pieters R.G.M. (2001), 'Het Oog Wil Originaliteit', *Adformatie* 42, 200.
- Woltman Elpers, J.L.C.M. (2003) 'Consumers' Moment-to-Moment Processing of Television Commercials', Dissertation, Ridderkerk: Labyrinth Publications
- 'Bijna helft Nederlanders mijdt reclame', *Adformatie*, , 26 september 2002.
- 'Consument wil geen ongevraagde reclame', *Adformatie*, 18 september 2003



## 4. De effectiviteit van overheidscampagnes: de resultaten van 5 jaar systematisch onderzoek

M. GERRITSEN en W.D. VAN DER NOORT

### SAMENVATTING

Het campagne-effectonderzoek van de Rijksvoorlichtingsdienst is een geavanceerde onderzoeksmethode om de effectiviteit van overheidscampagnes te evalueren. Dit onderzoek wordt sinds 1999 standaard toegepast op alle Postbus 51-campagnes en levert een schat aan gegevens op, die uniek is in de markt van media- en communicatieonderzoek.

Het onderzoek geeft niet alleen inzicht in hoeverre de verschillende overheidscampagnes in staat zijn de beleids- en communicatiestellingen (mede) te verwezenlijken (*verantwoording*), maar draagt ook in belangrijke mate bij aan de verdere professionalisering en verbetering van de overheidscommunicatie (*leereffecten*). Dit laatste is vooral te danken aan de systematische opbouw van een database met inmiddels meer dan 120 campagnes, die zeer waardevol is voor *benchmarking* en voor meta-analyses naar de factoren die de effectiviteit van campagnes bepalen.

In samenwerking met PointLogic is met geavanceerde multivariate analyses tevens een Media-tool ontwikkeld, waarmee campagnemanagers ondersteund worden in het bepalen van de optimale mediamix en die hen helpt realistische doelstellingen in relatie tot het beschikbare mediabudget te formuleren.

Dit artikel beschrijft de methodiek en de voorgeschiedenis van het continu *trackingonderzoek*, zoals het nu wordt toegepast. Tevens geven we de belangrijkste bevindingen en inzichten die het onderzoek inmiddels heeft opgeleverd.

### 1. INLEIDING

#### 1.1. Introductie

Voor de meeste televisiekijkers is het inmiddels een bekend beeld: tussen het bekeken televisieprogramma en het reclameblok verschijnt het logo en de karakteristieke tune van een Postbus 51-publiekscampagne. Uitingen van Postbus 51-publiekscampagnes zijn niet alleen te zien op televisie, maar bijvoorbeeld ook op de radio, in dagbladen en tijdschriften, als buitenreclame of op internet. Voordat zo'n campagne-uiting voor het eerst het publiek bereikt, is al een uitgebreid onderzoekstraject ingezet: het campagne-effectonderzoek<sup>1</sup> van de Rijksvoorlichtingsdienst (RVD) / Publiek en Communicatie is dan al ruim 2 maanden in volle gang. In dit artikel worden het doel en de achtergronden van dit onderzoek besproken, de onderzoeksmethode komt aan de orde en we besteden aandacht aan de nieuwste inzichten die vijf jaar campagne-effectonderzoek heeft opgeleverd.



### *1.2. Vroege effectstudies van publiekscampagnes door de rijksoverheid*

Niet altijd had campagne-effectonderzoek zo'n prominente plaats in publiekscampagnes van de rijksoverheid. Begin jaren negentig oordeelde de Algemene Rekenkamer (1991) dat de effectiviteit van campagnes te weinig werd onderzocht en dat de doelmatigheid van de bestede overheidsgelden daardoor niet voldoende bewezen kon worden. Dit negatieve rapport was de aanleiding voor de zoektocht naar een meer gestructureerde methode voor het onderzoeken van effectiviteit van overheidscampagnes.

Deze zoektocht resulteerde in een nieuwe onderzoeksmethode voor effectonderzoek, de zogenaamde Zilveren Standaard (Veldkamp, 1993). De Zilveren Standaard combineert twee onderzoeksontwerpen: enerzijds wordt onderzocht hoe de publieke opinie zich in de looptijd van de campagne ontwikkelt (een systeem van voor- en nametingen, ook wel het pre/post design genoemd) en anderzijds of mensen wel of niet aan de campagne-uitingen zijn blootgesteld (exposed/non-exposed design). Een voordeel van het combineren van deze twee onderzoeksontwerpen is dat het effect van de campagne kan worden geïsoleerd van het effect dat bijvoorbeeld overige media-aandacht voor het campagneonderwerp teweeg heeft gebracht. Wanneer het effect tussen voor- en name-ting bij mensen die zijn blootgesteld aan campagne-uitingen wordt gecorrigeerd voor het effect tussen voor- en nameting bij mensen die niet zijn blootgesteld aan de campagne, wordt het 'echte' effect van de campagne zichtbaar (Bronner et al, 2003).

### *1.3. Nieuwe eisen aan effectonderzoek*

Vanaf 1993 werd de Zilveren Standaard verplicht gesteld voor alle massamediale overheidscampagnes. Na enige tijd werden echter serieuze nadelen van de methode ontdekt. Het benodigde paneldesign was ingewikkeld en er bleek sprake te zijn van selectieve uitval van respondenten in de nametingen, waardoor de aangetoonde effecten van de campagne niet generaliseerbaar waren naar het algemeen publiek. Ook bleek het, door gebrek aan standaardisatie en de verschillende wijze waarop de ministeries de richtlijnen toepasten, moeilijk de verschillende effectonderzoeken met elkaar te vergelijken. Hierdoor bleven campagneoverstijgende leereffecten uit.

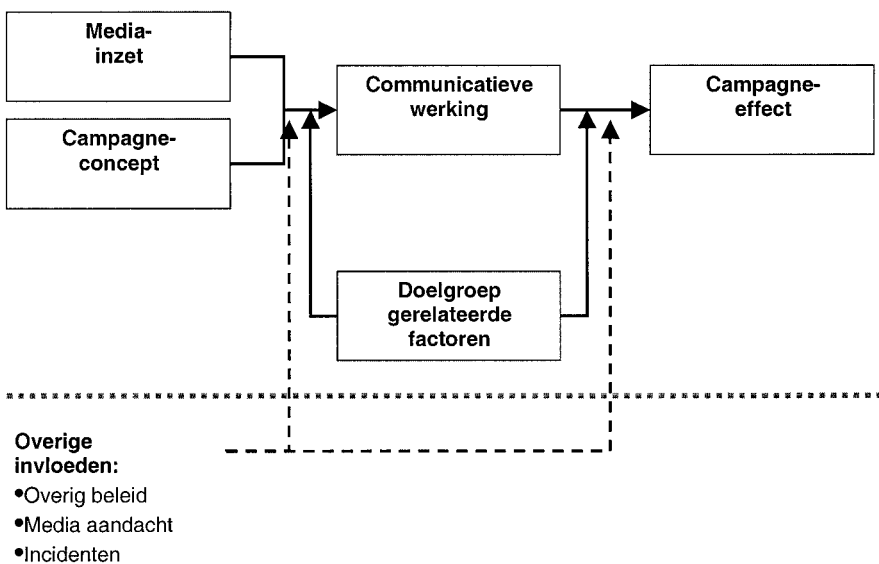
Daarnaast stegen de kosten voor voorlichtingscampagnes<sup>2</sup>. Hierdoor groeide onder campagnemanagers de behoefte aan beter inzicht in de doelmatigheid van de media-inzet en het campagneconcept. Dit zijn de twee belangrijkste factoren die de campagne-manager ter beschikking staan om uiteindelijk effect te bewerkstelligen en ze bepalen tevens een groot deel van de totale kosten van een campagne. Media-inzet verwijst naar alle media en middelen die worden ingeschakeld om het publiek met campagne-uitingen te confronteren. Ook gaat het hier om keuzes over welke media en middelen worden ingezet, gedurende welke periode en met welke frequentie. Met campagneconcept wordt bedoeld op de vertaling van beleidsdoelstellingen naar een communicatiestrategie en de creatieve uitwerking daarvan in spotjes en advertenties.

Verder groeide het besef dat de effectiviteit van campagnes niet alleen beïnvloed wordt door de kenmerken van de campagne zelf (media-inzet en campagneconcept), maar ook door kenmerken van de beoogde doelgroep. Hierbij valt te denken aan socio-demografische factoren (zoals geslacht, leeftijd en opleiding) en aan kenmerken van de doelgroep in relatie tot het campagneonderwerp (zoals interesse en betrokkenheid).

Ook is het belangrijk op te merken dat overheidsvoorlichting vrijwel nooit op zichzelf staat. Vaak maken voorlichtingscampagnes deel uit van een omvangrijker beleidstraject

en worden in combinatie met andere instrumenten (bijvoorbeeld wetgeving of subsidieregelingen) ingezet.

Bovenstaande overwegingen hebben geleid tot het opstellen van een communicatiemodel waarop het huidige campagne-effectonderzoek is gebaseerd (zie Figuur 1).



*Figuur 1. Communicatiemodel.*

Het model, zoals geschetst in Figuur 1, veronderstelt dus dat er twee belangrijke factoren zijn die het succes van een campagne uiteindelijk bepalen: het campagneconcept en de media-inzet. Media-inzet en campagneconcept zijn bepalend voor de communicatieve werking van de campagne, de mate waarin de boodschap van de campagne bij de burger overkomt en op de juiste manier begrepen en geapprecieerd wordt. Een goede communicatieve werking is geen garantie voor daadwerkelijke campagne-effecten maar is meestal toch een belangrijke voorwaarde voor succes.

#### *1.4. Het huidige campagne-effectonderzoek*

Bovenstaande overwegingen hebben geleid tot het opzetten van een gestandaardiseerd campagne-effectonderzoek naar de effectiviteit van publiekscampagnes door de rijksoverheid. Waar de Zilveren Standaard een meer experimenteel wetenschappelijke benadering van effectonderzoek was, staat bij het huidige campagne-effectonderzoek de gebruikswaarde voor verschillende partijen centraal (van Montfort & de Kleuver, 2001)<sup>3</sup>. Sinds 1999 is het campagne-effectonderzoek verplicht voor Postbus 51-campagnes en wordt, in opdracht van de verschillende ministeries, door RVD/Publiek en Communicatie uitgevoerd. Dit campagne-effectonderzoek biedt de volgende voordelen:

- Op een onderling vergelijkbare wijze kan verantwoording worden afgelegd aan de Tweede Kamer over de effectiviteit en doelmatigheid van Postbus 51-campagnes

- Door systematische kennisopbouw is het mogelijk om campagneoverstijgende analyses uit te voeren die inzicht geven in de relatie tussen media-inzet en campagneconcept enerzijds en campagne-effecten anderzijds
- Door systematische kennisopbouw is het mogelijk om indicatoren voor de effectiviteit van campagneresultaten te ontwikkelen (de zogenaamde benchmarks)
- Gezamenlijke uitvoering van campagne-effectonderzoek leidt tot kostenefficiëntie.

## 2. METHODE

### 2.1. *Het meetmodel: een systeem van voor-, tussen- en nametingen*

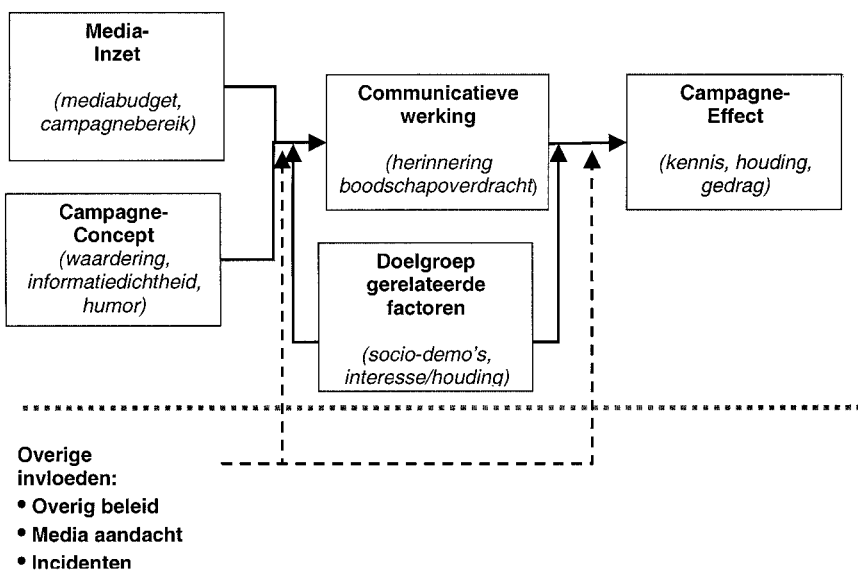
Het campagne-effectonderzoek voor Postbus 51-campagnes bestaat uit een voormeting, wekelijkse tussenmetingen tijdens de campagneperiode en een nameting. Tijdens voor-, tussen- en nametingen komen verschillende aspecten uit het hierboven beschreven communicatiemodel aan de orde;

- Het doel van voorlichtingscampagnes is het behalen van communicatie-effect op kennis, houding of gedrag(sintentie). Deze kennis-, houdings-, of gedrags(intentie) effecten worden geëvalueerd door de situatie voorafgaand aan de campagne (voormeting) te vergelijken met die na afloop van de campagne (nameting). Bovendien worden de belangrijkste effectvragen opgenomen in de tussenmeting om ook tijdens de campagne het verloop te kunnen volgen.
- De verschillende aspecten van communicatieve verwerking worden gemeten in tussenmetingen tijdens de looptijd van de campagne. Gemeten wordt in hoeverre de doelgroep de campagne heeft gezien (*herkenning van de uitingen*), onthouden (*herinnering*) en begrepen (*boodschapoverdracht*).
- Doelgroepgerelateerde kenmerken zoals socio-demografische factoren (geslacht, leeftijd, opleiding) en interesse en betrokkenheid bij het campagneonderwerp worden zowel in de voor- en nameting als in de tussenmeting opgenomen.
- Media-inzet wordt onderzocht aan de hand van factoren als campagne- en mediabereik en mediabudget.
- Het campagne-concept wordt onderzocht door te meten in hoeverre de campagne de juiste reacties heeft opgeroepen door de doelgroep de campagne te laten beoordelen op verschillende items (*waardering*, uitgesplitst in totale waardering en waardering op verschillende aspecten). Ook wordt de invloed van informatiedichtheid en het gebruik van humor van een campagne onderzocht.

In Figuur 2 wordt het eerder geschetste communicatiemodel nader ingevuld met de hierboven beschreven variabelen.

De voor- en nametingen dienen voor het onderzoeken van daadwerkelijk campagne-effect en zijn maatwerk, dus heel specifiek op een bepaalde campagne gericht. Campagne-effecten worden vastgesteld door het kennis-, houdings- en gedrag(sintentie) niveau van het algemeen publiek voorafgaand aan de campagne te vergelijken met die direct na afloop van de campagne.

De wekelijkse tussenmetingen dienen om de campagne te beoordelen op campagnebereik, herinnering, boodschapoverdracht en waardering en zijn (in tegenstelling tot de voor- en nametingen) zoveel mogelijk gestandaardiseerd om vergelijking tussen campagnes mogelijk te maken.



Figuur 2. Communicatiemodel met sleutelvariabelen.

## 2.2. Inrichting van het onderzoek

Het campagne-effectonderzoek voor Postbus 51-campagnes wordt uitgevoerd onder het algemeen publiek. Wekelijks worden minimaal 100 volwassenen van 18 jaar en ouder ondervraagd. Wanneer een campagne zich op specifieke doelgroepen richt (bijvoorbeeld jongeren in de leeftijd van 13 t/m 17 jaar) kunnen ook deze groepen worden ondervraagd.

Het onderzoek wordt hybride uitgevoerd, dat wil zeggen gebruik makend van meerdere dataverzamelingstechnieken. In de eerste plaats wordt gebruik gemaakt van een zogenaamd access panel van TNS NIPO (CAPI@home, een database van 70.000 personen)<sup>4</sup>. Vragenlijsten worden via internet of per modem verstuurd en het toonmateriaal (televisie- en radiospots, advertenties etc.) via CD ROM. Respondenten worden zo via hun eigen pc ondervraagd. Om ook diegenen zonder pc te kunnen ondervragen, worden daarnaast respondenten benaderd voor face-to-face onderzoek (CAPI). De verhouding CAPI@Home versus face-to-face bedraagt 75% tegen 25% (dit is conform de verhouding pc-/ niet pc-bezit 2002). Na weging van de resultaten van een specifieke campagne wordt ervoor zorggedragen dat deze representatief zijn voor de Nederlandse bevolking van 18 jaar en ouder.

Respondenten in zowel het CAPI@home als het face-to-face onderzoek krijgen (in beide methoden identieke) vragen over verschillende campagnes voorgelegd. Respondenten in beide methoden (dus ook in het face-to-face onderzoek!) vullen de vragenlijst geheel zelfstandig in. De enquêteur in het face-to-face onderzoek heeft voornamelijk een faciliterende rol. Naast campagnegerelateerde vragen wordt een aantal standaardvragen over mediaconsumptie, vertrouwen in de overheid, houding ten aanzien van Postbus 51 en achtergrondkenmerken voorgelegd.

In Tabel 1 wordt de opbouw van de vragenlijst zoals die verstuurd wordt aan respondenten schematisch weergegeven.

Tabel 1. Schematische weergave van de opbouw van een vragenlijst binnen het systeem van voor-, tussen- en nametingen.

Veldwerkweek	1	2	3	4	5	6	7	8	9	10	11	12	13	14
<b>Standaardvragen</b>	S	S	S	S	S	S	S	S	S	S	S	S	S	S
<b>Looptijd campagne A</b>	*	*	*	*	*	*	*	*						
Meting	T	T	T	T	T	T	T	T	N	N	N	N		
<b>Looptijd campagne B</b>					*	*	*	*	*	*	*	*		
Meting	V	V	V	V	T	T	T	T	T	T	T	T	N	N
<b>Looptijd campagne C</b>												*	*	*
Meting								V	V	V	V	T	T	T

S = standaardvragen

\* = looptijd van de campagne

V = voormeting

T = tussenmeting

N = nameting

In Tabel 1 is te zien is dat per campagne een voor- en nameting plaatsvindt en (gedurende de campagne) wekelijks tussenmetingen uitgevoerd worden. De voormeting vindt in principe in de 4 weken voorafgaand aan de campagne plaats en de nameting in de 4 weken direct na de campagne. Een respondent in veldwerkweek 9 zal volgens bovenstaand schema de standaardvragen, de nameting van campagne A, de tussenmeting van campagne B en de voormeting van campagne C, voorgelegd krijgen. Wanneer het aantal campagnes zo groot is dat het gesprek langer dan een uur zal gaan duren, wordt een extra steekproef getrokken. Dit leidt ertoe dat in de praktijk wekelijks gemiddeld 200 respondenten van 18 jaar en ouder worden ondervraagd.

In 1999 is het campagne-effectonderzoek verplicht gesteld voor Postbus 51-campagnes. Zo is in de loop der tijd een waardevolle database opgebouwd met resultaten van onder andere vragen naar bereik, herinnering, waardering en effect van campagnes. Ook zijn met behulp van deze database indicatoren voor de effectiviteit van campagnes ontwikkeld waarmee het mogelijk wordt campagnes onderling op aspecten als campagnebereik, waardering en houding te vergelijken: de zogenaamde benchmarks.

### 2.3. Nieuwe methodiek

Niet altijd heeft de dataverzameling van het campagne-effectonderzoek op boven beschreven wijze plaatsgevonden. In 2003 is de onderzoeksmethode ingrijpend gewijzigd. Waar nu gebruik wordt gemaakt van een hybride methode (met zowel een access panel als face-to-face onderzoek) werd voor 2003 alleen face-to-face onderzoek verricht. De reden voor deze wijziging is dat de bereidheid van mensen om mee te werken aan enquêtes de afgelopen jaren steeds verder is gedaald, waardoor de representativiteit van onderzoek onder druk is komen te staan.

Bij de overgang op de nieuwe methode is uitgebreid onderzoek verricht teneinde de verwachte trendbreuk in de database zo goed mogelijk in kaart te brengen en te cor-

rigeren (Veldkamp/ TNS NIPO, 2003). Gedurende 7 weken zijn in totaal 11 campagnes zowel via de oude methode als de nieuwe methode onderzocht.

De verschillen in uitkomsten tussen beide methoden bleken niet onaanzienlijk. Bij meer dan de helft van de uitkomsten was sprake van een significant verschil, terwijl beide methoden een representatief beeld pretenderen te geven.

Deze verschillen hebben slechts in beperkte mate te maken met de verandering in vraagmethodiek (self-completion zonder de aanwezigheid van een enquêteur). Immers, de rol van de enquêteur bij het face-to-face onderzoek was al beperkt: respondenten vulden de vragenlijst zelfstandig op de laptop in. De geconstateerde verschillen hangen vooral samen met systematische verschillen in de samenstelling van de steekproef, die ondanks weging resteren.

In de oude methodiek werden respondenten eerst telefonisch benaderd voor deelname aan het onderzoek. Dit leidt tot selectieve respons, zoals uit een testonderzoek onder weigeraars is komen vast te staan. Mensen met een positieve houding tegenover de overheid en Postbus 51 blijken meer geneigd om mee te doen met onderzoek over overheidscampagnes en waren daardoor in zekere mate oververtegenwoordigd in de steekproef. Deze mensen kijken ook vaker naar de publieke omroep, terwijl de ondervraagden in de nieuwe methodiek wat sterker gericht zijn op de commerciële zenders. Dit leidt er toe dat vooral bij de campagnes over maatschappelijke thema's zoals het klimaat, biologische producten en verkiezingen, de uitkomsten nu wat minder rooskleurig zijn. De betrokkenheid bij en interesse in dergelijke onderwerpen blijkt in de nieuwe methodiek minder hoog te zijn. Bij meer praktisch gerichte thema's, zoals Landelijk Telefoonnummer Politie, Belastingdienst en de BOB-campagne, zien we minder verschillen.

Op basis van deze analyse van verschillen tussen beide onderzoeksmethoden zijn ook de benchmarks in de database voor campagnebereik, waardering en houding aangepast. Hierdoor blijft, ook in de nieuwe methodiek, een realistische vergelijking mogelijk met andere overheidscampagnes.

### 3. RESULTATEN

#### *3.1. Hoe effectief zijn Postbus 51-publiekscampagnes?*

In hoeverre hebben Postbus 51-publiekscampagnes nu aantoonbare effecten op kennis, houding of gedrags(intentie) van het algemeen publiek? Deze vraag komt aan de orde in de Evaluatie Postbus 51-campagnes, een rapport dat RVD/Publiek en Communicatie jaarlijks uitbrengt voor de Tweede Kamer waarin de resultaten van het campagne-effectonderzoek naar alle Postbus 51-publiekscampagnes van dat jaar beschreven en geanalyseerd worden (RVD/Publiek en Communicatie 2003, RVD/Publiek en Communicatie, 2004).

#### *Kennis*

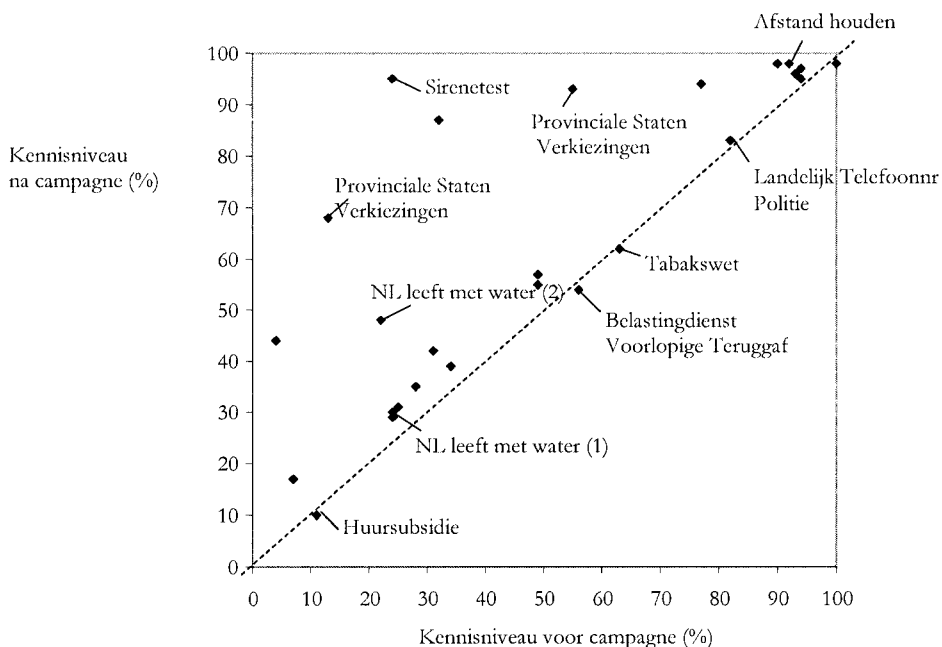
Vrijwel alle campagnes richten zich op het behalen van kenniseffecten en uit verschillende jaarevaluaties blijkt dat publiekscampagnes hier ook succesvol in zijn. Het gemiddelde kennisniveau van campagneonderwerpen in de periode 1999 tot en met

2002 steeg met 11%, van 49% bekendheid voorafgaand aan de campagne naar 60% na afloop. In 2003 waren de behaalde kenniseffecten nog groter: het gemiddeld kennisniveau nam toe van 49% naar 64% (RVD/Publiek en Communicatie, 2004).

Het gemiddeld kenniseffect zegt natuurlijk weinig over de effecten van individuele campagnes. Daarom is in Figuur 3 weergegeven in hoeverre de individuele campagnes uit 2003 kenniseffecten hebben weten te behalen. In deze figuur is het kennisniveau na afloop van de campagne afgezet tegen het kennisniveau voorafgaand aan de campagne. Een campagne die op de diagonaal scoort heeft na afloop van de campagne hetzelfde kennisniveau als voorafgaand aan de campagne: deze campagnes hebben dus geen kenniseffect bereikt.

In Figuur 3 is te zien dat de mate waarin kenniseffecten behaald worden, flink verschilt tussen campagnes. Zo bleek in 2003 de campagne ‘Sirenetest’ het meest succesvol. Deze campagne had als doel het algemeen publiek te laten weten dat het landelijke waarschuwings- en alarmeringsstelsel (opnieuw) maandelijks getest zou worden, en wel op de eerste maandag van de maand. In de voormeting wist 24% van het algemeen publiek al dat de sirenes voortaan maandelijks op de eerste maandag van de maand getest worden. In de nameting is dit opgelopen tot maar liefst 95%.

De campagne ‘Nederland leeft met water’ heeft twee campagneperioden gekend over twee verschillende deelthema’s (NL leeft met water 1 en 2). Het betreft hier een meerjaren campagne met Peter Timofeeff als ‘boegbeeld’ over het nieuwe waterbeleid van de overheid om ook in de toekomst de veiligheid te waarborgen en wateroverlast te voorkomen (*‘meer ruimte voor water’*).



Figuur 3. Kenniseffecten van 20 campagnes uit 2003.

In Figuur 3 is te zien dat in de eerste periode van de campagne (NL leeft met water (1)), een bescheiden kenniseffect is behaald, terwijl in de tweede periode (NL leeft met water (2)) een aanzienlijk sterker effect is behaald. In beide campagnes was dezelfde kennisdoelstelling geformuleerd. Hoe kan dit verschil verklaard worden? Uit het campagne-effectonderzoek van 'Nederland leeft met water 1' bleek de boodschap dat de overheid met een nieuwe aanpak voor waterbeheer zou komen, niet erg goed over te komen. Het kennisniveau op dit punt steeg licht van 24% naar 29%. Bij 'Nederland leeft met water 2' is, mede op basis van deze onderzoeksresultaten, een aangescherpt campagne-concept gemaakt. De kennis over de nieuwe aanpak van waterbeheer door de overheid was bij de start van de tweede periode, weer teruggezakt naar het niveau van voor 'Nederland leeft met water (1)'. Maar, door het aangescherpte campagneconcept heeft 'Nederland leeft met water 2' wel een duidelijk kenniseffect behaald: het kennisniveau stijgt van 22% naar 48%.

Maar niet altijd zien we kenniseffecten optreden. De campagne 'Tabakswet' was zeer succesvol in het communiceren van de nieuwe wetgeving ('recht op een rookvrije rookplek'). Daarnaast communiceert de campagne duidelijk, zowel in tekst als in beeld, dat meeroken zeer schadelijk voor de gezondheid is (de belangrijkste reden voor de wetgeving). Dit wordt door het algemeen publiek ook goed als boodschap herkend. Toch wordt er op dit aspect geen effect behaald: de kennis over schadelijkheid van meeroken is bij velen al aanwezig en de mate van gepercipieerde schadelijkheid wordt niet beïnvloed door de campagne.

#### *Houding*

Campagnes richten zich, naast het behalen van kenniseffecten, ook vrijwel zonder uitzondering op het behalen van houdingseffecten. Hierin zijn campagnes over het algemeen echter minder succesvol. In de afgelopen jaren werden veelal geen of slechts bescheiden houdingseffecten geconstateerd. Het beïnvloeden van de houding van mensen blijkt moeilijk te realiseren voor een relatief kortdurende interventie als een publiekscampagne. Steun voor deze veronderstelling wordt gevonden in het campagne-effectonderzoek naar de campagne voor de invoering van de euro. Deze campagne liep van het najaar 2000 tot en met december 2001 en had met ruim 12 maanden een beduidend langere looptijd dan de gebruikelijke 6 tot 8 weken. In de looptijd van de campagne was er een duidelijk stijgende trend in de ontwikkeling van houdingseffecten zichtbaar. Het behalen van houdingseffecten lijkt dus alleen weggelegd voor langlopende of met enige regelmaat terugkerende campagnes.

#### *Gedrag*

Campagnes richten zich beduidend minder op het behalen van gedrags(intentie)effecten. Gedrags(intentie)effecten lijken ook moeilijk te realiseren, maar in veel gevallen wordt toch een significante toename in gedrag(sintentie) gesignaleerd.

Net als bij de kenniseffecten is in het jaar 2003 een groter gedrags(intentie)effect te zien dan in eerdere jaren. In voorgaande jaren steeg het gemiddelde gedrags(intentie) niveau met gemiddeld 4 %, terwijl dit in 2003 met 9% steeg.

In 2003 zijn voor 6 van de 21 campagnes gedrags(intentie)doelstellingen geformuleerd. In 80% van de gevallen ging het om een gedragsdoelstelling voor een verkeerscampagne. Verkeerscampagnes worden vaak ingezet rond wet- en regelgeving en worden



ondersteund door (extra) handhaving van de verkeersregels. Zo werden tijdens de BOB-campagne extra alcoholcontrole's uitgevoerd en werd tijdens de campagne 'Fiets Verlicht' vaker gecontroleerd op het voeren van fietsverlichting. Het lijkt dan ook aannemelijk te veronderstellen dat campagnes die zich richten op het behalen van gedrags(intentie)effecten gebaat zijn bij de inzet van overige beleidsinstrumenten.

In de jaarevaluatie over het jaar 2002 (RVD/Publiek en Communicatie, 2003) is onderzocht in hoeverre gedragseffecten samenhangen met kenniseffecten. Hiervoor is een regressie-analyse uitgevoerd met gedragseffect als afhankelijke en kenniseffect als onafhankelijke variabele. Dit leverde een significant regressiemodel op met  $F(1, 22) = 8.055$ ,  $p = .01$  en  $R = .52$ ; 27% van de variantie in gedragseffect kan verklaard worden uit kenniseffect. Campagnes die erin zijn geslaagd kenniseffecten te realiseren, slagen er dus ook vaak in effecten op gedrags(intentie) te behalen.

Samengevat: Postbus 51-campagnes zijn succesvol in het realiseren van kenniseffecten bij het algemeen publiek van 18 jaar en ouder. Deze kenniseffecten blijken een sleutelrol te spelen bij het eventueel behalen van gedrags(intentie)effecten. Een belangrijke vraag die hieruit voortkomt is dan ook: welke factoren zijn van invloed op het behalen van (kennis)effecten? Deze vraag wordt behandeld in de volgende paragraaf.

### *3.2. Communicatieve werking en effect: de rol van herinnering en boodschapoverdracht*

In verschillende studies is de invloed van de communicatieve werking van een campagne op behaalde effecten onderzocht (zie bijvoorbeeld RVD/Publiek en Communicatie, 2002 en 2003).

Zoals eerder vermeld wordt de mate van communicatieve werking in het campagne-effectonderzoek gemeten aan de hand van 2 aspecten: herinnering en boodschapoverdracht (zie Figuur 2).

#### *Herinnering*

Bij herinnering wordt onderscheid gemaakt naar spontane en geholpen herinnering. Om de spontane herinnering te meten wordt de volgende vraag gesteld: 'Kunt u zich herinneren de afgelopen weken campagnes van Postbus 51 te hebben gezien of gehoord?'. En, als de respondent bevestigend antwoordt: 'Welke Postbus 51-campagnes van de afgelopen weken kunt u zich herinneren?'<sup>5</sup> Geholpen herinnering wordt als volgt gevraagd: 'herinnert u zich de afgelopen weken de campagne over [campagneonderwerp] te hebben gezien of gehoord?' Geholpen herinnering wordt uitgedrukt in een percentage dat verwijst naar het aantal respondenten dat deze vraag met 'ja' beantwoordt. Tijdens de campagneperiode wordt geholpen herinnering wekelijks gemeten, het percentage heeft betrekking op de week waarin de geholpen herinnering het hoogste punt bereikte.

Er zijn grote verschillen tussen campagnes wat betreft de gerealiseerde geholpen herinnering. In de periode 1999 tot en met 2002 bedroeg de gemiddelde geholpen herinnering 46%, terwijl de scores van de individuele campagnes uiteenliepen van 11% tot 96%. In 2003 bleek de gemiddelde geholpen herinnering te zijn gestegen naar 51%. Maar ook campagnes uit dit jaar vertonen grote onderlinge verschillen: de scores lopen uiteen van 18% (Nationale Ombudsman) tot 87% (Sirenetest).

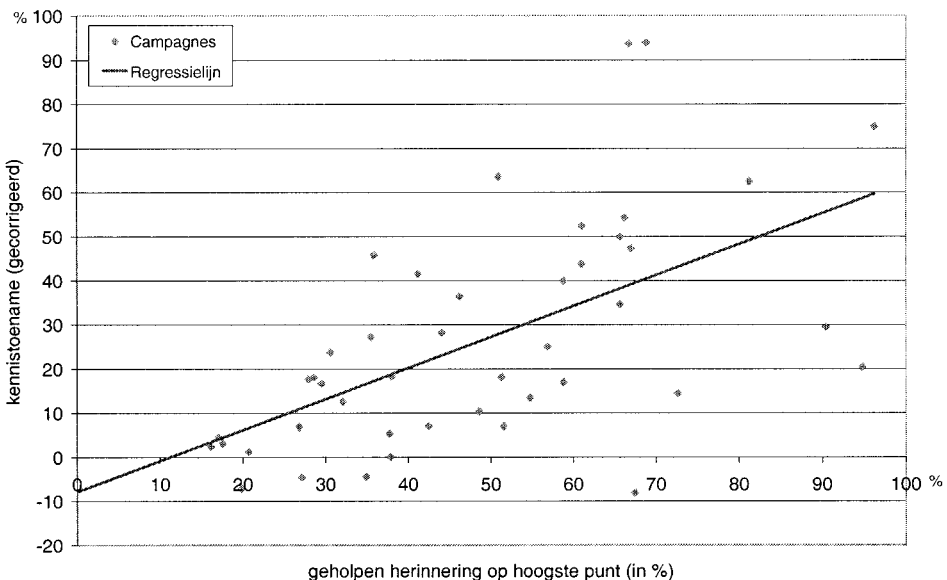
De mate waarin geholpen herinnering wordt gerealiseerd verschilt dus aanzienlijk tussen campagnes. Maar is geholpen herinnering ook echt bepalend voor het effect dat campagnes behalen? Om deze vraag te beantwoorden wordt in Figuur 4 de geholpen herinnering afgezet tegen het gerealiseerde kenniseffect.

Geholpen herinnering blijkt inderdaad een belangrijke voorspeller van (kennis)effect ( $r = .60$ ). Hoe hoger de geholpen herinnering, hoe sterker het effect van de campagne. Dus: om kenniseffecten te realiseren is het essentieel te zorgen voor een goede herinnering.

*Boodschapoverdracht*

Sinds begin 2003 wordt boodschapoverdracht gemeten, waarbij onderscheid wordt gemaakt naar spontane en geholpen boodschapoverdracht. Spontane boodschapoverdracht wordt onderzocht door aan diegenen die zich de campagne kunnen herinneren de volgende vraag te stellen: ‘Wat denkt u dat de belangrijkste boodschap van de campagne is, wat probeert de campagne u te vertellen?’<sup>6</sup>. Geholpen boodschapoverdracht wordt gemeten door aan diegenen die zich de campagne kunnen herinneren, de belangrijkste (deel)boodschap(en) voor te leggen en te vragen: ‘Kunt u aangeven in hoeverre u het eens of oneens bent met de volgende stelling(en)? Ik vind dat de campagne laat zien/horen dat ...[boodschap]’. Respondenten beantwoorden deze vraag op een 7-puntsschaal, waarbij 1 = helemaal niet mee eens en 7 = helemaal mee eens. Geholpen boodschapoverdracht wordt uitgedrukt in een percentage dat verwijst naar het aantal respondenten dat deze vraag met een score 6 of 7 heeft beantwoord.

Net als geholpen herinnering verschilt de mate van geholpen boodschapoverdracht sterk per campagne. De gemiddelde geholpen boodschapoverdracht in 2003 bedraagt



Figuur 4. De samenhang tussen geholpen herinnering en (kennis)effect op basis van 46 campagnes van 2001 en 2002.

59%, terwijl de scores van de individuele campagnes uiteen liepen van 34% (Klimaat) tot 84% (BOB, Alcohol in het Verkeer).

Uit nadere analyses naar boodschapoverdracht en effect bleek dat boodschapoverdracht een belangrijke voorspeller is voor effect ( $r = .52$ ): hoe beter de boodschapoverdracht, hoe sterker het effect van de campagne. Geholpen boodschapoverdracht blijkt dus in gelijke mate als geholpen herinnering bepalend voor het behalen van effect.

### *Samengevat*

Om met een publiekscampagne de gewenste effecten te behalen is het essentieel dat mensen zich de campagne goed kunnen herinneren en dat de boodschap goed overkomt. Maar welke factoren dragen nu bij aan een goede herinnering en boodschapoverdracht? In de volgende paragrafen wordt hier nader op ingegaan.

### *3.3. Media-inzet: de rol van mediabudget en campagnebereik*

Een basisvoorwaarde voor het realiseren van herinnering en een goede boodschapoverdracht (en in het verlengde daarvan campagne-effect) is dat het publiek door de campagne wordt bereikt. Mediabudget hangt hier nauw mee samen. Hoe meer budget beschikbaar is voor een campagne, hoe vaker campagne-uitingen te zien of te horen zullen zijn en hoe groter de kans dat de doelgroep door de campagne bereikt zal worden.

Het gemiddelde mediabudget per campagne in 2003 bedroeg gemiddeld € 475 290 en is ten opzichte van voorgaande jaren fors gedaald. In 2003 lag het gemiddelde mediabudget zo'n € 200 000 lager dan het jaar daarvoor, een daling van ruim 30%.

Analyses hebben aangetoond dat naarmate het media-budget stijgt, ook de geholpen herinnering toeneemt (RVD/Publiek en Communicatie, 2003). Opvallend is echter dat, hoewel de mediabestedingen in 2003 duidelijk zijn gedaald, de geholpen herinnering juist is toegenomen (van 49% in 2002 naar 51% in 2003). Mediabudget blijkt echter geen invloed te hebben op geholpen boodschapoverdracht. Dit lijkt logisch, aangezien verwacht kan worden dat de mate waarin campagne-uitingen worden ingezet geen invloed heeft op de mate waarin de campagne de boodschap goed en duidelijk overbrengt (RVD/Publiek en Communicatie, 2004).

Om campagnebereik te meten, worden alle uitingen van de campagne aan de respondent getoond. Al het toon- en geluidsmateriaal wordt op de computer afgespeeld. Zo kan bijvoorbeeld een tv-spot met beeld en geluid worden afgespeeld, waardoor de getoonde spot zoveel mogelijk overeenkomt met de op televisie uitgezonden spot. Per uiting wordt vervolgens gevraagd of men deze wel eens heeft gezien of gehoord. Respondenten die in een bepaalde week tenminste één van de voorgelegde uitingen hebben herkend, worden gerekend tot degenen die door de campagne zijn bereikt.

Ook het campagnebereik in 2003 is hoger dan in voorgaande jaren. In 2003 werd een gemiddeld bereik van 92% gerealiseerd, terwijl dit in de periode van 1999 tot en met 2002 met gemiddeld 85% beduidend lager lag. Ook hier geldt dat het hogere bereik niet veroorzaakt wordt door een hoger mediabudget. Het gemiddelde mediabudget is, zoals hierboven al werd besproken, in 2003 juist flink gedaald (RVD/Publiek en Communicatie, 2003; RVD/Publiek en Communicatie, 2004).

Uit voorgaande is gebleken dat hoewel de mediabestedingen in 2003 zijn teruggelopen, dit niet per definitie een negatieve invloed heeft op de communicatieve werking van campagnes. Ook met lagere mediabudgetten weten campagnes in 2003 een hogere herinnering te bewerkstelligen. Een mogelijke verklaring hiervoor is dat naast de media-inzet nog een andere factor bepalend is voor de mate van communicatieve werking van campagnes: het campagneconcept. In de volgende paragraaf wordt besproken hoe het campagneconcept van invloed is op de herinnering en boodschapoverdracht (en in het verlengde daarvan op effect).

### *3.4. Campagneconcept: de rol van waardering, informatiedichtheid en humor*

Zoals eerder beschreven wordt het campagneconcept onderzocht door de waardering van het publiek te meten. Tevens worden een aantal objectieve kenmerken van de campagne geregistreerd, zoals de informatiedichtheid en het gebruik van humor.

#### *Waardering*

De waardering voor een campagne geeft weer hoe het algemeen publiek het campagneconcept beoordeelt. Waardering is echter een complex begrip. In het campagne-effectonderzoek wordt waardering gemeten aan de hand van verschillende dimensies, te weten,

1. vormgeving/creatie (op deze dimensie scoren items als ‘grappig’, ‘opvallend’ en ‘mooi’)
2. inhoud (op deze dimensie scoren items als ‘duidelijk’, ‘informatief’ en ‘geloofwaardig’)
3. persoonlijke relevantie (bijvoorbeeld een item als ‘spreekt me aan’)
4. ‘tone of voice’ (hier gaat het om een item als ‘niet irritant’)

Respondenten krijgen de volgende vraag voorgelegd: ‘De [uitingen verschillende mediatypes<sup>7</sup>] behoren tot de campagne over [campagneonderwerp]. Wat vindt u van de campagne als geheel?’.

Respondenten waarderen de campagne aan de hand van 8 items behorend bij de genoemde dimensies op een 7-puntsschaal, waarbij 1 het meest negatief en 7 het meest positief is. Daarna wordt respondenten ook gevraagd een totaaloordeel over de campagne in zijn geheel te geven: ‘Welk rapportcijfer, van 1 tot en met 10, zou u willen geven aan de campagne over [campagneonderwerp]?’.

Het algemeen publiek blijkt de campagnes van 2003 goed te waarderen en geeft een ruime voldoende (een gemiddeld rapportcijfer van 7,0). Dit rapportcijfer ligt iets hoger dan het rapportcijfer van voorgaande jaren (het gemiddelde rapportcijfer over 98 campagnes van 1999 – 2002 was 6,8). Mensen beoordelen campagnes met name goed op de inhoudsdimensie: men vindt campagnes voornamelijk duidelijk (5,4), geloofwaardig (5,2) en informatief (5,2). Ook de opvallendheid (4,7) en aantrekkelijkheid (4,7) van campagneconcepten wordt in het algemeen positief beoordeeld.

De 8 afzonderlijke waarderingsitems blijken sterk samen te hangen met het algemene oordeel (het rapportcijfer) over de campagne: 98% van de verschillen in algemene waardering van campagnes kan verklaard worden uit de 8 waarderingsitems. Daarnaast is onderzocht waar een campagne aan moet voldoen om een hoge waardering te halen. Uit deze analyses bleek dat een campagne vooral duidelijk en opvallend moet zijn voor het behalen van een hoog rapportcijfer. Ook aspecten als ‘niet irritant’ en ‘spreekt me

aan' bleken hiervoor –hetzij in iets mindere mate- belangrijk te zijn. De relatie tussen het item grappig en algemene waardering bleek ingewikkelder te liggen. In een aparte studie is een curvilineair verband tussen waardering en de score op dit item aangetoond (zie Nelemans, 2003).

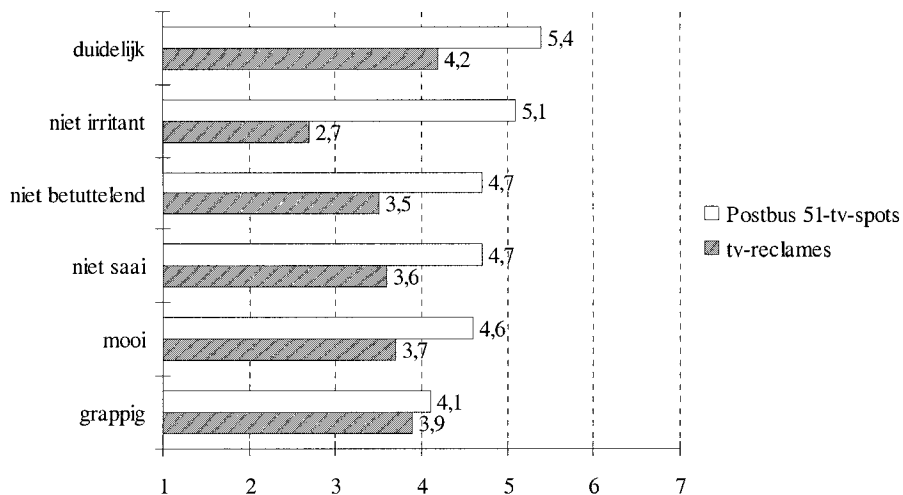
In onderzoek is aangetoond dat er een significante relatie bestaat tussen waardering (het rapportcijfer) en geholpen herinnering, waarbij  $r = .52$  en  $p < .01$ . Hoe hoger het rapportcijfer dat mensen geven, hoe beter ze zich de campagne kunnen herinneren (RVD/Publiek en Communicatie, 2003).

Ook blijkt de waardering voor een campagne bijzonder belangrijk voor de geholpen boodschapoverdracht ( $r = .81$ ,  $p < .01$ ). Naarmate de campagne een hoger rapportcijfer krijgt, neemt de geholpen boodschapoverdracht toe (RVD/Publiek en Communicatie, 2004).

Tenslotte: mensen waarderen Postbus 51-campagnes over het algemeen goed en beter dan commerciële reclames (zie figuur 5). Waardering blijkt een belangrijke voorspeller van zowel geholpen herinnering als geholpen boodschapoverdracht. Dus: waardering van een campagne blijkt bepalend te zijn voor hoe goed mensen zich die campagne kunnen herinneren en voor de mate waarin de boodschap van de campagne goed overkomt (RVD/Publiek en Communicatie, 2003).

*Informatiedichtheid en humor*

Campagnemanagers willen over het algemeen veel informatie kwijt aan het publiek. Maar voor massamediale campagnes komt het er juist op aan om de kernboodschap zo helder en beknopt mogelijk te houden. Er is een zeer sterke negatieve correlatie ( $r = -.71$ ,  $p < .01$ ) tussen de hoeveelheid informatie die de televisie-uiting bevat (*informatiedichtheid*) en de mate waarin de boodschapoverdracht door het publiek als goed wordt ervaren.



Figuur 5. Beoordeling van Postbus 51-spots versus tv-reclames in het algemeen.

Humor wordt steeds meer toegepast in overheidscampagnes, maar blijkt (zoals we zagen) het minst belangrijk voor een goede herinnering en waardering van de campagne. Juist bij het gebruik van humor is de zorg voor een lage informatiedichtheid extra belangrijk, omdat humor ‘cognitieve belasting’ is, waardoor het risico bestaat dat de humor afleidt van de boodschap. Uit nader onderzoek is gebleken (Nelemans, 2003), dat een goede integratie van humor met de boodschap belangrijk is. In dat geval komt de verwerkingscapaciteit, die wordt ingezet om de humor te begrijpen ook ten goede aan de boodschap. Als je de grap begrepen hebt, dan heb je in feite ook de boodschap begrepen.

### *3.5. Doelgroep gerelateerde factoren en effect: de rol van interesse en houding*

Vanuit de commerciële reclame weten we dat betrokkenheid van mensen bij het reclame- of campagneonderwerp en de persoonlijke relevantie van dat onderwerp bijzonder belangrijk is. Voor overheidscampagnes geldt hetzelfde: betrokkenheid en persoonlijke relevantie zijn belangrijk. Naast mediabudget en een goede waardering bevordert betrokkenheid bij het campagneonderwerp in belangrijke mate de herinnering en boodschapoverdracht. Daarnaast geldt voor overheidscampagnes dat ze ook maatschappelijk relevant moeten zijn. Dit komt tot uitdrukking in ‘goede zaak dat de overheid zich ermee bezighoudt’. Van alle houdingsitems ten aanzien van het campagneonderwerp is dit aspect het belangrijkste voor een goede werking van de campagne.

## 4. TOEPASSING VAN RESULTATEN

Het implementeren van de inzichten en leerervaringen in de praktijk is een vak apart; onderzoekers en potentiële gebruikers hebben vaak andere invalshoeken. Ondanks het feit dat communicatie-professionals onderschrijven dat reacties van doelgroepen niet altijd voorspelbaar zijn, wordt er in het algemeen weinig tijd en geld geïnvesteerd in bereiks- en effectonderzoek van communicatie. De vraag om accountability van communicatie-inspanningen klinkt wel, maar kennelijk niet luid genoeg. In onze ogen vormt de overheid een positieve uitzondering in het uitvoeren van onderzoek, maar ook daar vinden gebruikers het vaak duur, tijdrovend en moeilijk toepasbaar in de eigen praktijk.

Om de campagnemanagers zo goed mogelijk in staat te stellen de onderzoeksresultaten te gebruiken, worden de resultaten in verschillende producten en diensten toegepast:

- Per campagne ontvangt de campagnemanager schriftelijke rapportages van de voormeting, tussenmeting en nameting. De bevindingen en aanbevelingen uit het eindrapport worden gepresenteerd aan de campagnemanagers en vaak ook aan het betreffende reclame- en mediabureau.
- Vanaf de eerste campagneweek is de opbouw van bereik en effecten, online te volgen via de site van RVD/TNS NIPO, rechtstreeks toegankelijk voor de campagnemanager. Per week worden de grafieken aangevuld met de meest recente cijfers.
- Jaarlijkse bundeling en samenvatting van alle campagne-evaluaties met campagneoverstijgende analyses op de database en benchmarks, gepubliceerd als de Evaluatie Postbus 51-campagnes.
- Beschikbaarheid van een multimedia voorspellingstool, op basis van gemodelleerde data uit de database.

- Leerervaringen vormen input voor kennisbijeenkomsten voor communicatieadviseurs en cursussen vanuit de Academie voor Overheidscommunicatie.

Online rapportage maakt het mogelijk om gedurende de looptijd van een campagne de inzet bij te sturen. Als blijkt dat na twee of drie weken campagne voeren de opbouw van het campagnebereik (herkenning) achterblijft, volgt bijvoorbeeld het advies om de mediadruk op radio of televisie te verhogen of dagbladen in te zetten. Ook kan een uiting die onvoldoende impact blijkt te hebben, ingeruild worden voor een beter scorende variant.

Een daadwerkelijke ‘gefundeerde blik vooruit’ in de zin van voorspellingen doen voorafgaand aan een mogelijke campagne, vergt nog een extra stap. Hiervoor moeten de onderzoeksdata en leereffecten op een ‘bovencampagne’ niveau eenvoudig toegankelijk zijn.

Hiertoe is in 2002 in samenwerking met Pointlogic een Mediatool ontwikkeld op basis van data van ruim 100 campagnes uit de database. Uit de data zijn modellen en responscurves geschat, die zijn aangevuld met vakkennis en ervaring uit workshops met communicatie- en mediastrategen. De tool wordt regelmatig voorzien van een update met nieuwe cijfers en nieuwe functionaliteiten.

Deze multimediatool geeft een campagnemanager antwoord op vragen als:

- wat is het benodigde budget om mijn doelstellingen te realiseren?
- wat mag ik aan communicatie-effecten verwachten bij deze campagne?
- wat is een optimale verdeling van budget over de verschillende mediumtypen en hoe spreid ik de inzet in de tijd?

Input voor de tool zijn de uitgangspunten van de campagne, zoals ze in een communicatieplan omschreven worden; zaken als bekendheid van de doelgroep met het onderwerp, het budget en de campagnedoelstellingen. Bijvoorbeeld kennis van een nieuwe wet, het bevorderen van een positieve houding of gedragsbeïnvloeding (bijvoorbeeld bewust omgaan met energie of veilig verkeersgedrag). Door middel van een eenvoudig invoerscherm met schuifjes vult de campagnemanager zo de specifieke kenmerken van de campagne in.

De output is gebaseerd op deze campagnekenmerken, op basis van de modellen worden de optimale mediumtypemix en een inschatting van de stijging op de geformuleerde doelstellingen weergegeven. Hiermee gaan we dus van algemene benchmarks op databaseniveau naar gedifferentieerde benchmarks voor de campagne.

De tool wordt inmiddels gebruikt bij het opstellen van campagneplannen en helpt bij het kwantificeren van ‘SMART’ doelstellingen (dat wil zeggen: doelstellingen moeten **S**pecifiek, **M**eetbaar, **A**ceptabel, **R**ealistisch en **T**ijdgebonden zijn) en bij het onderbouwen van budgetten per mediumtype. Door verschillende scenario’s in te vullen wordt in een oogopslag zichtbaar wat het verwachte effect is van budgetwijzigingen of een andere mediumtypemix.

Natuurlijk blijft de tool ‘slechts’ een handig hulpmiddel. De output is richtinggevend en dient met vakkennis en ervaring geïnterpreteerd te worden. In gezamenlijke sessies met het reclame- en mediabureau dragen de inzichten bij aan een meer strategische invulling van de media en middelen.

## 5. DISCUSSIE

Het systematisch onderzoek naar de effectiviteit van overheidscampagnes heeft duidelijke inzichten opgeleverd in de werking van campagnes. Campagneherinnering is in de literatuur een enigszins omstreden effectindicator, maar blijkt in combinatie met een compacte, heldere en relevante boodschap toch een onmisbare voorwaarde voor effectieve voorlichtingscampagnes.

Postbus 51-campagnes presteren over het algemeen goed qua campagneherinnering, maar er zijn grote onderlinge verschillen, die voor een belangrijk deel (49% van de variantie) verklaard kunnen worden door het mediabudget, het campagneconcept en de betrokkenheid van de burger bij het onderwerp.

Het belang van *likeability*, de waardering van de campagne, voor het behalen van resultaten speelt ook voor overheidscampagnes een sleutelrol. Maar in tegenstelling tot wat reclamemakers soms lijken te denken, betekent dat niet dat het gebruik van humor daarvoor het aangewezen middel is.

Het is belangrijker dat de uitingen duidelijk en opvallend zijn en voorts informatief, geloofwaardig en niet irritant.

Voor de praktische toepassing van de resultaten is het vooral van belang dat de onderzoeksbevindingen een richtinggevende rol spelen in het hele proces van campagneontwikkeling. Naast de beschreven *tools* en activiteiten, zal de komende jaren het accent ook komen te liggen op een systematische aanpak van (kwantitatief en kwalitatief) pretestonderzoek. In plaats van alleen achteraf te evalueren, kan met de opgedane kennis nu ook vooraf de communicatieve werking goed voorspeld worden. Dit biedt uiteraard meer mogelijkheden voor tijdige bijsturing.

Tot slot nog een opmerking over het model. In vergelijking met de Zilveren Standaard is het met het huidige *pre post design* moeilijker om de werking van de campagne te isoleren van overige invloeden zoals media-aandacht en andere beleidsinstrumenten. Twee ontwikkelingen die voor de komende jaren voorzien zijn, kunnen hierin een verdere verbetering brengen. Ten eerste wordt momenteel gewerkt aan de opzet van systematische media-analyse, dat wil zeggen het vaststellen van de mate waarin en de wijze waarop de media over verschillende thema's berichten. Op die manier kan media-aandacht aan het verklaringsmodel worden toegevoegd. Ten tweede biedt het feit dat het onderzoek nu via een access panel uitgevoerd wordt, de mogelijkheid om het denkmodel van de Zilveren Standaard in de huidige opzet te integreren. Een eerste experiment (Bronner et al, 2003), waarbij het aantal (te verwachten) campagneconfrontaties per respondent zowel voorafgaand als na de campagne redelijk betrouwbaar wordt geschat, lijkt op dit punt goede vooruitzichten te bieden.

### NOTEN

1. Campagne-effectonderzoek is een andere benaming voor de meer bekende term tracking-onderzoek.
2. Met ingang van februari 1997 wordt Postbus 51-zendtijd ook ingekocht op de commerciële zenders. Hierdoor zijn de benodigde mediabudgetten aanzienlijk gestegen.



3. Potentiële gebruikers van (de resultaten) van het campagne-effectonderzoek zijn:
  - a. Campagnemanagers (om effectieve campagnes in te kunnen richten en het verantwoorden van bestede budgetten)
  - b. Tweede Kamer (om de effectiviteit van overheidsbeleid en doelmatigheid van overheids-gelden te bewaken)
  - c. Mediaplanners (om de media-inzet te optimaliseren)
  - d. Communicatiewetenschappers (geïnteresseerd in campagneoverstijgende leereffecten van voorlichting)
  - e. Reclamebureaus (voor kennis over de werking van eerdere campagneconcepten).
4. Het veldwerk wordt uitgevoerd door TNS NIPO, onder begeleiding van de unit Media & Monitoring van de RVD/Publiek en Communicatie.
5. De vraag naar spontane herinnering levert kwalitatieve data die waardevol zijn bij het beoordelen van de communicatieve werking van individuele campagnes, maar moeilijk bruikbaar zijn in de vergelijking met andere campagnes. In dit artikel wordt hier niet nader op ingegaan.
6. De kwalitatieve data van spontane boodschapoverdracht zullen hier niet worden behandeld, zie ook noot 5.
7. Bijvoorbeeld: tv-spot, radiospot, buitenreclame.

#### LITERATUUR

- Algemene Rekenkamer, 1991. Voorlichtingscampagnes van het Rijk. SDU, 1991.
- Bronner, F., Noort, W. van der, Ross, R., & Tchaoussoglou, C., 2003. The Netherlands live with water: exciting, efficient, effect research. Boston: Worldwide Readership Research Symposium, pp. 73-82.
- Montfort, P. van, & Kleuver, E. de, 2001. Trackingonderzoek naar overheidscampagnes. In: B. Dewez, P. et al (red). Overheidscommunicatie. De nieuwe wereld achter Postbus 51. Amsterdam: Boom, pp. 93-103.
- Nelemans, I. (2003). Humor in overheidscampagnes. Doctoraal scriptie, Universiteit Leiden.
- RVD/ Publiek en Communicatie, 2003. Evaluatie Postbus 51-campagnes 2002. Campagneverslagen, imago Postbus 51 en inzichten voor verbeteren van effectiviteit.
- RVD/ Publiek en Communicatie, 2004. Jaarevaluatie Postbus 51-campagnes 2003.
- Veldkamp, 1993. Naar een zilveren standaard; Een aanzet tot richtlijnen voor effectonderzoek naar campagnes van de rijksoverheid. Amsterdam, 1993.
- Veldkamp/TNS NIPO, 2003. Herijken van de benchmarks. Postbus 51-trackingonderzoek.

## 5. *Internetkannibalisatie in de dagbladen-industrie: mythe of realiteit?*\*

B. DELEERSNYDER, I. GEYSKENS, K. GIELENS en M.G. DEKIMPE

### SAMENVATTING

Heel wat bedrijven schuwen de mogelijkheden die het Internet biedt om hun producten rechtstreeks te verkopen aan hun klanten, voornamelijk uit angst voor kannibalisatieverliezen bij de bestaande distributiekanaalen. De vrees voor kannibalisatie leidt niet alleen tot grote verwarring over de bijdrage die het Internet als extra distributiekanaal kan leveren tot de totale inkomsten van de onderneming, maar geeft ook vaak aanleiding tot gespannen relaties met het bestaande distributienetwerk. Om te bepalen of kannibalisatieverliezen inderdaad optreden, bestudeerden we de invloed van het toevoegen van een Internetkanaal op de inkomsten van het gevestigde distributienetwerk in de dagbladenindustrie in het Verenigd Koninkrijk en Nederland. Concreet werd met behulp van een econometrische tijdreeksstechniek ("structural-break" analyse) bepaald of het introduceren van een "online" editie een permanente, negatieve verandering veroorzaakte in de inkomsten (uit circulatie en uit reclame) van de traditionele papieren krant. Uit analyse van de inkomsten van 85 individuele dagbladen blijken deze kannibalisatieverliezen eerder uitzonderlijk, waaruit we kunnen concluderen dat de algemene vrees voor Internetkannibalisatie in deze sector ongegrond is. Toch willen we erop wijzen dat het gevaar voor Internetkannibalisatie niet helemaal genegeerd mag worden. Wanneer het Internetkanaal immers gepositioneerd wordt als een substituuat voor de bestaande distributiekanaalen, is Internetkannibalisatie wel reëel.

### 1. DE VOOR- EN NADELEN VAN EEN "CLICK 'N BRICK" STRATEGIE

"Click'n brick" strategieën, i.e. een multi-kanaal aanpak waarbij distributie via een fysiek winkelnetwerk wordt aangevuld met een Internetkanaal, worden de voorbije jaren aanzienlijk aan populariteit. Voor traditionele producenten opent het Internet een nieuw kanaal, zowel naar hun bestaande als naar potentieel nieuwe klanten die ze voordien niet (efficiënt) konden bereiken. De consument wil namelijk steeds meer op elk ogenblik van de dag bediend worden en liefst via verschillende distributiekanaalen. Het aanbieden van het productassortiment via het Internet komt voor een groot stuk tege-

\* Dit artikel is gebaseerd op een hoofdstuk uit de dissertatie van Barbara Deleersnyder, getiteld "Marketing in turbulent times" (2003), die werd geschreven aan de Katholieke Universiteit van Leuven (promotor: M.G. Dekimpe), en op de studie van Deleersnyder, Geyskens, Gielens en Dekimpe die werd gepubliceerd in de *International Journal of Research in Marketing* (2002). Reproductie van de relevante delen uit dit artikel gebeurt met toestemming van Elsevier.

moet aan deze behoefte. Producenten hopen daarom via deze weg het bereik van hun markt aanzienlijk te vergroten. Sinds enige jaren is het Internet dan ook een sterk groeiend verkoopkanaal waarlangs producenten allerhande producten aanbieden aan de consument: van kleding en meubels tot reizen en bankdiensten. Webwinkelen wint in een steeds sneller tempo aan populariteit. In 2003 werden in Nederland voor een bedrag van 1.2 miljard euro producten gekocht via het Internet, wat een stijging van 32 procent ten aanzien van het vorige jaar betekende (Houtman 2004). Bovendien is het Internetpubliek in Nederland een potentieel interessante doelgroep, bestaande uit voornamelijk hoger opgeleiden en goede verdieners jonger dan 50.

Het toevoegen van een Internetkanaal wordt dan ook haast als een must voor zowat alle producenten beschouwd (Financieel Economische Tijd 2002). De vraag is echter of het toevoegen van een Internetkanaal inderdaad het gedroomde antwoord is op de steeds luider klinkende roep naar flexibele distributiekanaalen. Hoe effectief geld kan verdiend worden via een Internetkanaal is immers voor velen nog onduidelijk. Volgens de ene biedt een extra Internetkanaal de mogelijkheid tot het besparen van kosten en het stimuleren van omzet, voor de ander vormt een extra Internetkanaal juist een reële inkomstenbedreiging die enkel bijkomende kosten met zich brengt (Geyskens et al. 2002).

Via een Internetkanaal kunnen producenten rechtstreeks hun producten aanbieden aan de eindconsument in plaats van via fysieke winkels te opereren. Hierdoor zullen heel wat fysieke, schriftelijke en telefonische contacten tussen producenten en traditionale distributeurs wegvallen (Stellinga 2004). Dit kan een aanzienlijke reductie in het kostenplaatje betekenen. Zo omzeilen prijsvechters als Ryanair en Easyjet in de luchtvaartindustrie commissies van reisagenten door klanten rechtstreeks te laten boeken via Internet. Ook banken, waaronder ABN-AMRO en ING, besparen aanzienlijk in personeelskosten nu klanten steeds meer hun bankzaken zelf via het Internet behartigen.

Voor gevestigde bedrijven betekent een extra Internetkanaal echter niet automatisch extra inkomsten of minder kosten. Zo mogen belangrijke kostenposten eigen aan het Internet niet uit het oog verloren worden. Automatisering kan duurder blijken dan gedacht. Verder is in een multi-kanaal context een Internetkanaal uiteraard niet het enige distributiekanaal, zodat bestaande consumenten eenvoudig het ene kanaal voor het ander kunnen inruilen zonder dat de producent er in slaagt nieuwe klanten aan te trekken. Bestaande kanalen kunnen hierdoor klanten verliezen aan het nieuwe Internetkanaal en opstandig worden (Geyskens et al. 2002). Zo werd in het verleden de spijkerbroekfabrikant Levi's gedwongen zijn Internet verkoopactiviteiten te staken na hevig protest van zijn gevestigd distributienetwerk. Door het ontbreken van eenduidige onderzoeksresultaten over het effect van het inzetten van een extra Internetkanaal op de verkopen van bestaande distributiekanaalen, gaan velen er momenteel spontaan van uit dat Internetkannibalisatie onvermijdelijk optreedt (Ghosh 1998; Porter 2001).

De voornaamste doelstelling van deze studie is te achterhalen of deze vrees voor kannibalisatie gegrond is. Concreet willen we in dit onderzoek de volgende twee vragen beantwoorden: In welke mate heeft het toevoegen van een Internetkanaal een invloed op de inkomsten van het gevestigde distributienetwerk? Zijn deze potentiële kannibalisatieverliezen enkel tijdelijk of van permanente aard?

Om een antwoord op deze vragen te formuleren, analyseren we systematisch hoe de lange termijn omzetevolutie van gevestigde distributiekanaalen wijzigt na de introductie van een Internetkanaal aan de hand van een econometrische tijdreeksanalyse: de "structural-break" methode. Voor deze analyse werden tijdreeksgegevens verzameld over de inkomsten in de dagbladenindustrie in het Verenigd Koninkrijk en Nederland. Ook

voor krantenuitgevers was het Internet een nieuw en veelbelovend kanaal met een enorme groep potentiële klanten. Het Internet leek dé nieuwe manier om extra lezers te vinden. Maar hoe op die markt geld kon worden verdiend, was en is voor velen volstrekt onduidelijk (Wiersma 2004). Dagbladen zijn namelijk typische informatiegoederen die door hun aard enerzijds bijzonder interessant zijn om te verkopen via het Internet, maar anderzijds heel gevoelig zijn voor kannibalisatie.

## 2. INFORMATIEGOEDEREN EN DE DAGBLADENINDUSTRIE

Sectoren die geen fysieke producten aanbieden, maar digitaliseerbare informatiegoederen zoals kranten, software, en muziek, zijn het meest kwetsbaar voor potentiële kannibalisatiegevaren (Shapiro & Varian 1999). De “online” levering van zulke informatieproducten gebeurt doorgaans sneller en efficiënter dan bij de klassieke distributiekkanalen. Een Internetkanaal biedt daarenboven eenvoudige en uitgebreide zoekfaciliteiten die zeer aantrekkelijk zijn voor informatieproducten, en vergemakkelijkt de relatie met de klant. Deze eigenschappen maken dat het Internet een interessant distributiekanaal is voor de producent van informatieproducten.

In dit onderzoek leggen we ons toe op de dagbladenindustrie. Naast de bovenstaande eigenschappen die dagbladen delen met alle informatiegoederen, bezit deze sector een aantal bijkomende karakteristieken die haar bijzonder interessant maakt voor Internettransacties (hierop wordt later uitgebreider ingegaan). De keerzijde van de medaille is echter de grote vrees voor Internetkannibalisatie: één van de meest gehoorde bedenkingen die uitgevers van papieren media maakten bij het lanceren van een “online” tegenhanger voor hun gedrukte publicaties was of het Internet geen bedreiging vormde voor de inkomsten van hun “offline” publicatie. Traditioneel heeft de dagbladenindustrie twee afzonderlijke bronnen van inkomsten waartussen een wisselwerking bestaat (Abbrink & Van Ours 1994): circulatie-inkomsten verkregen uit de verkoop van het aantal gedrukte exemplaren en reclame-inkomsten. In 2003 bedroegen in Nederland de totale bruto-inkomsten uit abonnementen en losse verkoop 897 miljoen euro en uit advertenties 863 miljoen euro (CebucoNet 2004). Advertenties zorgen dus voor bijna 50% van de inkomsten van dagbladen.

Met betrekking tot de omzetinkomsten van de papieren editie, stelde men zich de vraag of men ongestraft extra geld kon creëren door het product op meerdere media aan te bieden zonder dat het kannibaliseert, of dat de mogelijkheid bestond om via een eigen Internetkanaal het bereik van een publicatie drastisch te vergroten. Zo blijkt uit Belgische gegevens dat nogal wat abonnees van De Standaard website nog nooit een papieren Standaard hebben aangeschaft. Een uitgebreide website blijkt daarom voor een krant vaak een ideaal middel om een jonger publiek aan te trekken, dat voordien nog nooit een abonnement op een papieren krant heeft genomen, maar dat hopelijk op latere leeftijd wel doet. Kranten en krantensites hoeven dus niet noodzakelijk concurrerende media te zijn, maar kunnen elkaar complementeren. Uit onderzoek van WebWereld (2004) blijkt dan ook dat bijna alle Nederlandse landelijke kranten in 2003 een flinke groei op het Internet hebben doorgemaakt. Grootste stijger in absolute aantallen is De Telegraaf. Hoewel exacte cijfers moeilijk te achterhalen zijn, heeft de telegraaf.nl dagelijks enkele tienduizenden bezoekers waarvan het aantal jaarlijks nog spectaculair toeneemt: de website van deze krant trok in januari 2004 vijfhonderdduizend bezoekers *meer* dan in januari 2003, een stijging van maar liefst 18,9 procent. Deze cijfers zijn

aanzienlijk, zeker wanneer we weten dat de dagelijkse oplage van de papieren krant in 2003 rond de 750.000 exemplaren schommelde.

Desondanks zijn niet alle berichten even positief. De Internetter is of wordt meestal geen abonnee op de papieren krant. Erger nog, momenteel leveren uitgevers een hevige strijd om het krimpende aantal lezers in de markt (Mediafacts 2004). Hierdoor vreest men dat elke toename in het aantal lezers van de Internetkrant afkomstig is (ten koste gaat) van het lezersbestand bij de bestaande krant (Abbring & Van Ours 1994). Het switchen van consumenten van de papieren naar de “online” editie kan daarenboven in de hand worden gewerkt door de volgende twee factoren. Op de eerste plaats kan het Internet veel sneller nieuws verspreiden dan de papieren krant, waarbij de snelheid waarmee het product ‘nieuws’ de consument kan bereiken de waarde ervan ten goede komt. In de nieuwsindustrie, waarbij alles draait om de snelheid waarmee de redactie haar klanten kan bereiken, wordt een Internetkrant hierdoor een heel aantrekkelijk alternatief (substituut) dat door lezers kan worden verkozen boven haar huidige papieren krant. Een andere belangrijke factor die Internetkannibalisatie kan stimuleren is het feit dat, in tegenstelling tot de papieren krant, de Interneteditie in het verleden vaak geheel gratis ter beschikking werd gesteld. Wanneer consumenten switchen van de papieren krant naar de gratis “online” editie betekent dit meteen ook een netto verlies voor de uitgever. Recent stellen uitgevers daarom in toenemende mate hun Internetkranten slechts tegen betaling ter beschikking.

Ook inkomsten uit reclame kunnen negatief evolueren na de introductie van de Interneteditie. Dit omdat enerzijds de inkomsten uit reclame positief gerelateerd zijn aan de oplage van de krant, zodat een daling in het aantal verkochte exemplaren als gevolg van Internetkannibalisatie ook een daling in de reclame-inkomsten kan veroorzaken. Anderzijds hebben Silk en collega's (1999) gewezen op de aantrekkelijkheid van het Internet als reclamemedium, en het gevaar dat hiermee gepaard gaat voor de traditionele media om haar adverteerders hieraan te verliezen. Op het Internet nemen advertentie-uitgaven dan ook sterk toe, hoewel het volume nog redelijk beperkt is. Momenteel wordt geschat dat dagbladen wereldwijd slechts twee procent van hun omzet uit Internetadvertenties halen. De rest wordt nog steeds via de papieren krant gegenereerd (Planetinternet 2004).

Ondanks hun vrees voor Internetkannibalisatie blijken uitgevers vaak het voortouw te nemen in het gebruik van het Internet als bijkomend distributiekanaal. De Europese Commissie (1996, p.1) erkende dat de ontwikkelingen met betrekking tot de handel over het Internet in deze sector toonaangevend kunnen zijn voor de ontwikkelingen die zich mogelijk later zullen ontplooiën in heel wat andere sectoren. De aantrekkelijkheid van het Internetkanaal in combinatie met haar gevoeligheid voor kannibalisatieverliezen maken van de dagbladenindustrie dan ook een bijzonder geschikte sector om te toetsen in welk mate het Internet de inkomsten van gevestigde distributiekanaalen kannibaliseert.

### 3. ONDERZOEKSMETHODE

Om te bepalen of de introductie van een Internetkanaal resulteert in een structurele verandering in de inkomsten van de gevestigde kanalen, bestudeerden we aan de hand van een “structural-break” analyse de circulatie- en advertentie-inkomsten van de papieren kranten van dagbladenuitgevers in het Verenigd Koninkrijk en Nederland. Deze

methode is een zeer aantrekkelijke tijdreeksstechniek die werd ontwikkeld om de potentiële invloed na te gaan van een unieke structurele gebeurtenis op de lange termijn evolutie van een tijdreeks (Ben-David & Papell 1995). In deze studie zal de techniek specifiek gebruikt worden om de invloed te bepalen van de introductie van een additioneel distributiekanaal, maar de methode biedt tevens interessante mogelijkheden om de invloed te bepalen van andere unieke, potentieel sterk ingrijpende, veranderingen.<sup>1</sup> Eerst bespreken we de intuïtie achter een “structural-break” analyse (3.1). Daarna lichten we de gebruikte data toe (3.2).

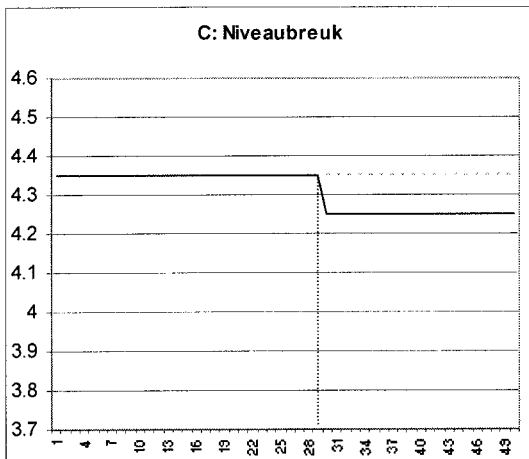
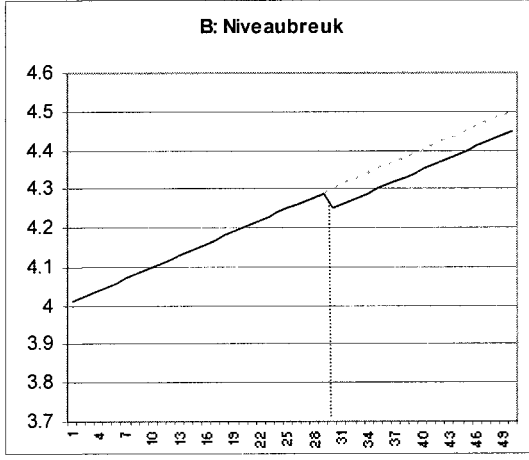
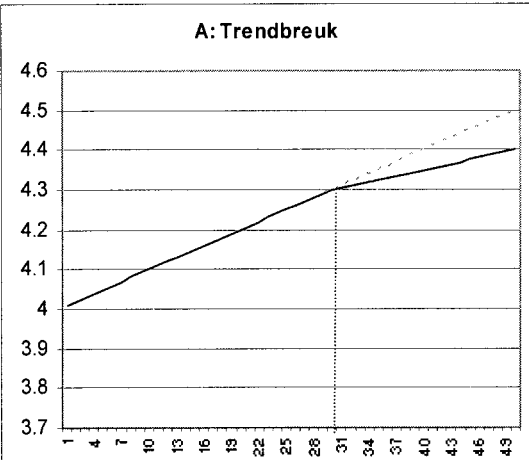
### 3.1. “Structural-break” analyse

Doorgaans zijn kanaalbeslissingen ingrijpende en eerder zeldzame gebeurtenissen die, eens ze genomen zijn, niet snel omkeerbaar zijn. Wanneer een onderneming een Internetkanaal heeft toegevoegd, is het vrij onwaarschijnlijk dat dit kanaal in dezelfde maand wordt afgevoerd/stopgezet, daar waar beslissingen over het dagelijkse beleid, zoals promotionele marketingacties, van week tot week kunnen worden aangepast. In de laatste decennia hebben marketingonderzoekers zich voornamelijk gefocust op deze frequente marketingacties (zie bijvoorbeeld het onderzoek naar de (lange termijn) impact van prijspromoties van Nijs et al. (2001), Steenkamp et al. (2004) en Srinivasan et al. (2000), waarbij uitgebreid gebruik wordt gemaakt van traditionele “unit-root” testen). De impact van meer *structurele* gebeurtenissen is tot op vandaag veel minder bestudeerd in de marketingliteratuur.

“Structural-break unit-root” analyse werd ontwikkeld om de invloed na te gaan van unieke structurele gebeurtenissen op de lange termijn evolutie van een tijdreeks (Ben-David & Papell 1995). Het toevoegen van een nieuw distributiekanaal door de producent is een dergelijke gebeurtenis, die mogelijk een *breuk* veroorzaakt in het toekomstige verloop van de inkomsten bij de andere, reeds bestaande distributiekanaalen. Wanneer deze introductie het verloop van de inkomsten ook na de periode waarin de introductie zich voordeed blijft beïnvloeden en de inkomsten dus niet opnieuw convergeren naar hun oorspronkelijk verwachte verloop (i.e. het verloop dat kan geëxtrapoleerd worden op basis van de historische tijdreekspatronen van vóór de gebeurtenis), is er sprake van een permanente of lange termijn impact als gevolg van de introductie van het Internetkanaal. Aan de hand van de “structural-break” analyse wordt met andere woorden bepaald of er een breuk is opgetreden in de evolutie van de inkomsten na het toevoegen van een Internetkanaal.

Die permanente verandering kan zich op twee manieren manifesteren: (i) een permanente verandering in de *groei* van de inkomsten (trendbreuk); en (ii) een permanente verandering in het *niveau* van de inkomsten (niveaubreuk). Lange termijn kannibalisatie doet zich voor wanneer een Internetkanaal resulteert in een *negatieve* trend- of niveaubreuk. Beide kannibalisatievormen worden geïllustreerd in Figuur 1 waar een producent een Internetkanaal toevoegt aan zijn gevestigd distributienetwerk in periode 30. Deze introductie veroorzaakt duidelijk kannibalisatieverliezen bij de gevestigde kanalen.

Wanneer deze kannibalisatieverliezen zich uiten als een *trendbreuk* (Figuur 1A), dan ligt de groei (weerspiegeld in de helling van de curve) na de introductie beduidend lager dan de lange termijn groei die we zouden verwachten indien er geen breuk was opgetreden in periode 30 (visueel voorgesteld met de stippellijn). De grootte van het overeenkomstige kannibalisatieverlies is dan het verschil tussen de werkelijke lange



Figuur 1. Negatieve trendbreuk (A) en niveaubreuk (B en C) in de tijdreeks.

termijn groeivoet na de breuk en de hypothetische verwachte lange termijn groeivoet wanneer er geen trendbreuk was opgetreden.

Kannibalisatieverliezen kunnen zich echter niet alleen uiten in de vorm van een verandering in de lange termijn groeivoet, maar kunnen zich tevens manifesteren in de vorm van een permanente daling in het gemiddelde niveau van de inkomsten. Bij een dergelijke *niveaubreuk* (Figuur 1B & 1C) stabiliseren de geobserveerde verkopen na de breuk op een lager niveau. De lange termijn groeivoet blijft daarbij ongewijzigd. Het overeenkomstige kannibalisatieverlies wordt nu bepaald als het verschil in het hypothetische lange termijn niveau van de inkomsten zonder de breuk (weerom voorgesteld door de stippellijn in Figuren 1B & 1C) en het werkelijke lange termijn niveau na de breuk.

De verliezen tengevolge van een trendbreuk in de inkomsten zullen steeds groter worden doorheen de tijd. Naarmate de tijd voortschrijdt, zullen de *geobserveerde* waarden bij een trendbreuk immers steeds verder afwijken van de *hypothetische* waarden zonder trendbreuk. Het kannibalisatieverlies bij een niveaubreuk daarentegen blijft in elke periode even groot. De potentiële kannibalisatiegevolgen op lange termijn bij een trendbreuk zullen daardoor groter worden dan die bij een niveaubreuk. Daarom is het belangrijk eerst te bepalen of er al dan niet een trendbreuk optreedt, om vervolgens, wanneer die niet optreedt, verder te zoeken naar een mogelijke niveaubreuk.

Om de aanwezigheid van kannibalisatieverliezen formeel te testen, moet nog rekening gehouden worden met twee bijkomende aspecten. Op de eerste plaats moet vooraf nagegaan worden of de inkomstenreeks van de gevestigde kanalen *stabiel* is of *evolueert*, los van de toevoeging van het Internetkanaal. Een stabiele inkomensreeks heeft namelijk een vast gemiddelde (of een deterministische trend/groeivoet) waar slechts tijdelijk van afgeweken wordt. Bij evoluerende verkopen bestaat er echter geen vast gemiddelde (of een deterministische trend), zodat de serie bij afwijkingen niet opnieuw naar het initiële niveau zal convergeren, maar permanent kan afwijken van het vorige niveau (of van de deterministische trendlijn). Dit onderscheid wordt traditioneel gemaakt op basis van “unit-root” testen, maar deze dienen aangepast te worden wanneer een “structural-break” aanwezig kan zijn in de series (zie verder).<sup>2</sup>

Een tweede belangrijk aspect dat in rekening wordt gebracht bij “structural-break” analyse is de *snellheid* (dynamiek) waarmee de inkomstenevolutie van vóór de breuk naar een nieuwe lange termijn evolutie na de breuk evolueert. Wanneer een structurele verandering zoals het toevoegen van een Internetkanaal een breuk genereert in de bestaande inkomstenevolutie, dan is het vrij onwaarschijnlijk dat de inkomsten onmiddellijk in de volgende periode stabiliseren op het nieuw (post-breuk) lange termijn evenwicht. Zo zijn lezers van dagbladen vaak voor een bepaalde periode gebonden aan hun abonnement, waardoor ze hun beslissing om dit contract op te zeggen niet van vandaag op morgen kunnen herzien. Daarom is het belangrijk dat de “structural-break” methode toelaat dat er meteen na de breuk tijdelijk een turbulente periode optreedt, en dat er enige tijd zal verlopen vooraleer de inkomsten op hun nieuwe lange termijn evolutie stabiliseren (Franses 1998).

Concreet wordt het volgende “structural-break unit-root” model (Perron 1989; 1994) geschat:

$$Y_t = \alpha + \beta T + [\theta DU_t + \gamma DT_t + \delta D(TB)_t] + \rho Y_{t-1} + \sum_{i=1}^k c_i \Delta Y_{t-i} + \sum_{s=2}^S \eta_s SD_{st} + \varepsilon_t \quad (1)$$



waarbij  $Y_t$  en  $Y_{t-1}$  de inkomsten bij de gevestigde kranten weergeven op tijdstip  $t$  en  $t-1$  en  $\Delta Y_t$  het verschil tussen beide is. De diverse  $\Delta Y_t$  worden toegevoegd om er voor te zorgen dat de foutenterm witte ruis wordt. De waarden van de  $c_i$  parameters zullen dan de hierboven besproken “adjustment” dynamiek weerspiegelen.  $T$  is een deterministische trend en  $SD_{st}$  is een set van seizoen dummy variabelen die de seizoensfluctuaties in de inkomsten weergeven.  $DU_t$ ,  $DT_t$  en  $D(TB)_t$  zijn drie “structural-break” variabelen die de verwachte trendbreuk of niveaubreuk kunnen oppikken, waarbij  $TB$  (“Time of Break”) de potentiële breukperiode is, of de datum waarop de breuk zich mogelijk heeft voorgedaan. In onze studie stellen we die datum (exogeen) vast als de datum waarop een bepaalde krant zijn Interneteditie introduceerde.  $DT_t$  neemt de waarde  $t+1$  aan als  $t \geq TB$ , en is nul anders. Deze variabele laat dus een verandering toe in de groeivoet vanaf het tijdstip van de breuk. De break variabele  $DU_t$  neemt de waarde 1 aan als  $t \geq TB$ , en is nul anders, waardoor deze een niveauverandering toelaat in de inkomstenreeks. Ten slotte neemt  $D(TB)_t$  de waarde 1 aan enkel op het ogenblik van de potentiële breuk (wanneer  $t = TB$ ), en is nul anders. Deze laatste variabele wordt toegevoegd om statistische redenen, maar heeft geen onmiddellijke interpretatie (Perron 1994). Zonder deze drie extra dummy variabelen herleidt vergelijking (1) zich tot een traditionele “unit-root” test, zoals o.a. gebruikt in Dekimpe en Hanssens (1995).

Op basis van de waarde van  $\rho$  wordt vergelijking (1) verder geanalyseerd zoals weergegeven (indien  $\rho < 1$ ), dan wel herschreven in “first-differences” (indien  $\rho$  in vergelijking (1) niet significant verschillend van 1 is). In beide gevallen wordt vervolgens in eerste instantie getest of  $\gamma$  (negatief) significant is. Dit laat toe om te testen of er een daling in de lange termijn groeivoet ontstaat ten gevolge van de Internetintroductie. Indien dit niet het geval is, wordt vervolgens getest of  $\theta$  significant (en negatief) is, om permanente niveaudalingen ten gevolge van de Internetintroductie te detecteren. Bemerkt dat we hier (in navolging van vroegere literatuur, zie bijvoorbeeld Perron 1989; 1994; Ben-David & Papell 1995) veronderstellen dat de structuurbreuk zich uit in de ligging van de deterministische trendlijn, en niet in veranderingen in de autoregressieve parameters van het model (i.e. de  $c_i$ ), het seizoenspatroon (i.e. de  $\eta_s$ ), of de variantie van de foutenterm. Voor een gedetailleerde beschrijving van de diverse stappen in de testprocedure, verwijzen we naar Deleersnyder et al. (2002). Een beknopt overzicht van deze procedure staat in Appendix 1.

### 3.2. Data

In deze studie maakten we gebruik van gegevens verzameld in de dagbladenindustrie in het Verenigd Koninkrijk en Nederland. Aangezien kannibalisatie ten gevolge van de toevoeging van een Internetkanaal zich manifesteert als een verlies bij de bestaande distributiekanaalen, werd onderzocht in welke mate de inkomsten uit de papieren krant *negatief* evolueerden nadat deze ook “online” beschikbaar werd gesteld.

Zoals eerder aangegeven (Sectie 2), hebben dagbladen twee afzonderlijke bronnen van inkomsten (Abbring & Van Ours 1994): circulatie-inkomsten, verkregen uit de verkoop van het aantal verkochte exemplaren, en reclame-inkomsten. In deze studie gebruikten we maandelijkse tijdreeksgegevens met betrekking tot de bestaande (“offline”) inkomsten van 85 individuele dagbladen, die elk een Interneteditie lanceerden in de voorbije 10 jaar. Informatie over het tijdstip waarop de gratis “online” editie werd geïntroduceerd (de potentiële breuk datum  $TB$  in de “structural-break” analyse) werd verkregen door contact op te nemen met elk van deze dagbladen.

De geanalyseerde gegevens hebben uitsluitend betrekking op het gevestigde distributiekanaal (de papieren editie). *Circulatie*-gegevens werden verkregen bij het Engelse ‘Audit Bureau of Circulation’ (ABC) voor 12 nationale en 55 regionale dagbladen in het Verenigd Koninkrijk. Deze circulatie-inkomsten zijn uitgedrukt als het gemiddelde aantal exemplaren van de papieren krant dat dagelijks wordt verkocht. In totaal beschikten we over 138 maandelijkse circulatie-gegevens over de periode januari 1990 tot en met juni 2001. Nederlandse *reclame*-inkomsten werden verkregen bij het ‘Bureau voor Budgetten Controle’ (BBC) voor 9 nationale en 9 regionale dagbladen die elk een “online” editie introduceerden in de afgelopen 10 jaar. Deze maandelijkse reclamedata zijn uitgedrukt in Nederlandse gulden (gedeflateerd) en beslaan de periode van januari 1990 tot en met oktober 2000 (130 observaties).<sup>3</sup>

Om bij de “structural-break” analyse de wijziging in de evolutie van de serie na de breuk te bepalen, hebben we voldoende observaties nodig na de verwachte breukdatum. Op basis van de observaties na de breukdatum wordt immers mede bepaald wat de nieuwe lange termijn evolutie is in de tijdreeks. Bij de circulatiegegevens hadden we gemiddeld 42 maanden na de introductie van de “online” editie beschikbaar (minimum = 20; maximum = 79). Bij de reclame-inkomsten waren gemiddeld 41 observaties na de introductie van de “online” editie beschikbaar (minimum = 6; maximum = 63).

## 4. RESULTATEN

De resultaten voor de circulatie- en advertentie-inkomsten staan samengevat in Tabel 1 en worden in detail beschreven in respectievelijk paragraaf 4.1 en 4.2.

### 4.1. *Circulatie*

De dagbladindustrie wordt al verscheidene jaren geteisterd door een aanhoudende daling in haar aantal lezers. Het is dan ook niet verwonderlijk dat de meeste tijdreeksen met betrekking tot het aantal verkochte exemplaren (circulatie-inkomsten) gekenmerkt werden door een significant dalende lange termijn trend, zowel voor als na de introductie van de “online” editie. Om van kannibalisatie te kunnen spreken, moeten we dan ook nagaan of en in welke mate de dalende lange termijn trend nog *versneld* wordt na de introductie van de Interneteditie. Wanneer we voor elk van de 67 dagbladen uit het Verenigd Koninkrijk het “structural-break” model schatten, bleek echter dat er bij slechts 5 kranten (i.e. 7% van het totaal aantal bestudeerde kranten) een significant negatieve trendbreuk is opgetreden na de introductie van een “online” editie (zie Tabel 1, kolom 2).<sup>4</sup> In afwezigheid van een significante trendbreuk, kan kannibalisatie zich echter ook manifesteren als een significante niveaudaling. Daarom werd bij de 52 kranten waar geen significante (positieve of negatieve) trendbreuk werd gevonden in een volgende stap nagegaan of een negatieve niveaubreuk was opgetreden na de introductie van de Interneteditie.<sup>5</sup> In slechts 4 van de 52 reeksen kon een significante niveaudaling gedetecteerd worden. Deze resultaten zijn weergegeven in kolom 3 van Tabel 1.

Tot dusver werd de “structural-break” analyse uitgevoerd voor elk van de dagbladen afzonderlijk. Wanneer we echter geïnteresseerd zijn in de algemene tendens in de industrie, kunnen de verkregen resultaten voor alle individuele dagbladen gecombineerd worden in een *meta-analyse*. In een meta-analyse worden de resultaten van een aantal vergelijkbare studies ‘gebundeld’. Zo kan met grotere betrouwbaarheid een alge-

Tabel 1. Lange termijn invloed van de introductie van een Internetkanaal (Deleersnyder et al. 2002, Tabel 1).

	Circulatie-inkomsten		Reclame-inkomsten	
	Invloed op lange termijn groeivoet	Invloed op lange termijn niveau	Invloed op lange termijn groeivoet	Invloed op lange termijn niveau
# bestudeerde dagbladen (n)	67	52	18	16
# negatieve breuk	35	26	11	7
# <b>significant</b> ( $p < .05$ ) negatieve breuk	5	4	1	0
# positieve breuk	32	26	7	9
Gemiddelde impact bij negatieve breuk	-.0014	-.0128	-.0065	-.0231
Gemiddelde impact bij positieve breuk	.0018	.0130	.0030	.0671
Algemeen gemiddelde	.0002	.0001	-.0030	.0276
Meta-analytische test voor kannibalisatie ( $H_0 =$ geen kannibalisatie-effect)	$p = .94$	$p = .58$	$p = .13$	$p = .98$

mene uitspraak worden gedaan over de gevolgen van de toevoeging van een “online” distributiekanaal (Rosenthal 1991). De resultaten verkregen op basis van de meta-analyse zijn daardoor ‘krachtiger’ dan de analyse van de individuele resultaten, en laten toe een uitspraak te doen die veralgemeenbaar is naar andere kranten die niet in onze steekproef zijn opgenomen.

Ten eerste kunnen we vaststellen dat voor een aanzienlijk aantal kranten (35) toch een negatieve coëfficiënt voor de trendbreuk werd gevonden, ook al is dit effect in bijna alle gevallen niet significant. Wanneer de werkelijke kannibalisatieverliezen eerder klein zijn, is het mogelijk dat dit effect in elke serie afzonderlijk moeilijk detecteerbaar is. De meta-analyse is nu in staat om de informatie op basis van deze ‘zwakkere’ individuele testen te combineren, en zo met grotere betrouwbaarheid te bepalen of kannibalisatie toch *collectief* aanwezig is in de circulatiegegevens. Concreet werd hiervoor de Rosenthal-test (1991) gebruikt. Deze meta-analytische test gaf opnieuw aan dat er geen collectief bewijs is voor een negatieve lange termijn *trendbreuk*: de nulhypothese dat er geen kannibalisatie optreedt, kon niet verworpen worden ( $p$ -waarde = 0.94). Wanneer we deze meta-analyse herhaalden met betrekking tot de *niveaudalingen*, konden we opnieuw vaststellen dat er geen kannibalisatie optrad ( $p$ -waarde = 0.58). Deze resultaten laten ons toe om te besluiten dat Internetkannibalisatie (in de jaren 1990-2001) afwezig was in de oplages van dagbladen in de UK.

#### 4.2. Reclame

De resultaten met betrekking tot de reclame-inkomsten zijn samengevat in Tabel 1, kolommen 4 en 5. In slechts 1 van de 18 reeksen werd een significant negatieve trendbreuk gedetecteerd.<sup>6</sup> Wanneer we vervolgens verder zochten naar mogelijke niveau-

dalingen in de evolutie van de reclame-inkomsten, konden we voor geen enkele krant een significante niveaudaling vinden.

Ondanks het feit dat slechts één Nederlands dagblad een significant negatieve trendbreuk vertoonde, vonden we ook hier een aanzienlijk aantal dagbladen met een niet-significante trendbreuk of niveaubreuk. Dit illustreert opnieuw het belang van een meta-analyse om met grotere betrouwbaarheid te bepalen of er al dan niet kannibalisatie optrad in de “offline” reclame-inkomsten bij dagbladen na introductie van de “online” editie. De meta-analyse resultaten toonden echter aan dat noch voor de trendbreuk, noch voor de niveaubreuk collectief bewijs bestaat voor Internetkannibalisatie: de nulhypothese die stelt dat er geen kannibalisatie optreedt, kon niet verworpen worden (de  $p$ -waarden zijn respectievelijk 0.13 en 0.98).

Kortom, ook bij de reclame-inkomsten is er geen empirisch bewijs te vinden voor kannibalisatie veroorzaakt door het Internetkanaal in de periode 1990-2000. We kunnen dus besluiten dat de kannibalisatievrees ook hier vooralnig ongegrond is.<sup>7</sup>

## 5. CONCLUSIES

Het Internet biedt producenten de mogelijkheid om rechtstreeks in contact te komen met de consument. De vele voordelen die elektronische handel biedt, in het bijzonder voor informatiegoederen, krijgen veel aandacht in de pers. Wanneer het er echter op aankomt om effectief een Internetkanaal toe te voegen aan hun bestaande distributiekanaalen, blijken bedrijven heel wat minder enthousiast: een Internet distributiekanaal wordt verwacht de verkopen bij de bestaande distributeurs te kannibaliseren. De angst voor deze kannibalisatieverliezen bij de huidige distributeurs leidde al tot hoog opgelopen conflicten tussen producenten en distributeurs. Ondanks de grote vrees voor kannibalisatie-effecten wanneer een nieuw kanaal wordt toegevoegd aan een gevestigd distributienetwerk, blijkt dat er tot op heden slechts heel beperkt onderzoek naar werd verricht.

De voornaamste doelstelling van dit onderzoek was te bepalen of deze kannibalisatievrees wel terecht is. Om dit te bestuderen verzamelden we tijdreeksgegevens met betrekking tot de inkomsten van 85 individuele dagbladen in het Verenigd Koninkrijk en Nederland, die allen een Interneteditie lanceerden in de afgelopen 10 jaar. Met een “structural-break” analyse werd vervolgens nagegaan of de introductie van een Interneteditie daadwerkelijk een lange termijn breuk veroorzaakte in de inkomsten. Onze resultaten zijn bemoedigend: we vonden slechts voor een beperkt aantal dagbladen (10) een permanente daling in het inkomstenverloop na de introductie van de Interneteditie. Ook de meta-analyse gaf aan dat permanente of lange termijn kannibalisatieverliezen in deze industrie globaal genomen afwezig zijn, zowel voor de circulatie-inkomsten als voor de reclame-inkomsten. Deze bevindingen lijken er op te wijzen dat uitgeverij er momenteel in slagen een nieuw publiek aan te boren met hun Internetedities. Uiteraard moeten we voorzichtig blijven met het veralgemenen van deze resultaten naar de toekomst. De huidige resultaten zijn namelijk gebaseerd op de assumptie dat er zich in de toekomst geen nieuwe, grote technologische verandering/voortgang voordoet op het vlak van Internetdistributie en –publiceren die er voor zorgt dat de “online” en “offline” lezers van de krant meer naar elkaar toegroeien (Silk et al. 2001).

Toch geven deze resultaten *niet* aan dat een Internetkanaal helemaal en in alle omstandigheden ongevaarlijk is voor de bestaande distributiekanaalen. De positionering

van het nieuwe “online” distributiekanaal is cruciaal en kannibalisatie kan toch optreden wanneer het Internetkanaal te veel als een substituuat wordt gepositioneerd voor de bestaande distributiekanaalen (Gilbert 2001; Silk et al. 2001). Ook onze resultaten bevestigen dat de positionering van het Internetkanaal uitermate belangrijk is. Wanneer we nagaan in welke mate kranten dezelfde inhoud aanbieden in hun gedrukte en “online” krant, merken we dat bij de dagbladen met een grotere overlap vaker kannibalisatieverliezen werden vastgesteld met betrekking tot de circulatie-inkomsten. Concreet was het percentage overlap (zoals gerapporteerd door de “webmaster” van elk van de dagbladen) bij de gekannibaliseerde dagbladen gemiddeld 78%, terwijl dit percentage bij de overige dagbladen gemiddeld slechts 45% was. Een formele *t*-test op het percentage overlap bij beide groepen gaf vervolgens aan dat deze gemiddelden inderdaad significant verschillen van elkaar.

Kortom, op basis van onze resultaten kunnen we besluiten dat de algemene vrees voor Internetkannibalisatie in de dagbladenindustrie niet gegrond is. Internetkannibalisatie mag dus geen reden zijn voor bedrijven om af te zien van de mogelijkheden die het medium biedt als alternatief distributiekanaal. Ook zijn deze resultaten geruststellend voor de bestaande distributiekanaalen in deze sector. Internet is immers niet noodzakelijk een substituuat, maar vormt een volwaardig, op zichzelf staand distributiekanaal dat, mits een adequate positionering, de belangen van een andere doelgroep dient.

Voor de praktijk van het marktonderzoek willen we er op wijzen dat de “structural-break” methode, die hier werd gebruikt om de invloed te bestuderen van het toevoegen van een Internetkanaal, ook bijzonder interessante mogelijkheden biedt om de invloed te bepalen van heel andere types van structurele veranderingen. Telkens wanneer bedrijven geconfronteerd worden met een grote externe verandering in hun markt, of ze nemen zelf een belangrijke beleidsbeslissing waarvan ze verwachten (of hopen) dat het een invloed heeft op hun bedrijfsactiviteiten, kan deze methode op eenvoudige wijze worden aangewend om de potentiële lange termijn invloed ervan effectief te meten. De techniek is bijzonder flexibel en laat toe dat de verandering die optreedt zich manifesteert in tal van verschillende vormen, zoals een positieve of negatieve verandering in de groeivoet (trendbreuk), een breuk in het gemiddelde niveau van de tijdreeks (of beide). Ook biedt de methode de mogelijkheid dat de breuk niet noodzakelijk onmiddellijk optreedt, maar slechts geleidelijk de evolutie van de bestudeerde tijdreeks verandert, en maakt ze onderscheid tussen enkel tijdelijke en permanente veranderingen als gevolg van deze interventie.

### Appendix 1: Bepalen van de lange termijn invloed van een structurele verandering

De “structural-break” analyse bestaat uit 3 opeenvolgende stappen, die voortvloeien uit het volgende “structural-break unit-root” model:

$$Y_t = \alpha + \beta T + \theta DU_t + \gamma DT_t + \delta D(TB)_t + \rho Y_{t-1} + \sum_{i=1}^k c_i \Delta Y_{t-i} + \sum_{s=2}^S \eta_s SD_{st} + \varepsilon_t \quad (A1)$$

waarbij:

- $Y_t$  de bestudeerde tijdreeks in periode  $t$ ;
- $\Delta Y_t = Y_t - Y_{t-1}$ . Deze extra autoregressieve termen worden toegevoegd om de fouten-term ( $\varepsilon_t$ ) witte ruis te maken en pikken de korte termijn dynamiek in de serie op. Het aantal van deze termen ( $k$ ) dat toegevoegd wordt in Eq. A1, wordt bepaald op basis van de procedure beschreven in Perron (1989; 1994);
- $T$  de deterministische trend;
- $SD_{st}$  een seizoen dummy variabele voor seizoen  $s$ ;
- $DU_t$  een step dummy variabele die de waarde 1 aanneemt vanaf de periode van de potentiële breuk  $TB$  (“Time of Break”), en 0 is in alle voorgaande perioden;
- $DT_t$  een trend dummy variabele die de waarde van de trend  $T$  aanneemt vanaf de periode van de potentiële breuk  $TB$ , en 0 is in alle voorgaande perioden;
- $D(TB)_t$  een puls dummy variabele die de waarde 1 aanneemt op het ogenblik van de potentiële breuk  $TB$ , en 0 is in alle andere perioden.

#### Stap 1: Is de tijdreeks stabiel of evoluerend?

In de eerste stap wordt bepaald of de serie stabiel is of evolueert. Formeel moet bepaald worden of in de tijdreeks een eenheidswortel (‘unit root’) aanwezig is, wat kan getest worden op basis van de parameter  $\rho$  in Eq. A1 (Perron 1989; 1994). Merk op dat de kritische waarden voor deze “unit-root” hypothese test afwijken van de normale kritische waarden, en kunnen gevonden worden in Perron (1989, Tabel VI.B).

(i) Indien  $\rho < 1$ , dan is de serie stabiel en kan men Eq. A1 vereenvoudigen door de autoregressieve term  $Y_{t-1}$  op te nemen in de andere autoregressieve termen  $\Delta Y_t$ , die de dynamiek in de serie na de breuk kunnen vatten. Het model wordt daarbij:

$$Y_t = \alpha + \beta T + \theta DU_t + \gamma DT_t + \delta D(TB)_t + \sum_{i=1}^{k+1} a_i Y_{t-i} + \sum_{s=2}^S \eta_s SD_{st} + \varepsilon_t \quad (A2)$$

(ii) Indien  $\rho = 1$ , dan is de serie evoluerend, en moet de tijdreeks  $Y_t$  getransformeerd worden tot  $\Delta Y_t$ , zodat de serie opnieuw stabiel wordt. Hierbij wordt het initiële model gereduceerd tot volgende vergelijking:

$$\Delta Y_t = \alpha + \gamma DU_t + \theta D(TB)_t + \sum_{i=1}^k a_i \Delta Y_{t-i} + \sum_{s=2}^S \eta_s SD_{st} + \varepsilon_t \quad (A3)$$

Dit onderscheid tussen beide soorten tijdreeksen blijft in elk van de volgende stappen de berekeningen beïnvloeden.

#### Stap 2: Is er een trendbreuk opgetreden als gevolg van de structurele verandering?

De parameterschattingen bij de break dummy variabelen geven het onmiddellijke

(korte termijn) effect weer van de structurele verandering op het verloop van de serie. Bij de “structural-break” analyse zijn we echter geïnteresseerd in de impact van de structurele verandering op de *lange termijn* evolutie van de tijdreeks. Om te bepalen of er een significante trendbreuk is opgetreden in het lange termijn verloop van de serie, wordt op basis van de parameterschattingen uit modellen A2 of A3 het verschil berekend tussen de lange termijn groeivoet *voor* en *na* de structurele verandering. Concreet berekent men het verschil in de *lange termijn* groeivoet voor de breuk ( $\Delta Y_{voor}$ ) en na de breuk ( $\Delta Y_{na}$ ) als volgt (Ben-David & Papell 1995):

(i) Stel  $\rho < 1$ :

$$\Delta Y_{na} - \Delta Y_{voor} = \frac{\beta + \gamma}{1 - \sum a_i} - \frac{\beta}{1 - \sum a_i} \quad (A4)$$

met  $\beta$ ,  $\gamma$  en  $a_i$  de parameterschattingen uit model A2.

(ii) Stel  $\rho = 1$ :

$$\Delta Y_{na} - \Delta Y_{voor} = \frac{\alpha + \gamma}{1 - \sum a_i} - \frac{\alpha}{1 - \sum a_i} \quad (A5)$$

met  $\alpha$ ,  $\gamma$  en  $a_i$  de parameter schattingen uit model A3.

De grootte van de structurele verandering op lange termijn wordt in beide gevallen berekend als  $\gamma/(1-\sum a_i)$ . Een trendbreuk treedt op wanneer deze verandering in de lange termijn groeivoet significant is.

In de afwezigheid van een significante verandering in de groei, kan een structurele verandering ook optreden onder de vorm van een niveaubreuk.

### Stap 3: Is er een niveaubreuk opgetreden als gevolg van de structurele verandering?

Om te bepalen of er een significante niveaubreuk optreedt als gevolg van de structurele verandering, wordt in het model uit stap 2 een gemeenschappelijke groeivoet of trend opgelegd vóór en na de potentiële breukdatum (aangezien in stap 2 gevonden werd dat er geen significante groeivoetverschillen waren). De verandering in het niveau van de serie kan daarbij bepaald worden na herschatting van de modellen A2 en A3:

(i) Stel  $\rho < 1$ :

$$Y_t = \alpha + \beta T + \theta DU_t + \delta D(TB)_t + \sum_{i=1}^{k+1} a_i Y_{t-i} + \sum_{s=2}^S \eta_s SD_{st} + \varepsilon_t \quad (A6)$$

(ii) Stel  $\rho = 1$ :

$$\Delta Y_t = \alpha + \theta D(TB)_t + \sum_{i=1}^k a_i \Delta Y_{t-i} + \sum_{s=2}^S \eta_s SD_{st} + \varepsilon_t \quad (A7)$$

waarbij de lange termijn verandering in het niveau vervolgens bepaald wordt door de grootte en significantie van  $\theta/(1-\sum a_i)$ .

Voor een meer gedetailleerde beschrijving van deze procedure verwijzen we naar Deleersnyder et al. (2002).

## NOTEN

1. De “structural-break” methode biedt nog talrijke andere mogelijkheden om de invloed te bepalen van structurele veranderingen in de markt. Zo analyseerden bijvoorbeeld Ben-David en Papell (1995) de invloed van wereldoorlog I en II, alsook de invloed van de grote depressie in de jaren '30, op de groei van het BBP in 16 verschillende landen. Een andere interessante mogelijkheid is het bepalen van de invloed van de recente prijzenoorlog op de inkomsten van de Nederlandse supermarkten, een gebeurtenis die uitvoerig aan bod kwam in de media. Deze structurele prijsverlaging van gemiddeld 11 procent op onze aankopen in de supermarkt startte in oktober 2003, en kan mogelijk een aanzienlijke verandering (breuk) veroorzaken in de evolutie van de huidige inkomsten (of marktaandeel) bij diverse Nederlandse distributeurs.
2. Om in de afwezigheid van een structurele verandering stabiele van evoluerende inkomstenreeksen van elkaar te kunnen onderscheiden wordt concreet gebruik gemaakt van volgende ‘unit root’ (eenheidswortel) test:  $Y_t = c + \rho Y_{t-1} + u_t$ , waarbij  $Y_t$  en  $Y_{t-1}$  respectievelijk de inkomsten van het gevestigde kanaal weergeven op tijdstip  $t$  en  $t-1$ ,  $c$  een constante is en  $u_t$  een fouteterm is. Merk op dat in dit model meestal diverse  $\Delta Y_t$  termen worden toegevoegd om ervoor te zorgen dat de fouteterm  $u_t$  witte ruis wordt. De inkomsten blijken stabiel wanneer de geschatte parameter  $\rho$  kleiner dan 1 is. Wanneer  $\rho$  gelijk aan 1 is, concluderen we dat de inkomsten evolueren overheen de tijd. Om beter te begrijpen wat het belang is van dit onderscheid, kunnen we een schok introduceren in het autoregressief model via de fouteterm  $u_t$ . Indien de tijdreeks *stabiel* is en dus  $\rho < 1$ , kunnen we het unit root model immers herschrijven als  $Y_t = [c/(1-\rho)] + u_t + \rho u_{t-1} + \rho^2 u_{t-2} + \dots$ . In dit model zal een schok  $u_t$  in de serie in elke periode in afnemende mate bijdragen tot een verandering in de gemiddelde waarde van de serie, en na verloop van tijd verdwijnt de impact van de initiële schok. De impact van een schok is dus tijdelijk. Wanneer een tijdreeks echter *evolueert* en  $\rho = 1$ , wordt het model na substitutie  $Y_t = [c + c + c + \dots] + u_{t-1} + u_{t-2} + \dots$  wat aangeeft dat een schok volledig in de daaropvolgende perioden aanwezig blijft. Iedere schok veroorzaakt dus een permanente of lange termijn verandering in de serie.
3. Deze reclame-inkomsten omvatten uitgaven voor merkgoederen en -diensten, maar *geen* personeelsadvertenties.
4. Hoewel we in deze studie mogelijke kannibalisatieverliezen onderzochten bij de “offline” (i.e. papieren) kranten, is het interessant om op te merken dat voor 10 Engelse dagbladen een positieve trendbreuk werd vastgesteld waarvoor de  $t$ -waarde  $> 1.64$  ( $p < .10$ ). Opvallend is daarbij dat de grootste positieve trendbreuk werd vastgesteld bij de *Financial Times*, de enige nationale krant in het Verenigd Koninkrijk die zich bijna exclusief focust op financieel en economisch nieuws. Mogelijk heeft het toevoegen van een Internetkanaal indirecte voordelen opgeleverd voor deze krant door bijvoorbeeld een positieve invloed op haar imago, waardoor de verkopen van de papieren editie zijn toegenomen.
5. Aangezien er 15 (5 negatieve en 10 positieve) kranten waren met een significante trendbreuk, werd voor slechts 52 dagbladen (van de 67 onderzochte) verder bestudeerd of er mogelijk een niveaudaling was opgetreden in de tijdreeks.
6. Voor één van de Nederlandse dagbladen in onze steekproef, namelijk het *NRC Handelsblad*, werd een significante *positieve* trendbreuk vastgesteld. Opnieuw blijkt dit het meest economisch georiënteerde dagblad te zijn.
7. Voorgaande resultaten werden op diverse manieren gevalideerd, maar onze uiteindelijke conclusies bleven onveranderd. Onder meer testten we de stabiliteit van de resultaten door de maandelijkse gegevens te aggregeren naar tweemaandelijkse gegevens en door na te gaan of kannibalisatie vaker voorkwam bij die kranten waar de Interneteditie vroeger werd geïntroduceerd.

## LITERATUUR

- Abbring, J.H., & J.C. Van Ours, 1994. Selling news and advertising space: The economics of Dutch newspapers. *De Economist* 142 (2), 151-170.



- Alba, J., J. Lynch, B. Weitz, C. Janiszewski, R. Lutz, A. Sawyer, & S. Wood, 1997. Interactive home shopping: Consumer, retailer, and manufacturer incentives to participate in electronic marketplaces. *Journal of Marketing* 61 (July), 38-53.
- Ben-David, D., & D.H. Papell, 1995. The great wars, the great crash, and steady state growth: Some new evidence about an old stylized fact. *Journal of Monetary Economics* 36, 453-475.
- Brynjolfsson, E., & M.D. Smith, 2000. Frictionless commerce? A comparison of Internet and conventional retailers. *Management Science* 46 (4), 563-585.
- Dekimpe, M.G., D.M. Hanssens, 1995. The persistence of marketing effects on sales. *Marketing Science* 14 (1), 1-21.
- Deleersnyder, B., I. Geyskens, K. Gielens, & M.G. Dekimpe, 2002. How cannibalistic is the Internet channel? A study of the newspaper industry in the United Kingdom and The Netherlands. *International Journal of Research in Marketing* 19, 337-348.
- Europese Commissie, 1996. *Strategic developments for the European publishing industry towards the year 2000*.
- Franses, P.H., 1998. *Time series models for business and economic forecasting*. Cambridge: Cambridge University Press.
- Frazier, G.L., 1999. Organizing and managing channels of distribution. *Journal of the Academy of Marketing Science* 27 (2), 226-240.
- Ghosh, S., 1998. Making business sense of the Internet. *Harvard Business Review* 76 (2), 126-135.
- Geyskens, I., K. Gielens, & M. G. Dekimpe, 2002. The Market Valuation of Internet Channel Additions. *Journal of Marketing* 66 (April), 102-119.
- Gilbert, C. (2001). *A dilemma in response: Beyond resource allocation. Examining the newspaper industry's response to the Internet*. Unpublished doctoral dissertation, Harvard Business School.
- Gilbert, A., & B. Bachelidor, 2000. The big squeeze. *Informationweek*, March 27, 47-56.
- Houtman, J., 2004. Online reclame groeit met 58 procent. *Emerce*, August 6.
- Mason, C.H., & G.R. Milne, 1994. An approach for identifying cannibalization within product line extensions and multi-brand strategies. *Journal of Business Research* 31, 163-170.
- Nijs, V.R., M.G. Dekimpe, J-B.E.M. Steenkamp, & D.M. Hanssens, 2001. The category demand effects of price promotions. *Marketing Science* 20 (1), 1-22.
- Perron, P., 1989. The great crash, the oil price shock, and the unit root hypothesis. *Econometrica* 57 (6), 1361-1401.
- Perron, P., 1994. Trend, unit root and structural change in macroeconomic time series. In B. Bhaskara Rao (ed.), *Cointegration for the applied economist*. New York, NY: St. Martin's Press, 113-146.
- Porter, M.E., 2001. Strategy and the Internet. *Harvard Business Review* 79 (3), 62-79.
- Reddy, S.K., S.L. Holak, & S. Bhat, 1994. To extend or not to extend: Success determinants of line extensions. *Journal of Marketing Research* 31 (May), 243-262.
- Rosenthal, R., 1991. *Meta-analytic procedures for social research*. Newbury Park, Sage.
- Shapiro, C., & H.R. Varian, 1999. *Information rules*. Boston: Harvard Business School Press.
- Silk, A.J., L.R. Klein, & E.R. Berndt, 1999. Restructuring in the U.S. advertising media industry. *Working paper 99-126, Graduate School of Business Administration, Harvard University*.
- Silk, A.J., L.R. Klein, & E.R. Berndt, 2001. The emerging position of the Internet as an advertising medium. *Netnomics* 3 (2), 129-148.
- Srinivasan, S., P.T.L. Popkowski Leszczyc, & F.M. Bass, 2000. Market share response and competitive interaction: The impact of temporary, evolving and structural changes in prices. *International Journal of Research in Marketing* 17 (4), 281-306.
- Steenkamp, J-B.E.M., V. Nijs, D.M. Hanssens, & M.G. Dekimpe, 2004. Competitive Reactions to Advertising and Promotion Attacks. *Marketing Science*, in druk.
- Stellinga, M., 2004. Verdienen aan het Internet; Business: de 2de Internetgolf. *Elsevier*, July 3.
- Wiersma, T., 2004. KPMG: Uitegevers blijven worstelen met Internet. *Emerce*, May 20.

## *6. Effecten van promoties op het huishoudelijke aankoopgedrag: analyses op basis van een consumenten panel*

A.L.J.M. LUIJTEN en M.M.L.R. NAGTZAAM

### SAMENVATTING

De laatste decennia is er een belangrijke verschuiving opgetreden in de manier waarop fabrikanten van merkartikelen hun communicatiebudget besteden: meer op retailers gericht ter ondersteuning van promotionele acties. Zowel merkentrouw als winkeltrouw staan onder druk zoals blijkt uit toegenomen merk- en winkelduplicatie. Ook in consumentenonderzoek wordt het toegenomen belang van het aspect prijs bij winkel- en merkkeuze onderstreept. In onderzoek is er derhalve veel aandacht voor effecten en effectiviteit van sales promotions. De meeste modellen zijn ontwikkeld voor het gebruiken van geaggregeerde data afkomstig uit retail panels. In dit artikel wordt daarentegen ingegaan op de effecten van promoties op het huishoudelijk aankoopgedrag. Vier mogelijk beoogde effecten worden onderscheiden: in hoeverre worden huishoudens door promoties aangezet tot veranderen van merk, veranderen van winkel, meer kopen dan gebruikelijk dan wel aankopen naar voren trekken. Daarnaast voorziet het model in een segmentatie van de kopers in de categorie op basis van intensiteit en feitelijke promotiegeneigdheid. Bij het toepassen van het model kan de insteek gekozen worden vanuit het perspectief van zowel de fabrikant van merkartikelen als, op categorieniveau, de retailer. Dit wordt geïllustreerd in de gepresenteerde case.

### 1. INLEIDING EN DOELSTELLING

In de marketing literatuur heeft het concept van de marketing mix en de daartoe behorende instrumenten, de 4 P's, veel aandacht gekregen. Het verwijst naar de vier klassen van instrumenten Product, Plaats (distributie), Prijs en Promotie.

Binnen dit artikel wordt de aandacht gericht op een deel van de instrumenten die behoren tot de klasse Promotie. Promoties ondersteunen in het algemeen een of meer van de volgende doelstellingen:

- Kennis bij de ontvanger over het product of dienst te laten toenemen (cognition, awareness).
- Het bijdragen aan het versterken of veranderen van de houding, attitude ten opzichte van het product of dienst (affection, conviction, liking, preference).
- Het bevestigen, versterken of veranderen van het gedrag met betrekking tot het product of dienst (conation, purchase).

Leeflang (1981) deelt in zijn leerboek de instrumenten met betrekking tot de promotie van goederen en diensten op in twee sub-klassen:

- De communicatie-mix (mass selling door advertising en merchandising, free publicity en public relations).
- De actie-mix (tijdelijke toename in “waar voor je geld” door sales promoties).

Binnen het ConsumerScan panel van GfK wordt het huishoudelijk winkel- en aankoopgedrag voor met name “fast moving consumer goods” geregistreerd. Door koppeling van gedetailleerde aankoopgedragsgegevens met bijvoorbeeld socio-demografische en psychografische achtergrondkenmerken kunnen markten en marktontwikkelingen goed beschreven en geanalyseerd worden. Uitgangspunt daarbij is telkens de centraal gestelde consument. Door op basis van externe bronnen vergaarde promotie-informatie achteraf te koppelen aan de ConsumerScan databases kan het effect bestudeerd worden van de op de consument gerichte sales promotie activiteiten die behoren tot de klasse van de actie-mix.

## 2. SALES PROMOTIES

Blattberg en Neslin (1990) omschrijven “sales promotions” als “action-focused marketing events whose purpose is to have a direct impact on the behaviour of the firm’s customers”. Teunter (2002) onderscheidt, in navolging van Blattberg en Neslin, drie belangrijke typen van verkooppromoties:

- Consumenten promoties door de fabrikanten direct aangeboden aan de consument
- Retailer promoties door de wederverkopers/detaillisten aangeboden aan de consument
- Trade promoties welke door de fabrikant aangeboden worden aan de wederverkopers/detaillisten

Wanneer het effect van de verkooppromoties op het aankoopgedrag van huishoudens bestudeerd wordt, vallen derhalve de trade promoties buiten beschouwing. De verkooppromoties op huishoudens gericht bestaan voornamelijk uit geadverteerde prijspromoties die eenvoudig vanuit folders en advertenties getraceerd kunnen worden. Met andere woorden het onderzoeksterrein wordt beperkt tot de effecten van op de consument gerichte prijs georiënteerde promoties op het aankoopgedrag. De doelstelling achter prijs georiënteerde promoties is veelal om de verkopen en marktaandeel op korte termijn te verhogen, terwijl bij niet-prijs georiënteerde promoties (premiums, sweepstakes) het doel luidt om de binding en trouw te bevorderen. (Transactie versus Relatie)

## 3. BEOOGDE EFFECTEN

In het onderzoek door Nijs, Dekimpe, Steenkamp en Hanssens (2001) wordt gesteld dat bij 95% van de verkooppromoties in Nederlandse supermarkten geen effect op lange termijn kan worden aangetoond. Er zijn echter wel duidelijke korte termijn effecten aantoonbaar. In de academische literatuur is in zekere mate overeenstemming over de volgende mogelijke effecten van verkooppromoties op het aankoopgedrag van consumenten:

- *Brand switching, merkenwissel*; een consument wordt door de promotie aangezet een ander merk te kopen dan gekocht zou zijn indien de promotie niet beschikbaar was.
- *Store switching, wisselen van aankoopplaats*; dit reactiemechanisme betekent dat de consument door de promotie aangezet wordt het product op een andere plaats te kopen dan gebruikelijk voor haar/hem is.

- *Stockpiling, voorraad vorming*; dit kan optreden als gevolg van twee verschillende mechanismen. Allereerst kan het zijn dat de consument de aankoopgebeurtenis naar voren trekt in de tijd (significante reductie van de tussenaankooptijd) terwijl het ook kan zijn dat er tijdelijk een significant groter volume tijdens promotie wordt gekocht. Beide mechanismen kunnen ook gelijktijdig optreden.
- *Increased consumption, category expansion, toegenomen consumptie*: hiermee wordt bedoeld op een permanente verhoging van de consumptie in de categorie door een consument ten gevolge van een promotie.

Het is, als gevolg van de variatie in het aankoopgedrag met betrekking tot zowel de gekochte hoeveelheid per aankoopgebeurtenis als ook de tijd tussen twee opeenvolgende aankoopgebeurtenissen, moeilijk om een onderscheid te maken tussen de twee laatstgenoemde effecten, voorraad vorming en toegenomen consumptie.

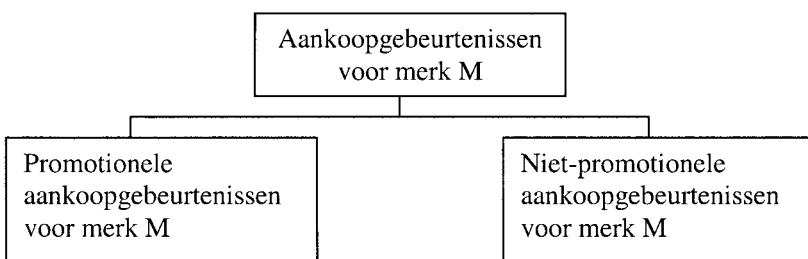
Teunter (2002) vond in slechts één van de zes door haar bestudeerde product categorieën een significant “category expansion” effect. Verder geldt dat, indien “category expansion” beschouwd mag worden als een lange termijn effect van promoties en de bewering van Nijs c.s. (2001) juist is dat 95% van de in Nederlandse supermarkten gehouden promoties geen lange termijn effect hebben, er geen onderscheid gemaakt hoeft te worden tussen de laatstgenoemde twee effecten. In het hier te beschouwen model worden de twee effecten samengevoegd onder de naam *categorie effect*, waarbinnen een onderscheid gemaakt kan worden tussen een *timing effect* en een *volume effect*, hierbij verwijzend naar de onder voorraadvorming genoemde twee mechanismen. Wanneer in dit artikel effecten geschat gaan worden, wordt afgezien van andere in de literatuur genoemde effecten zoals bijvoorbeeld categorie substitutie, complementaire effecten en effecten op de winkeldrukke (store traffic). Met categorie substitutie wordt bedoeld op het kopen van een product in promotie uit product categorie A in plaats van het kopen van een product uit product categorie B. Complementaire effecten treden op indien een promotie in categorie A de verkopen in categorie B stimuleren. Enige tijd geleden werd in een artikel, dat onder meer store traffic behandelde, aangetoond dat in de tijd dat het consumenten vertrouwen sterk daalde de gepercipieerde prijsgevoeligheid bij het selecteren van winkels in belang toenam. Het belang van service aspecten nam daarentegen af. Store traffic geoperationaliseerd door winkelduplicatie vertoonde ook een significante toename; zie Luijten and Nagtzaam (2003).

Bij het bepalen van de effecten van de op de consument gerichte prijspromoties zullen we ons richten op het aankoopgedrag van huishoudens binnen categorieën van “fast moving consumer goods”. Om maatstaven te kunnen ontwikkelen voor de vier te onderscheiden effecten (brand switching, store switching, category timing en category volume) zullen we de individuele aankoopgedrag historie voor panelleden gedurende een periode ter lengte van 52 weken bestuderen. Juist door gebruik te maken van data uit een panel van huishoudens over een langere periode zijn we in staat om het “switching”-gedrag in beeld te brengen.

#### 4. PROMOTIEGENEIGDHEID VAN HUISHOUDENS

Alvorens operationele definities van de maatstaven te geven, zal eerst het concept van de *aankoopgebeurtenis* behandeld worden. In tegenstelling tot Teunter (2002) wordt hier gesproken over aankoopgebeurtenissen (purchase occasions) in plaats van over winkelbezoeken (shopping trips) voor een categorie. Een winkelbezoek voor een cate-

gorie treedt op indien een huishouden gedurende een dag in een specifieke winkel binnen de categorie gekocht heeft. Maar aangezien we geïnteresseerd zijn in effecten van promoties voor merken (tijd en plaats afhankelijk) op het aankoopgedrag zullen we expliciet in beschouwing moeten nemen dat gedurende een winkelbezoek meer dan één merk binnen de categorie gekocht kan zijn. Een aankoopgebeurtenis (purchase occasion) voor een merk treedt op indien een huishouden tenminste één item van dat merk gedurende een dag bij een specifieke winkel gekocht heeft. We spreken over *promotionele activiteit* met betrekking tot een merk in een winkel indien tenminste één item van dat merk onder promotionele condities beschikbaar is. De aankoopgebeurtenissen voor een merk kunnen derhalve opgesplitst worden in *promotionele aankoopgebeurtenissen* en *niet-promotionele aankoopgebeurtenissen*. Een aankoopgebeurtenis voor een merk wordt promotioneel genoemd indien er promotionele activiteit is voor dat merk in die winkel. In andere gevallen spreken we over een niet-promotionele aankoopgebeurtenis.



Figuur 1. Aankoopgebeurtenissen voor merk M gedefinieerd.

Het op deze manier “vlaggen” van aankoopgebeurtenissen om aan te geven of al dan niet gebruik is gemaakt van aanbiedingen gebeurt achteraf in de ConsumerScan database op basis van gegevens afkomstig uit externe bronnen. Op deze wijze wordt de werkdruk voor panelleden niet opgevoerd zodat de uitbreiding in het meetmodel geen nadelig effect kan sorteren op het registratiegedrag en panelverloop.

Gebaseerd op bovenstaande uitsplitsing kan de aankoopshistorie van een huishouden in een jaar samengevat worden door de aankoopgebeurtenissen over de merken op te tellen tot:

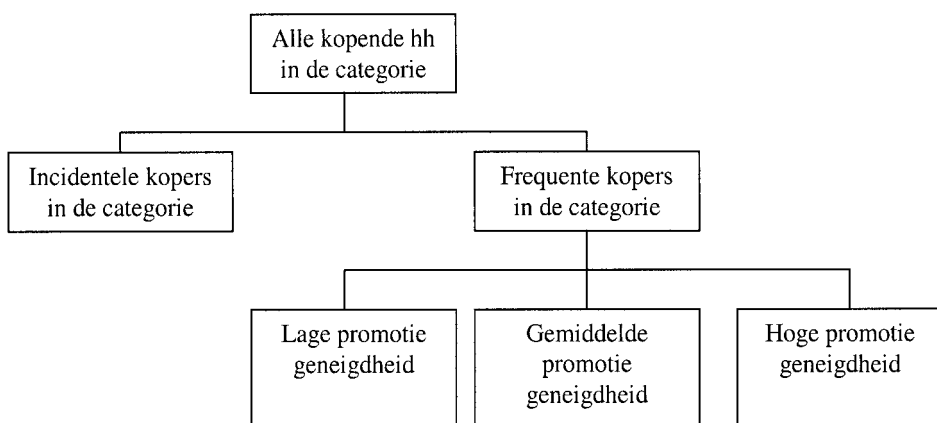
- aantal aankoopgebeurtenissen in de categorie.
- aantal aankoopgebeurtenissen onder promotionele condities in de categorie.
- aantal aankoopgebeurtenissen onder niet-promotionele condities in de categorie.

Opgemerkt wordt dat het aantal aankoopgebeurtenissen in de categorie zoals nu gedefinieerd groter of gelijk is aan het aantal winkelbezoeken voor de categorie. Met deze nieuwe variabelen kunnen huishoudens geclassificeerd op basis van de mate van feitelijke promotiegeneigdheid (deal proneness) binnen de categorie.

Allereerst kan de groep van kopende huishoudens binnen de categorie opgesplitst worden in twee deelgroepen gebaseerd op het aantal aankoopgebeurtenissen van de panelleden in de categorie in een jaar. De groep met weinig aankoopgebeurtenissen wordt aangeduid met *incidentele kopers*, terwijl de rest aangeduid wordt met *frequente*

*kopers*. De frequente kopers in de categorie kunnen verder opgesplitst worden in drie deelgroepen op basis van het aandeel van de aankoopgebeurtenissen onder promotionele condities binnen de categorie. Het segment met het laagste aandeel zal aangeduid worden als kopers met een lage promotiegeneigdheid. Voorts worden onderscheiden kopers binnen de categorie met een gemiddelde promotiegeneigdheid en tot slot de groep met een hoge promotiegeneigdheid.

De exacte grenzen voor het opdelen van de categoriekopers in de vier deelgroepen kunnen bepaald worden op basis van controle van de relevante verdelingen. Dit is mede de reden dat uitgegaan wordt van data over het aankoopgedrag van panelleden over een periode ter lengte van een jaar om op deze wijze de garantie te hebben van voldoende informatiedichtheid. De voorgestelde segmentatie wordt onderstaand gevisualiseerd:



*Figuur 2. Classificatie van kopers in de categorie.*

Voor marketing managers is het belangrijk het profiel van de onderscheiden kopersgroepen te kennen in termen van socio-demografie, psychografie, koopgedrag en mediagedrag.

## 5. HET VASTSTELLEN VAN DE BEOOGDE EFFECTEN

Voor ieder van de promotionele aankoopgebeurtenissen voor de onderscheiden merken zal bepaald moeten worden welke van de beoogde effecten concreet gerealiseerd zijn. De niet-promotionele aankoopgebeurtenissen worden onveranderd gelaten.

Voor iedere promotionele aankoopgebeurtenis wordt gedefinieerd of er sprake is van brand switching (B), store switching (S), category timing (T) en category volume effecten (V). Deze effecten kunnen al dan niet gelijktijdig optreden. Voor ieder van de beoogde effecten wordt derhalve een dummy variabele gedefinieerd. Deze variabelen nemen de waarde 1 aan indien het beoogde effect kan worden vastgesteld en de waarde 0 indien dit niet het geval is. Dit mondt uit in 16 verschillende promotionele response patronen, waaronder 4 pure effecten, 11 gecombineerde effecten en 1 response patroon waarbij geen enkel effect optreedt. In dit laatste geval treedt derhalve normaal aankoopgedrag op onder promotionele condities. De 16 mogelijke combinaties zijn onderstaand weergegeven.

Tabel 1. Overzicht van promotionele effecten.

B	S	V	T	#effecten	soort
0	0	0	0	0	geen
1	0	0	0	1	puur
0	1	0	0	1	puur
0	0	1	0	1	puur
0	0	0	1	1	puur
1	1	0	0	2	gecombineerd
1	0	1	0	2	gecombineerd
1	0	0	1	2	gecombineerd
0	1	1	0	2	gecombineerd
0	1	0	1	2	gecombineerd
0	0	1	1	2	gecombineerd
1	1	1	0	3	gecombineerd
1	1	0	1	3	gecombineerd
1	0	1	1	3	gecombineerd
0	1	1	1	3	gecombineerd
1	1	1	1	4	gecombineerd

Om de promotionele aankoopgebeurtenissen volgens bovenstaand schema te kunnen classificeren dient normaal aankoopgedrag met betrekking tot de beoogde effecten te worden gedefinieerd. Het normale aankoopgedrag zal worden gedefinieerd op basis van alle aankoopgebeurtenissen op het niveau van ieder afzonderlijk huishouden. Promotie effecten worden vastgesteld indien een huishouden tijdens een promotionele aankoopgebeurtenis afwijkt van het normale aankoopgedrag. Deze aanpak sluit aan bij die welke beschreven is door Wilborts (2002).

#### *Brand switching (B)*

In het hier te beschouwen model wordt over brand switching gesproken indien het merk gekocht tijdens de promotionele aankoopgebeurtenis verschillend is van het door het huishouden meest geprefereerde merk. Het meest geprefereerde merk is binnen het hier gepresenteerde model gedefinieerd als het merk met het hoogste volume aandeel gerekend over alle aankoopgebeurtenissen binnen het beschouwde jaar. In wetenschappelijke publicaties wordt het favoriete merk soms gedefinieerd als het merk met het hoogste exponentieel gewogen gemiddelde van de in het verleden gedane aankopen. (o.a. Teunter (2002)) Teunter gebruikt voorts alleen de niet-promotionele aankoopgebeurtenissen om het favoriete merk te bepalen.

#### *Store switching (S)*

Er is sprake van store switching indien gedurende een promotionele aankoopgebeurtenis het product onder promotie gekocht is bij een andere retailer/winkel(formule) dan de meest favoriete retailer/winkel(formule). De meest geprefereerde winkel wordt gedefinieerd als de winkel met het hoogste volume aandeel voor de categorie over alle aankoopgebeurtenissen in het jaar.

#### *Categorie effecten (V, T)*

De in de literatuur aangetroffen effecten voorraadvorming/stockpiling en categorie uitbreiding worden hier samen genomen onder de term categorie effecten. Voorraadvor-

ming wordt echter beschouwd als veruit het belangrijkste categorie effect. Derhalve wordt binnen het hier te presenteren model verondersteld dat het categorie effect kan optreden via twee verschillende mechanismen. De eerste mogelijkheid luidt dat door een promotie een consument een aankoopgebeurtenis naar voren trekt in de tijd (Timing effect), terwijl de tweede mogelijkheid is dat ten gevolge van de promotie een veel grotere hoeveelheid dan normaal wordt aangeschaft (Volume effect). Aangezien er veel variatie bestaat in individuele tussenaankooptijden en de gekochte hoeveelheid per aankoopgebeurtenis, ook in het geval dat er geen promoties aanwezig zouden zijn, dient met de bestaande spreiding in gedragskenmerken rekening gehouden te worden. De variabiliteit met betrekking tot de tijd tussen twee successieve aankoopgebeurtenissen werd bijvoorbeeld door Ehrenberg (1972) gemodelleerd in zijn theorie over herhalingsaankoopgedrag door uit te gaan van exponentieel verdeeld veronderstelde tussenaankoop tijden.

Een volume effect (V) tijdens een promotionele aankoopgebeurtenis wordt verondersteld aanwezig te zijn indien het volume gekocht tijdens de promotionele aankoopgebeurtenis significant groter is dan de gemiddeld gekochte hoeveelheid over alle aankoopgebeurtenissen in een jaar.

Een timing effect (T) wordt verondersteld aanwezig te zijn indien de tussenaankooptijd voorafgaand aan een winkelbezoek, waarbij een promotionele aankoopgebeurtenis plaatsvindt, significant korter is dan de gemiddelde tussenaankooptijd.

Significantie voor het volume effect wordt hier gedefinieerd als het overschrijden van het gemiddeld gekocht volume plus één keer de standaarddeviatie van het gekochte volume op basis van alle aankoopgebeurtenissen in een jaar. De voorafgaande tussenaankooptijd voor een winkelbezoek voor de categorie, waarbij tenminste één merk in promotie gekocht werd, wordt als significant korter beschouwd indien deze tijd kleiner is dan de gemiddelde tussenaankooptijd minus één keer de standaarddeviatie op basis van alle beschikbare winkelbezoeken binnen de categorie in het beschouwde jaar. De heterogeniteit in de populatie met betrekking tot het aankoopgedrag maakt het nodig de benodigde parameters voor deze effecten op individueel niveau te schatten.

## 6. PROMOTIEGENEIGDHEID EN BEOOGDE EFFECTEN: BEPERKINGEN EN OPERATIONALISERING

In de hierboven weergegeven definities van meest geprefereerd merk binnen de categorie, meest belangrijke winkel(formule) voor de categorie en gemiddeld volume en standaarddeviatie van het gekochte volume, wordt uitgegaan van alle aankoopgebeurtenissen binnen de categorie voor een panellid in een jaar. Deze maatstaven weerspiegelen binnen dit model “normaal” aankoopgedrag aangezien promoties binnen productvelden permanent beschikbaar zijn. Door voor promotionele aankoopgebeurtenissen de kenmerken merk/brand, winkel/store en volume te vergelijken met hetgeen normaal verwacht mag worden bij het panellid kunnen de effecten B, S en V bepaald worden. Voor het berekenen van de maatstaven die samenhangen met de tussenaankooptijd zijn beperkingen aanwezig vanwege de mogelijkheid van het kopen van meerdere merken binnen de categorie tijdens één winkelbezoek voor de categorie. Derhalve is uitgegaan van de tussenaankooptijden van opeenvolgende winkelbezoeken in de categorie. Door de periodelengte van de tijd voorafgaand aan een winkelbezoek in de categorie met tenminste één promotionele aankoopgebeurtenis te vergelijken met de



afgeleide statistische maatstaven is het mogelijk het timing effect T te bepalen. Echter, in het geval van een promotionele aankoopgebeurtenis tijdens het éérste winkelbezoek voor de categorie kan het timing effect niet worden vastgesteld vanwege het ontbreken van de informatie voor het vaststellen van de aan deze gebeurtenis voorafgaande tussenaankooptijd. In verband met de heterogeniteit in het aankoopgedrag worden de beslissingsregels op individueel niveau toegepast, zowel voor het classificeren van de kopers binnen de categorie als het bepalen van de aanwezigheid van promotionele effecten binnen de promotionele aankoopgebeurtenissen.

In verband met de classificatie van de kopers in de categorie berekenen we per koper over het jaar:

- aantal aankoopgebeurtenissen in de categorie.
- aantal aankoopgebeurtenissen onder promotionele condities in de categorie.
- het aandeel van de aankoopgebeurtenissen onder promotionele condities binnen de categorie.

Zoals eerder aangegeven zijn op basis van inspectie van de eerste en derde verdeling de afkappgrenzen voor de uiteindelijke classificatie in vier deelgroepen bepaald. (zie ook figuur 2) De eerste opsplitsing in incidentele en frequente kopers is gebaseerd op het totaal aantal aankoopgebeurtenissen in de categorie. De verdere opsplitsing van de frequente kopers naar de mate van promotiegeneigdheid in drie deelgroepen gebeurt op basis van het aandeel van de aankoopgebeurtenissen onder promotionele condities.

Aangezien er voor de incidentele kopers in de categorie weinig aankoopdata beschikbaar zijn is het voor huishoudens behorende tot deze categorie niet mogelijk om het normale aankoopgedrag te bepalen. De door hen gedane aankopen onder promotionele condities worden allemaal geclassificeerd als promotionele aankopen zonder enig beoogd effect.

Voor de frequente kopers in de categorie zijn de noodzakelijke data aanwezig om normaal aankoopgedrag te kunnen definiëren, derhalve kunnen al de promotionele aankoopgebeurtenissen van deze kopersgroepen aan één van de zestien mogelijke combinatie van vier beoogde effecten worden toegekend.

Een algemene slotopmerking kan worden gemaakt: het beschreven model is uitsluitend toepasbaar indien er sprake is van voldoende beschikbare informatie. Voor laagfrequent gekochte categorieën is het wellicht zinvol om aankoopgebeurtenissen te vlaggen in termen van wel of niet onder promotionele condities gekocht, doch ten gevolge van de lage aankoopfrequentie zijn er te weinig gegevens voor handen om de parameters van het normale aankoopgedrag te kunnen schatten.

## 7. EFFECTEN VAN SALES PROMOTIES: EEN CASE STUDIE

In deze paragraaf worden resultaten getoond van een uitgevoerde studie naar de promotie effecten in een markt voor verzorgingsproducten. Zowel de product categorie (markt) als de daarbinnen opererende merken zullen niet herkenbaar worden weergegeven. Volstaan zal worden met een globale schets van deze markt in de periode 2002-Q4 t/m 2003-Q3. Binnen het beschouwde jaar kocht 96 % van de Nederlandse huishoudens tenminste één keer in de categorie. De gemiddelde aankoopfrequentie bedraagt 10.34, zodat globaal gesteld kan worden dat gemiddeld eens per 5 weken in de categorie gekocht wordt. Echter er is sprake van heterogeniteit in het aankoopgedrag. Gemiddeld besteedt een kopend huishouden € 26.60 in het beschouwde jaar in deze markt,

dus gemiddeld € 2.57 per aankoopgebeurtenis. Binnen deze markt nemen de drie grootste A-merken ruim 40% van de markt op volumebasis voor hun rekening. De vier grootste private labels zijn in deze markt goed voor meer dan 25% volume aandeel.

Bij het bepalen van de effecten van sales promoties op het aankoopgedrag kunnen we ons in eerste instantie beperken tot informatie op basis van het uitsluitend vlaggen van de aankoopgebeurtenissen. Binnen deze meest algemene analyse wordt derhalve nog geen gebruik gemaakt van de kopersclassificatie en de effectbepaling. De analyse kan uitgevoerd worden vanuit de doelstelling om merken met elkaar te vergelijken dan wel retailers op categorieniveau met elkaar te vergelijken. Bij het vergelijken van de merken zullen deze aangeduid worden met NB voor de A-merken/National Brands en met PL voor de huismerken/Private Labels. Een eerste indruk van het effect van de gevoerde promoties op de verkopen van merken wordt gegeven door aan te geven welk deel van de verkopen in termen van aankoopgebeurtenissen, volume en bestedingen gedaan wordt onder promotionele condities.

Tabel 2. Het belang van promoties voor de afzet van merken.

<b>Verkopen onder promotionele condities</b>			
<b>Merk</b>	<b>% van totaal</b>		
	<b>Aankoopgeb</b>	<b>Vol</b>	<b>Best</b>
<b>NB1</b>	41.4%	55.8%	49.7%
<b>NB2</b>	44.3%	56.0%	52.4%
<b>NB3</b>	43.5%	55.5%	48.7%
<b>PL1</b>	5.0%	15.2%	11.1%
<b>PL2</b>	12.2%	25.0%	17.9%
<b>PL3</b>	12.5%	16.3%	17.3%
<b>Totaal</b>	<b>15.3%</b>	<b>27.6%</b>	<b>27.4%</b>

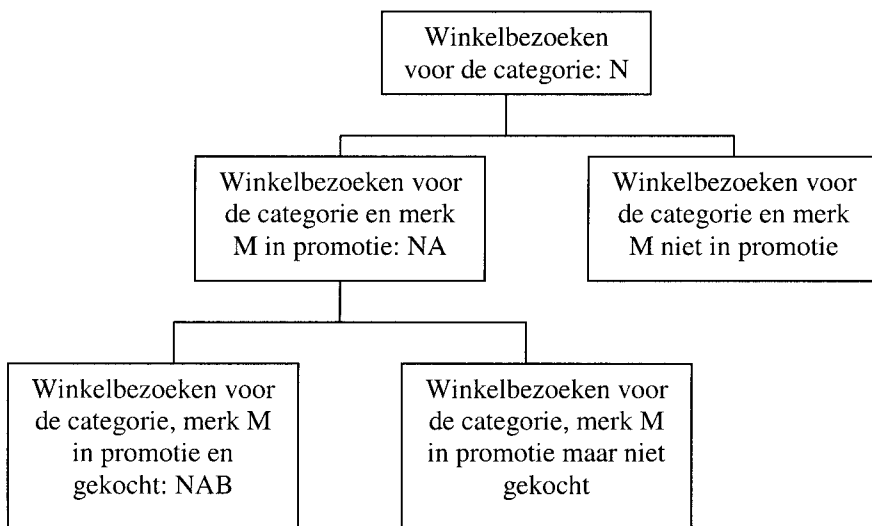
De drie beschouwde A-merken zijn voor hun afzet in zeer sterke mate afhankelijk van promoties. Ruim 55% van het merkvolume wordt gerealiseerd onder promotionele condities. Een eerdere studie binnen deze categorie liet een volume effect van ruim 50% zien, zodat geconstateerd kan worden dat zeer waarschijnlijk de promotiedruk in deze markt verder is opgevoerd.

De afhankelijkheid van retailorganisaties van promoties in de product categorie wordt getoond door een soortgelijk overzicht op te stellen:

Tabel 3. Het belang van promoties voor de categorie afzet van retailers.

<b>Verkopen onder promotionele condities</b>			
<b>Merk</b>	<b>% van totaal</b>		
	<b>Aankoopgeb</b>	<b>Vol</b>	<b>Best</b>
<b>R1</b>	30.9%	53.6%	48.0%
<b>R2</b>	21.2%	36.5%	32.5%
<b>R3</b>	23.8%	41.6%	40.0%
<b>R4</b>	23.2%	39.7%	36.7%
<b>Totaal</b>	<b>15.3%</b>	<b>27.6%</b>	<b>27.4%</b>

Op basis van de twee overzichten kan worden geconcludeerd dat de beschouwde retailers in belangrijke mate afhankelijk zijn van de promoties van de A-merken, aangezien deze merken een aanmerkelijk groter promotie-aandeel kennen dan de private labels. Om promotionele druk en promotiegebruik voor merken te kunnen bepalen introduceren we het volgende overzicht:



*Figuur 3. Classificatie van winkelbezoeken binnen de categorie.*

Promotionele druk van een merk kan gedefinieerd worden als dat deel van de afgelegde winkelbezoeken in de categorie waarbij het merk onder promotionele condities beschikbaar is. De effectiviteit van de gevoerde promoties zou gemeten kunnen worden aan de hand van het promotionele gebruik. Vanwege de algemene verkrijgbaarheid van A-merken en de beperkte beschikbaarheid van PL's is het vergelijken van de promotionele druk alleen zinvol voor A-merken.

Verwijzend naar figuur 3 worden de grootheden als volgt gedefinieerd:

$$\text{Promotionele druk in \%} = \{NA / N\} * 100$$

$$\text{Promotioneel gebruik in \%} = \{NAB / NA\} * 100$$

*Tabel 4. Promotionele maatstaven voor merken.*

Merk	Promotie maatstaven	
	Druk	Gebruik
NB1	12.9	44.6
NB2	9.5	39.1
NB3	4.6	50.3
PL1		43.1
PL2		43.6
PL3		33.8

Uit bovenstaand overzicht valt af te leiden dat bij ongeveer 27% van de winkelbezoeken in de categorie er tenminste één van de top drie A-merken in promotie beschikbaar is. In termen van effectiviteit doen de promoties voor PL's nauwelijks onder voor die van de NB's.

De promotiedruk kan op een andere, eenvoudigere manier worden weergegeven door het opstellen van een promotiekalender. Dit is een driedimensionale kalender met als assen:

- Merken
- Afzetkanalen
- Tijd gemeten in weken.

Het vlaggen van aankoopgebeurtenissen vindt plaats op basis van de promotiekalender. Hierin staat aangegeven wanneer en waar een merk onder promotionele condities beschikbaar is. Voor illustratiedoeleinden worden uit de driedimensionale kubus tweedimensionale selecties gemaakt.

Tabel 5. Een selectie van gegevens uit de promotiekalender.

Alle Afzetkanalen	Week												
Merk	1	2	3	4	5	6	7	8	9	10	11	12	13
NB1	2	8	2	3		3	3	6	2	4	3	2	1
NB2	1	1	1	1	5	2		1	1	2	3		5
NB3	1	1	1	3		4	1	2	2	1	1	1	3

R1	Week			
Merk	2	4	7	10
NB1	1			1
NB2		1		
NB3			1	

Alle Merken	Week												
Afzetkanaal	1	2	3	4	5	6	7	8	9	10	11	12	13
R1		2		1			1	1	1	2		1	
R2		1							1				
R3	1	2			1	2		1			1	1	2
R4	1		1		1	1	2	1	1	1	1	1	1

NB1	Week				
Afzetkanaal	2	7	8	9	10
R1	1				1
R2	1			1	
R3	1		1		
R4		1			

In het bovenste deel van de tabel, waarbij geaggregeerd is over alle afzetkanalen, zien we het aantal winkelformules weergegeven dat in specifieke weken promotie voert voor de onderscheiden A-merken. In de tweede deeltabel is voor het afzetkanaal R1 aangegeven in welke weken en voor welk merk promotie gevoerd is. We hebben ons daarbij beperkt tot de data over de eerste dertien weken van de beschouwde periode. In het onderste deel van de tabel zijn de dimensies afzetkanaal en merk omgewisseld. De promotiedruk zoals gemeten binnen tabel 4 is fijnmaziger aangezien daar rekening ge-

houden wordt met verschillen in kwaliteit van de distributie. Het verschil tussen de informatie uit tabel 5 en die in tabel 4 zou aangeduid kunnen worden met *numerieke promotiedruk* versus *gewogen promotiedruk* (op basis van aankoopgebeurtenissen).

Tot nu toe is uitsluitend gebruik gemaakt van de externe data met betrekking tot promoties (vlaggen) en zijn de kernelementen van het gepresenteerde model buiten beschouwing gebleven. Deze elementen van het model zijn:

- het classificeren van de kopers binnen de categorie
- het toekennen van beoogde effecten aan promotionele aankoopgebeurtenissen.

De classificatie en verdere typering van de kopers wordt onderstaand gepresenteerd.

Tabel 6. *Kopersgroepen voor de categorie.*

Kopers categorie		% van categorie totaal			Koopintensiteit <sup>1</sup>		Prijs Index <sup>2</sup>
Frequentie	Promotiegeneigdheid	Kopers	Vol	Best	Vol	Best	
Incidenteel	-	27.1%	10.5%	10.6%	38.6%	39.0%	100.8%
Frequent	Laag	45.7%	53.1%	51.1%	116.2%	111.9%	96.3%
Frequent	Gemiddeld	21.2%	27.7%	29.2%	130.6%	137.7%	105.5%
Frequent	Hoog	6.0%	8.7%	9.1%	145.9%	152.2%	104.4%
<b>Totaal</b>		<b>100.0%</b>	<b>100.0%</b>	<b>100.0%</b>	<b>100.0%</b>	<b>100.0%</b>	<b>100.0%</b>

1. De koopintensiteit meet het geïndexeerde gemiddelde verbruik in een koperscategorie ten opzichte van het gemiddelde verbruik over alle kopers in de categorie. Deze grootte is te bepalen op zowel volume- als bestedingenbasis.
2. De prijsindex geeft aan of de onderscheiden koperscategorieën gemiddeld meer of minder betalen per volume-eenheid dan gemiddeld over alle kopers in de categorie. De grootte kan bepaald worden door 100 keer het bestedingsaandeel te delen door het bijbehorende volume-aandeel.

De kopersgroepen worden gedefinieerd op basis van het aantal aankoopgebeurtenissen in de categorie in een jaar (frequentie) en het belang daarbinnen van de aankoopgebeurtenissen onder promotionele condities (promotiegeneigdheid). Bovenstaand zijn onder meer de relatieve verdelingen over de vier klassen weergegeven voor zowel de dimensies kopers, volume en bestedingen. Hieruit wordt afgeleid dat met het groter worden van de frequentie en promotiegeneigdheid de koopintensiteit/gebruiksintensiteit op huishoudensbasis toeneemt. Met andere woorden grote kopers maken relatief vaker gebruik van promoties. We zien voorts in de tabel dat de kopers met de lage promotiegeneigdheid gemiddeld een lagere prijs per volume eenheid neertellen, hetgeen veroorzaakt wordt door een relatief hogere voorkeur voor standaard lager geprijsde PL's.

Het is algemeen bekend dat in markten van fast moving consumer goods de meeste kopers in de categorie meer dan één merk kopen en ook meerdere retailers daarbij bezoeken. Op het moment dat men profielen van merkkopers, dan wel profielen van categorie kopers bij retailers onderling met elkaar vergelijkt, dient men rekening te houden met het feit dat er duplicatie aanwezig is. Op het moment dat in deze markt de merkkopers getypeerd worden op basis van het aankoopgedrag van de merkkopers in de categorie vinden we een bevestiging van een eerdere conclusie dat de PL's in deze markt een relatief sterkere positie hebben bij de categoriekopers met een lage promotiegeneigdheid.

Voor de kopers van NB1 is in tabel 7 een schets van het koopgedrag weergegeven

Tabel 7. Aankoopgedrag van de kopers van NB1 (National Brand 1) beschreven.

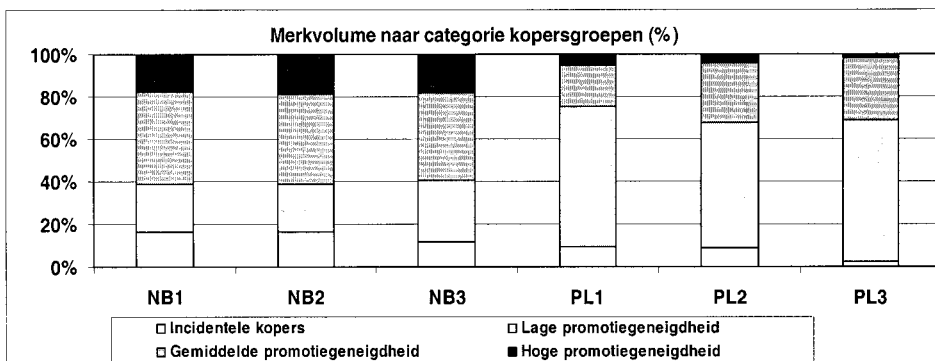
Kopers categorie		Kopers NB1 kolom %	Kopersaandeel binnen cat. kopers	NB1 Vol Kolom %	Volume aandeel NB1 promo	SOR NB1 <sup>1</sup>	
Frequentie	Promotiegeenheid					Vol	Best
Incidenteel	-	24.3%	38.4%	16.4%	67.7%	62.0%	62.4%
Frequent	Laag	26.5%	24.8%	22.3%	18.5%	31.0%	35.3%
Frequent	Gemiddeld	36.6%	73.8%	43.4%	58.8%	40.3%	41.6%
Frequent	Hoog	12.6%	90.3%	17.9%	84.0%	46.3%	48.7%
<b>Totaal</b>		<b>100.0%</b>	<b>42.8%</b>	<b>100.0%</b>	<b>55.8%</b>	<b>40.8%</b>	<b>42.6%</b>

1. Het "share of requirement", aangeduid met SOR, meet de mate waarin een kopers(deel)groep van een merk zich voorziet in de totale categoriebehoefte door middel van het merk.. In dit voorbeeld geldt dat 40.8% van het categorie volume aangeschaft door de kopers van NB1 betrekking heeft op het merk NB1.

waarbij de merkkopers getypeerd zijn op basis van hun aankoopgedrag op categorie-niveau.

In de eerste kolom zien we bijvoorbeeld dat 24.3% van de merkkopers van NB1 op productcategorie niveau gekarakteriseerd wordt als incidentele categorie koper. In de tweede kolom worden de kopersaandelen weergegeven: 42.8% van de categorie kopers heeft tenminste één keer NB1 gekocht gedurende het jaar, 90.3% van de categorie kopers met een hoge promotiegeenheid heeft tenminste één keer NB1 gekocht. In de daarop volgende kolommen zien we de uitsplitsing van het totale merkvolume over de koperstypologie en per onderscheiden klasse zien we het aandeel daarbinnen dat onder promotionele condities verkocht is. De merkbinding, trouw van de merkkopers aan het merk, is in de laatste twee kolommen onder de kop SOR (share of requirement) weergegeven.

Eerder werd opgemerkt dat de PL's voor hun afzet sterk afhankelijk zijn van huishoudens met een lage promotiegeenheid. Door de informatie zoals weergegeven in tabel 7 per aspect tussen merken te vergelijken, kan dit nog duidelijker worden gemaakt. In figuur 4 is bijvoorbeeld het volume aandeel per merk over de kopersgroepen voor de categorie weergegeven.



Figuur 4. Merkvolume naar kopersgroepen, relatieve verdelingen.

Hieruit blijkt dat huiskopmers en A-merken een totaal verschillende structuur kennen met betrekking tot de afhankelijkheid van kopersgroepen. De NB's doen ongeveer 60% van hun merkvolume bij de kopersgroepen die zich kenmerken door een gemiddelde of hoge promotiegeenheid. Voor de PL's is de afhankelijkheid van deze kopersgroepen

aanmerkelijk minder. Meer dan 65% van het volume kan worden afgezet bij incidentele kopers en categoriekopers met een lage promotiegeneigdheid. Dit zal mede ingegeven worden door het lagere prijsniveau voor PL's.

Het moge duidelijk zijn dat het bovenstaande ook uitgewerkt kan worden naar andere kenmerken en ook voor de categoriekopers bij verschillende afzetkanalen. In het kader van dit artikel zal de overstap gemaakt worden naar de decompositie van de verkopen onder promotionele condities.

In de tabellen 2 en 3 is al weergegeven wat het belang van promoties voor merkvolume en categorievolume bij retailers is. Voor de aankopen onder promotionele condities worden binnen het beschreven model bepaald of vier onderscheiden effecten (B, S, V en T) wel of niet voorkomen. Dit resulteerde uiteindelijk in 16 verschillende combinaties van effecten. De effecten zijn bepaald zowel op merkniveau over alle afzetkanalen heen als ook op categorieniveau voor onderscheiden retailers. Op deze wijze ontstaat zowel voor brand management als accountmanagement belangrijke informatie. Beide gezichtspunten zullen gepresenteerd worden.

Tabel 8. Merkvolume onder promotionele condities naar promotie effect.

Promotie effecten	Merk						
	Totaal	NB1	NB2	NB3	PL1	PL2	PL3
<b>Geen effect</b>	36.4%	43.0%	34.3%	22.2%	29.6%	34.8%	52.3%
<b>T</b>	2.3%	3.0%	2.0%	1.0%	0.0%	2.3%	6.3%
<b>V</b>	14.1%	16.8%	11.3%	10.6%	13.4%	20.1%	10.3%
<b>VT</b>	1.6%	1.5%	1.9%	0.1%	3.6%	5.3%	0.5%
<b>S</b>	5.2%	6.8%	6.0%	4.6%	0.0%	0.2%	0.1%
<b>ST</b>	0.6%	1.1%	0.5%	0.1%	0.0%	0.0%	0.0%
<b>SV</b>	3.4%	4.6%	2.6%	4.1%	0.0%	0.9%	0.7%
<b>SVT</b>	0.8%	1.0%	0.7%	0.7%	2.1%	0.4%	0.0%
<b>B</b>	12.1%	6.4%	13.5%	20.8%	27.6%	9.1%	12.9%
<b>BT</b>	1.5%	1.3%	1.7%	2.6%	0.6%	1.2%	0.2%
<b>BV</b>	8.4%	5.5%	11.5%	14.0%	9.8%	8.1%	0.0%
<b>BVT</b>	0.8%	0.7%	0.9%	1.2%	0.0%	0.7%	0.0%
<b>BS</b>	5.9%	3.5%	6.9%	6.8%	6.9%	4.3%	12.9%
<b>BST</b>	0.9%	0.3%	0.7%	2.4%	0.7%	0.5%	2.7%
<b>BSV</b>	5.0%	3.7%	4.8%	7.5%	4.9%	10.0%	1.1%
<b>BSTV</b>	0.9%	0.6%	0.6%	1.4%	0.9%	2.1%	0.0%
<b>Totaal</b>	<b>100.0%</b>	<b>100.0%</b>	<b>100.0%</b>	<b>100.0%</b>	<b>100.0%</b>	<b>100.0%</b>	<b>100.0%</b>
<b>Geen effect</b>	36.4%	43.0%	34.3%	22.2%	29.6%	34.8%	52.3%
<b>Brand switching</b>	35.4%	22.1%	40.7%	56.7%	51.3%	36.0%	29.8%
<b>Store switching</b>	22.7%	21.6%	22.8%	27.4%	15.4%	18.3%	17.5%
<b>Volume effect</b>	35.0%	34.5%	34.4%	39.6%	34.6%	47.5%	12.6%
<b>Timing effect</b>	9.4%	9.6%	9.0%	9.5%	7.8%	12.5%	9.8%

Naast de zestien afzonderlijke combinaties worden in het onderste deel van de tabel ook de samenvoegingen gepresenteerd van alle effecten waarbij brand switching, store switching enz. betrokken zijn. Ook wordt de regel geen effect herhaald, aangezien daarbij tijdens promotie voor het merk het normaal te verwachten koopgedrag vertoond werd door de kopende huishoudens. Voor NB1 geldt dat bij 43% van het merkvolume, dat onder promotie wordt verkocht, geen enkel beoogd effect aantoonbaar is. Voor PL3 ligt dit percentage zelfs boven de vijftig procent. Brand switching effecten en volume

effecten zijn de belangrijkste aantoonbare promotie effecten in deze markt voor verzorgingsproducten. Voor de merken NB2, NB3 en PL1 zien we dat bij de verkopen onder promotionele condities er een meer dan gemiddeld brand switching effect optreedt. De bereidheid om deze merken onder promotionele condities te kopen en daarbij af te wijken van het belangrijkste merk is hoog. Verder valt op dat prijs promoties voor de PL's minder store switch genereren dan de promoties voor de NB's. Het timing effect is overduidelijk het kleinste effect. De enorme promotiedruk in het veld en de switchbereidheid stellen de consument in staat om in principe permanent onder promotionele condities te kunnen kopen. Het belang van de promotionele verkopen voor de afzet van de NB's toont wel aan dat er geen duidelijk prijsbeleid is. Het lijkt erop dat door de nadruk op promoties voor de NB's de prijsafstand tot de PL's structureel verkleind moet worden.

Tabel 9. *Categorievolume onder promotionele condities bij retailers naar promotie effect.*

Promotie effecten	Afzetkanaal				
	Totaal	R1	R2	R3	R4
<b>Geen effect</b>	36.4%	42.2%	31.0%	38.4%	39.5%
<b>T</b>	2.3%	2.4%	2.8%	2.4%	3.5%
<b>V</b>	14.1%	10.0%	17.3%	17.9%	14.7%
<b>VT</b>	1.6%	1.0%	2.2%	0.8%	3.0%
<b>S</b>	5.2%	3.1%	3.7%	6.5%	8.4%
<b>ST</b>	0.6%	0.5%	1.3%	0.4%	0.0%
<b>SV</b>	3.4%	2.8%	3.1%	6.2%	2.0%
<b>SVT</b>	0.8%	1.1%	0.7%	1.4%	0.8%
<b>B</b>	12.1%	16.7%	11.1%	7.8%	6.2%
<b>BT</b>	1.5%	2.2%	1.4%	0.9%	0.4%
<b>BV</b>	8.4%	8.2%	10.5%	6.3%	8.1%
<b>BVT</b>	0.8%	0.3%	1.5%	1.0%	1.0%
<b>BS</b>	5.9%	4.0%	5.2%	6.1%	6.2%
<b>BST</b>	0.9%	0.6%	0.8%	1.0%	1.0%
<b>BSV</b>	5.0%	4.4%	6.3%	2.3%	5.2%
<b>BSTV</b>	0.9%	0.6%	1.2%	0.7%	0.0%
<b>Totaal</b>	<b>100.0%</b>	<b>100.0%</b>	<b>100.0%</b>	<b>100.0%</b>	<b>100.0%</b>
<b>Geen effect</b>	36.4%	42.2%	31.0%	38.4%	39.5%
<b>Brand switching</b>	35.4%	37.0%	38.0%	26.1%	28.0%
<b>Store switching</b>	22.7%	17.1%	22.2%	24.5%	23.7%
<b>Volume effect</b>	35.0%	28.3%	42.7%	36.6%	34.8%
<b>Timing effect</b>	9.4%	8.5%	11.7%	8.5%	9.8%

Vanuit de optiek van het managen van categorieën zien we dat er duidelijke verschillen bestaan tussen winkel formules met betrekking tot het voeren van een niet effectief promotiebeleid. Als voorbeeld kan gekeken worden naar retailer 1. Uit tabel 3 kan worden afgeleid dat deze retailer meer dan 50% van het categorievolume realiseert onder promotionele condities. Bij meer dan 40% van dit volume kan geen enkel beoogd effect aangetoond worden. Het laagste store switching percentage toont ook aan dat men niet in staat is om op basis van deze categorie de winkelformule aantrekkelijk te maken voor categoriekopers met een andere winkelvoorkeur.



## 8. SLOT

In dit artikel is een model ontwikkeld voor de decompositie van de promotionele verkopen naar beoogde effecten. Het model onderscheidt zich van andere modellen op dit onderzoeksterrein doordat niet uitgegaan wordt van data op basis van geaggregeerde winkelverkopen. In het hier gepresenteerde model staat centraal wat prijspromoties voor merken bij verschillende retailers betekenen voor het aankoopgedrag van de individuele consumenten. Worden mensen door promoties aangezet tot het wisselen van merk en/of winkel en wat betekenen promoties voor de timing en het volume? Door aggregatie van de resultaten in verschillende richtingen is men enerzijds in staat een typologie te maken van categoriekopers op basis van feitelijke promotiegeneigdheid en kunnen anderzijds zowel merken als retailers onderling met elkaar vergeleken worden met betrekking tot het gerealiseerd hebben van te beoogen effecten.

Het nu gepresenteerde model is onder andere doorgerekend voor de in de voorgaande paragraaf geschetste categorie van frequent gekochte verzorgingsproducten. Opmerkelijk is dat de A-merken in deze markt sterk afhankelijk zijn van hun promotionele inspanningen: ongeveer 55% van het totale merkvolume in een jaar wordt gerealiseerd onder promotionele condities. Ook de getoonde promotiekalender geeft aan dat de A-merken bijna permanent wel ergens onder promotionele condities beschikbaar zijn. De A-merken zijn voor de afzet ook sterk afhankelijk van de categoriekopers met een gemiddelde tot hoge promotiegeneigdheid. Door de promoties wordt de prijsafstand tussen de A-merken en winkelmerken dusdanig klein dat, gegeven de afhankelijkheid van A-merken van specifieke deelgroepen categoriekopers, de gemiddeld betaalde prijzen per volume-eenheid weinig verschillen tussen de onderscheiden deelgroepen. Het lijkt erop dat de hoogste doelstelling van de fabrikanten van merkartikelen in deze markt luidt het verdedigen van het marktaandeel op volumebasis. De enorme promotiedruk in deze markt nodigt consumenten uit om met name de A-merken onder promotionele condities te kopen. Dat daarbij de consument wordt gesponsord wordt voor lief genomen. Voor de marktleider geldt bijvoorbeeld dat bij 43% van het promotionele volume geen enkel beoogd effect kon worden aangetoond. De permanente aandacht voor promoties haalt het prijsbeeld voor met name de A-merken onderuit. Vanuit het standpunt van een retailer is het interessant na te gaan hoe de categorie feitelijk beheerd is en of er onderscheid is ten opzichte van de concurrentie. Voor een grote retailer zien we dat meer dan 50% van het categorievolume bij deze retailer verkocht is onder promotionele condities. Voor meer dan 40% van dit promotionele volume is geen enkel effect aantoonbaar, terwijl de retailer er duidelijk niet in slaagt met promoties in de categorie mensen naar de winkel te trekken. Tot slot kan worden opgemerkt dat ook in deze studie de omvang van het Brand Switching effect veel geringer is dan in een aantal eerdere publicaties werd aangegeven. De hier gepubliceerde resultaten, maar ook die voor een andere onderzochte product categorie, liggen voor het aspect merkenwissel in de orde van grootte van de resultaten die Van Heerde, Gupta en Wittink (2003) publiceerden.

Voor toekomstig onderzoek is het interessant na te gaan of er relaties te vinden zijn tussen "perceived deal proneness" en de nu ontwikkelde classificatie. Ook het onderzoek naar het optreden van cross-category promotie effecten is een interessante ontwikkelingsrichting welke met name voor retailers impact kan hebben op het te voeren promotiebeleid in ruime zin.

## LITERATUUR

- Blattberg, R.C. en S.A. Neslin (1990) *Sales Promotion: Concepts, Methods, Strategies*. Englewood Cliffs, NJ: Prentice Hall.
- Ehrenberg, A.S.C. (1972) *Repeat Buying: Theory and Applications*. North-Holland.
- Heerde, H.J. van, S. Gupta en D.R. Wittink (2003) "Research Notes and Communications: Is 75% of the Sales Promotion Bump due to Brand Switching? No, only 33% is." *Journal of Marketing Research*, 40, 4, p. 481-491.
- Leeftang, P.S.H. en P.A. Beukenkamp (1981) *Probleemgebied Marketing: een Managementbenadering*. Leiden/Antwerpen: H.E. Stenfert Kroese B.V.
- Luijten, A.L.J.M. en M.M.L.R. Nagtzaam (2003) "Dalend consumentenvertrouwen en effecten op het koopgedrag van Food". *Tijdschrift voor Marketing*, jaargang 37, nummer 7/8.
- Nijs, V.R., M.G. Dekimpe, J-B.E.M. Steenkamp en D.M. Hanssens (2001) "The Category-Demand Effect of Price Promotions". *Marketing Science*, 20, 1, p. 1-22.
- Teunter, L.H. (2002) *Analysis of Sales Promotion Effects on Household Purchase Behavior*. Erasmus Research Institute of Management (ERIM), Rotterdam, ERIM Ph.D. Series, 16
- Wilborts, L. (2002) "De Verkoopiek Ontmaskerd". Niet gepubliceerde master thesis, Universiteit van Tilburg/Gfk PanelServices Benelux B.V.

## *7. Het meten en bewaken van de essentie van een merk*

J. VAN REKOM, G. JACOBS, P.W.J. VERLEGH en T. VAN STEEKELENBURG

### SAMENVATTING

Het intact houden van de essentie van een merk is een punt van constante zorg voor merkenmanagers. Dit geldt met name bij herpositionering van merken en bij het toevoegen van nieuwe producten onder een bestaande merkenparaplu. De eigenschappen die de essentie van een merk vormen bepalen of een consument veranderingen in een merkbeeld accepteert. Dit artikel introduceert een methode om greep te krijgen op de essentie van een merk. Het kernargument is dat die eigenschappen waarvan de consument gelooft dat die de meeste andere eigenschappen van dat merk veroorzaken de essentie van dat merk vormen. Dit wordt getest aan de hand van een onderzoek bij een wasmiddelenmerk. De hier voorgestelde methode is, in tegenstelling tot de in de praktijk van het marktonderzoek dominerende correlatieve methoden in staat, causale verbanden tussen merkeigenschappen vast te leggen. Het wordt daardoor mogelijk, met veel grotere accuratesse nog niet bestaande eigenschappen toe te voegen aan een merk en na te gaan in hoeverre de consument die eigenschappen bij het merk vindt passen.

### 1. WAAROM MERKESENTIE?

Succesvol merkenmanagement vereist het continu bewaken van de markt en tijdige aanpassingen van het merk aan de ontwikkelingen in de markt. De kracht van een merk ligt in het imago dat huidige en potentiële kopers van dat merk hebben (Keller, 2003; Krishnan, 1996). Daarin ligt besloten hoe consumenten reageren op de marketingspanningen van aanbieders.

Leidende auteurs op het gebied van merkmeerwaarde (Aaker, 1996; Keller, 2003; Riezebos, 2003) benaderen het imago van merken als een min of meer samenhangend patroon van associaties die mensen bij dat merk hebben. Aaker (1996) geeft de identiteit van het merk Kodak bijvoorbeeld weer met maar twee woorden: "eenvoud" ("simplicity"), die vooral blijkt uit de producteigenschappen, en "familie", een aspect dat Kodak ondersteunt in haar communicatie-uitingen en het beeldmateriaal daarbij. Essentiële eigenschappen spelen een cruciale rol in het creëren en in stand houden van merkmeerwaarde, en zouden daarom de rode draad moeten vormen voor de hele communicatie in woord en daad over een merk. Succesvolle merken blijken in de regel hun kerneigenschappen door alle tijden heen vast te houden. Een ideale aanpassing van de merkpositionering zou zelfs moeten uitmonden in een versterking van de betekenis van een merk in plaats van de verwatering ervan (Keller, 2003: 634). Hier speelt het dilem-

ma voor de marketing-manager, waar dit artikel op ingaat: hoe kun je een merk aanpassen aan de veranderende eisen van de consument, zonder dat er afbreuk wordt gedaan aan de essentie van een merk?

Bij de herpositionering van een merk dient dan ook bijzonder voorzichtig met die essentie worden omgesprongen, opdat het merk zich niet vervreemdt van haar bestaande klanten. Een klassiek voorbeeld is het “New Coke” fiasco van Coca Cola. Toen in 1985 het management met veel publiciteit bekend had gemaakt dat de smaak van Coca Cola was verbeterd, vonden veel Amerikanen dat Coca Cola niet meer hetzelfde was. Dit leidde tot veel emotionele reacties in de media, inclusief een hitparadehit onder de titel “Coke was it” (Oliver, 1985).

Het merk dat in dit hoofdstuk onderzocht zal worden is een wasmiddelenmerk. Om redenen van betrouwbaarheid zal het hier als Merk X door het leven gaan. De producent had de indruk dat Merk X zijn aantrekkingskracht op de jongere generatie had verloren en wilde het merk herpositioneren, zonder daarbij de grote groep trouwe consumenten voor het hoofd te stoten. Een complicerende factor daarbij was dat consumenten zich niet altijd bewust zijn van reële verschillen tussen concurrerende merken (cf., Rosa et al., 1999). Dit is het geval bij veel grootschalig geproduceerde producten die al geruime tijd in de markt zijn. De twee grootste producenten van wasmiddelen, Procter & Gamble en Unilever, gaven zelf in de pers te kennen dat ook zij zelf vrijwel geen verschillen in de prestaties van de diverse merken wasmiddelen zagen (Baltesen, 2001). De kansen om die merken nog van elkaar te laten verschillen leken daardoor vooral te liggen op het terrein van de merkpersoonlijkheid. Het hieronder besproken onderzoek is dan ook toegespitst op aan de merken toegeschreven persoonlijkheidseigenschappen.

Hoe kunnen we in marktonderzoek vaststellen welke persoonlijkheidseigenschappen de essentie van een merk vormen? Er is tot nu toe geen onderzoeksmethode beschikbaar om de meer essentiële van de meer perifere eigenschappen te onderscheiden. Het antwoord hierop zal gezocht moeten worden in de manier waarop eigenschappen met elkaar samenhangen. Dat zal in paragraaf 2 uitgewerkt worden. De volgende vraag is welke eigenschappen aan een merk toegevoegd kunnen worden zonder de essentie ervan aan te tasten. Het beantwoorden van deze vraag vereist gegevens over de invloed van eventuele nieuwe eigenschappen op de perceptie van bestaande merkeigenschappen. In paragraaf 3 komt een methode aan bod, die op deze vraag een antwoord kan geven. Hiermee willen wij een uitweg bieden uit het dilemma waar de marketingmanager vaak voor staat: hoe kan een merk aangepast worden aan de veranderende eisen van de consument, zonder afbreuk te doen aan de essentie van het merk?

## 2. HOE MERKEIGENSCHAPPEN MET ELKAAR SAMENHANGEN

### 2.1. *De rol van causale relaties*

De eerste vraag is hoe consumenten centrale en minder centrale eigenschappen onderscheiden, en hoe deze in relatie tot elkaar staan. Volgens Murphy en Medin (1985) houden mensen er “naïeve theorieën” op na over hoe de wereld in elkaar zit (zie ook Morel, 2002). Twee of meer eigenschappen van een begrip (bijvoorbeeld een diersoort,

of een merk) hangen samen als mensen er een eigen theorie op nahouden waarmee ze kunnen uitleggen *waarom* die eigenschappen samen gaan (cf. Pinson, 1986). Twee eigenschappen van een kat zijn bijvoorbeeld, dat hij scherpe klauwen heeft en dat hij in bomen kan klimmen. Die eigenschappen hangen samen: een kat kan in bomen klimmen, omdat hij scherpe klauwen heeft. Het verband tussen twee eigenschappen is dus niet een simpele, algemene associatie, maar een *causale* verklaring waarom de ene eigenschap van de andere afhangt (Murphy & Medin, 1985). Aan de hand van hun theorieën leggen mensen aan zichzelf en anderen uit wat begrippen voorstellen, en waarom dingen zijn zoals ze zijn. De causale relaties waar mensen in geloven hoeven niet per se wetenschappelijke basis te hebben: het gaat om wat mensen zelf denken, niet om een onderliggende wetenschappelijke realiteit.

De onderlinge samenhangende eigenschappen vormen zo een rijk geschakeerde structuur (Ahn, 1998). Positieve causale relaties versterken de samenhang van de eigenschappen van een persoon of merk, negatieve causale relaties doen juist afbreuk aan die samenhang. In dit netwerk zullen sommige eigenschappen meer andere eigenschappen veroorzaken, en zijn er ook eigenschappen die niet zoveel invloed zullen hebben op andere eigenschappen. Een verandering in een eigenschap die weinig andere eigenschappen veroorzaakt zal niet veel effect hebben op het imago van het product. Een verandering in een eigenschap die veel andere eigenschappen veroorzaakt zal daarentegen doorwerken in het hele merkbeeld (Murphy & Medin, 1985). Bijvoorbeeld: vuil verwijderen, schuimen en kleren fris laten ruiken zijn alledrie eigenschappen van wasmiddelen. We weten, of nemen aan, dat een wasmiddel kleren fris laat ruiken omdat een wasmiddel vuil verwijdert. Als iemand ons zou vertellen dat hij een middel had om kleren fris te laten ruiken zonder het vuil te verwijderen, dan zouden we ons afvragen of het wel om een wasmiddel ging, en niet om een soort parfum. “Vuil verwijderen” is dus een kerneigenschap van wasmiddelen, omdat die eigenschap zo nauw verbonden is aan de andere informatie die wij over wasmiddelen hebben. Een verandering in een centrale eigenschap kan een compleet ander totaalbeeld opleveren, terwijl een verandering in een perifere eigenschap veel minder effect heeft (Asch, 1946). Bijvoorbeeld, als een wasmiddel niet schuimt, maar wel vuil verwijdert en kleren fris laat ruiken, zal het nog steeds als wasmiddel gelden. Daarentegen is het de vraag of een middel dat kleren fris laat ruiken en flink schuimt, maar het vuil niet verwijdert, nog wel een wasmiddel is. “Schuimen” blijkt hieruit gezien te worden als een perifere eigenschap, terwijl waar “Vuil verwijderen” een centrale eigenschap is, zonder welke een wasmiddel geen wasmiddel meer is.

Hoe kunnen we verklaren dat mensen veel meer aandacht geven aan veroorzakende eigenschappen dan aan de gevolgen daarvan? De reden is dat veroorzakende eigenschappen veel beter helpen te verklaren waarom dingen zijn zoals ze zijn. Als mensen zich een beeld vormen van andere mensen, dan zijn veroorzakende eigenschappen veel nuttiger om andere mensen te begrijpen. Veroorzakende eigenschappen maken het namelijk mogelijk om ook gedrag in nieuwe situaties te voorspellen. Op grond hiervan formuleerde Ahn (1998) haar “causale status hypothese”: “Wanneer een aantal eigenschappen causaal samenhangt, zullen mensen meer afgaan op eigenschappen die de andere eigenschappen veroorzaken dan op eigenschappen die het gevolg zijn van andere eigenschappen.” Deze hypothese heeft zij onderzocht en bevestigd voor producten als televisietoestellen en spiegels, en voor dieren, zoals mieren.

## 2.2. Causale status en wat mensen als “de essentie” zien

Essentiële eigenschappen zijn die eigenschappen die een product voor de consument maken tot wat het is (Morel, 2002). In Ahn's (1998) benadering is essentie een kwestie van gradatie, en de mate waarin een eigenschap essentieel is varieert met mate waarin een eigenschap andere eigenschappen veroorzaakt. De mate waarin een eigenschap als essentieel wordt gezien vloeit voort uit de mate waarin die eigenschap andere eigenschappen lijkt te veroorzaken. Wanneer wij dit toepassen op merken, komen wij tot onze definitie van merkessentie:

**Merkessentie:** Hoe meer een eigenschap als de oorzaak van andere eigenschappen van een merk wordt gezien, hoe meer die eigenschap de essentie van dat merk vormt.

Het grote voordeel van deze conceptualisatie van essentie in termen van eigenschappen die de meeste andere eigenschappen veroorzaken is, dat zij rekening houdt met de menselijke neiging om naar oorzaken van eigenschappen te zoeken, zonder dat het veronderstellen van een daadwerkelijke, “objectieve” essentie nodig is.

### 2.2.1. Huidige en potentiële merkeigenschappen

De eerste stap in het voorbereiden van een eventuele herpositionering van een merk is het vaststellen van de mate waarin consumenten een eigenschap als de essentie van een merk zien. De benadering van Ahn (1998) maakt het mogelijk om vast te stellen *waarom* een bepaalde eigenschap als essentieel wordt gezien. Het in kaart brengen van de oorzaak-gevolgrelaties laat zien welke andere eigenschappen volgens de consument afhankelijk zijn van een specifieke essentiële eigenschap. In principe zouden eigenschappen die goed aansluiten bij de bestaande structuur van oorzaak-gevolgrelaties goed moeten passen bij wat consumenten als de essentie van een merk zien. Met “goed aansluiten” wordt bedoeld dat deze nieuwe eigenschappen mede door de bestaande eigenschappen veroorzaakt zouden kunnen worden, waardoor zij de evaluatie van de bestaande eigenschappen zelfs zouden kunnen versterken. Met andere woorden, de mate waarin nieuwe eigenschappen passen in hetzelfde patroon van oorzaak-gevolgrelaties waarin eigenschappen van elkaar afhankelijk zijn biedt inzicht in hoe nieuwe eigenschappen met de huidige merkessentie samen kunnen gaan. Een nieuwe eigenschap als een “werkt ook goed bij lage temperaturen” zou op die manier goed kunnen aansluiten op een eigenschap als betrouwbaarheid: als consumenten een merk betrouwbaar vinden is het goed werken bij lage watertemperatuur voor hen een extra bevestiging van hun idee over het merk. Ideaal gesproken zouden causale relaties hier twee kanten op kunnen werken: doordat Merk X nu ook bij lage temperaturen heel goed wast, wordt het nog betrouwbaarder. Potentieel nieuwe eigenschappen waar consumenten een conflict met bestaande eigenschappen zien kunnen beter niet opgevoerd worden. Zo heeft Harley Davidson weinig plezier beleefd aan de introductie van hun parfummerk. Het idee dat het stoere en mannelijke Harley merk gekoppeld wordt aan parfum was voldoende om een belangrijke groep bestaande klanten te vervreemden van het merk (Haig, 2003).

Eigenschappen die afbreuk doen aan bestaande merkeigenschappen kunnen afbreuk doen aan de waarde van een merk, omdat zij conflicteren met de attributen die het merk

maken tot wat het is. Wanneer mensen bijvoorbeeld denken dat iemand die jolig is daarom ook minder zorgzaam zal zijn, dan zou het toevoegen van de eigenschap “jolig” de essentie van dat merk als “zorgzaam” behoorlijk kunnen ondergraven. Het is daarom van het grootste belang, te weten te komen welke merkeigenschappen welke andere in de perceptie van de consument veroorzaken. Zo is het voor een ziektekostenverzekeraar bijvoorbeeld niet voldoende om te weten, dat een nieuw toe te voegen eigenschap “bereikbaar” door de klanten als passend bij de kerneigenschappen “vriendelijk” en “betrouwbaar” wordt gezien. Pas wanneer die verzekeraar weet, dat “bereikbaar” bij het merk past, omdat continue bereikbaarheid een gevolg is van zowel „vriendelijkheid“ als „betrouwbaarheid“, dan weet die ook, dat de nieuwe eigenschap “bereikbaar” een perifere eigenschap is, die in een reclamecampagne door de kerneigenschappen ingekleurd moet worden: “Ons verzekeringsbedrijf is vriendelijk en betrouwbaar. Daarom vinden wij het vanzelfsprekend dat wij dag en nacht eenvoudig te bereiken zijn”. Doordat de marketer weet, welke eigenschap welke andere veroorzaakt, is het beter mogelijk een overtuigende, samenhangende positionering voor een merk te creëren.

### 3. DE GEBRUIKTE ONDERZOEKSMETHODE

#### 3.1. *Kwalitatief vooronderzoek*

Om de belangrijkste persoonlijkheidseigenschappen van Merk X te achterhalen zijn in totaal twaalf kwalitatieve interviews gehouden. Hiervoor zijn twee locaties gekozen: een plattelandsgemeente en één van de drie grote steden. Respondenten kregen de vraag: “Stel dat Merk X een mens zou zijn, wat voor mens zou het dan zijn?” De interviews zijn op cassettebandjes opgenomen en naderhand letterlijk uitgetypt. Alleen eigenschappen die door minimaal de helft van de twaalf respondenten genoemd zijn, in totaal zeven, zijn in het verdere onderzoek meegenomen: “betrouwbaar”, “zorgzaam”, “vriendelijk”, “betrokken”, “behulpzaam”, “nuchter” en “eerlijk”. De linker kolom in het bovenste deel van Tabel 1 geeft aan, door hoeveel van de respondenten elk van deze eigenschappen is genoemd.

Onafhankelijk van dit vooronderzoek heeft een marktonderzoeksbureau zestien andere respondenten geïnterviewd uit het marktsegment waar Merk X in de nabije toekomst op wilde mikken. Deze respondenten werd gevraagd hoe zij Merk X op dit moment zagen en hoe zij Merk X graag zouden willen zien. In totaal werden vijf gewenste eigenschappen door minimaal de helft van die respondenten genoemd: “origineel”, “modern”, “vrolijk”, “levendig” en “expressief” (zie linkerkolom in onderste deel van Tabel 1). Deze vijf eigenschappen vormden samen met de zeven eigenschappen uit ons eigen kwalitatieve onderzoek de input voor onze vragenlijst. In de vragenlijst zijn de eigenschappen in willekeurige volgorde door elkaar gemengd.

In de analyses zullen we eerst kijken naar hoe elk van de zeven bestaande persoonlijkheidseigenschappen van Merk X door elk van de andere eigenschappen veroorzaakt wordt en in hoeverre zij de essentie van Merk X vormen. Daarna zullen we kijken naar de aansluiting van de mogelijke nieuwe vijf persoonlijkheidseigenschappen op de bestaande. Alle analyses zijn gebaseerd op het bestaande marktsegment van trouwe gebruikers van Merk X.

### 3.2. Opzet van de vragenlijst voor kwantitatief onderzoek

Om de mate waarin elke eigenschap elk van de andere veroorzaakte vast te kunnen stellen is een vragenlijst opgesteld. De respondenten kregen elke combinatie van eigenschappen voorgeschoteld, zodat we voor alle twaalf eigenschappen de relatieve mate waarin ze elk van de andere veroorzaakten konden vaststellen. Voor elke eigenschap is een pagina uitgetrokken, waarin de mate is vastgesteld waarin deze afhangt van elk van de andere eigenschappen. Figuur 1 geeft een voorbeeld van hoe dat gebeurd is.

Wilt U hieronder aangeven in welke mate de hieronder gesuggereerde verbanden juist zijn? Kruis bij elke stelling het antwoord aan dat het meest van toepassing is.

Merk X is **betrouwbaar**, omdat Merk X ...

	<b>Nee, integendeel</b>	<b>Heeft niets met elkaar te maken</b>	<b>Juist, maar gaat niet op voor Merk X</b>	<b>Ja, juist</b>
zorgzaam is	0	0	0	0
vriendelijk is	0	0	0	0
betrokken is	0	0	0	0
behulpzaam is	0	0	0	0
nuchter is	0	0	0	0
eerlijk is	0	0	0	0

Figuur 1. Voorbeeld van hoe oorzaak-gevolgrelaties gevraagd zijn.

Wanneer respondenten “ja, juist” aankruisten, werd voor deze relatie een “1” genoteerd. Wanneer respondenten aankruisten “heeft niets met elkaar te maken”, dan werd hiervoor een “0” genoteerd. Het is echter ook denkbaar, dat mensen het in principe eens zijn met de voorgestelde relatie, maar dat zij vinden dat dat om een of andere reden voor Merk X niet opgaat. Daarom is in de vragenlijst, en in Figuur 1, de antwoordmogelijkheid opgenomen: “juist, maar gaat niet op voor Merk X”. Voor de berekeningen van de mate waarin een eigenschap de andere eigenschappen veroorzaakt is dit als “0” gecodeerd, maar als respondenten deze antwoordmogelijkheid aankruisen, dan kan dat betekenen dat hier voor Merk X een kans ligt om de samenhang tussen de merkeigenschappen te versterken – mits in vervolgonderzoek goed onderzocht wordt waarom een op zich logische relatie kennelijk voor Merk X niet op gaat. Als mensen tenslotte “nee, integendeel” aankruisten, is dat als “-1” gecodeerd, omdat in dit geval de ene eigenschap afbreuk doet aan de andere. Om de mate waarin een eigenschap de andere eigenschappen veroorzaakt uit te kunnen rekenen zijn gegevens over alle combinaties van eigenschappen nodig: respondenten moeten voor elk van de twaalf eigenschappen de relaties met alle 11 andere eigenschappen, in totaal  $12 * (12-1) = 132$  relaties invullen. Bij  $k$  attributen geldt, dat het aantal in te vullen relaties gelijk is aan  $k * (k-1)$ .



Op basis van deze data wordt per eigenschap de totale mate van “oorzakelijkheid” berekend. In feite is dat voor elke eigenschap de som van de antwoorden (-1, 0 of 1) op alle voorgelegde relaties, waarin die eigenschap als de mogelijke oorzaak wordt voorgelegd. Op dezelfde manier is ook de mate van “gevolg zijn” uitgerekend: als de som van de antwoorden op alle voorgelegde relaties, waarin de betreffende eigenschap als mogelijk gevolg van een andere eigenschap wordt voorgelegd. Zo kan met deze methode per respondent per eigenschap een totale “oorzakelijkheid” worden berekend, en een totale mate van “gevolg zijn”.

De hier ontwikkelde definitie en operationalisatie van merkessentie zijn nieuw. In het verleden hebben Michel en Ambler (1999) merkessentie benaderd door na te gaan of een merk nog wel hetzelfde merk zou zijn als een bepaalde eigenschap weg zou worden gelaten. Ook Ahn (1998) testte haar causale status hypothese met de controlevraag, of bij het wegvallen van een centrale eigenschap een bepaalde diersoort nog wel diezelfde diersoort zou zijn. Zoals hierboven beredeneerd is, zou het wegvallen van een heel centrale eigenschap inderdaad een ander merkbeeld opleveren. De benadering van Michel en Ambler (1999) zou dus wel in zekere mate moeten convergeren met de hier gepresenteerde benadering van essentie. Daarom is als controlevraag in de vragenlijst opgenomen: “Zou Merk X nog steeds Merk X zijn als Merk X niet meer ..... zou zijn?” Respondenten konden deze vraag beantwoorden op een 7-puntsschaal, van “1” = ja, nog steeds Merk X tot “7” = nee, absoluut geen Merk X meer. Deze schaal geeft de mate weer waarin een eigenschap essentieel is voor Merk X. Het antwoord “1” = ja, nog steeds merk X” betekent dat de betreffende eigenschap absoluut niet essentieel is. Het antwoord “7 = nee, absoluut geen Merk X meer” betekent, dat de eigenschap heel essentieel is, omdat het merk kennelijk niet zonder deze eigenschap kan (vgl. Ahn, 1998; Michel & Ambler, 1999). Om er zeker van te zijn dat de in het vooronderzoek gevonden eigenschappen wel voldoende relevant waren in de ogen van respondenten is zijn twee items per eigenschap opgenomen om dit te controleren: “Hoe belangrijk is deze eigenschap voor u bij het kiezen van een wasmiddel?” (1 = volkomen onbelangrijk, 7 = heel belangrijk), en “Wanneer u aan wasmiddelen denkt, hoeveel betekent deze eigenschap dan voor u?” (1 = betekent absoluut niet veel voor mij, 7 = betekent heel veel voor mij”).

### 3.3. Steekproef

De vragenlijst is per post verstuurd naar 1000 gebruikers van merk X uit het respondentenpanel van een marktonderzoeksbureau. Om de respons te stimuleren was er een loterij aan de enquête verbonden: drie inzenders wonnen een heel jaar gratis wassen, en één respondent een ballonvaart boven Nederland. Uiteindelijk zijn 146 vragenlijsten terug ontvangen (14,6 % respons).

## 4. RESULTATEN

### 4.1. Validatie van de methode

De uitkomsten van de schriftelijke enquête vertonen duidelijke overeenkomsten met die van het vooronderzoek: de eigenschappen, die in het vooronderzoek naar voren

Tabel 1. Gemiddelden en standaardafwijkingen van huidige en potentieel nieuwe eigenschappen van Merk X\*

<i>Huidige eigenschappen</i>	Hoe vaak genoemd in vooronderzoek (N = 12)	Is merk X zonder deze eigenschap nog Merk X?***	Aantal eigenschappen dat deze eigenschap veroorzaakt****
Betrouwbaar	9	5,04 (1.96) <sup>a</sup>	4,95 (1,76) <sup>a</sup>
Zorgzaam	8	4,47 (1.81) <sup>b</sup>	4,11 (2,06) <sup>b</sup>
Vriendelijk	7	3,80 (1.57) <sup>c</sup>	2,23 (2,21) <sup>c</sup>
Betrokken	7	4,26 (1.80) <sup>b</sup>	3,19 (2,36) <sup>d</sup>
Behulpzaam	7	4,82 (1.80) <sup>a,b</sup>	3,79 (2,16) <sup>e</sup>
Nuchter	6	3,96 (1.69) <sup>c</sup>	2,00 (2,37) <sup>c</sup>
Eerlijk	6	4,67 (1.80) <sup>b</sup>	3,73 (2,28) <sup>e</sup>
Totaal gemiddelde		4.43 (1.82)	3.43 (1.66)

<i>Potentiële nieuwe eigenschappen</i>	Hoe vaak genoemd in vooronderzoek (N = 16)	Is merk X zonder deze eigenschap nog Merk X?***	Aantal eigenschappen dat deze eigenschap veroorzaakt****
Vrolijk	10	3,42 (1.48) <sup>a</sup>	1,05 (2.64) <sup>a</sup>
Levendig	9	3,43 (1.41) <sup>a</sup>	1,49 (2.62) <sup>b</sup>
Origineel	12	3,39 (1.49) <sup>a</sup>	2,30 (3.96) <sup>c</sup>
Modern	12	3,63 (1.60) <sup>a</sup>	2,03 (3.96) <sup>b,c</sup>
Expressief	9	3,31 (1.46) <sup>a</sup>	0,57 (2.80) <sup>a</sup>
Totaal gemiddelde nieuwe eigenschappen		3.44 (1.49)	1.50 (2.41)

\* De getallen tussen haakjes zijn de standaardafwijkingen.

\*\* 1 = "ja, nog steeds merk X" (= niet essentieel), 7 = "absoluut geen merk X meer" (= zeer essentieel). Waarden met hetzelfde superscript zijn niet significant verschillend ( $\alpha = .05$ ).

\*\*\* In deze berekening zijn alleen de huidige eigenschappen van het merk meegenomen. De eigenschap in de meest linkse kolom is hier de oorzaak. Waarden met hetzelfde superscript zijn niet significant verschillend ( $\alpha = .05$ ).

\*\*\*\* In deze berekening zijn alle 11 andere eigenschappen meegenomen (huidige en potentiële). De eigenschap in de meest linkse kolom is hier de oorzaak. Waarden met hetzelfde superscript zijn niet significant verschillend ( $\alpha = .05$ ).

kwamen als de persoonlijkheidseigenschappen van Merk X werden door de respondenten als belangrijk beschouwd (gemiddelde van "belang" 4,95; gemiddelde van "betekent veel voor me": 5,19, beiden op een 7-punts schaal). Dit betekent dat de eigenschappen die in het vooronderzoek als persoonlijkheidseigenschappen van Merk X geïdentificeerd zijn ook voor de grotere populatie zinvol zijn. De huidige eigenschappen worden ook stuk voor stuk als essentiële gepercipieerd dan de nog niet bestaande eigenschappen. Zij veroorzaken gemiddeld meer eigenschappen dan de nog niet bestaande eigenschappen: de minst essentiële huidige eigenschap scoort 3,80 op een 7-puntsschaal, terwijl de meest essentiële nog niet bestaande eigenschap 3,63 scoort

(Tabel 1). Dit ondersteunt de validiteit van de in het vooronderzoek geselecteerde huidige eigenschappen. “Betrouwbaar” blijkt verreweg de meest essentiële eigenschap van Merk X te zijn, die de meeste andere eigenschappen veroorzaakt. Respondenten zijn van mening dat de “betrouwbaarheid” van Merk X meer andere eigenschappen veroorzaakt (gemiddeld 4,95) dan welke andere eigenschap dan ook. “Betrouwbaarheid” is ook de eigenschap, waarvoor het meest geldt, dat Merk X niet langer Merk X zou zijn als deze eigenschap weg zou vallen (5,04 op een 7-punts schaal, Tabel 1).

Hoe hangen nu de beoordelingen van deze eigenschappen onderling samen? De rangordecorrelatie tussen de mate waarin Merk X zonder de huidige eigenschappen kan en de gemiddelde aantallen andere eigenschappen die de huidige eigenschappen van Merk X veroorzaken (de derde en vierde kolom in Tabel 1) bedraagt 0,86 ( $p < 0,01$ ). Dit resultaat is bemoedigend. Het zou echter zo positief kunnen uitvallen omdat bij het berekenen van correlaties tussen gemiddelden heterogeniteit onder de respondenten wordt “uitgemiddeld”. Daarom hebben wij een verfijndere analyse toegevoegd. Elke respondent heeft voor alle huidige en potentieel nieuwe eigenschappen beoordelingen gegeven, en in principe is het daardoor mogelijk voor elke respondent apart de samenhang vast te stellen. Wij zijn geïnteresseerd in de gemiddelde samenhang, over de hele populatie, tussen de mate waarin een eigenschap de andere eigenschappen veroorzaakt en de mate waarin Merk X zonder elk van die eigenschappen kan. Als eerste analysestap zijn voor elke individuele respondent de correlaties berekend tussen de mate waarin de respondent vindt dat Merk X zonder die eigenschap nog steeds Merk X was (de essentie), het aantal eigenschappen dat elke eigenschap veroorzaakte en het aantal eigenschappen waarvan elke eigenschap het rechtstreekse gevolg was. Om tot een overzicht over de hele groep respondenten te komen is op elke correlatie een Fisher-r-to-Z transformatie toegepast. Hiervan is per paarsgewijze correlatie het gemiddelde over alle respondenten berekend. Tenslotte is dit gemiddelde teruggetransformeerd door het toepassen van een inverse Fisher-Z transformatie (zie Rosenthal, 1991: 86). Zo zijn de gemiddelde correlaties voor onze hele groep respondenten berekend, die in Tabel 2 staan. De zeven huidige merkeigenschappen (het bovenste deel van Tabel 1) vormen de basis voor de in Tabel 2 berekende correlaties.

*Tabel 2. Correlatiematrix van de eigenschappen  
(gemiddelden van de correlaties per respondent)*

	Onmisbaarheid van eigenschap voor Merk X	Mate waarin eigenschap oorzaak is
Mate waarin eigenschap oorzaak is	.49*	
Mate waarin eigenschap gevolg is	-0.09	-0.20*

\* Deze correlaties zijn significant, met  $p < 0.001$ .

De mate waarin een eigenschap de andere eigenschappen van Merk X veroorzaakt hangt duidelijk samen met de mate waarin consumenten vinden dat Merk X zonder die eigenschap Merk X niet meer is: de correlatie bedraagt 0,49 (Tabel 2). Hoe meer een eigenschap de andere eigenschappen van een merk veroorzaakt, hoe meer deze eigenschap als essentieel voor dat merk gezien wordt. De mate van essentie is duidelijk gekoppeld aan de mate waarin een eigenschap de oorzaak van andere eigenschappen is (de correlatie in Tabel 2 is 0.49), en niet de mate waarin de eigenschap een gevolg is – die laatste correlatie is niet significant en zelfs negatief (-0.09) (Tabel 2). Er is dus sprake van een zekere mate van convergentie tussen Michel en Ambler's (1999) maat voor essentie ("zonder deze eigenschap zou Merk X Merk X niet meer zijn) en de mate waarin een eigenschap de andere eigenschappen van Merk X lijkt te veroorzaken. De resultaten laten duidelijk zien dat de mate waarin een eigenschap de andere eigenschappen van een merk veroorzaakt niet identiek is aan de mate waarin een eigenschap onmisbaar is voor Merk X: de gevonden correlatie van 0,49 is veel kleiner dan de correlatie van 1,00 die we zouden vinden als beide benaderingen identiek zouden zijn. Hoe onmisbaar een eigenschap is hangt dus wel samen met hoe essentieel die eigenschap is, maar "essentieel zijn" is niet precies hetzelfde als "onmisbaar zijn".

Een ander relatief nadeel van "onmisbaarheid" als benadering van essentie is, dat respondenten het kennelijk soms moeilijk vinden om onderscheid te maken tussen meer en minder onmisbare elementen. In de rechter kolom van Tabel 1 zien we, dat de scores van de verschillende eigenschappen via de causale benadering verder uit elkaar liggen dan die bij de onmisbaarheidsscores. De causale benadering van essentie, die in dit artikel voorgesteld wordt, lijkt veel scherper de "echt essentiële" eigenschappen er uit te filteren. Dit verschil wordt nog veel scherper wanneer we kijken naar eventueel nieuw toe te voegen eigenschappen.

#### *4.2. Analyse van de samenhang tussen eigenschappen onderling*

Tot zover hebben de analyses in dit hoofdstuk betrekking op de vraag, welke eigenschappen de consument als het meest essentieel ziet, en met name op hoe dat te beïnvloeden is vanuit de manier waarop die eigenschappen samenhangen. Nu op deze vraag een antwoord is gegeven richten we ons op de tweede vraag: Hoeveel speelruimte heeft het management om nieuwe eigenschappen aan een merk toe te voegen? In wezen is de aanpak die wij voorstellen heel eenvoudig: De eigenschappen die, in de ogen van de consument, de meest essentiële huidige eigenschappen in stand kunnen houden of kunnen versterken, of door de huidige essentiële eigenschappen veroorzaakt zouden kunnen worden, zijn het makkelijkst door te voeren (zie ook Keller, 2003: 638). Wanneer het management overweegt om nieuwe eigenschappen toe te voegen aan een bestaand merkbeeld, zoals bij Merk X het geval was, is het raadzaam een overzicht te maken van hoe deze eigenschappen zich verhouden tot de meest essentiële eigenschappen van dat merk. Voor dit concrete voorbeeld geeft Tabel 3 dat overzicht.

De getallen in Tabel 3 zijn proporties. Links boven in de tabel staat bijvoorbeeld 0,77 bij "Merk X is betrouwbaar (eerste kolom) omdat merk X zorgzaam is (bovenste getal in die kolom). 77 % van de respondenten die die vraag beantwoord hebben vulden hier "ja juist" in. Wanneer er ook respondenten waren, die "nee integendeel" invulden, dan zijn de getallen gecorrigeerd tot netto proporties. Bij de vraag "Merk X is betrouwbaar

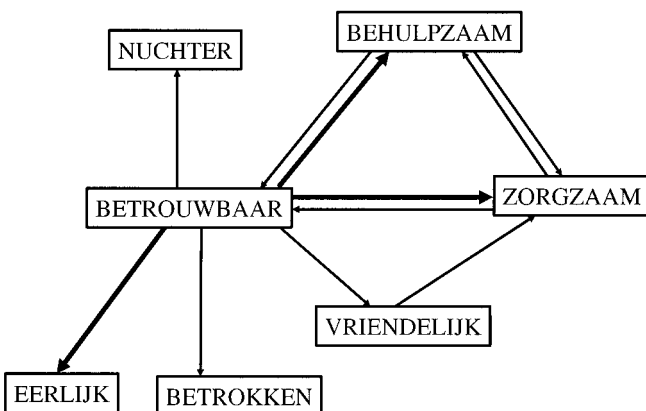
Tabel 3. Gemiddelde scores van de causale relaties tussen de huidige en potentieel nieuwe eigenschappen van Merk X.

Merk X is:	Huidige eigenschappen					Potentieel nieuwe eigenschappen						
	betrouwbaar	zorgzaam	vriendelijk	betrokken	behelpzaam	nuchter	eerlijk	vrolijk	leven dig	origineel	modern	expressief
omdat Merk X: ....												
betrouwbaar is		0,85	0,77	0,79	0,86	0,74	0,85	0,51	0,35	0,59	0,67	0,39
zorgzaam is	0,77		0,72	0,68	0,76	0,51	0,67	0,36	0,27	0,46	0,56	0,30
vriendelijk is	0,34	0,43		0,40	0,39	0,33	0,37	0,49	0,28	0,27	0,34	0,26
betrokken is	0,52	0,58	0,54		0,58	0,47	0,50	0,31	0,23	0,38	0,43	0,26
behelpzaam is	0,73	0,76	0,67	0,56		0,50	0,62	0,40	0,37	0,44	0,46	0,22
nuchter is	0,31	0,37	0,30	0,30	0,29		0,40	0,17	0,17	0,14	0,19	0,12
eerlijk is	0,70	0,59	0,53	0,68	0,57	0,61		0,39	0,37	0,44	0,42	0,32
vrolijk is	0,05	0,05	0,18	0,05	0,04	0,03	0,08		0,25	0,09	0,15	0,19
levendig is	0,11	0,08	0,22	0,13	0,06	0,02	0,05	0,28		0,14	0,17	0,23
origineel is	0,09	0,15	0,21	0,15	0,18	0,27	0,25	0,26	0,26		0,30	0,28
modern is	0,24	0,15	0,22	0,27	0,11	0,13	0,09	0,27	0,24	0,25		0,20
expressief is	-0,02	0,04	0,06	0,00	-0,02	-0,02	0,02	0,18	0,16	0,02	0,08	

• De respondenten konden aangeven of ze het met een voorgestelde oorzaak-gevolg relatie eens waren (score '1'), of ze vonden of er geen verband was (score '0') of dat ze vonden dat de voorgestelde oorzaak het bijbehorende gevolg juist verhinderde (score '-1'). Een score van 0,85 is de netto uitkomst van het percentage mensen dat het met deze relatie eens was minus het aantal mensen dat het er niet mee eens was. De getallen in deze tabel zijn de gemiddelden daarvan over alle respondenten. De relatieskerkten in deze tabel verschillen significant van nul (t-test, tweezijds,  $p \leq 0,05$ ), met uitzondering van de cursief gedrukte waarden.

(eerste kolom) omdat Merk X levendig is (het op drie na onderste getal in de eerste kolom van Tabel 3), vulden 15 % van de respondenten die deze vraag beantwoordden in “ja, juist” en 4,5 % “nee integendeel. Netto levert dat 10,5 % op, vandaar het getal 0,11 in Tabel 3.

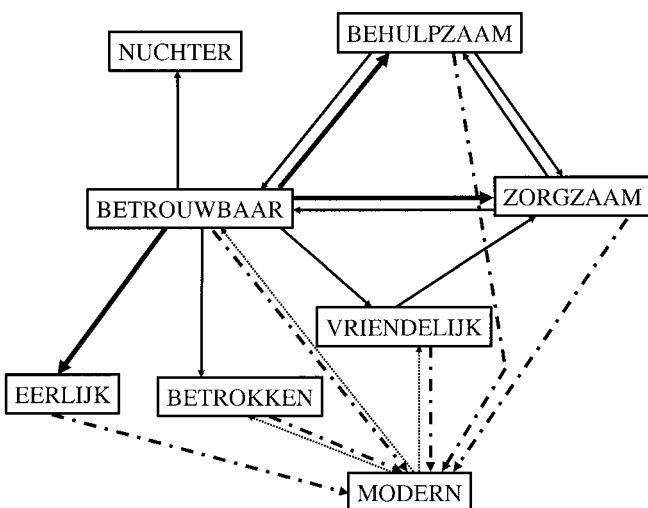
Tabel 3 bevestigt dat “betrouwbaarheid” een krachtig stempel drukt op het algemene beeld dat de consument heeft van Merk X. De enige relaties waar 80 % of meer van de consumenten het mee eens is zijn alledrie relaties waar “betrouwbaarheid” als oorzaak fungeert: namelijk de relaties tussen betrouwbaarheid en “zorgzaam”, “behulpzaam” en “eerlijk” (zie eerste regel in tabel 3). De drie dikkere pijlen in Figuur 2 geven dat weer. De dunnere pijlen in Figuur 2 staan voor de relaties waar minimaal 70 procent van de respondenten “ja, juist” geantwoord heeft.



*Figuur 2. De meest dominante causale relaties tussen de eigenschappen van Merk X.*

Figuur 2 laat zien dat “betrouwbaarheid” centraal is bij het uitleggen waarom Merk X is zoals het is, op de voet gevolgd door “behulpzaam”, “zorgzaam” en “eerlijk”. De meer perifere eigenschappen “nuchter”, “vriendelijk” en “betrokken” worden als gevolgen van deze vier centrale aspecten gezien. Het grote voordeel van de hier gepresenteerde benadering van merkessentie is, dat niet alleen duidelijk wordt welke eigenschappen het meest essentieel zijn, maar ook waarom die het meest essentieel zijn. Doordat de onderlinge samenhang van de eigenschappen wordt blootgelegd is precies te zien welke eigenschap van welke andere eigenschap afhangt.

Dit inzicht in hoe de eigenschappen van het merk onderling samenhangen, samen met de informatie hoe essentieel elk van die eigenschappen is, geeft belangrijke informatie over de speelruimte die het management heeft bij het introduceren van nieuwe eigenschappen voor merk X. Van de vijf eigenschappen die het management overwogen heeft lijken “modern”, “origineel” en “vrolijk” het best bij de al aanwezige eigenschappen te passen. Tabel 3 laat zien dat ruim 50 % van de respondenten het erover eens is dat Merk X ook vrolijk, origineel en modern kan zijn, omdat het betrouwbaar is. Deze drie eigenschappen worden dan ook alledrie gezien als gevolgen van de meest centrale eigenschappen. De grootste voorzichtigheid zou betracht moeten worden wanneer het management “expressief” of “levendig” als nieuwe eigenschap zou willen neerzetten.



*Figuur 3. De relaties van “modern” met de bestaande eigenschappen.*

Het kwadrant linksonder in Tabel 3 laat zien de respondenten niet echt in een aansluiting tussen “expressief” en de meest essentiële eigenschap van Merk X, “betrouwbaarheid”, lijken te geloven. Het opvoeren van eigenschappen die niets toevoegen aan de essentie van een merk brengt een inherent verwateringsgevaar met zich mee (Keller, 2003). Nieuwe eigenschappen die niet de essentie van een merk ondersteunen maar wel andere, minder essentiële eigenschappen (mede) veroorzaken kunnen immers op termijn andere eigenschappen essentiëler doen overkomen dan de oorspronkelijke eigenschappen. Dit is vanuit het oogpunt van het handhaven van merkmeewarde niet aan te bevelen. Bij goed merkenmanagement zou immers de kern van merk zoveel mogelijk versterkt moeten worden (Keller, 2003).

Uitgaande van Tabel 3 zouden wij het management van Merk X daarom aangeven, dat zij de eigenschappen “levendig” en “expressief” beter niet naar voren kunnen brengen. De eigenschappen “origineel” en “modern” kunnen goed aan het merk kunnen worden toegevoegd. “Modern” is de meest geschikte kandidaat-eigenschap om toe te voegen aan Merk X. Tabel 3 laat zien, dat deze eigenschap de meest essentiële huidige eigenschap van Merk X, “betrouwbaar”, duidelijk meer ondersteunt dan “origineel” (0,24 i.p.v. 0,09 in de meest linkse kolom in Tabel 3). Figuur 3 laat zien, hoe “modern” aansluit op de al bestaande eigenschappen. De gestreepte lijnen in figuur 3 stellen relaties van “modern” met de zeven bestaande eigenschappen voor waar meer dan 30 % van de respondenten het al mee eens was, de fijn gestippelde lijnen werden al door minstens 20 % van de respondenten onderschreven.

## 5. DISCUSSIE

De opzet van deze studie was het vinden van een manier om essentie van een merk in kaart te brengen. Ons kernargument bij het operationaliseren van het begrip “merk-essentie” is, dat die eigenschappen, waarvan de consument gelooft dat die de meeste

andere eigenschappen van dat merk veroorzaken de essentie van dat merk vormen. De uitkomsten in Tabel 2 ondersteunen dit kernargument. Hiermee is voor het eerst aangetoond dat de door Ahn (1998) ontwikkelde “causale status hypothese” uit de cognitieve psychologie ook van toepassing is op merken. In deze benadering is essentie geen kwestie van “alles of niets”, maar een kwestie van gradatie. De hier voorgestelde methode maakt het mogelijk om na te gaan, of de ene eigenschap essentiëler is dan de andere. Erg nuttig is daarbij de gedetailleerde informatie over waarom de ene eigenschap essentiëler is dan de andere – het is nu mogelijk het hele onderliggende patroon van causale relaties in kaart te brengen.

Deze gedetailleerde informatie over causale relaties heeft het mogelijk gemaakt, “modern” – en niet “origineel” – als de meest geschikte eigenschap voor de nieuwe positionering van Merk X eruit te filteren. Maar Tabel 3 geeft vooral ook inzicht in hoe “modern” in een reclamecampagne in te bedden is. De rechter bovenhoek van Tabel 3 laat zien, in hoeverre respondenten het eens waren met relaties waar gesuggereerd werd dat nog niet bestaande eigenschappen het gevolg waren van de huidige eigenschappen van Merk X. Die gegevens bieden meerdere aangrijpingspunten, hoe de nieuwe eigenschappen aan de hand van de al bestaande eigenschappen te onderbouwen zijn. Tabel 3 laat bijvoorbeeld zien dat Merk X ook “modern” is omdat het al betrouwbaar is – deze relatie heeft de hoogste score in de rechter bovenhoek van tabel 3 (0,67).

Methodologisch gaat deze studie een stap verder dan Ahn (1998): De hier voorgestelde methode bepaalt empirisch het associatiepatroon dat mensen erop na houden. Deze benadering wijkt af van eerder cognitief psychologisch onderzoek waar de causale status hypothese getoetst wordt door de causale relaties tussen de eigenschappen te manipuleren en vervolgens het effect van deze manipulaties te toetsen (Ahn, 1998; Ahn et al., 2000, Sloman et al. 1998). Onze benadering maakt het mogelijk het patroon van causale relaties tussen merkeigenschappen in kaart te brengen. Het feit dat wij de causale status hypothese konden bevestigen aan de hand van patronen die bijvoorbeeld door figuur 2 in kaart gebracht worden, opent perspectieven voor het in kaart brengen van de “mentale blauwdrukken” die consumenten in hun hoofd hebben, en die zo nuttig zijn voor het opstellen van merkenbeleid. Tot nu toe was dat een van de grote bottlenecks bij merkenonderzoek (Keller, 2003: 430).

Toch zijn er nog belangrijke vragen voor vervolgonderzoek blijven liggen. Ten eerste heeft de gebruikte onderzoeksmethode als nadeel dat de uitvoering ervan in het veld veel inspanning vereist en ongeschikt is voor telefonische enquêtes. Respondenten wordt namelijk gevraagd een hele matrix van causale relaties in te vullen. Een uitdaging voor toekomstig onderzoek, is hier methoden voor te vinden die dezelfde informatie geven, maar toch veel respondentvriendelijker overkomen. Een tweede vraag voor vervolgonderzoek betreft het verband tussen “essentie” en “onderscheid”. Het is erg verleidelijk om “essentie” aan “onderscheid” te koppelen. Er zijn in de literatuur bij tijd en wijle felle discussies gaande over het verschil tussen “essentie” en “onderscheid” (zie bijvoorbeeld Sayer, 1997). Merk X zou een aantal van deze eigenschappen wel eens kunnen delen met de naaste concurrenten. De onderscheidende eigenschappen hoeven niet per se de meest oorzakelijke eigenschappen te zijn. Om terug te komen op het eerder gegeven voorbeeld van de verzekeraar: deze kan zich door de relatief perifere eigenschap “bereikbaar” met succes tegen zijn concurrenten afzetten. Het zou heel



zinnig zijn om dit voor merken te onderzoeken, te meer omdat bij merkenmanagement het onderscheid met concurrerende merken een centrale plaats inneemt.

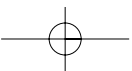
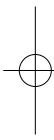
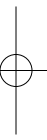
## 6. CONCLUSIES

Onze benadering van merkessentie helpt managers en marktonderzoekers vast te stellen welke eigenschappen de essentie van een merk vormen. Dit geeft hen handvatten voor de communicatie over het merk: het is makkelijker om de meest essentiële eigenschappen van een merk te identificeren en vooral ook om doelgericht om te gaan met waarom de consument die eigenschappen zo essentieel vindt. Een heel praktische toepassing is het toevoegen van nog niet bestaande, of althans in de positionering van het merk nog niet uitgewerkte eigenschappen. In de vragenlijst worden nieuwe eigenschappen eenvoudig tussen al bestaande beoordeeld en niet pontificaal geïntroduceerd als mogelijke productwijziging. Deze manier van voorleggen vermijdt alle risico op emotionele reacties die daaruit voort kunnen vloeien. Dan kan, zonder dat ook maar de suggestie van een productwijziging of herpositionering gemaakt is, de acceptatie van mogelijke nieuwe merkeigenschappen verkend worden, en kan alvast bekeken worden of die nieuwe eigenschappen in de ogen van de consument consistent zijn met wat ze al over dat merk menen te weten. Hierdoor is een gedetailleerd en verfijnd aftasten van de positioneringsperspectieven voor een merk mogelijk. Dit geeft de hier gepresenteerde methode een duidelijke meerwaarde. Merk X is zelf een voorbeeld van een situatie waarin dit laatste erg plezierig is. Het merk kent nog een groep trouwe gebruikers, maar om het enthousiasme van de nieuwere generaties te winnen wordt revitalisatie overwogen (cf., Wansink, 1997). De hier beschreven methode vormt een hulpmiddel om het terrein voor een dergelijke revitalisatie te verkennen zonder de bestaande groep trouwe gebruikers voor het hoofd te stoten. Een groot voordeel van de kennis van de causale associatiepatronen tussen de eigenschappen onderling, is dat in reclamecampagnes heel concreet kan worden uitgelegd waarom een merk is zoals het is. De onderzoeksresultaten bij Merk X bieden de eerste aanwijzing dat de hier gepresenteerde benadering wel eens een antwoord kunnen geven op Keller's (1993; 2003) vraag, hoe we de samenhang en de consistentie van het imago van een merk in verschillende marktsegmenten en over langere tijdsperiodes kunnen managen.

## LITERATUUR

- Aaker, David A., 1996. *Building strong brands*, New York: The Free Press.
- Aaker, J.L., 1997. Dimensions of brand personality. *Journal of Marketing Research*, 34, 347-356.
- Ahn, W.K., 1998. Why are different features central for natural kinds and artifacts? The role of causal status in determining feature centrality. *Cognition*, 69, 135-178.
- Ahn, W.K., N.S. Kim, M.E. Lassaline, en M.J. Dennis, 2000. Causal status as a determinant of feature centrality. *Cognitive Psychology*, 41, 361-416.
- Asch, S.E., 1946. Forming impressions of personality. *Journal of Abnormal and Social Psychology*, 41, 258-290.
- Baltesen, F., 2001. Alle zeepmerken zijn even goed! *NRC-Handelsblad*, 3 November 2001.
- Davies, G., Chun, R., Vinhas da Silva, R., Roper, S., 2003. *Corporate Reputation and competitiveness*. London: Routledge.

- De Chernatony, L., 2001. *From brand vision to brand evaluation. Strategically building and sustaining brands*. Oxford: Butterworth Heinemann.
- Haig, M., 2003. *Brand Failures*. London: Kogan Page.
- Keller, K. L., 1993. Conceptualizing, measuring and managing customer-based brand equity. *Journal of Marketing*, 57 (January), 1-22.
- Keller, K.L., 2003. *Strategic Brand Management: Building, measuring and managing brand equity*. 2<sup>nd</sup> edition. Upper Saddle River (NJ): Prentice Hall/Pearson Education.
- Krishnan, H. S., 1996. Characteristics of memory associations: a consumer-based brand equity perspective, *International Journal of Research in Marketing*, 13, 389-403.
- Michel, G. en T. Ambler, 1999. Establishing brand essence across borders. *Journal of Brand Management*, 6, 333-345.
- Morel, K. P. N., 2002. Wat is het? Een motor, een auto, of eh...? Consumentencategorisatie van hybride producten. In: A.E. Bronner, P. Dekker, J.C. Hoekstra, E. de Leeuw, Th. Poiesz, K. de Ruyter, A. Smidts (red): *Ontwikkelingen in het marktonderzoek. Jaarboek 2002 Markt-OnderzoekAssociatie*,
- Murphy, G.L. en D. L. Medin, 1985. The Role of Theories in Conceptual Coherence. *Psychological Review*, 92, 289-316.
- Oliver, T., 1986. *The Real Coke, the Real Story*. New York: Random House.
- Pinson, C., 1986. An implicit product theory approach to consumers' inferential judgments about products", *International Journal of Research in Marketing*, 3, 19-38.
- Putnam, H., 1977. Is Semantics Possible? In: Schwartz, S.P. (ed.): *Naming, Necessity and Natural Kinds*. Ithaca (N.Y.): Cornell University Press.
- Riezebos, R., 2003. *Brand management, a theoretical and practical approach*. Upper Saddle River: Prentice Hall.
- Rosa, J.A., Porac, J.F., Runser-Spanjol, J. & Saxton, M.S., 1999. Sociocognitive dynamics in a product market. *Journal of Marketing*, 63, (special issue), 64-77.
- Rosenthal, R. 1991. *Meta-analytic procedures for social research*. Applied Social Research Methods Series, Vol. 6. Beverly Hills: Sage.
- Sayer, A., 1997. Essentialism, social constructionism, and beyond. *The Sociological Review*, 45, 453-487.
- Sloman, S.A., B.C. Love and W.K. Ahn, 1998. Feature centrality and conceptual coherence. *Cognitive Science*, 22, 189-228.
- Wansink, B., 1997. Making Old Brands New. *American Demographics*, 19, 12, 53-58.



## 8. *Marketing Maatwerk*

B.G.C. DELLAERT\*

### SAMENVATTING

Dit hoofdstuk biedt een overzicht van recente ontwikkelingen in het onderzoek op het vlak van marketing maatwerk. Op basis van dit overzicht worden de condities verkend waaronder producenten en consumenten bereid zullen zijn om gezamenlijk tot marketing maatwerk te komen. In de analyse wordt met name het perspectief van de consument als eindafnemer in het productieproces centraal gesteld.

### 1. INLEIDING

Maatwerk speelt een groeiende rol binnen allerlei deelgebieden van de marketing. Nog niet zo lang geleden was het zeer kostbaar en tijdrovend om verschillende producten en diensten aan te bieden voor elke individuele consument. Snelle ontwikkelingen in de ICT en de productietechnologie hebben maatwerk echter meer en meer toegankelijk gemaakt voor fabrikanten en dienstverleners. Zo gebruikt de firma Nike bijvoorbeeld een combinatie van flexibele fabricagetechnologie en moderne communicatiemiddelen om de consument de mogelijkheid te geven om zelf de kleuren en materialen van zijn of haar eigen sportschoenen uit te kiezen op het Internet.

Toch zal maatwerk waarschijnlijk niet voor alle marketingvraagstukken de beste oplossing bieden. Het doel van dit hoofdstuk is daarom om te verkennen onder welke condities het het meest waarschijnlijk is dat producenten en consumenten gezamenlijk tot marketing maatwerk zullen komen. Bij deze analyse wordt met name het perspectief van de consument -als eindafnemer in het productieproces- centraal gesteld. De werkdefinitie van marketing maatwerk die het uitgangspunt vormt voor de analyse is: Het benutten van een voortbrengingsketen met als doel producten en diensten te genereren die door elke consument afzonderlijk kunnen worden gedefinieerd, afgenomen en geconsumeerd.

\* Dit hoofdstuk is een verkorte versie van de inrede die de auteur heeft uitgesproken op 12 september 2003 aan de Universiteit Maastricht, ter aanvaarding van het ambt van hoogleraar Economische Wetenschappen en Bedrijfskunde aan de Faculteit der Economische Wetenschappen en Bedrijfskunde.

## 2. DE ORGANISATIE EN MAATSCHAPPELIJKE CONCEPTUALISATIE VAN MAATWERK

Een belangrijke conceptuele grondslag van veel op maatwerk gerichte voortbrengingsprocessen is het denken in modulaire concepten. Modules zijn min of meer zelfstandige componenten die binnen een bepaalde structuur relatief snel en gemakkelijk kunnen worden gecombineerd om zo op een flexibele wijze een groter geheel te kunnen opbouwen. In een grondige analyse van de grote rol die modulaair denken al lange tijd speelt in de cultuur van de V.S. noemt Blair (1988) een aantal interessante en prikkelende voorbeelden van modulaire systemen die in de Amerikaanse cultuur zijn ontstaan en die in loop van de tijd ook in de rest van de wereld zijn aangeslagen: Wolkenkrabbers, die zijn opgebouwd uit modulaire verdiepingen, en die naar believen tot steeds grotere hoogte kunnen worden uitgebreid, Jazz muziek, met zijn vaste schema's waarbinnen vrijelijk kan worden geïmproviseerd en samengespeeld, en ook het moderne academisch curriculum, dat is opgebouwd uit studieblokken die kunnen worden gecombineerd om gezamenlijk een gehele opleiding te vormen. Gemeenschappelijk aan deze voorbeelden is dat zij een vast omlijnde basisstructuur hebben waarbinnen in grote vrijheid met componenten kan worden geschoven en gecombineerd.

Ook op het vlak van de productietechnologie dragen modulaire structuren in hoge mate bij aan de flexibiliteit en snelheid van productie (Sanchez 1999). Ontwerpers houden rekening met het feit dat producten in componenten moeten kunnen worden geproduceerd en machines worden zo ontwikkeld dat ze in de productie snel van het ene onderdeel over kunnen schakelen op het andere onderdeel (bijv., Anderson 1997). Met name de rol van de digitalisering van producten en diensten is hierbij van groot belang, omdat computersturing en computer gebaseerde informatie uitwisseling het mogelijk maakt om grote delen van de productie en dienstverlening te standaardiseren en uit te besteden.

Organisatiestructuren passen zich aan deze ontwikkelingen aan, en ook op dit vlak is het verschijnsel modularisering terug te vinden, met name in de verschuiving van traditionele verticale, soms zelfs geheel geïntegreerde voortbrengingsketens, naar open netwerkstructuren bestaande uit meerdere bedrijven die onderling een aantal vaste waarden, productie afspraken en technologische kennis delen (bijv. Achrol en Kotler 1999). In dergelijke netwerkstructuren vormen de bedrijven als het ware organisatorische modules die dankzij geavanceerde ICT naar aanleiding van specifieke klantenvragen snel en flexibel kunnen samenwerken in een groot aantal verschillende verbanden. Deze netwerkstructuren breiden zich in toenemende mate ook uit in de richting van integratie van klanten als actieve participanten in het productieproces (bijv. Prahalad en Ramaswamy 2000).

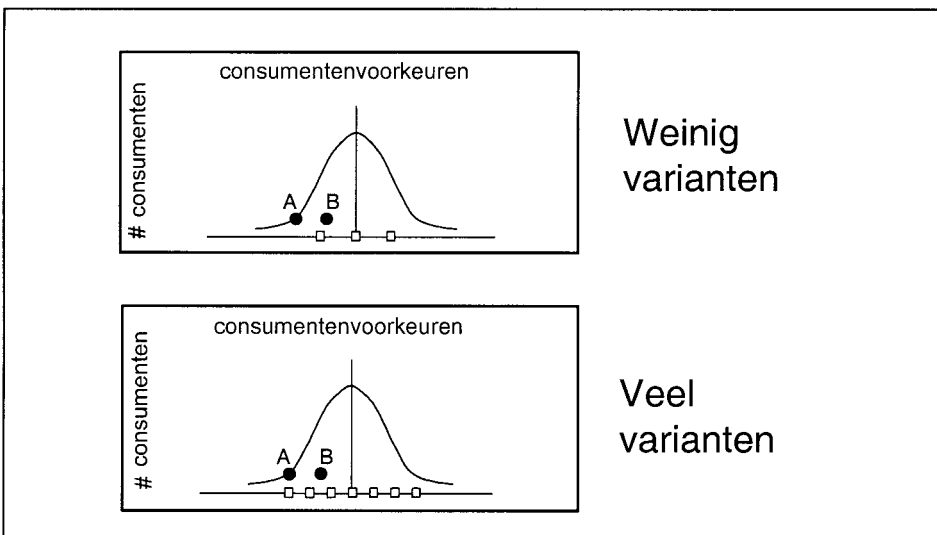
Tot slot zijn ook maatschappelijk gezien modulaire structuren te herkennen die de hedendaagse flexibiliteit van de productie reflecteren en ondersteunen. Deze structuren zijn bijvoorbeeld zichtbaar in trends als individualisering, door Beck omschreven als de institutionalisering van het moderne individualisme (Beck en Beck-Gernsheim, 2002), en in Ritzer's McDonaldisering (Ritzer 1998), als synoniem voor een hedendaags efficiënt systeem voor de organisatie van consumptie. In beide trends is er sprake van (zeer verschillende) onafhankelijk te variëren componenten –respectievelijk het individu en de productiehandeling— die dankzij een systeem van maatschappelijke regels en ondersteund door de technologie toch flexibel en efficiënt kunnen samengaan. Ook dergelijke maatschappelijke ontwikkelingen vonden veelal hun oorsprong in de V.S. maar bleken ook elders in de wereld succesvol te zijn.

### 3. WAT ZIJN DE VOORDELEN VAN MARKETING MAATWERK?

#### 3.1. Inleiding

Hoewel het tot stand brengen van marketing maatwerk dus aansluit bij brede maatschappelijke trends en bovendien organisatorisch en technologisch gezien steeds gemakkelijker wordt, is het nog de vraag of consumenten behoefte hebben aan maatwerkproducten en -diensten, en zo ja, of deze door de markt zullen worden voortgebracht.

Een belangrijk voordeel van maatwerk voor de consument is dat het er voor kan zorgen dat producten en diensten dicht bij de smaak van elk individu kunnen aansluiten. Ter illustratie kunnen bijvoorbeeld de preferenties van alle individuen voor een bepaald product (bijvoorbeeld PC's) op een rechte lijn worden weergegeven die de markt met alle modelvarianten representeert oplopend in kwaliteit en prijs. Wanneer daar de consumenten 'A' en 'B' op worden geplaatst op de locatie van hun 'ideale' product, dan kunnen A en B naarmate er meer verschillende producten beschikbaar zijn, producten kopen die dicht bij hun voorkeur liggen. Als er werkelijk maatwerk wordt geboden, dan kan iedere consument zelfs het product kopen dat precies op zijn of haar voorkeurslocatie ligt.



Figuur 1. Productvoorkeuren van consumenten en maatwerk.

Ook voor producenten zijn er voordelen aan marketing maatwerk verbonden zijn. Eén van de grootste voordelen is dat producenten lagere voorraadkosten hebben, omdat er alleen nog op bestelling wordt geleverd (bijv. Tseng en Piller 2003). Alle producten die worden gemaakt zijn in principe al verkocht voordat ze gemaakt zijn, waardoor het aanhouden van nog niet verkochte voorraden minder noodzakelijk wordt. Bovendien kan maatwerk naar verwachting de concurrentiedruk tussen producenten verkleinen (bijv. Pine, Peppers en Rogers 1995). De gedachte achter deze redenering is dat produ-

cent en klant gezamenlijke kennis opbouwen die door de klant maar deels overdraagbaar is aan andere producenten. Hierdoor wordt het voor klanten lastiger om naar andere producenten over te stappen, hetgeen vervolgens de concurrentiedruk verlaagt.

### *3.2. Is een maatwerk benadering superieur aan traditionele massaproductie?*

Een belangrijke voorwaarde voor marketing maatwerk is dat het meten en verwerken van consumentenpreferenties wordt geïntegreerd in de voortbrengingsketen (Wind and Rangaswamy 2001). Conceptueel moet hiertoe de traditionele splitsing van het waardecreatie proces in enerzijds een voortbrengingsdeel en anderzijds een consumptie- of 'waardevernietigings'-deel worden doorbroken. In plaats daarvan dient de consument te worden gezien als een actieve partner in de op maatwerk gerichte marketing- en productieprocessen.

Consumenten kunnen hun rol van partner in verschillende stadia van de voortbrengingsketen vervullen. Hierbij kunnen globaal vijf functionele stadia in de voortbrenging van een product of dienst worden onderscheiden: R&D, Facilitering, Productie, Consumptie en Nazorg. Bij het tot stand brengen van maatwerk kan de consument in elk van deze stadia actief zijn. Bijvoorbeeld in de productontwikkeling door nieuwe ideeën aan te reiken (von Hippel 2001), in de productie door het moment en de componenten van productie te bepalen (Pine en Gilmore 2000), in de facilitering door de methode en locatie van betaling en aflevering te kiezen, en in de nazorg door suggesties ter aanpassing aan te reiken, bijv. in gebruikersgemeenschappen (Fattah 2002).

Bij het onderzoeken van de vraag of, en zo ja in welke mate, consumenten en producenten baat hebben bij deze nieuwe invulling van de keten en de introductie van een op maatwerk gericht voortbrengingsproces, is het begrip 'marketing efficiency' van Wenerfelt (1984) een bruikbaar concept. In zijn analyse poneert Wenerfelt dat op termijn alleen die marketingbenaderingen in de markt overleven die tenminste Pareto optimaal zijn ten opzichte van andere benaderingen. Dat wil zeggen dat er alleen marketingbenaderingen in de markt kunnen bestaan waaraan geen enkel aspect verbeterd kan worden zonder dat andere aspecten van deze benaderingen verslechteren. Dit betekent bijvoorbeeld dat in de markt een flexibele, snelle en dure productie methode wel naast een reguliere en goedkope productie methode kan overleven, maar niet naast een tweede flexibele, snelle productie methode die wel goedkoop is.

Marketing maatwerk moet daarom tenminste meerwaarde voor consumenten en producenten gezamenlijk opleveren ten opzichte van meer traditionele marketingsystemen wil het als bedrijfsaanpak kunnen overleven in de markt. Meer concreet houdt dit veelal in dat géén van de betrokken partijen deel zal willen nemen aan een op maatwerk gerichte marketing benadering wanneer er voor elk afzonderlijk niet tenminste een gelijke opbrengst is te behalen als bij de reeds bestaande marketingbenaderingen. Zowel de keuze van producenten als van consumenten om deel te nemen aan op maatwerk gerichte voortbrengingsketens vormen daarom een interessant onderzoeksterrein met belangrijke implicaties.

## 4. DE CONSUMENTENKEUZE VOOR MARKETING MAATWERK

### 4.1. De keuze om deel te nemen aan marketing maatwerk

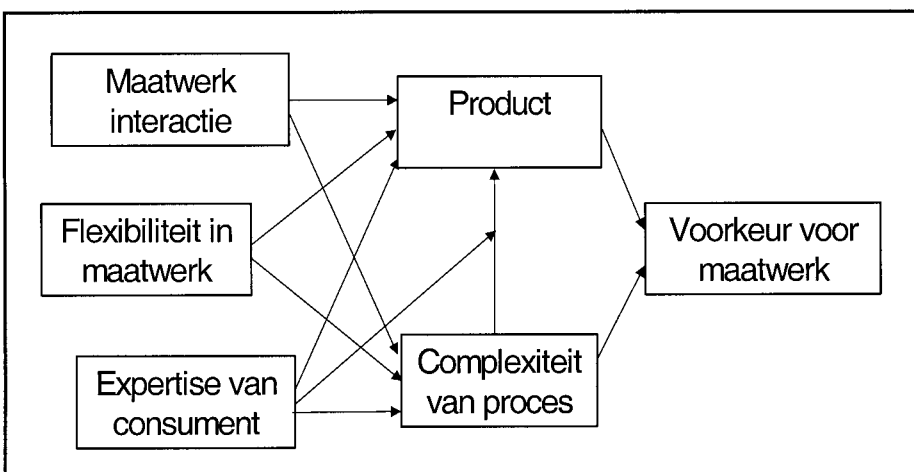
In enkele recente onderzoeken zijn de consumentenkeuze om deel te nemen aan marketing maatwerk nader onder de loep genomen. Uitgangspunt in deze analyses zijn meestal nutsmodellen van consumentenkeuzegegedrag (bijv. McFadden 1986). In deze modellen wordt de kans dat consumenten een bepaald product aanschaffen of een bepaalde dienst gebruiken uitgedrukt als een functie van de kenmerken van het product of de dienst, de kenmerken van concurrerende producten of diensten, de kenmerken van de consument en de kenmerken van mogelijk relevant omgevingsvariabelen, zoals bijvoorbeeld de technologische infrastructuur. De kans dat consumenten voor marketing maatwerk kiezen kan met een dergelijk model worden uitgedrukt. Naarmate de kenmerken van marketing maatwerk aantrekkelijker worden ten opzichte van andere reeds bestaande marketing alternatieven, neemt de kans toe dat voor maatwerk wordt gekozen.

Bij het analyseren van de consumentenkeuze voor marketing maatwerk is het daarom belangrijk om de voor- en nadelen van maatwerk voor consumenten goed in kaart te brengen. Het blijkt dat er naast het voordeel van een product dat beter aansluit bij de eigen smaak, voor consumenten ook nadelen aan marketing maatwerk zijn verbonden, m.n. in de vorm van mogelijke additionele kosten en onzekerheden die horen bij het kopen van producten en diensten op maat. Een van de belangrijkste niet-financiële nadelen voor consumenten die zijn verbonden aan marketing maatwerk zijn de zogenaamde complexiteitskosten. Complexiteit vormt een nadeel voor consumenten omdat het menselijk brein beperkt is in zijn cognitieve capaciteiten en meestal niet in staat is alle aspecten van een beslissing te overzien (Simon and Newell 1972). Bij de aanschaf van producten en diensten op maat moeten consumenten echter noodgedwongen een zeer groot aantal kenmerken vergelijken en beoordelen, vaak zonder dat ze alle gevolgen van hun keuzes kunnen kennen (bijv. Huffman en Kahn 1998). Een dergelijke complexiteit wordt door consumenten als onaantrekkelijk gezien en verlaagt de kans dat ze voor een maatwerk traject zullen kiezen.

In twee recente projecten is de rol van complexiteit in de consumentenkeuze om wel of niet voor een maatwerkeraanpak te kiezen nader onderzocht. In de eerste studie is een nutsmodel ontwikkeld om de invloed te schatten van verschillende kenmerken van de marketing maatwerk benadering (zoals bijvoorbeeld het aantal modules waarbinnen de consument een keuze kan maken), de deskundigheid van consumenten ten aanzien van het product, de complexiteit van het maatwerkproces en de verwachte aantrekkelijkheid van het product dat kan worden samengesteld, op de kans dat consumenten een bepaald systeem van maatwerk zullen accepteren (Dellaert en Stremersch 2005). Figuur 2 illustreert de structuur van dit model.

Dit model is getoetst op consumentengegevens die zijn verzameld via het Internet. Respondenten stelden in denkbeeldige situaties hun ideale PC samen op verschillende websites die leken op de bestaande Dell website. Op basis van recent ontwikkelde modelbenaderingen en schattingstechnieken (zie bijvoorbeeld Ashok, Dillon en Yuan 2002 en Train 2003) kan het gelijktijdig effect van deze verschillende kenmerken op de keuzes van worden consumenten bepaald en bovendien worden nagegaan in hoeverre de smaken van verschillende consumenten onderling afwijken.





*Figuur 2. De structuur van de consumentenvoorkeur voor maatwerk.*

Uit de resultaten van dit project blijkt dat de complexiteit van het maatwerkproces inderdaad het verwachte negatief effect heeft op de kans dat consumenten er aan deel zullen nemen. Bovendien blijkt dat complexiteit ook een negatief effect heeft op het oordeel van de consument over het product dat zij in het maatwerkproces samenstellen. Hoe complexer de taak hoe minder aantrekkelijk men het product vindt. Tenslotte blijkt dat de deskundigheid van de consument deze effecten beïnvloedt. Deskundige consumenten hebben een grotere waardering voor extra keuzemogelijkheden en zijn minder gevoelig voor complexiteit.

In een tweede project is verder in detail onderzocht of andere onderliggende latente factoren naast complexiteit mede bepalen of consumenten maatwerk aantrekkelijk vinden (Dellaert, Goebel en Dabholkar 2004). De productcategorie die in deze studie is gebruikt is kleding. Er zijn verschillende kledingzaken op het Internet waar kleding op maat kan worden gekocht, zoals bijvoorbeeld de firma Lands End (Ives and Piccoli 2003). Uit de resultaten blijkt dat in lijn met recente uitbreidingen van het zogeheten Technology Acceptance Model (Davis, Bagozzi en Warshaw 1989), naast complexiteit of gebruiksgemak, ook het plezier in het gebruiken van een maatwerkproces een belangrijke rol speelt. Bovendien maken consumenten zich zorgen over de mate waarin zij het proces kunnen beheersen (Dabholkar 1996). Ook blijkt dat in de context van marketing maatwerk het concept 'bruikbaarheid' doeltreffend kan worden gemeten in termen van de verwachte aantrekkelijkheid van het product dat de consument kan samenstellen.

#### *4.2. De productkeuze in een marketing maatwerk context*

Naast de keuze om al dan niet deel te nemen in marketing maatwerk, is ook een tweede type consumentenkeuze van belang, namelijk de keuze van een specifiek product of een specifieke dienst binnen de maatwerk context. Recentelijk zijn econometrische modellen ontwikkeld om dergelijke keuzes te kunnen beschrijven en voorspellen in een duidelijk omlijnd maatwerkproces (Liechty, Ramaswamy en Cohen 2001).

Eerder onderzoek heeft aangetoond dat de productkeuze van consumenten echter kan

afhangen van de wijze waarop de vergelijking tussen producten wordt gepresenteerd (bijv. Kahneman en Tversky 1984, Payne 1982). Het is dan ook aannemelijk dat verschillen in de processen om consumentenvoorkeuren in kaart te brengen en de wijze waarop bijvoorbeeld maatwerk modules worden vergeleken ook zullen leiden tot verschillen in productkeuzes.

Recent onderzoek bevestigt deze verwachting (Huber, Ariely en Fischer 2001). Een terrein waarop met name veel onderzoek is gedaan is de rol van persoonlijke aanbevelingen voor producten die zo goed mogelijk bij de voorkeuren van de consument moeten aansluiten (bijv. Diehl, Kornish en Lynch 2003). Bijvoorbeeld websites als [activebuyersguide.com](http://activebuyersguide.com) maken dit type van diensten mogelijk. In de eerste plaats blijkt uit dit onderzoek dat aanbevelingen consumenten enorm kunnen helpen bij het kiezen van die producten die zo dicht mogelijk tegen hun voorkeuren aanliggen. Bovendien kost het ze minder moeite om dergelijke producten te selecteren (Häubl en Trifts 2000). Daarnaast blijkt er een aantal verschuivingen in de voorkeuren plaats te vinden. Zo blijkt uit een recent onderzoek dat de voorkeuren van consumenten significant beïnvloed worden door de kenmerken waarover zij naar hun voorkeur zijn gevraagd (Häubl en Murray 2003). Ook is bekend dat consumenten in een lijst met aanbevelingen andere heuristische gebruiken dan wanneer zij op basis van toeval gerangschikte producten onder ogen krijgen (Häubl, Dellaert, Murray en Trifts 2004).

## 5. DE VERDELING VAN MEERWAARDE TUSSEN PRODUCENT EN CONSUMENT

Wanneer er door de benutting van marketing maatwerk daadwerkelijk meerwaarde in de keten tot stand wordt gebracht, dan doet zich vervolgens de vraag voor hoe die gerealiseerde meerwaarde wordt verdeeld. Zal de meerwaarde bijvoorbeeld tot gevolg hebben dat de consument voor hetzelfde geld en dezelfde moeite een aantrekkelijker product krijgt, of wordt de meerwaarde eerder vertaald in een hogere winstmarge voor de producent omdat deze het maatwerkproduct tegen een hogere prijs kan verkopen?

Een veel gehoord argument voor marketing maatwerk is dat het de concurrentie tussen voortbrengingsketens vermindert (Pine, Peppers en Rogers 1995). De belangrijkste reden hiervoor is dat er leereffecten plaatsvinden, waarbij consumenten en producenten gaandeweg expliciete en impliciete kennis verwerven die moeilijk overdraagbaar is en die transacties tussen consument en ketenpartners vergemakkelijken en verbeteren. In dit geval wordt met name de positie van de producent ten opzichte van de consument versterkt, omdat de overstapkosten voor de consument toenemen naarmate er meer maatwerk wordt geboden. Denk bijvoorbeeld aan de moeite die het (nu al) kost om van verzekering of bank te veranderen en hoe die nog zou worden vergroot als u bij elke organisatie opnieuw een individuele op maat gemaakte website zou moeten samenstellen, uw persoonlijke contactpersoon zou moeten bijpraten en uw persoonlijke polis- of hypotheekvoorwaarden zou moeten vaststellen.

Deze strategische interactie tussen consumenten en producenten in marketing maatwerk kan ook worden bestudeerd met behulp van speltheorie. Een dergelijke analyse laat zien dat het leveren van maatwerk inderdaad de concurrentiedruk kan verminderen, maar alleen onder bepaalde voorwaarden (Syam en Dellaert 2001). Wanneer namelijk elk bedrijf in een markt maatwerk aan zou bieden voor elk kenmerk van de producten, zou dit de concurrentie sterk vergroten in plaats van verkleinen. De reden is dat in dat

geval elk product rechtstreeks concurreert met elk ander product (de producten bevin- den zich allemaal direct op het 'ideaal' punt van de consumer, zie figuur 1). Er is geen enkel competitief voordeel meer, omdat alle producenten alle consumenten precies op maat kunnen bedienen. Daarom is een belangrijke conclusie dat bedrijven maatwerk slechts voor een deel van alle productkenmerken zullen aanbieden. Bovendien zullen zij dit zoveel mogelijk voor dezelfde kenmerken doen, om zo ruimte voor concurrentie- voordeel op de andere kenmerken te kunnen bewerkstelligen.

Een dergelijke analyse laat verder zien dat ondanks dit concurrentievoordeel tussen producenten de prijs van maatwerkproducten voor consumenten grotendeels gelijk zal zijn aan reeds bestaande producten in de markt. De reden hiervoor is dat de producent bij het leveren van maatwerk afhankelijk is van de input van de consument. Wanneer de prijs van maatwerkproducten te hoog wordt zal de consument niet bereid zijn om de benodigde productspecificaties door te geven aan de producent en in plaats daarvan kiezen voor een goed standaard alternatief in de markt.

## 6. CONCLUSIE

### 6.1. *Samenvatting*

Op basis van dit overzicht van verschillende aspecten van marketing maatwerk en hoe consumenten en producenten daarmee omgaan, kan worden gesteld dat marketing maatwerk een waardevol en realistisch marketing model vormt in een groot aantal ver- schillende contexten. Consumenten en producenten kunnen profiteren van de invoering van marketing maatwerk en hebben dan ook verschillende redenen om eraan mee te werken. Tegelijkertijd blijken er echter ook factoren te zijn die de mogelijke rol van marketing maatwerk begrenzen. Het belang van deze factoren vormt deels een open veld voor verder onderzoek en deels een min of meer logische beperking van de rol van marketing maatwerk, waarvan het relatief belang in de markt nog moet worden vastge- steld. In de volgende paragrafen zal tot slot een aantal van deze factoren nader worden toegelicht.

### 6.2. *Smaken verschillen, maar niet altijd*

Het succes van maatwerk staat of valt grotendeels met de aanwezigheid van heterogeni- teit in de vraag. Als ieder individu écht het liefst in een groene overall zou lopen, dan valt er voor hightech flexibele productiesystemen weinig eer te behalen. Gelukkig is het zo dat op de meeste terreinen smaken verschillen en dat consumenten bijvoorbeeld hun eigen identiteit in hun kleding tot uitdrukking willen brengen.

Maar er zijn grenzen aan de behoefte om uniek te zijn. Het is waarschijnlijk dat voor sommige producten en diensten de consumentenvoorkeuren niet zo ver uiteen zullen lopen, bijvoorbeeld omdat zij grotendeels biologisch of sociaal-cultureel bepaald zijn. Ook wanneer er sterke (sociale) netwerk-externaliteiten optreden bij consumptie kan differentiatie in maatwerk onaantrekkelijk zijn voor consumenten. Op het gebied van kleding bestaat er voor de consument bijvoorbeeld het risico dat de kleding die hij of zij zelf samen stelt niet in de mode is waardoor het (deels) beoogde effect om anderen te imponeren teniet wordt gedaan. Coördinatie van de smaak kan in een dergelijke situ- atie een belangrijke taak van de producent zijn.

### 6.3. Gedeelde consumptie

Een tweede aandachtspunt is dat een groot aantal producten en diensten wordt geconsumeerd door meerdere consumenten tegelijk. Bijvoorbeeld in het gezin zijn smaken vaak verschillend en is het niet mogelijk of wenselijk dat een product op maat wordt gemaakt naar de smaak van slechts één gezinslid (Aribarg, Arora en Bodur 2002). Dergelijke conflicten doen zich ook op een groter schaal voor bij projecten zoals de inrichting van de openbare ruimte waar conflicten tussen de belangen van verschillende gebruikers haast onvermijdelijk zijn. In een recent onderzoek wordt bezien in hoeverre dergelijke conflicten deels kunnen worden beperkt door elke gebruiker (in dit geval toeristen in de stad) individuele adviezen te geven over waar en wanneer verschillende activiteiten het best kunnen worden ondernomen (Han, Dellaert, Timmermans en van Raaij 2004). In deze adviezen wordt rekening gehouden met de activiteiten van andere gebruikers waardoor de totale gebruikscapaciteit en de beleving van de ruimte kan worden verbeterd.

### 6.4. Wees of wordt geïnspireerd

Behalve de coördinatie tussen gebruikers onderling, kan ook de rolverdeling tussen producent en consument een reden zijn om maatwerk te beperken. Bijvoorbeeld op het vlak van productinnovatie is het soms de vraag of de consument de meest geschikte partij is om deze tot stand te brengen. De deskundigheid van de producent op het terrein van een bepaald product is vaak groter dan die van de consument en dit is een van de redenen waarom producenten veelal de rol van innovator op zich nemen. Een alternatieve benadering is om alleen deskundige consumenten te betrekken bij innovaties en op die manier horizontale in plaats van verticale specialisatie in de keten tot stand te brengen (von Hippel 2002).

In beide gevallen geldt echter dat het merendeel van de consumenten waarschijnlijk niet in staat is, of de behoefte heeft, om zelf innovaties tot stand te brengen. Misschien dat een schilderij van Rembrandt op maat nog voor te stellen is, of wie weet zelfs wel gerealiseerd, voor een Mondriaan is dat erg moeilijk. In beide gevallen bepaalt bovendien de kunstenaar vóóraf de stijldefinities en niet de koper zelf. Op een meer abstract niveau bezien is het haast per definitie een voorwaarde voor inspiratiebronnen dat deze niet door het individu zelf kunnen worden gedefinieerd.

Op dit vlak ontrolt zich tegenwoordig een interessant spanningsveld. Hoewel de consument geïnspireerd en uitgedaagd wil worden, wil hij of zij tegelijkertijd zoveel mogelijk greep houden op deze uitdagingen. De hedendaagse consument begeeft zich daarom in zeer gecontroleerde 'speel' omgevingen waarbinnen de uitdaging of inspiratie vaak zeer beperkt gedefinieerd is en eendimensionaal van karakter (Eco 1986). De vele extreme sporten, themaparken en verre reizen die tegenwoordig worden geconsumeerd vormen in zekere zin daarom toch modulaire ervaringen die door de consumenten naar believen kunnen worden geschakeld en die passen binnen zeer strak en geavanceerd systeem van informatiesystemen, logistieke systemen en institutionele kaders.

### 6.5. Categorijsatie en communicatie

Een laatste aspect van de voortbrengingsketen dat een mogelijke beperking vormt van marketing maatwerk is de centralisatie en afstemming in de voortbrengingsketen die

van belang is bij de communicatie over de producten of diensten die worden voortgebracht. Eén van de belangrijkste middelen om een krachtig merk op te bouwen is bijvoorbeeld de consistentie in de uitingen over dit merk (Erdem en Swait 1998). Een dergelijke consistentie vergroot de herkenbaarheid en geloofwaardigheid van het merk. Hoewel de firma Nike de consument in staat stelt om zelf zijn of haar nieuwe sportschoenen vorm te geven, moet het merk daarbij tegelijkertijd wel een gemeenschappelijk imago zien te handhaven voor al zijn producten. De merknaam en sommige stijlelementen kunnen daarom niet als maatwerk worden aan geboden en moeten gemeenschappelijk zijn voor alle producten van hetzelfde merk, hoe flexibel deze producten wat pasvorm, materiaal en kleurstelling betreft ook mogen worden aangeboden. Als Nike het toe zou staan dat ook Mike en Spike hun eigen 'swosh' zouden ontwerpen en op hun sportschoenen zetten, dan zou het waarschijnlijk met het merk snel gedaan zijn...

## LITERATUUR

- Achrol, Ravi S. en Philip Kotler (1999). "Marketing in the Network Economy", *Journal of Marketing* 63 (Special Issue), 146-163.
- Anderson, David M. (1997). "Agile Product Development for Mass Customization", Chicago: Irwin.
- Aribarg, Anocha, Neeraj Arora en H. Onur Bodur (2002). "Understanding the Role of Preference Revision and Concession in Group Decisions", *Journal of Marketing Research* 39, (August), 336-349.
- Ashok, Kalidas, William R. Dillon en Sophi Yuan (2002). "Extending Discrete Choice Models to Incorporate Attitudinal and Other Latent Variables", *Journal of Marketing Research* 39 (February), 31-46.
- Blair, John G. (1988). "Modular America: Cross-Cultural Perspectives on the Emergence of an American Way", New York: Greenwood Press.
- Beck, Ulrich en Elisabeth Beck-Gernsheim (2002). "Individualization", London: Sage Publications.
- Dabholkar, Pratibha (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 (1), 29-51.
- 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), 982-1004.
- Dellaert, Benedict G.C. en Stefan Stremersch (2005). "Marketing Mass Customized Products: striking the balance between utility and complexity", *Journal of Marketing Research* 42 (May), forthcoming.
- Dellaert, Benedict G.C., Jens Goebel en Pratibha Dabholkar (2004). "Consumer Decisions to Use Online Mass Customized Retailing", working paper, departement Marketing, Faculteit der Economische Wetenschappen en Bedrijfskunde, Universiteit Maastricht.
- Diehl, Kristin, Laura Kornish en John G. Lynch (2003). "Smart Agents: When Lower Search Costs for Quality Information Increase Price Sensitivity", *Journal of Consumer Research* 30, (June), 56-71.
- Eco, Umberto (1986). *Travels in Hyperreality*. London: Picador.
- Erdem, Tülin en Joffre Swait (1998). "Brand Equity as a Signaling Phenomenon", *Journal of Consumer Psychology* 7 (2), 131-157.
- Fattah, Hassan M. (2002). *P2P: How Peer-to-Peer Technology is Revolutionizing the Way We Do Business*. Chicago: Dearborn.

- Han, Qi, Benedict G.C. Dellaert, W. Fred van Raaij en Harry J.P. Timmermans (2004). "Tourist Activity Planning in Congested Urban Tourism Environments: Towards a Game-Theoretic Model and Decision Support System", in Geoffrey.I. Crouch, Rick R. Perdue, Harry J.P. Timmermans en Muzzafer Uysal eds. *Consumer Psychology of Tourism, Hospitality and Leisure*, 3<sup>rd</sup> edition, 91-102.
- Häubl, Gerald, Benedict G.C. Dellaert, Kyle B. Murray en Valerie Trifts (2004). "Buyer Behavior in Personalized Shopping Environments: Insights from the Institute for Online Consumer Studies". In C.M. Karat, J. Blom, en J. Karat eds., *Designing Personalized User Experiences for eCommerce*, Amsterdam: Kluwer, 145-160.
- Häubl, Gerald en Kyle Murray (2003). "Preference Construction and Persistence in Digital Market Places: The Role of Electronic Recommendation Agents", *Journal of Consumer Psychology* 13 (1), 149-159.
- Häubl, Gerald en Valerie Trifts (2000). "Consumer Decision Making in Online Shopping Environments: The Effects of Interactive Decision Aids", *Marketing Science* 19 (1), 4-21.
- Huber, Joel, Dan Ariely en Gregory Fischer (2001). "Expressing Preferences in a Principal-Agent Task: A Comparison of Choice, Rating and Matching", Working paper, Department of Marketing, Fuqua School of Business, Duke University.
- Huffman, Cynthia en Barbara E. Kahn (1998). "Variety for Sale: Mass Customization or Mass Confusion", *Journal of Retailing* 74 (4), 491-514.
- Ives, Blake en Gabriele Piccoli (2003). "Custom Made apparel and Individualized Service at Land's End", *Communications of the Association for Information Systems* 11, 79-93.
- Kahneman, Daniel en Amos Tversky (1984). "Choices, Values and Frames", *American Psychologist* 39, 341-350.
- Liechty, John, Venkatram Ramaswamy en Steven H. Cohen (2001). "Choice Menus for Mass Customization: An Experimental Approach for Analyzing Customer Demand with an Application to a Web-Based Information Service", *Journal of Marketing Research* 38 (May), 183-197.
- McFadden, Daniel (1986). "The Choice Theory Approach to Market Research", *Marketing Science* 5 (4), 275-298.
- Newell, Allen en Herbert A. Simon (1972). *Human Problem Solving*, Englewood Cliffs, N.J.: Prentice-Hall.
- Payne, John (1982). "Contingent Decision Making: A Review and Discussion of Issues", *Psychological Bulletin* 73, 221-230.
- Prahalad, C. K. en Venkatram Ramaswamy (2000). "Co-opting Customer Competence", *Harvard Business Review* 78 (1), 79-87.
- Pine, Joseph B. en James H. Gilmore (2000). *Markets of One: Creating Customer-Unique Value through Mass Customization*. Boston: HBS Press.
- Pine, B. Joseph II, Don Peppers en Martha Rogers (1995). "Do You Want to Keep Your Customers Forever?", *Harvard Business Review* 73 (2), 103-114.
- Ritzer, George (1998). *The McDonaldization Thesis*, London: Sage Publications.
- Sanchez, Ron (1999). "Modular Architecture in the Marketing Process", *Journal of Marketing* 63(Special Issue), 92-111.
- Syam, Niladri B. en Benedict G.C. Dellaert (2001). "Consumer-Producer Interaction: A Strategic Analysis of the Market for Customized Products", *Review of Marketing Science* 1 (1).
- Train, Kenneth E. (2003). *Discrete Choice Methods with Simulation*. Cambridge, UK: Cambridge University Press.
- Tseng, Mitchell M. en Frank T. Piller (2003). *The Customer Centric Enterprise: Advances in Mass Customization and Personalization*. Berlin: Springer Publishers.
- Von Hippel, Eric (2001). "Perspective: User Toolkits for Innovation", *Journal of Product Innovation Management* 18 (4), 247-257.
- Von Hippel, Eric (2002). "Horizontal Innovation Networks: By and For Users", MIT Sloan School of Management Working Paper 4366-02 (June).

Wernerfelt, Birgir (1994). "An Efficiency Criterion for Marketing Design", *Journal of Marketing Research* 31 (4), 462-471.

Wind, Jerry en Arvind Rangaswamy (2001). "Customerization: The Next Revolution in Mass Customization", *Journal of Interactive Marketing* 15 (1), 13-32.

## *9. De rol van imago-congruentie in de adoptie van mobiel bankieren*

M. KLEIJNEN, K. DE RUYTER en T.W. ANDREASSEN

### SAMENVATTING

In dit artikel besteden we aandacht aan de impact van imago-congruentie op de attitude van consumenten ten opzichte van mobiel bankieren en de intentie om dergelijke diensten te gebruiken. Daarbij wordt ook de rol van de gebruikssituatie (privé versus publiek, en vrienden versus collega's) onderzocht. De resultaten van twee experimenten laten zien dat imago-congruentie niet alleen een direct effect heeft op attitude en intentie, maar dat er ook een interactieeffect is tussen imago-congruentie en de gebruikssituatie. Tegen de verwachtingen in worden consumenten met een lage imago-congruentie meer door hun omgeving beïnvloed dan consumenten met een hoge imago-congruentie. Het lijkt erop dat zolang consumenten overeenkomsten zien tussen hun eigen imago en het imago van de mobiele dienst, signalen uit de omgevingen genegeerd worden.

### 1. INLEIDING

Per definitie redden innovatieve diensten het vaker niet dan wel in de huidige markt (Christensen 1997; Moore 2002). Een aanzienlijke hoeveelheid marketingliteratuur heeft zich gericht op de factoren die ervoor zorgen dat een innovatie door consumenten geadopteerd wordt. Hoewel het belang van technologische kenmerken in deze literatuur sterk wordt benadrukt (Dabholkar 1996; Davis, Bagozzi, en Warshaw 1989), hebben verschillende auteurs aangegeven dat consument-gerelateerde factoren minstens even belangrijk zijn (Dabholkar en Bagozzi 2002; Meuter et al. 2000). Recente literatuur laat zien dat veel consumenten hun keuze laten afhangen van imago (Plouffe, Vandembosch, en Hulland 2001; Venkatesh en Davis 2000). Terwijl deze studies zich alleen richten op het imago van de dienst- of productinnovatie, suggereert Schlosser (2002) dat het gebruik van mobiele technologie ook afhankelijk is van het imago van de consument zelf. Daarbij gaat het erom dat het adopteren van nieuwe producten of diensten gezien wordt als een vorm van zelfexpressie, het maken van een statement dat laat zien over welke persoonlijkheid de consument beschikt. In deze studie richten wij ons op het gebruik van mobiele diensten, en mobiel bankieren in het bijzonder. Dit zijn diensten die geleverd worden via mobiele telefoons of zogenaamde 'hand-held' computers, waardoor de diensten op elk moment en elke plaats gebruikt kunnen worden (Balasubramanian, Peterson, en Jarvenpaa 2002; Müller-Veerse et al. 2001). Het belang van imago bij adoptie van mobiele diensten zien we ook terug in de praktijk. Toen Nokia bijvoorbeeld de oude 9110 Communicator verving voor de Communicator 9210i waren reclames vooral gericht op het imago en de status die een mobiele-dienstengebruiker



geniet als hij dit nieuwe model zou gaan gebruiken. In plaats van de traditionele benadering waarbij vooral veel aandacht werd besteed aan technische aspecten, richtte Nokia zich dus vooral op het communiceren van de “sociale” voordelen zoals de match tussen het imago van de Nokia communicator en de levensstijl van de consument. Waarom deze strategieverandering? Onderzoek heeft aangetoond dat consumenten vaak de voorkeur geven aan merken en winkels die een imago hebben dat consistent is met hun eigen imago/zelfbeeld (Sirgy 1986; Sirgy 1982; Sirgy 1985). In de literatuur wordt dit fenomeen de imago-congruentie hypothese genoemd (Grubb en Grathwohl 1967). In onze studie suggereren we dat imago-congruentie een cruciale rol speelt bij de adoptie van diensten door consumenten. Daarnaast kijken we niet alleen naar imago, maar ook naar de gebruikssituatie. Verschillende studies hebben aangetoond dat de situatie waarin een consument zich bevindt een invloed heeft op zijn of haar beslissingsproces. Graeff (1997) bijvoorbeeld concludeert dat bepaalde producten of merken typisch in openbare of publieke situaties gebruikt worden, omdat deze een bepaald imago uitstralen. Daarom nemen we de gebruikssituatie mee in onze studie. In veel marketingonderzoek wordt er onderscheid gemaakt tussen privé- en publieke situaties (Schmitt en Shultz II 1995), wat zeker van belang zal zijn in het geval van mobiele diensten. Door de flexibiliteit van mobiele diensten heeft de consument volledige controle over het moment en de plaats waar deze diensten gebruikt worden, hetgeen impliceert dat deze diensten in zowel privé- als publieke situaties gebruikt kunnen worden. In deze paper onderzoeken we 1) de rol van imago-congruentie in het adoptieproces van een innovatieve dienst, en 2) het effect van de gebruikssituatie op de attitude en gebruikssintentie ten opzichte van een innovatieve dienst. Als eerste geven we een kort literatuuroverzicht. Daarna bespreken we ons eerste experiment dat zich richt op de effecten van imago-congruentie en de invloed van privé- versus publieke situaties. Een tweede experiment werd uitgevoerd om de bevindingen van ons eerste experiment te verifiëren en de invloed van de gebruikssituatie verder uit te diepen. Tot slot worden de implicaties van onze resultaten besproken.

## 2. OVERZICHT VAN DE LITERATUUR

Zowel in de literatuur over de marketing van diensten als over adoptie wordt aandacht besteed aan het concept imago. Hiervan doen we hieronder kort verslag.

Volgens Olson (1972) kan imago een bron van informatie zijn die de keuzes van consumenten sterk beïnvloedt. Het imago van een product of dienst is een signaal naar de consument dat leidt tot het vormen van percepties en attitudes ten opzichte van dit product of dienst. In de literatuur over adoptie is imago vaak impliciet opgenomen in de factor relatief voordeel (oftewel de toegevoegde waarde van innovatie in vergelijking met andere producten of eerdere versies). Verschillende auteurs (Moore en Benbasat 1991; Tornatzky en Klein 1982) hebben echter aangetoond dat imago een unieke factor is die consumenten duidelijk onderscheiden van andere, bijvoorbeeld technische voordelen. Terwijl eerdere onderzoeken zich vooral richten op het imago van producten of diensten, richten wij ons op de overeenkomsten tussen dit imago en het imago van de consument, oftewel imago-congruentie, een aspect dat tot dusver onderbelicht is gebleven in de adoptieliteratuur.

Eén van de eerste onderzoekers die aandacht besteedde aan het concept imago-congruentie was Martineau. Hij stelt dat consumenten een voorkeur hebben voor een win-

kel die een persoonlijkheid uitstraalde die gelijk is aan het beeld dat de consument heeft van hem/haarzelf (Bellenger, Steinberg, en Stanton 1976). Met name de attitude en gebruiksintenties lijken afhankelijk te zijn van de mate waarin het imago van een product of dienst gelijkenis vertoont met het imago van de consument zelf, ook wel het *werkelijke* zelfbeeld genoemd. Deze fit wordt in de literatuur de *imago-congruentie hypothese* genoemd (Grubb en Grathwohl 1967). Ondanks de intuïtieve aannemelijkheid achter deze relatie zijn de bevindingen met betrekking tot dit verband niet eenduidig (Hogg, Cox, en Keeling 2000). Daarnaast is onderzoek naar deze relatie in dienstenmarketing zeer beperkt. De rol van imago-congruentie lijkt echter belangrijker dan ooit in de context van innovatieve diensten die geassocieerd worden met life-style producten en persoonlijke expressie (zoals mobiele telefoons). Omdat diensten per definitie niet tastbaar zijn, is het voor consumenten moeilijk om ze te evalueren. Signalen met betrekking tot bijvoorbeeld betrouwbaarheid of toegevoegde waarde worden dan cruciaal, met name als er sprake is van een innovatieve dienst en informatie zeer beperkt is. Daarom richt deze studie zich vooral op de rol van imago-congruentie bij het vormen van attitudes en gebruiksintenties met betrekking tot innovatieve diensten. Zowel attitude als gebruiksintentie worden erkend als belangrijke variabelen in het adoptieproces. Volgens Graeff (1997) kan de perceptie van een consument van zijn eigen imago gebruikt worden om zowel de attitude als de gebruiksintentie te beïnvloeden: door overeenkomsten te creëren tussen het eigen imago en dat van het product of de dienst zal de attitude en gebruiksintentie ten opzichte van dit product of dienst positief beïnvloed worden. Deze bevindingen zijn door verschillende auteurs bevestigd (Martin en Belizzi 1982; Sirgy en Ericksen 1992). Daarom nemen wij attitude en intentie op als afhankelijke variabelen in onze studie.

Tot slot worden gebruikssituaties meegenomen. Onderzoek heeft aangetoond dat deze de beslissingen van consumenten aanzienlijk beïnvloeden (Belk, Balm, en Mayer 1982). Vooral als we praten over imago-congruentie, is het belangrijk om rekening te houden met de situatie waarin de consument zich bevindt. Wanneer consumenten omringd zijn door anderen willen ze vaak een bepaald imago tentoonstellen (Graeff 1997; Hogg et al. 2000). In dit onderzoek bekijken we daarom ook de rol van de gebruikssituatie, waarbij we zowel directe effecten als interactie-effecten met imago-congruentie in ogenschouw nemen.

### 2.1. Hypotheses

In veel onderzoek wordt aandacht besteed aan de invloed van imago-congruentie op de keuzes die consumenten maken (Heath en Scott 1998; Hogg et al. 2000; Martin en Belizzi 1982). Uit deze studies kunnen we afleiden dat imago-congruentie niet over het hoofd gezien mag worden bij het bestuderen van het adoptieproces. Consumenten zien de overeenkomsten met hun eigen imago als een signaal, waarop ze hun keuze om een innovatie wel of niet te adopteren baseren. Dit is met name van belang voor mobiele diensten. Door hun unieke voordelen (Kleijnen, de Ruyter, en Wetzels 2004; Müller-Veerse 2000) zijn mobiele diensten de meest persoonlijke vorm van diensten die “op afstand” (via een elektronisch kanaal) geleverd worden. Consumenten vinden het dan ook belangrijk dat dergelijke diensten aansluiten bij hun eigen imago. Imago-congruentie lijkt dus een essentiële factor te zijn voor het vormen van een attitude en gebruiksintentie ten opzichte van mobiele diensten, wat leidt tot de volgende hypothese:

H<sub>1</sub>: Consumenten die een hoge imago-congruentie waarnemen zullen innovatieve, mobiele diensten positiever evalueren dan consumenten die een lage imago-congruentie waarnemen, met betrekking tot:

H<sub>1a</sub>: attitude

H<sub>1b</sub>: gebruiksintentie

In het verleden is imago vaak gedefinieerd als een multi-dimensioneel construct (Dittmar en Drury 2000; Sirgy, Grewal, en Mangleburg 2000). Hierbij worden vaak twee dimensies herkend: het *werkelijke* zelfbeeld – hoe een individu zichzelf daadwerkelijk ziet – en het *ideale* zelfbeeld – het ideaal zoals iemand zichzelf graag zou willen zien (zie Sirgy 1982 voor een overzicht).

Onderzoek heeft aangetoond dat sommige voorkeuren meer samenhangen met het werkelijke zelfbeeld, terwijl andere meer samenhangen met het ideale zelfbeeld (Onkvisit en Shaw 1987). Welk zelfbeeld het belangrijkste is hangt af van de situatie waarin de consument zich bevindt. Met andere woorden, het “masker” dat iemand opzet hangt af van de sociale omstandigheden waarin hij of zij zich bevindt (Sirgy et al. 2000). In eerder onderzoek worden merken, producten en diensten omschreven als typisch “privé” of “publiek” op basis van de situatie waarin ze meestal gebruikt worden. Bij mobiele diensten ligt deze scheidingslijn niet zo eenvoudig. Het belangrijkste kenmerk van mobiele diensten is juist dat ze op elk moment en in elke situatie gebruikt kunnen en zullen worden. Daarnaast heeft onderzoek door Graeff (1997; 1996) aangetoond dat consumenten sterke overeenkomsten zien tussen hun werkelijke en ideale imago. Verschillen in effecten van deze zelfbeelden op attitude en intentie zijn daarom nauwelijks aanwezig. Daarom suggereren wij het volgende:

H<sub>2</sub>: bij het adopteren van innovatieve, mobiele diensten zullen werkelijke en ideale imago-congruentie een even sterk (positief) effect hebben op:

H<sub>2a</sub>: attitude

H<sub>2b</sub>: gebruiksintentie

Naast imago-congruentie bestuderen we ook de rol van de gebruikssituatie in de adoptie van mobiele diensten. Onderzoek op het gebied van zowel psychologie als marketing geeft aan dat gebruikssituaties een belangrijke verklaring kunnen zijn voor het gedrag van consumenten (Schmitt en Shultz II 1995). Ook sociale identiteitstheorie (Tajfel and Turner 1986) suggereert dat consumenten geneigd zijn om hun keuzes te baseren op hun sociale omgeving. Ze kiezen voor bepaalde producten of diensten om zo een bepaalde status of de goedkeuring van deze omgeving te verkrijgen (Chau 1996; Graeff 1997).

In het verleden hebben onderzoekers vaak onderscheid gemaakt tussen privé and publieke situaties (Dolich 1969; Graeff 1997; Graeff 1996). In een privé-situatie zijn mensen niet omringd door andere personen en zullen ze dus ook geen sociale druk ervaren. In een publieke situatie daarentegen worden mensen voor een groot deel beïnvloed door de mensen om hen heen. In het algemeen worden mobiele telefoons en gerelateerde diensten gezien als innovatief, trendy, en sociaal. Dit impliceert dat de publieke opinie over mobiele diensten positief is. Daarom zullen consumenten in een publieke setting waarschijnlijk een positiever beeld hebben over dit soort diensten. Dit leidt tot de volgende hypothese:

H<sub>3</sub>: Consumenten in een publieke situatie zullen innovatieve, mobiele diensten positiever evalueren dan consumenten in een privé-situatie, met betrekking tot:

H<sub>3a</sub>: attitude

H<sub>3a</sub>: gebruiksintentie

Hoewel verschillende onderzoeken zich gericht hebben op het begrip imago-congruentie, is er relatief weinig bekend over variabelen die de effecten van imago-congruentie versterken of verzwakken (Graeff 1996). Volgens Schmitt en Shultz II (1995) is het bestuderen van gebruikssituaties vooral belangrijk voor producten en diensten die gezien worden als “imagoversterkers”. In een privé-situatie zijn consumenten alleen, waardoor de behoefte om een bepaald imago uit te stralen veel minder sterk is dan in een publieke situatie. We verwachten dat consumenten in een publieke situatie eerder de behoefte voelen om een bepaald imago uit te stralen dan in de privé-situatie. Indien consumenten overeenkomsten zien tussen hun eigen imago en dat van de mobiele dienst, zullen ze deze dienst eerder gebruiken in de publieke situatie om zo een bepaald signaal (imago) naar de omgeving af te geven. Met andere woorden:

H<sub>4</sub>: consumenten in een publieke situatie die een hoge imago-congruentie waarnemen zullen innovatieve, mobiele diensten positiever evalueren dan consumenten die zich in een privé-situatie bevinden en een hoge imago-congruentie waarnemen, met betrekking tot:

H<sub>4a</sub>: attitude

H<sub>4b</sub>: gebruiksintenties

Om onze hypothesen te toetsen werd een experiment uitgevoerd. De uitvoering en resultaten worden hierna besproken.

### 3. EXPERIMENT 1

#### 3.1. Onderzoekscontext

Volgens Barwise en Strong (2002) is het mobiele kanaal zeer persoonlijk. Consumenten dragen hun mobiele telefoon vaak bij zich en deze is vaak bedoeld voor individueel gebruik (in tegenstelling tot bijvoorbeeld computers, die vaak voor gemeenschappelijk gebruik zijn). De toekomstvoorspellingen voor mobiele dienstverlening zijn dan ook optimistisch, waarbij de aandacht vooral uitgaat naar mobiel bankieren en aandelen verhandelen (Barnard et al. 2001; Viquez 2001). Verschillende banken hebben hier al op ingespeeld. De Postbank ([www.postbank.nl](http://www.postbank.nl)) bijvoorbeeld heeft 500.000 mobiele telefoons gratis verstrekt om het gebruik van mobiel bankieren aan te moedigen. Op dit moment valt het daadwerkelijk gebruik echter tegen. Hoewel sommige consumenten wel interesse tonen in het nieuwe medium (Adriaenssens 2001), staat mobiel bankieren wat de meeste mensen betreft nog in de kinderschoenen. Dit kan liggen aan het feit dat bedrijven zich altijd gericht hebben op het promoten van de functionele voordelen in plaats van de sociale voordelen. In ons onderzoek toetsen we daarom of imago-congruentie kan bijdragen aan een positieve evaluatie van mobiele dienstverlening.

### 3.2. Experimentele opzet

We hebben een zogenaamd “between-subjects” experiment uitgevoerd om onze hypothesen te toetsen, waarbij de gebruikssituatie gemanipuleerd werd. Als eerste werd respondenten gevraagd om vragen te beantwoorden met betrekking tot het imago van de bank, het werkelijke zelfbeeld, en hun ideale zelfbeeld. Vervolgens werd de situatie gepresenteerd (zie figuur 1), waarna de vragen werden gesteld over de attitude en de gebruikssintentie met betrekking tot mobiel bankieren in die situatie. Op basis van een kwalitatieve pretest met 20 respondenten werden kleine aanpassingen in de scenario’s aangebracht.

<p><b>Basis Scenario</b></p> <p>De afgelopen jaren heeft u iedere maand wat geld opzij gelegd en dat is inmiddels een aardig bedrag geworden. Omdat uw reguliere spaarrekening weinig rente oplevert, heeft uw financiële adviseur voorgesteld dat u het geld gaat beleggen in aandelen. U hebt dit advies opgevolgd.</p> <p><b>Situatie Manipulatie</b></p> <p><i>Privé:</i> Nu stel u voor dat u ’s avonds thuis voor de TV zit, er is niks speciaals op de TV. U ontvangt een SMS van uw financiële adviseur. Hij adviseert u om meteen actie te ondernemen en bepaalde aandelen te kopen vanwege de huidige (gunstige) stand van de aandelen op de beurs. U kunt de transactie meteen uitvoeren via uw mobiele telefoon. Als u de transactie via uw mobiele telefoon uitvoert, ontvangt u later een SMS met de details van uw transactie.</p> <p><i>Publiek:</i> Nu stel u voor dat u in een bar bent, omgeven door vreemde mensen. Deze mensen kent u verder niet. U ontvangt een SMS van uw financiële adviseur. Hij adviseert u om meteen actie te ondernemen en bepaalde aandelen te kopen vanwege de huidige (gunstige) stand van de aandelen op de beurs. U hebt de mogelijkheid de transactie meteen uit te voeren via uw mobiele telefoon. Als u de transactie via uw mobiele telefoon uitvoert, ontvangt u later een SMS met de details van uw transactie. De mensen om u heen zouden zien dat u de transactie uitvoert via uw mobiele telefoon.</p>
---

Figuur 1. Scenario Privé versus Publieke Situatie.

### 3.3. Vragenlijst

De items met betrekking tot het imago zijn gebaseerd op Graeff (1997) en domeinspecifieke literatuur op het gebied van mobiele dienstverlening (zoals Müller-Veerse et al. 2001). Ook deze items zijn weer geëvalueerd door 20 respondenten. De uiteindelijke dimensies worden gepresenteerd in tabel 1, evenals de vragen die gebruikt zijn om attitude en gebruikssintentie te meten.

### 3.4. Steekproef

Respondenten werden geworven door middel van straatinterviews in verschillende steden in Nederland, gedurende een periode van 14 dagen op verschillende tijdstippen en

Tabel 1. Vragenlijst Items.

Construct	Items	Alpha Studie 1	Alpha Studie 2	Bron
Imago	• Helemaal niet carrière gericht – heel sterk carrière gericht	Bank: .87 Werkelijk: .86 Ideaal: .85	Bank: .75 Werkelijk: .72 Ideaal: .75	Graeff
	• ouderwets – modern			
	• onsuccesvol – succesvol			
	• lokaal – globaal			
	• relaxed – functioneel			
	• inefficiënt – efficiënt			
	• langzaam – snel			
	• conventioneel – trendy			
	• onbekwaam – invloedrijk			
	• volger – pionier			
Attitude	• slecht – goed	.96	.95	Dabholkar en Bagozzi
	• plezierig – onplezierig			
	• nadelig – voordelig			
	• gunstig- ongunstig			
Intentie	• onwaarschijnlijk – waarschijnlijk	.97	.97	Bruner II en Hensel / Dabholkar en Bagozzi
	• niet aannemelijk – aannemelijk			
	• onmogelijk – mogelijk			
	• onzeker -zeker			
	• absoluut niet gebruiken – absoluut wel gebruiken			

dagen. Respondenten konden alleen deelnemen aan het onderzoek als ze in bezit waren van een mobiele telefoon en enigszins bekend waren met mobiele technologie en het reguliere aandelen verhandelen. De uiteindelijke dataset bestond uit 209 respondenten die willekeurig toegewezen waren aan een scenario (104 privé versus 105 publiek). Verder bestond de steekproef uit 49% vrouwen en 51% mannen. Wat betreft leeftijd, was de groep als volgt verdeeld: < 25 jaar: 26%, 25-40 jaar: 26%, 41-55 jaar: 26%, en >55 jaar: 22%.

## 4. RESULTATEN

### 4.1. Manipulatiecheck

Als eerste werd een manipulatie check (7-puntsschaal met de polen privé en publieke situatie) uitgevoerd om te kijken of de respondenten de scenario's goed begrepen hadden. De resultaten geven aan dat de twee scenario's voldoende van elkaar verschilden ( $F_{1,207} = 1269.126$ ;  $p < .001$ ).

### Imago-congruentie

Imago-congruentie werd gemeten met behulp van een Euclidean distance model:

$$D = \sqrt{\sum_{i=1}^n (P_i - S_i)^2}$$

D = verschil tussen het imago van de mobiele dienst en het eigen imago

$i$  = imago-dimensie

$P_i$  = perceptie van de mobiele dienst gemeten met imago-dimensie  $i$

$S_i$  = perceptie van eigen imago gemeten met imago-dimensie  $i$

Een klein (groot) verschil impliceert een hoge (lage) imago-congruentie.

Als eerste hebben we de relatie onderzocht tussen de verschillende types imago-congruentie (werkelijk en ideaal) en de afhankelijke variabelen in de verschillende situaties. Tabel 2 geeft de correlaties weer tussen imago-congruentie (werkelijk en ideaal) en attitude en gebruiksintenties. Op basis van deze resultaten kunnen we concluderen dat zowel het werkelijke als het ideale zelfbeeld een significante invloed heeft op attitude ( $H_{1a}$ ) and gebruiksintenties ( $H_{1b}$ ).

Tabel 2. Correlaties privé versus publiek

	Privé (n = 104)		Publiek (n = 105)	
	Werkelijke imago-congruentie	Ideale imago-congruentie	Werkelijke imago-congruentie	Ideale imago-congruentie
<i>Attitude</i>	-.83**	-.80**	-.36**	-.39**
<i>Intentie</i>	-.75**	-.74**	-.36**	-.40**

\*\* Significant at 0.01 level (2-tailed).

Bij het vergelijken van de correlaties op basis van de  $T_2$  statistic (Williams 1959) bleek er geen significant verschil te bestaan tussen de correlaties voor werkelijke en ideale imago-congruentie (alle  $p$ -waarden  $> .05$ ), wat leidt tot het accepteren van  $H_2$ .

Gezien deze resultaten en de hoge correlatie tussen voor werkelijke en ideale imago-congruentie ( $r = 0.92$ ,  $p < .01$ ), werd besloten om de gezamenlijke effecten verder te onderzoeken. In de volgende analyses werd het gemiddelde voor werkelijke en ideale imago-congruentie gebruikt als covariaat.

#### 4.3. MANCOVA privé versus publiek

Nadat de assumpties getoetst en gevalideerd waren, hebben we een MANCOVA gebruikt om onze hypothesen te testen. De resultaten zijn in Tabel 3 gerapporteerd.

De resultaten duiden op een significant interactieeffect ( $V = .237$ ;  $F_{2,204} = 31.725$ ;  $p < .001$ ) evenals significante directe effecten voor situatie ( $V = .059$ ;  $F_{2,204} = 6.366$ ;  $p = .002$ ) en imago-congruentie ( $V = .515$ ;  $F_{2,204} = 108.229$ ;  $p < .001$ ). Om onze hypothesen verder te onderzoeken hebben we een ANCOVA uitgevoerd (Tabel 4). Daarnaast geeft figuur 2 een grafische presentatie van onze bevindingen.

Zoals gebruikelijk beginnen we onze discussie met het significante interactieeffect. De resultaten geven aan dat de relatie significant is ( $F_{1,208} = 64.054$ ;  $p < .001$ ,  $\eta^2 = .230$ ), maar nadere inspectie van figuur 2 laat zien dat de interactie niet is zoals we verwacht hadden. Het zijn de respondenten met een lage imago-congruentie die worden beïnvloed door hun omgeving, in plaats van zoals verwacht de respondenten met hoge ima-

Tabel 3. Resultaten MANCOVA: Privé versus Publiek.

Privé versus Publiek				
Effect	Pillai-Bartlett trace (V)	F <sub>2,204</sub>	p-waarde	Partial $\eta^2$
<i>Directe effecten</i>				
Situatie	.059	6.366	.002	.059
Imago-congruentie	.515	108.229	< .001	.515
<i>Two-way interactie</i>				
Situatie * congruentie	.237	31.725	< .001	.237

R<sup>2</sup> adjusted Attitude = .568R<sup>2</sup> adjusted Intentie = .432

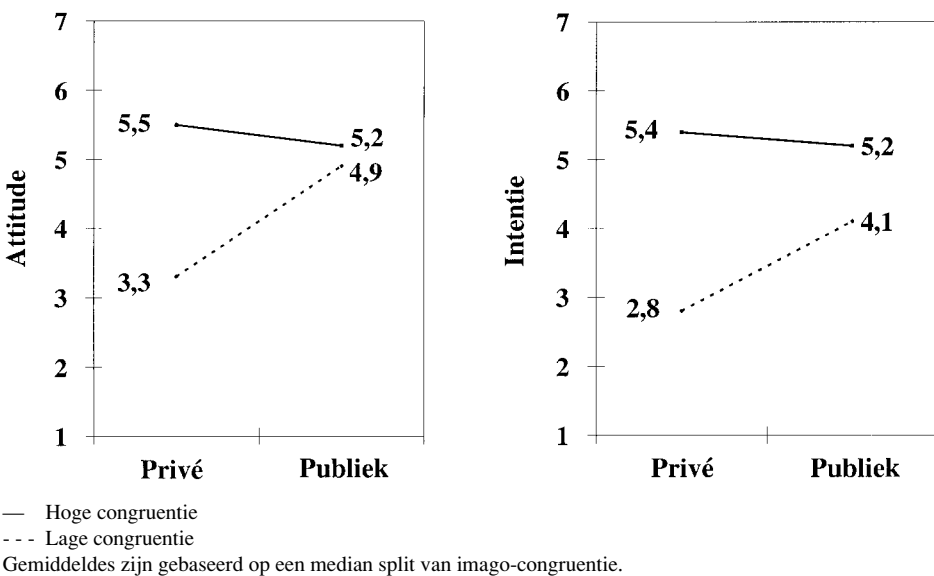
Tabel 4. Resultaten Omnibus ANCOVA: Privé versus Publiek.

Privé versus Publiek				
Afhankelijke variabele	Effect	F <sub>1,208</sub>	p-waarde	Partial $\eta^2$
<i>Attitude</i> (R <sup>2</sup> adj. = .568)				
Directe effecten	Situatie	12.284	.001	.057
	Imago-congruentie	186.849	< .001	.477
Two-way interactie	Situatie * congruentie	64.054	< .001	.230
<i>Intention to Adopt</i> (R <sup>2</sup> adj. = .432)				
Directe effecten	Situatie	5.044	.026	.024
	Imago-congruentie	125.536	< .001	.380
Two-way interactie	Situatie * congruentie	24.727	< .001	.108

go-congruentie (H4a). Verder laten we zien dat beide directe effecten voor gebruikssituatie (H3a;  $F_{1,208} = 12.284$ ;  $p = .001$ ,  $\eta^2 = .057$ ) en imago-congruentie (H1a;  $F_{1,208} = 186.849$ ;  $p < .001$ ) ook significant zijn. Deze bevindingen worden ook ondersteund door additionele One-way ANOVAs die uitgevoerd werden met attitude als afhankelijke variabele.

Voor gebruikssintenties vinden we vergelijkbare resultaten. Ook hier is het interactie-effect ( $F_{1,208} = 64.054$ ;  $p < .001$ ) significant (Tabel 4), maar tegen onze verwachtingen in (H4b) is het ook hier de respondent met lage imago-congruentie die beïnvloed werd. Tevens vinden we significante directe effecten voor imago-congruentie (H1b:  $F_{1,208} = 125.536$ ;  $p < .001$ ) en gebruikssituatie (H4b:  $F_{1,208} = 5.044$ ;  $p = .026$ ). Ook deze resultaten worden bevestigd door ANOVAs, deze keer met gebruikssintentie als afhankelijke variabele. De gevonden resultaten waren niet geheel volgens verwachting. Om de resultaten verder uit te diepen werd een tweede experiment uitgevoerd, waarbij met name de publieke situatie nader belicht werd.





Figuur 2. Effect van Gebruikssituatie en Imago-Congruentie (Privé versus Publiek).

## 5. VRIENDEN VERSUS COLLEGA'S

In dit experiment richten we ons vooral op de publieke omgeving. Leden van een sociale groep worden gedefinieerd op basis van hun overeenkomsten en verschillen met een bepaalde groep (Graeff 1997; Tajfel en Turner 1986). Wij kijken naar twee verschillende groepen, namelijk vrienden en collega's. In andere studies worden vaak alleen vrienden meegenomen als relevante groep (bijvoorbeeld Grewal, Mehta, en Kardes 2000). Bij mobiel bankieren is het van belang om beide groepen in beschouwing te nemen, aangezien de scheidingslijn tussen sociale en zakelijke tijdsbesteding bij mobiele diensten moeilijk te maken is (May 2001). Juist door het feit dat ze overal en op elk moment gebruikt kunnen worden, zijn er meerdere groepen die een invloed kunnen hebben. Vrienden worden gezien als mensen die je goed kent, terwijl collega's meer oppervlakkige kennissen zijn. Daarom verwachten we dat de invloed van vrienden veel sterker zal zijn dan de invloed van collega's. Bovendien zijn de meeste interacties via het mobiele kanaal nog altijd gebaseerd op sociale (vrienden)netwerken. Dit leidt tot de volgende hypothese:

H<sub>5</sub>: Consumenten in een "vrienden"-situatie zullen innovatieve, mobiele diensten positiever evalueren dan consumenten in een "collega"-situatie, met betrekking tot:

H<sub>5a</sub>: attitude

H<sub>5b</sub>: gebruiksententie

Ook hier verwachten we een interactieeffect tussen imago-congruentie en de gebruikssituatie. Op basis van de resultaten van het vorige experiment verwachten we dat consumenten met een lage imago-congruentie eerder beïnvloed worden door hun omge-

ving dan mensen met een hoge imago-congruentie. Een reden hiervoor kan zijn dat consumenten de behoefte voelen om een bepaald imago te laten zien aan hun omgeving (door het gebruik van mobiele diensten) zodat ze geaccepteerd worden door diezelfde omgeving. Herr et al. stellen dat signalen uit de omgeving belangrijker zijn dan eventuele negatieve informatie die de consument zelf bezit. In een zakelijke omgeving is het vaak belangrijk om een bepaald imago neer te zetten. Consumenten met een lage imago-congruentie zullen waarschijnlijk toch gebruik gaan maken van mobiele diensten omdat collega's dan een positiever beeld van hen krijgen. Aangezien vrienden mensen zijn die je goed kent, is het niet nodig om je anders voor te doen dan je bent. Onze hypothese luidt dan ook als volgt:

H<sub>6</sub>: Consumenten in een “collega”-situatie die een lage imago-congruentie waarnemen zullen innovatieve, mobiele diensten positiever evalueren dan consumenten in een “vrienden”-situatie die een lage imago-congruentie waarnemen, met betrekking tot:

H<sub>6a</sub>: attitude

H<sub>6b</sub>: gebruiksintentie

## 6. EXPERIMENT 2

Het tweede experiment is op dezelfde manier opgezet en uitgevoerd als het eerste. De situaties zijn weergegeven in figuur 3. De steekproef bestond in dit geval uit 102 respondenten voor het vriendschapscenario en ook 102 voor het collega-scenario. Verder bestond de groep uit 49% vrouwen en 51% mannen en zijn de leeftijdscategorieën als volgt verdeeld: < 25 jaar: 25%, 25-40 jaar: 27%, 41-55 jaar: 25%, >55 jaar: 23%.

## 7. RESULTATEN

### 7.1. Manipulatiecheck

Ook in dit experiment werden manipulatiechecks uitgevoerd om de verschillen tussen de scenario's te toetsen. De resultaten laten zien dat deze verschillen groot genoeg zijn ( $F_{1,202} = 489.394$ ;  $p < .001$ ).

### 7.2. Imago-congruentie

Om onze resultaten van het vorige experiment te valideren kijken we ook hier naar de correlaties tussen de verschillende types imago-congruentie (werkelijk en ideaal) en de afhankelijke variabelen in de verschillende situaties (Tabel 5).

Zowel het werkelijke als het ideale zelfbeeld hebben een significante invloed (H<sub>1</sub>; H<sub>2</sub>) op de attitude en gebruiksintentie. Ook hier zijn er geen significante verschillen tussen de correlaties van het werkelijke en het ideale zelfbeeld (H<sub>2</sub>). De correlatie tussen werkelijk en ideaal zelfbeeld was in dit geval .90 ( $p < .01$ ). Ook hier gebruiken we daarom het gemiddelde van imago-congruentie in de verdere analyses.

**Basis Scenario**

De afgelopen jaren heeft u iedere maand wat geld opzij gelegd en dat is inmiddels een aardig bedrag geworden. Omdat uw reguliere spaarrekening weinig rente oplevert, heeft uw financiële adviseur voorgesteld dat u het geld gaat beleggen in aandelen. U hebt dit advies opgevolgd.

**Situatie Manipulatie**

*Vrienden:* Nu stel u voor dat u in een bar bent met een paar goede vrienden. Deze vrienden kennen u erg goed en er is geen reden om u anders voor te doen dan u bent. U ontvangt een SMS van uw financiële adviseur. Hij adviseert u om meteen actie te ondernemen en bepaalde aandelen te kopen vanwege de huidige (gunstige) stand van de aandelen op de beurs. U hebt de mogelijkheid de transactie meteen uit te voeren via uw mobiele telefoon. Als u de transactie via u mobiele telefoon uitvoert, ontvangt u later een SMS met de details van uw transactie. Uw vrienden zouden dan zien dat u de transactie uitvoert via uw mobiele telefoon.

*Collega's:* Nu stel u voor dat u in een bar bent met een paar nieuwe collega's. Deze collega's kennen u nog niet goed en u vindt het belangrijk om een goede indruk achter te laten. U ontvangt een SMS van uw financiële adviseur. Hij adviseert u om meteen actie te ondernemen en bepaalde aandelen te kopen vanwege de huidige (gunstige) stand van de aandelen op de beurs. U hebt de mogelijkheid de transactie meteen uit te voeren via uw mobiele telefoon. Als u de transactie via uw mobiele telefoon uitvoert, ontvangt u later een SMS met de details van uw transactie. Uw collega's zouden dan zien dat u de transactie uitvoert via uw mobiele telefoon.

Figuur 3. Scenario Vrienden versus Collega's.

Tabel 5. Correlaties vrienden versus collega's.

Vrienden (n= 102)		Collega's (n = 102)	
Werkelijke imago-congruentie	Ideale imago-congruentie	Werkelijke imago-congruentie	Ideale imago-congruentie
-.77**	-.77**	-.78**	-.79**
-.76**	-.74**	-.74**	-.75**

\*\* Significant at 0.01 level (2-tailed).

### 7.3. MANCOVA vrienden versus collega's

Na het testen van de assumpties hebben we weer een MANCOVA uitgevoerd. De resultaten zijn samengevat in tabel 6. Ook hier vinden we weer een significant interactieeffect, evenals significante directe effecten. De ANCOVAs zijn gerapporteerd in Tabel 7 en de grafische weergave in figuur 4.

Net als in studie 1 vinden we een significant interactieeffect ( $F_{1,200} = 33.214$ ;  $p < .001$ ) voor de respondenten die een lage imago-congruentie waarnemen. Echter, ook in dit

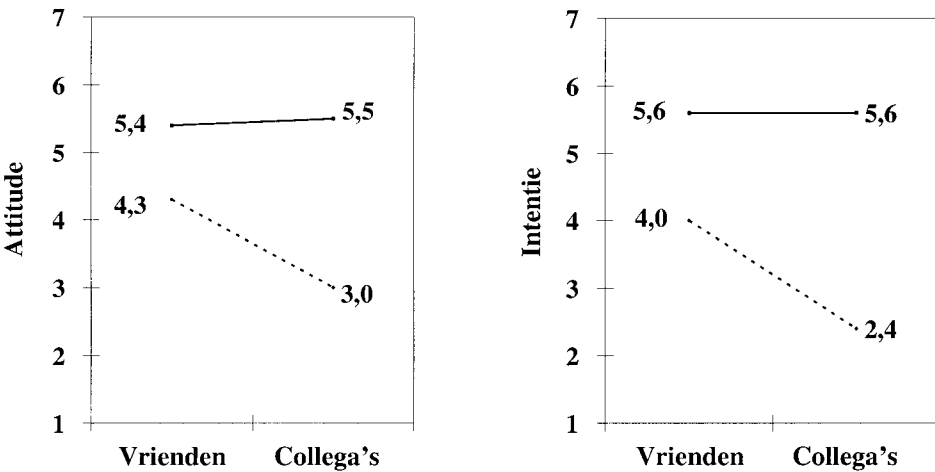
Tabel 6. Resultaten MANCOVA: Vrienden versus Collega's.

Vrienden versus Collega's				
Effect	Pillai-Bartlett trace (V)	F <sub>2,199</sub>	p-waarde	Partial η <sup>2</sup>
<i>Directe effecten</i>				
Situatie	.049	5.151	.007	.049
Imago-congruentie	.667	198.864	< .001	.667
<i>Two-way interactie</i>				
Situatie * congruentie	.143	16.607	< .001	.143

R<sup>2</sup> adjusted Attitude = .420  
 R<sup>2</sup> adjusted Intention = .482

Tabel 7. Resultaten Omnibus ANCOVA: Vrienden versus Collega's.

Vrienden versus Collega's				
Afhankelijke variabelen	Effect	F <sub>1,200</sub>	p-waarde	Partial η <sup>2</sup>
<i>Attitude</i>				
Directe effecten	Situatie	10.245	.002	.049
	Imago-congruentie	373.004	< .001	.651
Two-way interactie	Situatie * congruentie	33.214	< .001	.162
<i>Intention to Adopt</i>				
Directe effecten	Situatie	5.243	.023	.026
	Imago-congruentie	290.981	< .001	.593
Two-way interactie	Situatie * congruentie	23.296	< .001	.104



Note:  
 — hoge congruentie  
 --- lage congruentie  
 Gemiddeldes zijn gebaseerd op een median split van imago-congruentie.

Figuur 4. Effect van Gebruikssituatie en Imago-congruentie (Vrienden versus Collega's).

geval zijn de resultaten verrassend: de attitude is positiever voor respondenten die omringd worden door vrienden, en niet door collega's (H6a). Ook vinden we weer significante directe effecten voor imago-congruentie (H1a;  $F_{1,200} = 373.004$ ;  $p < .001$ ) en de situatie (H5a;  $F_{1,200} = 10.245$ ;  $p = .002$ ). Deze bevindingen worden weer ondersteund door de resultaten van additionele one-way ANOVAs.

Voor intenties werd precies hetzelfde patroon gevonden (zie Tabel 6 & 7), waardoor we hypothese 6b niet kunnen accepteren. Hypothese 1b en 5b worden wel bevestigd.

## 8. DISCUSSIE

De resultaten van onze twee experimenten tonen aan dat imago-congruentie een belangrijke rol speelt bij het adopteren van mobiel bankieren. Een interessante conclusie is het feit dat zowel het werkelijke als het ideale zelfbeeld een gelijkwaardige rol lijken te spelen in het verklaren van adoptie, ongeacht de gebruikssituatie. Verder kunnen we ook eerdere bevindingen met relatie tot de gebruikssituaties nuanceren: de uitkomsten tonen aan dat vrienden een belangrijke rol spelen in het vormen van een attitude en de gebruikssintentie, en dan met name wanneer de imago-congruentie laag is. Hoewel dit niet onze verwachting was, is dit wel te verklaren. Wanneer consumenten een lage imago-congruentie waarnemen, is het waarschijnlijk dat ze op zoek gaan naar andere signalen. Herr et al. (1991) concludeerden uit hun onderzoek dat de invloed van de omgeving vaak leidt tot een positieve evaluatie, zelfs wanneer uitgebreide informatie beschikbaar is die tot minder positieve evaluaties zou moeten leiden. In dit geval is het signaal van de vrienden sterker dat het signaal van de lage imago-congruentie.

### 8.1. Suggesties voor verder onderzoek

Ieder onderzoek kent zijn beperkingen, wat leidt tot enkele suggesties voor verder onderzoek. Ten eerste is dit onderzoek uitgevoerd in een "laboratorium"-omgeving. De realiteit is de daadwerkelijke toetsing van onze theorie. Daarnaast is verder onderzoek naar de verschillende types zelfbeeld nodig. Ons onderzoek is slechts een eerste stap in het ontrafelen van de effecten van diverse zelfbeelden (zie Sirgy 1982 voor een overzicht). Verder zijn persoonlijkheidskenmerken niet opgenomen in deze studie. Sommige consumenten zijn van nature minder geneigd om nieuwe dingen uit te proberen. Dit soort consumenten zal wellicht sneller bereid zijn een innovatie te adopteren wanneer ze overeenkomsten zien tussen hun eigen imago en dat van de innovatie. Tot slot, in het verleden heeft adoptieonderzoek zich vooral gericht op de functionele aspecten van innovaties (bijvoorbeeld Davis, 1989), waarbij socio-emotionele aspecten slechts impliciet aan bod komen (Rogers, 1995). De relevantie van imago-congruentie in de huidige studie geeft aan dat er behoefte is aan onderzoek dat verder gaat dan de puur functionele voordelen. Vooral de rol van sociale netwerken in de consumentenmarkt is een gebied dat onderbelicht is gebleven en meer aandacht verdient.

### 8.2. Managementimplicaties

Er zijn verschillende managementimplicaties aan onze studie verbonden. Ten eerste suggereren de resultaten dat het gebruik van diverse gebruikssituaties als een promotietactiek zinvol kan zijn, met name voor consumenten die een lage imago-congruentie

waarnemen. In plaats van de nadruk te leggen op functionele aspecten, zou het zinvoller zijn om de socio-emotionele aspecten van het gebruik van mobiele diensten te benadrukken. Daarnaast is viral marketing een goede mogelijkheid om mobiele diensten te adverteren. Consumenten die een innovatie zoals mobiel bankieren voorstellen aan andere consumenten zijn waarschijnlijk een effectieve manier om mensen over te halen dit soort diensten te adopteren. Een mogelijkheid is bijvoorbeeld een mobiele-dienstgebruiker in de gelegenheid stellen een vriend(in) een gratis introductieabonnement aan te bieden. Tot slot, gezien de enorme impact van imago-congruentie op attitude en intentie, moeten marketers zeker overwegen hun markt te segmenteren op basis van dit criterium. In de telefoonindustrie is dit zeker niet ongebruikelijk: de Nokia 5100 met zijn toetsenbord dat er uitziet als een 'wasbordje' is duidelijk gericht op mensen met een sportief zelfbeeld. Dit soort strategieën kan ook toegepast worden op mobiele diensten.

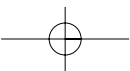
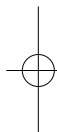
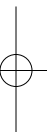
## LITERATUUR

- Adriaenssens, Charles (2001), "The potential for m-commerce in retail financial services," in *Mobile commerce – connecting with the future*, Mike Elysée en Charlie Henderson en Nicholas A Hopwood en John Machin en Malcolm Marshall, Eds. London, UK: Technology Publishing Ltd.
- Balasubramanian, Sridhar, Robert A. Peterson Peterson, en Sirikka L. Jarvenpaa (2002), "Exploring the implications of m-commerce for markets and marketing," *Journal of the Academy of Marketing Science*, 30 (4), 348-61.
- Barnard, Chris, Elena Carreri, Pim Bilderbeek, Barbara Blesio, Daniele Bonfanti, en Tim Sheedy (2001), "Banking on mobile: prospects for Western Europe," *IDC Research Report*, 1-58.
- Barwise, Patrick en Colin Strong (2002), "Permission-based mobile advertising," *Journal of Interactive Marketing*, 16 (1), 14-24.
- BBC (2000), "A simple guide to m-commerce," *BBC News Business Section*, Februari 18, 2004 World Wide Web: [www.thefeature.com.http://news.bbc.co.uk/1/hi/business/992292.stm](http://www.thefeature.com.http://news.bbc.co.uk/1/hi/business/992292.stm).
- Belk, R.W., K.D. Balm, en R.N. Mayer (1982), "Developmental recognition of consumption symbolism," *Journal of Consumer Research*, 9 (June), 4-17.
- Bellenger, Danny N., Earle Steinberg, en Wilbur W. Stanton (1976), "The congruence of store image and self image: as it relates to store loyalty," *Journal of Retailing*, 52 (1), 17-32.
- Bruner II, Gordon C. en Paul J. Hensel (1998), *Marketing scales book: a compilation of multi-item measures* (2 ed.). Chicago, Illinois USA: American Marketing Association.
- Chau, Patrick Y.K. (1996), "An empirical assessment of a modified technology acceptance model," *Journal of Management Information Systems*, 13 (2), 185-204.
- Christensen, Clayton M. (1997), *Innovator's Dilemma: When New Technologies Cause Great Firms to Fail*: Harvard Business School Publishing.
- Dabholkar, Pratibha A. (1996), "Consumer evaluation of new technology-based self-service options: an investigation of alternative models of service quality," *International Journal of Research in Marketing*, 13 (1), 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 the Academy of Marketing Science*, 30 (3), 184-201.
- Davis, F.D. (1989), "Perceived usefulness, perceived ease of use, en usage of information technology: A replication," *MIS Quarterly*, 16 (2), 319-39.
- Davis, Fred D., R.P. Bagozzi, en P.R. Warshaw (1989), "User acceptance of computer technology: a comparison of two theoretical models," *Management Science*, 35, 982-1003.
- Dittmar, Helga en John Drury (2000), "Self-image – is it in the bag? a qualitative comparison between 'ordinary' and 'excessive' consumers," *Journal of Economic Psychology*, 21, 109-42.

- Dolich, Ira J. (1969), "Congruence relationships between self images en product brands," *Journal of Marketing Research*, 6 (February), 80-84.
- Graeff, Timothy R. (1997), "Consumption situations and the effect of brand image on consumers' brand evaluations," *Psychology & Marketing*, 14 (1), 49-70.
- (1996), "Using promotional messages to manage the effects of brand and self-image on brand evaluations," *Journal of Consumer Marketing*, 13 (3), 4-18.
- Grewal, R., R. Mehta, en F.R. Kardes (2000), "The role of the social-identity function of attitudes in consumer innovativeness en opinion leadership," *Journal of Economic Psychology*, 21, 233-52.
- Grubb, Edward L. en Harrison L. Grathwohl (1967), "Consumer self-concept, symbolism en market behavior: A theoretical approach," *Journal of Marketing*, 31 (October), 22-27.
- Heath, Adam P. en Don Scott (1998), "The self-concept and image congruence hypothesis," *European Journal of Marketing*, 32 (11/12), 1110-23.
- Herr, Paul M., Frank R. Kardes, en John Kim (1991), "Effect of word-of-mouth and product-attribute information on persuasion: an accessibility-diagnostics perspective," *Journal of Consumer Research*, 17, 454-62.
- Hogg, Margaret K., Alastair Cox, J., en Kathy Keeling (2000), "The impact of self-monitoring on image congruence and product/brand evaluation," *European Journal of Marketing*, 34 (5/6), 641-66.
- Jarvenpaa, Sirikka L., Karl Reiner Lang, Yoko Takeda, en Virpi Kristiina Tuunainen (2003), "Mobile commerce at cross roads," *Communications of the ACM*, 46 (12), 41-45.
- Kleijnen, Mirella H.P., J.C. de Ruyter, en Martin G.M. Wetzels (2004), "Consumer Adoption of Wireless Services: Discovering the rules, while playing the game," *Journal of Interactive Marketing*, 18 (2), 51-61.
- Martin, Warren S. en Joseph Belizzi (1982), "An analysis of congruous relationships between self-images and product images," *Journal of the Academy of Marketing Science*, 10 (4), 473-89.
- Martineau, Pierre (1958), "The personality of the retail store," *Harvard business review*, 36 (January-February), 47-55.
- May, Paul (2001), *Mobile commerce – breakthroughs in application development: opportunities, applications and technologies of wireless business*. Cambridge UK: Cambridge University Press.
- Meuter, M.L., A.L. Ostrom, R.I. Roundtree, en M.J. Bitner (2000), "Self-service technologies: understanding customer satisfaction with technology-based service encounters," *Journal of Marketing*, 64 (July), 50-64.
- Moore, Gary C. en Izak Benbasat (1991), "Development on an instrument to measure the perceptions of adopting an information technology innovation," *Information Systems Research*, 2 (3), 192-222.
- Moore, Geoffrey A. (2002), *Crossing the Chasm: Marketing and Selling High-Twch Products to Mainstream Consumers*. New York, NY: HarperCollins Publishers.
- Müller-Veerse, Falk (2000), "Mobile commerce," *Research Report Durlacher Research Ltd*, May 20, 2001, World Wide Web: <http://www.durlacher.com>.
- Müller-Veerse, Falk, Bodo Kohlenbach, Dirk Bout, Shum Singh, Gennady Golub, Jukka Häyrynen, Sasu Laitinen, en Erkko Autio (2001), "UMTS – an investment perspective," *Research Report Durlacher Research Ltd. in cooperation with Eqvitec and HuT*, December 20, 2001, World Wide Web: <http://www.durlacher.com>.
- Olson, Jerry C. (1972), "Cue Utilization of the Quality Perception Process: A Cognitive Model and an Empirical Test," *unpublished doctoral dissertation, Purdue University*.
- Onkvisit, Sak en John Shaw (1987), "Self-concept and image congruence: some research and managerial implications," *Journal of Consumer Marketing*, 4 (1), 13-23.
- Plouffe, C.R., M. Vandenbosch, en J. Hulland (2001), "Intermediating technologies and multi-group adoption: a comparison of consumer and merchant adoption intentions toward a new electronic payment system," *Journal of Product Innovation Management*, 18 (2), 65-81.

- Rogers, E.M. (1995), *Diffusion of Innovations* (4 ed.). New York, NY: The Free Press.
- Schlosser, Francine K. (2002), "So, how do people really use their handheld devices? An interactive study of wireless technology use," *Journal Organizational Behavior*, 23, 401-23.
- Schmitt, Bernd H. en Clifford J. Shultz II (1995), "Situational effects on brand preferences for image products," *Psychology & Marketing*, 12 (5), 433-46.
- Sirgy, J.M. en M.K. Ericksen (1992), "Employed females clothing preference, self-image congruence, and career anchorage," *Journal of Applied Social Psychology*, 22, 408-22.
- Sirgy, M. J. (1986), *Self Congruity: Toward a Theory of Personality and Cybernetics*. New York: Praeger Publishers.
- (1985), "Using self-congruity and ideal congruity to predict purchase motivation," *Journal of Business Research*, 13, 195-206.
- Sirgy, M. Joseph (1982), "Self-concept in consumer behavior: a critical review," *Journal of Consumer Research*, 9, 287-300.
- Sirgy, M.J., D. Grewal, en T Mangleburg (2000), "Retail environment, self-congruity, and retail patronage: An integrative model and a research agenda," *Journal of Business Research*, 49, 127-38.
- Tajfel, Henri en John C. Turner (1986), "The Social Identity Theory of Intergroup Behavior," in *Psychology of Intergroup Relations*, S. Worchel en W.G. Austin, Eds. Chicago.
- Tornatzky, L.G. en K.J. Klein (1982), "Innovation Characteristics and Innovation Adoption-Implementation: A Meta-Analysis of Findings," *IEEE Transactions on Engineering Management*, 29 (1), 28-45.
- Venkatesh, Viswanath en Fred D. Davis (2000), "A theoretical extension of the technology acceptance model: four longitudinal field studies," *Management Science*, 46 (2), 186-204.
- Viquez, Frank (2001), "Mobile Commerce: analyzing emerging opportunities in technologies, applications, markets and trends," *Allied Business Intelligence Inc. Strategic Research Report*, 1-225.
- Williams, E.J. (1959), "The comparison of regression variables," *Journal of the Royal Statistical Society, Series B*, 21, 396-99.





## 10. Medicijnen, Marketing en Modellen

P.S.H. LEEFLANG, J.E. WIERINGA en D.R. WITTINK

### SAMENVATTING

In Nederland wordt jaarlijks bijna een half miljard Euro uitgegeven aan marketinginspanningen voor geneesmiddelen<sup>1</sup>. Het bedrag dat in Amerika jaarlijks hieraan wordt besteed ligt ergens tussen de 10 en 20 miljard dollar<sup>2</sup>. Het meten van de effecten van marketinginspanningen op afleverhoeveelheden van recepten krijgt de laatste jaren meer aandacht. In deze bijdrage geven we aan welke spelregels bij het modelleren van de effecten van marketinginspanningen in acht moeten worden genomen. We illustreren dit aan de hand van twee recente studies, die door het Centraal Planbureau (de Laat et al., 2002) en door ons zijn uitgevoerd<sup>3</sup>. We laten zien dat de keuze voor een modelspecificatie verstrekkende gevolgen heeft voor de conclusies over de effecten van marketinginspanningen.

### 1. INLEIDING

De aandacht voor het bouwen van modellen voor markten is in de afgelopen decennia zowel in de academische wereld als in de bedrijfspraktijk sterk toegenomen<sup>4</sup>. Van het wetenschappelijk onderzoek dat in de afgelopen dertig jaar op dit vakgebied is uitgevoerd valt te leren dat het belangrijk is om bij de specificatie van dergelijke modellen aandacht te besteden aan een aantal aspecten. Zo zal het model gebaseerd moeten zijn op een adequate theorie die het gedrag van afnemers beschrijft. Verder zal men zeer zorgvuldig de variabelen moeten kiezen die het afnemersgedrag kunnen verklaren. Elementen die hierbij een rol spelen zijn eenvoud, volledigheid en robuustheid<sup>5</sup>. Andere belangrijke keuzes betreffen de aggregatieniveaus waarop de variabelen gedefinieerd worden. Aggregatie heeft diverse aspecten:

1. aggregatie van het vraagniveau;
2. aggregatie van de onafhankelijke variabelen,
3. aggregatie van de variabelen in de tijd, enzovoort.

Keuzes ten aanzien van het aggregatieniveau bepalen welke variabelen over welke tijdperioden gemeten worden. Deze keuze bepaalt op zijn beurt de schattingsmethoden en de resultaten die met een model verkregen worden. Het is ook belangrijk dat modeluitkomsten getoetst worden met behulp van validatiecriteria.

In deze bijdrage besteden we aandacht aan de keuze voor een modelspecificatie (paragraaf 2). We doen dit aan de hand van enkele modellen die recentelijk ontwikkeld zijn om de effecten van marketinginspanningen op de afleverhoeveelheden van receptgeneesmiddelen te bepalen. Deze modellen zijn onder meer ontwikkeld door het Centraal Planbureau (CBP) (de Laat et al., 2002) en zullen in paragraaf 3 besproken wor-

den. Na publicatie van deze modellen hebben Leeftang en Wittink (2002) kritiek op deze modellen geformuleerd. Deze kritiek is samengevat in paragraaf 4. In een studie, die deels gefinancierd is door Nefarma<sup>6</sup>, is deze kritiek empirisch onderbouwd en zijn tal van varianten van de CBP-modellen geschat. Enkele van deze modellen zullen in paragraaf 5 besproken worden. In paragraaf 6 geven we aan welke modellen mogelijkheden bezitten om de effecten van marketinginspanningen op het voorschrijfgedrag van medicijnen te bepalen.

## 2. MODELLEN SPECIFICEREN

In de loop der tijd zijn er diverse criteria geformuleerd waaraan de specificatie van modellen moet voldoen. Het zou te ver voeren om daar in deze bijdrage op in te gaan en we volstaan met te verwijzen naar de literatuur waarin deze criteria besproken worden<sup>7</sup>. Hier willen we stil staan bij de keuzes met betrekking tot de eerder genoemde aggregatieniveaus.

De vraag naar een product kan op diverse niveaus gemodelleerd worden op:

- het niveau van de individuele afnemer;
- segmentniveau; en
- marktniveau.

Daarbij kan de vraag betrekking hebben op

- het individuele merk;
- de productcategorie (“product class” en/of “product form”).

In veel modellen probeert men de effecten van marketinginspanningen over merken te poolen. Dat wil zeggen dat men de gegevens die de vraag naar een merk in de tijd weergeven over verschillende merken combineert. Stel dat men per merk de beschikking heeft over  $T$  waarnemingen en men deze waarnemingen ‘pooft’ over  $n$  merken dan werkt men met  $nT$  waarnemingen. Het aantal waarnemingen bij modellen waarin men over cross-secties (hier merken) poolt is dus (veel) groter dan wanneer men op individueel merkniveau een model schat.

Het poolen van merken is alleen toegestaan onder bepaalde voorwaarden. Deze komen er op neer dat de met elkaar corresponderende parameters voor de gepoolde merken bij benadering dezelfde grootte bezitten. Statistische testen bepalen of over de cross-secties gepoold mag worden.

De conclusies over de rol die marketinginstrumenten spelen worden deels beïnvloed door de keuze van de onafhankelijke variabelen die verondersteld worden de vraag te bepalen. Ook hier moet men keuzes maken ten aanzien van de aggregatieniveaus. Zo zou men de vraag naar een merk kunnen verklaren met behulp van de reclame-uitgaven die gedaan zijn in een medium  $i$  (tijdschriften). Maar men kan de reclame-uitgaven ook aggregeren over alle media  $i=1, \dots, m$ , of zelfs alle marketinguitgaven aggregeren. Voorwaarde is dan wel dat de diverse marketinguitgaven min of meer hetzelfde effect hebben op de vraag.

Bij het specificeren van de onafhankelijke variabelen is het van belang dat deze variabelen voldoende fluctuaties vertonen en niet zodanig gecorreleerd zijn dat het onmogelijk

lijk is betrouwbare effecten te schatten. Daarnaast moet de modelbouwer nagaan of de betreffende variabele werkelijk het koopgedrag kan beïnvloeden. Zo is het opnemen van de prijs in vraagvergelijkingen voor receptgeneesmiddelen hoogst twijfelachtig omdat de prijs niet bepaald wordt door de farmaceutische bedrijven; in Nederland is immers sprake van een strenge prijsregulering. Fabrikanten kunnen derhalve de prijs niet gebruiken om de vraag te beïnvloeden. Bovendien is het onduidelijk in hoeverre artsen en apothekers de prijzen meenemen in de keuze tussen alternatieve medicijnen die verschillen in effectiviteit of bijverschijnselen. We komen op de relevantie van prijs terug in paragraaf 4.

Een ander aggregatievraagstuk heeft betrekking op de keuze van de tijdsperiode die in een model gebruikt wordt: gebruikt men weekgegevens of gegevens die betrekking hebben op een maand, een kwartaal, enzovoort. Voor frequent aangekochte (consumptie)goederen ligt het gebruik van weekgegevens voor de hand. In het algemeen is het logisch dat de periode waarop de gegevens betrekking hebben gerelateerd is aan de aankoopfrequentie. Door met dynamische effecten rekening te houden kunnen korte- en lange(re) termijneffecten bepaald worden. Des te groter de dynamiek des te groter zullen de verschillen tussen deze effecten zijn.

In de navolgende paragrafen zullen we op deze specificatievraagstukken nader ingaan aan de hand van modellen die de vraag naar receptgeneesmiddelen bepalen.

### 3. MODELLEREN VAN MARKETINGINSPANNINGEN VOOR MEDICIJNEN

Er zijn diverse modellen ontwikkeld om de effecten van marketinginspanningen op het voorschrijfgedrag van medicijnen te bepalen. We kunnen deze modellen classificeren conform tabel 1.

De marketingmixmodellen die geen of enig gedragsdetail bevatten (IA en IB) worden geschat met over individuele voorschrijvers (artsen, specialisten) geaggregeerde gegevens. De aggregatie kan plaats vinden tot op het niveau van het individuele merk of tot op het niveau van de productcategorie (aggregatie over merken). Diffusiemodellen bieden de mogelijkheid om aan de hand van geaggregeerde gegevens op een indirecte wijze enig gedragsdetail te determineren. Zo kan de invloed van marketing op de probeer-voet (trial rate), en op het percentage herhalingsaankopen worden gemeten. Het effect van mond-tot-mond reclame is ook bij uitstek geschikt om bepaald te worden in diffusiemodellen.

In dit artikel concentreren we ons op modellen waarmee marketing mix effecten op verkopen bepaald worden (I), die geen gedragsdetail bevatten (A) en die of de vraag naar een individueel merk (i) of de over merken gepoolde / geaggregeerde vraag (ii) specificeren. We zullen thans twee voorbeelden van dergelijke modellen geven. Deze voorbeelden zijn voor het verloop van dit verhaal relevant.

Rizzo (1999) schat een gepoold model voor de hypertensie submarkt. Op basis van Amerikaanse jaarlijkse data over de periode 1988—1993 van 46 producten wordt een

Tabel 1. Classificatie van modellen.

**I Modellen waarmee marketing mix effecten op verkopen bepaald worden**A Modellen die geen gedragsdetail bevatten<sup>8</sup>(i) modellen die de vraag naar een *specifiek* merk specificeren:

Montgomery &amp; Silk (1972), Parsons &amp; Vanden Abeele (1981), Leeflang, Mijatovic, &amp; Saunders (1992), Dekimpe &amp; Hanssens (1999), Berndt et al. (1995, 1997, 2003), Shankar (2003), Steenburgh (2003)

(ii) modellen die de over merken *geaggregeerde* of *gepoolde* vraag modelleren:

Rizzo (1999), de Laat et al. (2002), Neslin (2001), Wittink (2002)

(iii) modellen die de vraag naar een *productcategorie* in kaart brengen:

Rosenthal et al. (2002), Berndt et al. (1995, 1997, 2003), Wosńska (2002)

## B Modellen met enig gedragsdetail: diffusiemodellen

Lilien, Rao, &amp; Kalish (1981), Rao &amp; Yamada (1988), Hahn et al. (1994)

## C Modellen met veel gedragsdetail

Gönül et al. (2001), DeSarbo et al. (2002), Manchanda &amp; Chintagunta (2003), Narayan, Manchanda, &amp; Chintagunta (2003), Akçura, Gönül, &amp; Petrova (2004), Manchanda, Rossi, &amp; Chintagunta (2004)

**II Modellen waarin specifieke marketingproblemen behandeld worden**

- invloed van generieke merken: Hurwitz & Caves (1988)
- prijsbepaling: Berndt (2002)
- levenscycli: Bauer & Fischer (2000), Fischer, Leeflang, & Wittink (2003)
- over-the-counter drugs: Akçura, Gönül, & Petrova (2004)

relatie gelegd tussen de afleverhoeveelheden en verklarende variabelen. In het model wordt de vraag naar een merk  $i$  op tijdstip  $t$  gemodelleerd als een functie van prijs, het aantal artsenbezoeken, door prijzen en artsenbezoeken van concurrerende producten, en door de leeftijd van het geneesmiddel.

$$\ln q_{it} = \alpha_0 + (\alpha_1 + \alpha_2 \ln d_{it} + \alpha_3 \ln D_{it}) \ln p_{it} + \alpha_4 \ln d_{it} + \alpha_5 (\ln d_{it})^2 + \alpha_6 \ln D_{it} + \alpha_7 (\ln D_{it})^2 + \beta X_{it} + \varepsilon_{it}, \quad (1)$$

waar voor merk  $i$  in jaar  $t$ ,

$q_{it}$  = de afleverhoeveelheid;

$d_{it}$  = de uitgaven aan artsenbezoeken (detailing);

$D_{it}$  = een gediscoteerde cumulatie van uitgaven aan artsenbezoeken tot op tijdstip  $t-1$ , berekend als volgt:  $D_{it} = \rho D_{it-1} + d_{i,t-1}$  waar  $\rho \in (0,1)$  de verdisconteringvoet is;

$p_{it}$  = de prijs;

$X_{it}$  = een vector van variabelen, waaronder prijzen en uitgaven aan artsenbezoeken van concurrerende producten van product  $i$ , leeftijd en leeftijd<sup>2</sup> van product  $i$ , en jaardummy's.

De belangrijkste bevindingen van Rizzo zijn de volgende

1. De prijs heeft een significant en negatief effect op de verkopen van een product;
2. artsenbezoeken vergroten brand loyalty (dit wordt afgeleid uit positieve coëfficiënten van de beide interacties van prijs met de uitgaven aan artsenbezoeken);

3. de uitgaven aan artsenbezoeken hebben ook een direct effect op de verkopen;
4. de verkopen nemen aanvankelijk toe met de leeftijd van een product, maar nemen af naarmate het product ouder wordt.

In deze specificatie wordt onderscheid gemaakt tussen het huidige effect van de uitgaven aan verkoopbevordering en (een gediscoteerde) cumulatie van eerdere uitgaven, wat recht doet aan de eerdere bevindingen van Montgomery en Silk (1972), en in lijn is met kennis over het langetermijneffect van tijdschriftenreclame in andere productcategorieën (Dekimpe & Hanssens 1999). In dit model wordt het diffusieproces van nieuwe producten gemodelleerd door de lineaire en de kwadratische trendterm (leeftijd, respectievelijk het kwadraat van leeftijd). Rizzo neemt aan dat de verkoopbevorderingsinspanningen de prijselasticiteit beïnvloeden, en daarnaast een direct effect hebben op de afleverhoeveelheid. De eerste aanname lijkt in het licht van de prijsregulering (zie paragraaf 4) weinig zinvol voor de Nederlandse situatie.

Een tweede model is door de Laat et al. (2002) ontwikkeld voor het Centraal Plan Bureau; we zullen naar dit model dan ook refereren als “het CPB model”. De Laat et al. (2002) analyseren op basis van maandelijks Nederlandse gegevens het effect van farmaceutische marketing op de afleverhoeveelheden. De auteurs gaan uitgebreid in op het fenomeen prijsregulering. Desondanks wordt een vrijwel ongewijzigde versie van het model van Rizzo (1999) geschat. De aanpassingen bestaan uit het toevoegen van vertraagde endogenen ( $q_{i,t-1}$  en  $q_{i,t-2}$ ), en product-specifieke intercepten. De motivatie van de auteurs voor het toevoegen van vertraagde endogenen is dat het model rekening moet houden met gewoontes in het voorschrijfgedrag. Het model is als volgt gespecificeerd:

$$\ln q_{it} = \gamma_i + \rho_1 q_{i,t-1} + \rho_2 q_{i,t-2} + (\alpha_1 + \alpha_2 \ln mes_{it}) \ln p_{it} + (\alpha_3 + \alpha_4 \ln mes_{it}) \ln pc_{it} + \alpha_5 \ln mef_{it} + \alpha_6 \ln mes_{it} + \alpha_7 \ln mefc_{it} + \alpha_8 \ln mesc_{it} + \beta X_{it} + \varepsilon_{it} \quad (2)$$

waar, voor merk  $i$  in maand  $t$ :

$q_{it}$  = de afleverhoeveelheid;

$\gamma_i$  = het merkspecifieke intercept;

$mes_{it}$  = een gediscoteerde cumulatie van uitgaven aan verkoopbevordering (“marketing expenditures stock”) tot op tijdstip  $t-1$ , berekend als volgt:

$$mes_{it} = \rho mes_{i,t-1} + mef_{i,t-1} \text{ waar } \rho = 0.55 \text{ de verdisconteringsvoet is}^9;$$

$mef_{it}$  = de huidige uitgaven aan verkoopbevordering (“marketing expenditures flow”);

$p_{it}$  = de prijs;

$pc_{it}$  = de gemiddelde prijs van concurrerende producten;

$mesc_{it}$  = de gediscoteerde cumulatie van uitgaven aan verkoopbevordering door concurrenten;

$mefc_{it}$  = de huidige uitgaven aan verkoopbevordering door concurrenten;

$X_{it}$  = een vector van variabelen waaronder leeftijd van het product, maand- en jaardummy’s, en dummy’s voor beleidsveranderingen.

Uit de schattingsresultaten van de Laat et al. (2002) blijkt dat zowel de met een periode vertraagde endogene variabele als de met twee perioden vertraagde endogene variabele een belangrijke bijdrage leveren aan de verklaring van het verloop van de afleverhoe-

veelheden. Feitelijk specificeren de Laat et al. (2002) een AutoRegressief model met eXogene variabelen van orde 2 (een ARX(2) model, zie bijvoorbeeld Harvey, 1993). Het model is geschat met behulp van zeventuizend maandelijkse gegevens uit de periode 1994-1999 die betrekking hebben op elf deelmarkten en 140 merken. In tabel 2 zijn de schattingsresultaten van het model van de Laat et al. weergegeven.

Tabel 2. Schattingsresultaten van het CPB-model, vergelijking (2).

Variabele	Parameter	Schatting
vertraagde endogene periode $t-1$	$\rho_1$	0,68 <sup>a</sup>
vertraagde endogene periode $t-2$	$\rho_2$	0,14 <sup>a</sup>
Prijs	$\alpha_1$	-0,18 <sup>b</sup>
interactie prijs en marketinguitgaven	$\alpha_2$	0,02 <sup>a</sup>
prijs concurrenten	$\alpha_3$	0,10
interactie prijs concurrenten en marketinguitgaven	$\alpha_4$	-0,01 <sup>b</sup>
marketinguitgaven (flow)	$\alpha_5$	0,01 <sup>a</sup>
marketinguitgaven (stock)	$\alpha_6$	0,04 <sup>a</sup>
marketinguitgaven (flow) concurrentie	$\alpha_7$	0,01 <sup>a</sup>
marketinguitgaven (stock) concurrentie	$\alpha_8$	-0,03 <sup>a</sup>

<sup>a</sup>  $p < .05t$

<sup>b</sup>  $p < .10$

Uit tabel 2 kunnen we opmaken dat de eigen-merk prijselasticiteit (absoluut gezien) laag is. De interactie-effecten tussen prijs en de promotie-uitgaven zijn eveneens erg klein.

Op basis van dit model worden door de Laat et al. (2002) ook de langetermijneffecten van promotie-uitgaven op de verkopen bepaald. Door de eigen en de kruiselingse effecten van promoties te vergelijken vindt men dat het totale effect van de verkoopbevordering op de verkopen kan worden ontleed in zestig procent marktvergroting ('market making') en veertig procent merksubstitutie ('market stealing'). Deze bevinding kan ook als volgt worden samengevat: de verkoopbevordering voor een bepaald merk geneesmiddel zal per saldo tot een hogere afzet van het geneesmiddel leiden en slechts voor een deel ten koste gaan van de afzet van een concurrerend merk. Promotie kan zowel een verandering in het voorschrijven van geneesmiddelen als een verhoging in het voorschrijven bevorderen.

#### 4. KRITIEK OP HET CPB-MODEL

Leeflang en Wittink (2002) formuleren diverse kritiekpunten op het model van de Laat et al. (2002):

- (1) Er wordt gepoold over 140 merken uit 11 verschillende markten, waarbij alle responsparameters gelijk verondersteld worden. Of pooling toegestaan is wordt niet getoetst.
- (2) Voor elke markt en voor elk merk wordt verondersteld dat de concurrentie dezelfde effecten heeft op de verkopen. Ook deze veronderstelling wordt niet getoetst.
- (3) Er wordt verondersteld dat de effecten van de diverse uiteenlopende vormen van verkoopbevordering eveneens gelijk zijn, zowel voor het korte- als het lange-

termijneffect. Bestaande marketingkennis leert ons echter dat deze veronderstellingen onjuist zijn (zie bijvoorbeeld Parsons & Vanden Abeele 1981, en Leeftang, Mijatovic, & Saunders 1992, Dekimpe & Hanssens 1999).

- (4) De interpretatie van de coëfficiënten van de gecumuleerde variabelen is problematisch. Via de vertraagde endogene variabelen worden reeds vertraagde effecten van de exogene variabelen in het model gebracht. Een effect van een vertraagde exogene variabele wordt dan op niet triviale wijze verdeeld over meerdere coëfficiënten, waardoor het moeilijk wordt betekenis toe te kennen aan de schattingsresultaten. De voordelen van het toevoegen van vertraagde endogenen (dit leidt tot een verbetering van de fit) wordt hierdoor voor een belangrijk deel teniet gedaan.
- (5) Door vertragingen van de endogene variabele wordt impliciet aangenomen dat alle variabelen dezelfde dynamische structuur bezitten, hetgeen niet in overeenstemming is met de werkelijkheid (Montgomery & Silk 1972). Bovendien wordt aangenomen dat alle prijzen en parameters uit het verleden effect hebben op de vraag naar een geneesmiddel.
- (6) Er wordt geen rekening gehouden met het verloop van de diffusie van een nieuw geneesmiddel op de markt
- (7) De verklaring die de Laet et al. (2002) geven voor het gevonden verband tussen de prijselasticiteit en promotie-uitgaven is niet de enig mogelijke. Leeftang en Wittink (2002) presenteren een alternatieve verklaring. Het gevonden 'verband' zou te maken kunnen hebben met het min of meer parallelle verloop van prijselasticiteiten én promotie-uitgaven in de tijd. Onderzoek van andere productcategorieën laat zien dat de eerste kopers van een nieuw product het meest geïnteresseerd zijn in het product en daarom het minst prijsgevoelig zijn. Voor een nieuw medicijn is het normaal dat de patienten die de meeste baat kunnen hebben van een verandering het eerst aan de beurt zullen zijn. Voor zover de prijs al een rol speelt bij de keuze tussen alternatieve medicijnen zal die rol nog geringer zijn als de lopende behandeling niet effectief is. Naarmate een nieuw medicijn een tijd op de markt is geweest, is er minder onzekerheid over de effectieve werking en andere aspecten. Dit leidt tot beter geïnformeerde beslissingen door artsen. De mogelijkheid bestaat dat de prijs vervolgens een, nog steeds beperkte, rol zal spelen. Tegelijkertijd is het zo dat de introductie van een nieuw geneesmiddel in de beginjaren sterk ondersteund wordt door verkoopbevorderingsinspanningen. Naarmate artsen beter geïnformeerd raken, vermindert de noodzaak voor verkoopbevordering. Dit leidt ertoe dat een hoge intensiteit van promotionele activiteiten in de tijd parallel zal lopen aan een lage prijselasticiteit (dichtbij nul), en dat een vermindering in promoties samenvalt met een hogere (meer negatieve) prijselasticiteit. .
- (8) Statistische toetsen waarmee de uitkomsten gevalideerd kunnen worden, ontbreken.

Daarnaast is de conclusie dat zestig procent van het marketingeffect het gevolg is van groei van de markt een erg hoog percentage wanneer we dit vergelijken met marketing-effecten in andere productcategorieën. Bovendien is interpretatie hiervan moeilijk. Hoe kan het zijn dat de markt groeit als gevolg van een marketingactiviteit die gericht is op *doctoren* terwijl die pas in actie kunnen komen nadat een *patiënt* een afspraak heeft gemaakt?



We hebben hiermee geïllustreerd dat in het model van de Laet et al. (2002)

1. de “aggregatie” (hier “pooling”) niet (statistisch) ondersteund wordt;
2. onafhankelijke variabelen zoals de concurrentie en de uiteenlopende vormen van verkoopbevordering geaggregeerd worden zonder dat getoetst wordt of dit gerechtvaardigd is;
3. de aggregatie van de variabelen in de tijd door middel van de keuze van én stockvariabelen én vertraagde endogene variabelen problematisch is.

In het model van de Laet et al. (2002) wordt de prijsvariabele opgenomen. Prijzen van geneesmiddelen komen, kortweg gesteld, niet door een afstemming van vraag op aanbod tot stand maar worden gereguleerd middels het Geneesmiddelen Vergoedingsstelsel. De Wet Geneesmiddelenprijzen bepaalt de maximumprijs die fabrikanten aan hun geneesmiddel mogen geven<sup>10</sup>. Bovendien vertonen prijzen relatief weinig fluctuaties. Al met al verdient het geen aanbeveling om de prijs als onafhankelijke variabele en in een interactieterm op te nemen.

## 5. EMPIRISCHE VALIDATIE VAN HET CPB-MODEL EN ALTERNATIEVE SPECIFICATIES

De modellen (1) en (2) en enkele varianten, waarbij we de vertraagde endogene variabelen uit (2), al dan niet in combinatie met de prijs hebben weggelaten zijn door ons geschat op het niveau van de *individuele merken*. We maken gebruik van gegevens die afkomstig zijn van dezelfde bronnen als die van de Laet et al. (2002). We bestuderen 72 merken<sup>11</sup> op de vijf belangrijkste deelmarkten die in de Laet et al. (2002) onderzocht zijn. Op deze markten worden geneesmiddelen tegen maagzuur, hypertensie, cholesterol, depressiviteit en astma / allergieën aangeboden. Per deelmarkt en per merk hebben we een unieke verzameling concurrerende merken. Dat impliceert dat de effecten van inspanningen van concurrenten van markt tot markt, en per merk, verschillen. De door ons bestudeerde tijdsperiode omvat die van de Laet et al. (2002).

In tabel 3 hebben we de schattingsresultaten samengevat van

1. het Rizzo-model = vergelijking (1);
2. het CPB-model = vergelijking (2);
3. het CPB-model zonder prijs en vertraagde endogene variabelen.

Elke cel van Tabel 3 correspondeert met een coëfficiënt van een van de modellen. De modellen verschillen in het aantal parameters. Bijgevolg stellen de modellen dus ook verschillende eisen aan het aantal waarnemingen dat per merk beschikbaar moet zijn om betrouwbare schattingsresultaten te kunnen verkrijgen. Het Rizzo-model is geschat met behulp van de gegevens van 64 merken; de modellen van de Laet et al. met behulp van 72 merken. Omdat ieder model per merk geschat is, bevat iedere cel in feite 64 of 72 waarden. Voor de overzichtelijkheid is ervoor gekozen een aantal kengetallen van de verdeling van deze waarden te presenteren. Het gemiddelde van de schattingen voor elke coëfficiënt is aangegeven doormiddel van  $\mu$ . Negentig procent van alle waarden ligt in het interval begrensd door  $p_{0.05}$  en  $p_{0.95}$ <sup>12</sup>. Ten slotte geeft  $f$  aan welke fractie van de schattingen significant is en het juiste teken heeft. De schattingen voor de coëfficiënten van de jaar- maand- markt- en andere dummy's zijn weggelaten. Wel hebben

we de schattingen van de parameters die bij de (product)leeftijd en het kwadraat van de leeftijd van het geneesmiddel horen samengevat weergegeven.

De BIC'-waarde is een getal waarmee we de modellen onderling kunnen vergelijken, en is opgebouwd uit twee delen<sup>13</sup>. Een deel is altijd kleiner dan nul en is meer negatief naarmate het model beter de data beschrijft. Een tweede deel is altijd positief en is groter naarmate meer parameters gebruikt worden. Op deze manier wordt er rekening gehouden met de 'fit' van een model, maar ook hoeveel parameters geschat worden. Dit laatste is zinvol omdat de fit niet vermindert door meer verklarende variabelen in het model op te nemen. In deze zin heeft een beter model een kleinere (d.w.z. meer negatieve) BIC'-waarde.

Bij bestudering van tabel 3 vallen een drietal zaken op. Ten eerste, de fractie significante parameterschattingen met het verwachte teken is doorgaans niet groot. Voor alle modellen geldt dat de fractie significante schattingen met het verwachte teken het grootst is voor de coëfficiënten van de leeftijdsvariabelen. Er is een verband tussen afleverhoeveelheid en de fase van de levenscyclus waarin een merk zich bevindt. Het is opmerkelijk dat we met behulp van model (2) (het ongewijzigde CPB-model) voor alle 72 merken niet één interactie-effect (tussen prijs en verkoopbevordering) vinden dat significant en positief is en op reductie van de prijsgevoeligheid door promotie zou duiden.

Ten tweede, de parameterschattingen verschillen nogal tussen de merken. De invloed van de variabelen op het verloop van de afleverhoeveelheden van de verschillende merken is derhalve tamelijk heterogeen te noemen. De resultaten versterken het vermoeden dat de verschillende merken niet gepoold mogen worden. We zullen dit formeel onderzoeken.

Ten derde, in geen enkel geval zijn de parameterschattingen van de prijs- en verkoopbevorderingsvariabelen consequent groter of kleiner dan nul. Samen met de relatief grote spreiding betekent dit dat er geen eenduidig effect aan deze variabelen valt toe te kennen. Zo zijn er merken die een sterk negatieve prijselasticiteit laten zien, maar ook zijn er merken die een sterk positieve prijselasticiteit vertonen. Hetzelfde geldt voor de effecten van verkoopbevordering. De merkspecifieke schattingen van coëfficiënten van elk van de drie modellen zijn zodanig verschillend dat er geen eenduidig beeld van de invloed van de uitgaven aan verkoopbevordering te constateren valt. Ook de interactiecoëfficiënten zijn niet systematisch positief of negatief, waardoor de invloed van promotionele activiteiten op de prijsgevoeligheid niet te duiden is.

De parameterwaarden van de leeftijdsvariabelen laten wel een duidelijk beeld zien: de coëfficiënten van de kwadratische leeftijdsterm zijn overwegend negatief, wat voor een concaaf verloop van de afleverhoeveelheidskromme zorgdraagt. De schattingen voor de coëfficiënt van de lineaire leeftijdsterm zijn altijd positief. Dat is conform de verwachting.

Vergelijking van model (2) met model (3) laat zien dat het aantal merken waarvoor de effecten van marketinguitgaven én het goede teken bezitten (positief) en significant zijn groter is in model (3). Negentien procent van de merken bezit een significante parameter voor de stock variabele. Het aantal merken met een significante parameter voor de flow marketinguitgaven is in model (3) evenwel weer kleiner dan in model (2).

Tabel 3. Schattingsresultaten van model (1), model (2) en model (2) zonder prijzen vertraagde endogene variabelen (model 3).

	Model (1)	Model (2)	Model (3)
Coëfficiënt van	(Rizzo)	(De Laat et al.)	(De Laat et al. zonder prijs en vertragingen)
Prijs	$\mu^a = -0.02$ $\rho_{0.05}^b = -1.48$ $\rho_{0.95}^c = 3.18$ $f^d = 0.13$	$\mu = -0.20$ $\rho_{0.05} = -1.27$ $\rho_{0.95} = 0.94$ $f = 0.09$	—
Interactie prijs en marketing-uitgaven (stock)	$\mu = 0.41$ $\rho_{0.05} = -0.57$ $\rho_{0.95} = 2.86$ $f = 0.12$	$\mu = -0.1139$ $\rho_{0.05} = -1.1238$ $\rho_{0.95} = 0.2474$ $f = 0$	—
Interactie prijs en marketing-uitgaven (flow)	$\mu = -0.22$ $\rho_{0.05} = -1.79$ $\rho_{0.95} = 0.34$ $f = 0.05$	—	—
Marketing-uitgaven (stock)	$\mu = -0.68$ $\rho_{0.05} = -2.85$ $\rho_{0.95} = 1.20$ $f = 0.05$	$\mu = -0.05$ $\rho_{0.05} = -0.70$ $\rho_{0.95} = 0.16$ $s = 0.02$	$\mu = 0.07$ $\rho_{0.05} = -0.17$ $\rho_{0.95} = 0.63$ $f = 0.19$
Marketing-uitgaven (flow)	$\mu = 0.14$ $\rho_{0.05} = -0.82$ $\rho_{0.95} = 1.92$ $f = 0.14$	$\mu = 0.00$ $\rho_{0.05} = -0.02$ $\rho_{0.95} = 0.04$ $f = 0.09$	$\mu = -0.02$ $\rho_{0.05} = -0.25$ $\rho_{0.95} = 0.03$ $f = 0.02$
Kwadraat van marketing-uitgaven (stock)	$\mu = 0.04$ $\rho_{0.05} = -0.03$ $\rho_{0.95} = 0.19$ $f = 0$	—	—
Kwadraat van marketing-uitgaven (flow)	$\mu = -0.01$ $\rho_{0.05} = -0.10$ $\rho_{0.95} = 0.03$ $f = 0.17$	—	—
Productleeftijd	$\mu = 0.20$ $\rho_{0.05} = 0.06$ $\rho_{0.95} = 0.46$ $f = 0.76$	$\mu = 0.14$ $\rho_{0.05} = 0.05$ $\rho_{0.95} = 0.29$ $f = 0.91$	$\mu = 0.15$ $\rho_{0.05} = 0.07$ $\rho_{0.95} = 0.29$ $f = 0.96$
Kwadraat van productleeftijd	$\mu = -0.0008$ $\rho_{0.05} = -0.0031$ $\rho_{0.95} = 0.0010$ $f = 0.57$	$\mu = -0.0003$ $\rho_{0.05} = -0.0008$ $\rho_{0.95} = 0.0000$ $f = 0.46$	$\mu = -0.0004$ $\rho_{0.05} = -0.0018$ $\rho_{0.95} = 0.0000$ $f = 0.51$
# merken	64	72	72
BIC'	-31971	-26532	-24842

<sup>a</sup>  $\mu$  = het gemiddelde van de parameterschattingen;

<sup>b</sup>  $\rho_{0.05}$  = het vijfde percentiel van de parameterschattingen;

<sup>c</sup>  $\rho_{0.95}$  = het vijfennegentigste percentiel van de parameterschattingen;

<sup>d</sup>  $f$  = de fractie van de parameterschattingen dat significant is én het verwachte teken heeft.

In paragraaf 2 hebben we aangegeven dat het poolen van waarnemingen over merken alleen toegestaan is onder bepaalde voorwaarden. We hebben statistische testen uitgevoerd om na te gaan of de over individuele merken gepoolde waarnemingen per deelmarkt al dan niet gepoold mogen worden. De resultaten voor model (2) (het CPB-model) staan vermeld in tabel 4<sup>14</sup>.

Tabel 4. Waarden van de Chow-test per deelmarkt.

Deelmarkt	F-waarde (Chow-test)	Kritieke waarde
Maagzuurremmers	17,1	1,3
Hypertensie	7,8	1,1
Cholesterol	9,3	1,3
Depressie	56,6	1,3
Astma	99,3	2,0

Voor alle vijf markten wordt de kritieke waarde ruimschoots overschreden. Met andere woorden: poolen over individuele merken per deelmarkt is niet toegestaan. De gepoolde resultaten uit het CPB-model zijn dus verkregen op een wijze die niet is toegestaan. Daarbij zijn de resultaten in het CPB-model *ook* nog eens gepoold over deelmarkten hetgeen, vanzelfsprekend, ook niet is toegestaan.

Wat betekent dit nu voor de resultaten die gebaseerd zijn op het CPB-model? We zullen dit voor drie belangrijke conclusies nagaan.

1. In het rapport van de Laar et al. wordt geconcludeerd dat de gemiddelde (pre-marketing) prijselasticiteit ( $\alpha_1$  in (2)) negatief en statistisch significant is. Wij vinden (zie de bovenste cel in de kolom voor model (2) in tabel 3) dat slechts 9 procent van alle merken een significant negatieve prijselasticiteit bezit. Ons gemiddelde ( $\mu = -0.20$ ) is van dezelfde orde van grootte als  $\hat{\alpha}_1$  in het CPB-model (zie tabel 2). Door de resultaten per *merk* te bestuderen ontstaat een genuanceerder beeld dan wanneer we gepoolde of gemiddelde resultaten bestuderen.
2. De Laar et al. (2002) concluderen dat de post-marketing prijselasticiteit ( $\alpha_1 + \alpha_2 \ln mes_{it}$ ) nul is. De negatieve pre-marketing prijselasticiteit (zie hierboven) wordt opgeheven door de positieve interactie tussen prijs en marketinguitgaven (zij schatting van de interactiecoëfficiënt is gelijk 0.02, zie tabel 2). Wij vinden voor *geen enkel merk* een significante interactie tussen prijs en marketinguitgaven-stock ( $mes_{it}$ ). Post- en pre-prijselasticiteiten zijn, zo concluderen wij, even groot. “Gemiddeld” vinden we overigens een negatieve waarde (!) voor de interactie tussen prijs en marketinguitgaven (zie tabel 3, het gemiddelde van de coëfficiënten van de interactie tussen prijs en marketinguitgaven is gelijk aan -0.11 voor het model van de Laar et al.).
3. Een derde conclusie van de Laar et al. (2002) is dat marketinguitgaven het voorschrijfgedrag beïnvloeden (zie tabel 2: de schattingen voor de coëfficiënten van marketinguitgaven zijn gelijk aan 0.01 en 0.04 voor de flow, respectievelijk de stock variabele). Wij vinden dit voor slechts 2 procent van de merken wanneer het om gecumuleerde effecten gaat (stock) en voor 9 procent wanneer directe effecten (flow) beschouwd worden wanneer we het model van de Laar et al. per merk schatten (zie de cellen die corresponderen met de marketinguitgaven in tabel 3). Gegeven

de omvang van de bestedingen aan de promotie van geneesmiddelen lijkt ons dit geen realistisch beeld. Ook vanuit dit oogpunt zullen we vraagtekens moeten stellen bij de door het CPB gebruikte modellen.

We moeten dus concluderen dat door uitspraken te doen op een ander aggregatieniveau (individuele merken versus totaalmarkt) totaal andere conclusies getrokken worden. Hiermee moge het belang van de keuze voor een adequaat aggregatieniveau geïllustreerd zijn.

## 6. MEDICIJNEN, MARKETING EN MODELLEN: HOE NU VERDER?

Naast de onder meer in tabel 3 aangeduide modellen hebben we modellen ontwikkeld waarin we de effecten van marketinginspanningen voor de diverse marktinstrumenten bepaald hebben. Voor meer dan driekwart van alle merken vinden we significante effecten van de inzet van artsenbezoekers op de afleverhoeveelheden van geneesmiddelen. Voor de helft van alle merken vinden we positieve, significante effecten van advertenties en voor meer dan twintig procent van alle merken kunnen we een significant, positief effect van direct mail activiteiten vaststellen.

Deze resultaten en het gegeven dat er, zelfs per deelmarkt, niet over individuele merken gepoold mag worden geeft aan dat er sprake is van een grote heterogeniteit tussen merken. Deze heterogeniteit valt waarschijnlijk voor een belangrijk deel toe te schrijven aan de fase van de levenscyclus waarin een geneesmiddel verkeert. In de eerste fasen (introductie en groei) zijn de promotie-inspanningen hoog en neemt de afzet toe. Na enige tijd (volwassenheid- en verzadigingsfase) neemt de afzet nauwelijks toe en nemen de promotie-inspanningen (in het algemeen) sterk af. Voor nieuwe geneesmiddelen is de relatie tussen afzet en promotie-inspanningen geheel anders dan voor producten die enige tijd op de markt zijn.

We hebben voor “nieuwe” merken, dat zijn merken die na 1 januari 1994 geïntroduceerd zijn, diffusiemodellen ontwikkeld en geschat. Met behulp van deze modellen vinden we hoe en hoeveel de diffusie van een nieuw geneesmiddel versneld kan worden door de inzet van verkoopbevorderingsinstrumenten.

In andere, nog te ontwikkelen, modellen zullen we interacties tussen marketinginspanningen en variabelen als de entree-orde, het aantal concurrenten, de groeivoet, enzovoort expliciet in beschouwing nemen. Tevens zullen we aandacht moeten besteden aan de endogeniteit die er bestaat tussen marketinguitgaven en afleverhoeveelheden.

Op deze wijzen zullen we, zo verwachten we, meer inzicht met modellen verkrijgen in de relaties tussen marketinginspanningen en het voorschrijfgedrag van medicijnen.

### NOTEN

1. Essers (2003)
2. Zie bijvoorbeeld Breitstein (2002)
3. Leeflang, de Weerd, Wieringa, Wittink, & Woltil (2004)
4. Zie bijvoorbeeld Leeflang, & Wittink (2000a, 2000b), de special van IJRM, vol. 17, nr. 2-3
5. Zie bijvoorbeeld Leeflang, Wittink, Wedel, & Naert (2000)

6. Nederlandse Vereniging van de Research-georiënteerde Farmaceutische Industrie.
7. Zie Leeflang et al. (2000)
8. Zie voor meer achtergrond met betrekking tot deze indeling Leeflang et al. (2000, Ch. 4.3 en Ch. 10)
9. Deze waarde van de disconteringsvoet maximaliseert de 'fit' van het model.
10. Zie Leeflang et al. (2004, H. 3)
11. In ons onderzoek bestuderen we alleen de verkopen en marketinginspanningen voor "specialités" (merken) en laten we de generieke ("eigen") merken geneesmiddelen buiten beschouwing.
12. Merk op dat we het hier niet hebben over betrouwbaarheidsintervallen: het gaat hier om percentielen van de empirische verdeling van de coëfficiënten.
13. We gebruiken hier de variant die Raftery (1995) voorstelt voor lineaire modellen. De BIC-waarden zijn telkens berekend op basis van alle individueel geschatte modellen.
14. Om ervoor te zorgen dat onze resultaten vergelijkbaar zijn met die van de Laet et al. (2002) hebben we dezelfde observationele periode als de Laet et al. gebruikt (januari 1994 – december 1999)

## LITERATUUR

- Akçura M.T., F.F. Gönül, E. Petrova (2004) Consumer Learning and Brand Valuation: An Application on Over-the-Counter Drugs, *Marketing Science*, forthcoming.
- Bauer, H.H., & M. Fischer (2000), Product life cycle patterns for pharmaceuticals and their impact on R&D profitability of late mover products, *International Business Review*, 9, pp 703-725.
- Berndt, E.R. (2002), Pharmaceuticals in U.S. Health Care: Determinants of Quantity and Price, *Journal of Economics Perspectives*, 16(4), 45-66.
- Berndt, E.R., L. Bui, D.R. Reiley, G.L. Urban (1995), Information, marketing, and pricing in the U.S. Antiulcer drug market, *The American Economic Review*, vol. 85, pp. 100—105.
- Berndt, E.R., L.T. Bui, D. H. Reiley and G.L. Urban (1997), The Roles of Marketing, Product Quality and Price Competition in the Growth and Composition of the U.S. Anti-Ulcer Drug Industry, in *The Economics of New Products* (Timothy F. Bresnahan and Robert J. Gordon, eds.), Vol 58, Chicago, pp. 277-322.
- Berndt, E.R., R.S. Pindyck, and P. Azoulay (2003), Consumption Externalities and Diffusion in Pharmaceutical Markets: Antiulcer Drugs, *The Journal of Industrial Economics*, Vol. LI(2) pp. 243—269.
- Breitstein, J. (2002), Spending Hits a Wall, *Pharmaceutical Executive* te verkrijgen via <http://www.pharmexec.com/pharmexec/article/articleDetail.jsp?id=29975>.
- Dekimpe, M.G., & Hanssens, D.M. (1999), Sustained Spending and Persistent Response: A New Look at Long-Term Marketing Profitability, *Journal of Marketing Research*, 36, 397-412.
- DeSarbo, W.S., Degeratu, A.M., Ahearne, M.J. & Saxton, M.K. (2002), Disaggregate Market Share Response Models, *International Journal of Research in Marketing*, 19, 253-266.
- Essers, G.T.J.M. (2003), Een gewaardeerde maar partijdige bezoeker, informatievoorziening door artsbezoekers getoetst, *Medisch Contact*, 58(26-27), 1079-1082, 27 juni 2003.
- Families USA (2002), "Profiting from Pain: Where Prescription Drug Dollars Go", te verkrijgen via <http://www.familiesusa.org>.
- Fischer M., P.S.H. Leeflang, & D.R. Wittink (2004), Time to Peak Sales for New Brands, Working paper UCLA/University of Groningen
- Gönül, F.F., Carter, F., Petrova, E. & Srinivasan, K. (2001), Promotion of Prescription Drugs and Its Impact on Physicians' Choice Behavior, *Journal of Marketing*, 65, 79-90.
- Hahn, M., Park, S., Krishnamurthi, L. & Zoltners, A.A. (1994), Analysis of new product diffusion using a four-segment trial-repeat model, *Marketing Science*, 13, 224-247.

- Harvey, A.C. (1993), *Time Series Models*, New York: Harvester Wheatsheaf.
- Hurwitz, M.A. & Caves, R.E. (1988), Persuasion or information? Marketing and the shares of brand name an generic pharmaceuticals, *Journal of Law and Economics*, 31, 299-320.
- Laat, de E.F., Windmeijer, F. & Douven, R. (2002), *How does pharmaceutical marketing influence doctors' prescription behavior?*, Den Haag: CPB Netherlands' Bureau for Economic Policy Analysis.
- Leeflang, P.S.H., Mijatovic, G.M. & Saunders, J. (1992), Identification and estimation of complex multivariate lag structures: a nesting approach, *Applied Economics*, 24, 273-283.
- Leeflang, P.S.H., Weerde, de H., Wieringa, J.E., Wittink, D.R. & Woltil, M. (2004), *Onderzoek naar de effecten van marketinginspanningen op de afleverhoeveelheden van receptgeneesmiddelen*, Working Paper, Rijksuniversiteit Groningen, te verkrijgen via <http://som.rug.nl/wieringa.pdf>.
- Leeflang, P.S.H. & D.R. Wittink (2000a), Building Models for Marketing Decisions: Past, Present and Future, *International Journal of Research in Marketing*, vol. 17, pp. 105-126.
- Leeflang, P.S.H. & D.R. Wittink (2000b), Models for Marketing Decisions: Postscriptum, *International Journal of Research in Marketing*, vol. 17, pp. 237-253.
- Leeflang, P.S.H., & D.R. Wittink (2002), Promotie van geneesmiddelen, *Economisch Statistische Berichten*, 87, nr. 4352, 228-231.
- Leeflang, P.S.H., D.R. Wittink, M. Wedel, Ph. A. Naert (2000), *Building Models for Marketing Decisions*, Kluwer Academic Publishers, Dordrecht/Boston.
- Lilien, G.L., Rao, A.G. & Kalish, S. (1981), Bayesian estimation and control of detailing effort in a repeat purchase diffusion environment, *Management Science*, 27, 493-506.
- Manchanda, P., & Chintagunta, P.K. (2003), Responsiveness of Physician Prescription Behavior to Salesforce Effort: An Individual Level Analysis, Working Paper, University of Chicago.
- Manchanda, P., Rossi, P.E. & Chintagunta, P.K. (2004), Response Modeling with Non-Random Mix Variables: Physician Prescription Behavior and Detailing, *Journal of Marketing Research*, forthcoming.
- Montgomery, D.B., & Silk, A.J. (1972), Estimating the effects of market communications expenditures, *Management Science*, 18, B-485-B-501.
- Narayan, S., Manchanda, P. & Chintagunta, P.K. (2003), The Informative versus Persuasive Role of Marketing Communication in New Product Categories: An Application to the Prescription Antihistamines Market, Working paper, University of Chicago ([http://papers.ssrn.com/sol3/papers.cfm?abstract\\_id=472881](http://papers.ssrn.com/sol3/papers.cfm?abstract_id=472881)).
- Neslin, S.A. (2001), ROI Analysis of Pharmaceutical Promotion (RAPP), Presentation to the Association of Medical Publications, May 22, 2001 (<http://www.rxpromoroi.org/rapp/index.html>).
- Parsons, L.J., & Vanden Abeele, P. (1981), Analysis of Sales Call Effectiveness, *Journal of Marketing Research*, 18, 107-113.
- Raftery, A.E. (1995), Bayesian Model Selection in Social Research (with discussion by Andrew Gelman, Donald B. Rubin and Robert M. Hauser), in *Sociological Methodology 1995* (ed. Peter V. Marsden), Blackwell, Oxford, U.K. pp. 111-196.
- Rao, A.G., & Yamada, M. (1988), Forecasting with a repeat purchase diffusion model, *Management Science*, 34, 734-752.
- Rizzo, J.A. (1999), Advertising and competition in the ethical pharmaceutical industry: The case of antihypertensive drugs, *Journal of Law and Economics*, 42, 89-116.
- Rosenthal, M.B., Berndt, E.R., Donohue, J.M., Frank, R.G. & Epstein, A.M. (2002), Promotion of prescription drugs to consumers, *The New England Journal of Medicine*, 346, 498-505.
- Shankar, V. (2003), The Roles of the Product Life Cycle and Market Dominance in Marketing Expenditures of Products, Working paper, University of Maryland.
- Steenburgh, T.J. (2003), Market-Source Decompositions of Demand, Working paper, Harvard University.

- Wittink, D.R. (2002), Analysis of ROI for Pharmaceutical Promotion (ARPP), Presentation to the Association of Medical Publications, September 8, 2002 (<http://www.rxpromoroi.org/arpp/index.html>).
- Wosińska (2002), Just What the Patient Ordered? Direct-to-Consumer Advertising and the Demand for Pharmaceutical Products, Harvard Business School Marketing Research Papers No. 02-04.



# 11. Percepties van energieverbruik van huishoudelijke apparaten

G. SCHUITEMA en L. STEG

## SAMENVATTING

Kennis over het energieverbruik huishoudelijke apparaten is belangrijk om efficiënt energie te besparen. Deze kennis is niet altijd accuraat. In deze studie werd nagegaan welke factoren dergelijke (mis)percepties van het energieverbruik van huishoudelijke apparaten beïnvloeden. Ten eerste werd nagegaan of kenmerken van de apparaten van invloed zijn. Nagegaan werd of de perceptie van het energieverbruik van huishoudelijke apparaten afhankelijk is van de zichtbaarheid, grootte, status, geluid, bezit en waargenomen noodzakelijkheid van de apparaten. Grootte bleek de belangrijkste voorspeller van de perceptie van het energieverbruik te zijn. Deze heurstiek is over het algemeen accuraat, maar leidt soms tot mispercepties (bv. zonnepaneel, koffiezetapparaat). Ten tweede werd nagegaan of de perceptie van het energieverbruik van huishoudelijke apparaten beïnvloed wordt door individuele kenmerken. Geslacht, inkomen, huishoudtype, woonomgeving en milieubesef bleken geen invloed te hebben op de perceptie van het energieverbruik van huishoudelijke apparaten.

## 1. INLEIDING

Volgens het Kyoto Protocol moet Nederland tussen 2008 en 2012 8% minder broeikasgassen uitstoten ten opzichte van 1990. In Nederland neemt de hoeveelheid uitgestoten broeikasgassen echter alleen maar toe, voornamelijk door een toename van de uitstoot van CO<sub>2</sub> (CBS, 2004). Ongeveer 11% van de totale uitstoot van CO<sub>2</sub>-emissies wordt door huishoudens geproduceerd (Energie in cijfers, 2004a). Het totale energieverbruik van huishoudens bestaat voor 54% uit indirect energieverbruik en voor 46% uit direct energieverbruik (Vringer & Blok, 2000). Direct energieverbruik is de hoeveelheid elektriciteit, gas en brandstof die huishoudens daadwerkelijk zelf *in* hun huishouden consumeren. Indirect energieverbruik verwijst naar de hoeveelheid energie die nodig is om goederen te maken, te vervoeren en uiteindelijk te verwerken als afval (voor een overzicht van indirect energieverbruik, zie Vringer & Blok, 2000).

Het directe energieverbruik van huishoudens is de afgelopen 15 jaar met ongeveer 2% toegenomen (CBS, 2004; Steg, 1999; Energie in cijfers, 2004b). Deze toename is vooral toe te schrijven aan een toename van het elektriciteitsverbruik in Nederland (CBS, 2004). Steg (1999) verklaart deze toename ten eerste doordat er steeds meer elektrische apparaten in huishoudens komen en, ten tweede, doordat er steeds meer apparaten continue stroom verbruiken doordat zij een *stand-by* functie hebben. Een toename van het aantal huishoudelijke apparaten verhoogt overigens niet alleen het directe energiever-

bruik, maar ook het indirecte energieverbruik van huishoudens. Veel mensen lijken zich niet bewust te zijn van hun indirecte energieverbruik. Een apparaat dat niet wordt gebruikt kost bijvoorbeeld ook energie omdat het geproduceerd en vervoerd moet worden. Vooral bij de aanschaf van apparaten zouden mensen rekening kunnen houden met hun indirecte energieverbruik, door bijvoorbeeld te kijken naar het materiaal waarmee het apparaat gemaakt is (energie voor productie en afvalverwerking) en uit welk land het apparaat komt (energie voor vervoer). In dit onderzoek beperken we ons tot het directe energieverbruik van huishoudens.

Hoewel veel Nederlanders zich bewust zijn van de invloed van huishoudelijke apparaten op het milieu door hun energieverbruik, blijken zij niet bereid om huishoudelijke apparaten de deur uit te doen. Men blijkt daarentegen wel bereid te zijn om huishoudelijke apparaten minder te gebruiken (Gatersleben & Vlek, 1998). Om hun energieverbruik efficiënt te kunnen verminderen, door bepaalde apparaten minder of anders te gebruiken, is het belangrijk dat mensen inzicht hebben in de hoeveelheid energie die apparaten daadwerkelijk verbruiken.

Het energieverbruik van huishoudelijke apparaten kan op veel verschillende manieren gedefinieerd worden. Een veel gebruikte definitie van het energieverbruik is *de hoeveelheid energie die een apparaat gedurende een uur gebruikt terwijl het 'aan' dan wel 'stand-by' staat* (het vermogen van een apparaat, gemeten in kilowattuur (kWh)). Energie besparen door middel van gedragsverandering zou dan inhouden dat mensen vooral apparaten die veel energie verbruiken als ze 'aan' staan (groot vermogen) minder gebruiken. Vanzelfsprekend levert het minder gebruiken van apparaten met weinig vermogen een geringere energiebesparing op. Daarnaast kan energie worden bespaard door apparaten met minder vermogen aan te schaffen.

Een andere definitie van energieverbruik is *de gemiddelde hoeveelheid energie die een apparaat in een gemiddeld gezin verbruikt in een bepaald tijdsbestek*, bijvoorbeeld per jaar. Het energieverbruik is dan niet alleen meer afhankelijk van het vermogen, maar ook van het gebruik van het apparaat. Een apparaat met weinig vermogen dat continu energie verbruikt, heeft per jaar een hoog gemiddeld verbruik. Een apparaat met een groot vermogen dat zeer weinig gebruikt wordt heeft, daarentegen, per jaar een laag gemiddeld energieverbruik. In dit geval is het zinvol om energie te besparen door apparaten minder vaak te gebruiken. Ook de aanschaf van apparaten met minder vermogen (vooral van apparaten die continu of zeer frequent worden gebruikt) zal energiebesparing opleveren. In dit onderzoek is voor deze laatste definitie van energieverbruik gekozen.

De perceptie die mensen hebben van het energieverbruik van huishoudelijke apparaten blijkt niet altijd accuraat te zijn (Baird & Brier, 1981). Zo is bijvoorbeeld het energieverbruik van een koffiezetapparaat met een warmhoudplaatje vrij hoog. Als mensen dit niet weten, zullen zij waarschijnlijk hun koffie niet in een thermosfles doen om energie te besparen. Daarentegen is het ook mogelijk dat mensen bewust een scheerapparaat of elektrische tandenborstel heel weinig of niet gebruiken, terwijl deze apparaten vrijwel geen invloed hebben op het energieverbruik van een huishouden. Deze studie richt zich op factoren die dergelijke (mis)percepties van het energieverbruik van huishoudelijke apparaten kunnen verklaren.

## 2. CATEGORISATIE

De belangrijkste aanname van dit onderzoek is dat (mis)percepties worden veroorzaakt doordat mensen voortdurend alles wat ze waarnemen categoriseren. Categorisatie is een belangrijk hulpmiddel om de vele en complexe stimuli, waar mensen elke dag mee te maken hebben, te verwerken. Door deze stimuli in overzichtelijke groepen, categorieën, op te delen zijn mensen beter in staat om snel en accuraat te reageren.

Mensen blijken vrijwel alles te categoriseren. Zo worden onder andere personen, gebeurtenissen (Fiske & Taylor, 1984) en objecten (Baird & Brier, 1981) in categorieën opgedeeld. Bij categorisatie worden stimuli, zoals huishoudelijke apparaten, in groepen ingedeeld. Stimuli die als gelijk worden beschouwd worden in één categorie ingedeeld (Rosch, Mervis, Gray, Johnson, & Boyes-Bream, 1976). Met andere woorden, op basis van kenmerken van objecten bepalen mensen welke objecten samen in één categorie belanden. Omdat mensen vaak snel beslissingen moeten nemen is het onmogelijk om alle kenmerken van objecten naast elkaar te leggen en uitgebreide afwegingen te maken over het indelen van objecten in categorieën. Om snel beslissingen te kunnen nemen maken mensen gebruik van heuristieken, waarbij ze letten op een beperkt aantal kenmerken. Kenmerkend voor het gebruik van heuristieken is dat het snelle beslisprocedures zijn, die weinig informatieverwerking vergen (O'Keefe, 1990). Het nemen van snelle beslissingen heeft veel voordelen en is vaak noodzakelijk om te overleven.

Hoewel het snel nemen van beslissingen een voordeel is, is een nadeel van het gebruik van heuristieken dat er vertekeningen op kunnen treden. Mensen zijn geneigd om alle objecten binnen één categorie dezelfde eigenschappen toe te schrijven. Bijvoorbeeld, op basis van hun aanschafprijs heeft iemand een aantal apparaten gecategoriseerd: dure apparaten in de ene categorie en goedkope apparaten in een andere categorie. Als deze persoon nu uitspraken moet doen over de kwaliteit van de apparaten kan hij/ zij bijvoorbeeld zeggen dat alle apparaten in de categorie 'duur' van goede kwaliteit zijn en de apparaten in de categorie 'goedkoop' van slechte kwaliteit. Nu kan dit over het algemeen een goede en bruikbare heuristiek zijn om snelle uitspraken te doen, maar het is goed mogelijk dat deze uitspraak niet altijd waar is. Immers, er zijn ook goedkope apparaten van goede kwaliteit en dure apparaten van slechte kwaliteit.

Er zijn veel verschillende modellen die beschrijven hoe mensen heuristieken gebruiken bij het maken van categorieën (zie voor een overzicht Berretty, Todd, & Martignon, 1999). Voor deze studie wordt uitgegaan van het *Categorization by Elimination* model (CBE). Het CBE model gaat ervan uit dat mensen zo min mogelijk kenmerken gebruiken om categorieën van objecten te maken, zodat de categorisatie zo efficiënt mogelijk is (Berretty et al., 1999).

Dit onderzoek gaat na welke heuristieken mensen gebruiken bij het schatten van het energieverbruik van huishoudelijke apparaten. Met andere woorden, op basis van welke kenmerken categoriseren mensen huishoudelijke apparaten en gebruiken ze deze kenmerken bij het schatten van het energieverbruik van de apparaten? We verwachten dat mensen huishoudelijke apparaten categoriseren met behulp van heuristieken, waarna ze aan elke categorie een bepaalde mate van energieverbruik koppelen. Mispercepties van het energieverbruik van huishoudelijke apparaten ontstaan naar verwachting doordat niet elk apparaat uit elke categorie evenveel energie verbruikt. In de volgende paragraaf wordt hier dieper op in gegaan. Daarnaast wordt nagegaan of individuele kenmerken van invloed zijn op de perceptie van het energieverbruik. Verwacht wordt

dat individuele factoren invloed hebben op de perceptie van het energieverbruik omdat niet iedereen dezelfde kennis over en betrokkenheid bij energieverbruik heeft.

### *2.1. Invloed van kenmerken van huishoudelijke apparaten op de perceptie van het energieverbruik*

Op basis van het CBE model wordt verwacht dat mensen slechts een klein aantal kenmerken gebruiken bij het inschatten van het energieverbruik van huishoudelijke apparaten (hypothese 1). In dit onderzoek wordt nagegaan welke kenmerken van belang zijn. Omdat categoriseren een snel proces is, is het waarschijnlijk dat mensen zich bij het inschatten van het energieverbruik van huishoudelijke apparaten vooral baseren op hele duidelijke en gemakkelijk waarneembare kenmerken van deze apparaten. Immers, gebruik maken van kenmerken die minder in het oog springen zou teveel tijd en informatieverwerking vergen. In paragraaf 2.1.1 worden drie kenmerken besproken die gemakkelijk waar te nemen zijn en die mensen zouden kunnen gebruiken bij het inschatten van het energieverbruik van huishoudelijke apparaten: de zichtbaarheid, het geluid dat huishoudelijke apparaten maken en de grootte. In paragraaf 2.1.2 worden drie kenmerken besproken die (mis)percepties van het energieverbruik van huishoudelijke apparaten kunnen verklaren door het ontstaan van cognitieve dissonantie (Festinger, 1954). Dit zal vooral het geval zijn voor apparaten die mensen aantrekkelijk of noodzakelijk vinden om te bezitten. Als mensen aan de ene kant een apparaat heel aantrekkelijk vinden, maar aan de andere kant zich realiseren dat dit apparaat (veel) energie verbruikt, kan er cognitieve dissonantie ontstaan. Mensen proberen over het algemeen cognitieve dissonantie te reduceren. Eén manier om dit te doen is door de perceptie van het energieverbruik van huishoudelijke apparaten aan te passen. In paragraaf 2.1.2 wordt ingegaan op de relatie tussen het energieverbruik van huishoudelijke apparaten en het bezit van een apparaat, de waargenomen noodzaak van dit bezit en de status die mensen toekennen aan huishoudelijke apparaten.

#### *2.1.1 Zichtbaarheid, geluid en grootte*

We verwachten dat mensen het energieverbruik van apparaten met gemakkelijk waar te nemen kenmerken hoger inschatten dan het energieverbruik van de apparaten met minder opvallende kenmerken. Het eerste kenmerk dat van belang zou kunnen zijn is de *zichtbaarheid* van een apparaat. Verwacht wordt dat mensen het energieverbruik van zichtbare apparaten (bijvoorbeeld een televisie) hoger inschatten dan het energieverbruik van minder of niet zichtbare huishoudelijke apparaten (bijvoorbeeld een boiler) (hypothese 2).

Een ander kenmerk dat invloed zou kunnen hebben op de perceptie van het energieverbruik van huishoudelijke apparaten is de hoeveelheid geluid dat een apparaat maakt. De verwachting is dat mensen denken dat apparaten die veel geluid maken (en dus meer opvallen) meer energie verbruiken dan apparaten die weinig energie verbruiken (hypothese 3).

Tenslotte zou ook de grootte van de apparaten van invloed kunnen zijn op de perceptie van het energieverbruik van huishoudelijke apparaten. In hun studie vroegen Baird en Brier (1981) mensen om een aantal huishoudelijke apparaten in te delen naar energieverbruik. Hieruit bleek dat mensen bij het maken van de indeling vooral letten op de grootte van de apparaten. In overeenstemming met dit onderzoek wordt verwacht dat mensen denken dat grote apparaten veel energie verbruiken en dat kleine apparaten weinig energie verbruiken (hypothese 4).

### 2.1.2 Bezit, noodzaak en status

Een vierde kenmerk dat van invloed zou kunnen zijn op de perceptie van het energieverbruik van huishoudelijke apparaten is de bekendheid die mensen met het apparaat hebben. In een studie liet Wesp (2000) mensen de grootte van bekende objecten (bijvoorbeeld een videorecorder) en onbekende objecten (bijvoorbeeld een stuk hout ter grootte van een videorecorder) inschatten. Hieruit bleek dat mensen een betere inschatting van de grootte van onbekende objecten konden maken dan van bekende objecten. Mensen schatten bekende objecten niet consequent groter of kleiner dan onbekende objecten, maar de *fout* die ze maakten was groter.

Bij het inschatten van het energieverbruik van huishoudelijke apparaten zou de bekendheid met de apparaten een rol kunnen spelen. Op basis van de cognitieve dissonantie theorie van Festinger (1954) verwachten we dat mensen die een apparaat bezitten, en er dus bekend mee zijn, het energieverbruik van dit apparaat lager inschatten dan mensen die dit apparaat niet bezitten. Aan de ene kant realiseren mensen zich dat het bezit van huishoudelijke apparaten invloed heeft op het milieu en het totale energieverbruik. Aan de andere kant zijn mensen niet bereid het aantal apparaten dat zij hebben te verminderen (Gatersleben & Vlek, 1998). Deze twee cognities zijn niet in overeenstemming met elkaar, zodat er cognitieve dissonantie ontstaat. Over het algemeen proberen mensen cognitieve dissonantie te reduceren door één van beide cognities aan te passen. In dit geval is het niet waarschijnlijk dat mensen afstand van huishoudelijke apparaten doen. Het is gemakkelijker om de perceptie van het effect van de huishoudelijke apparaten op het milieu en dus de perceptie van het energieverbruik van apparaten aan te passen. Om de ontstane cognitieve dissonantie te reduceren gaan mensen dus naar verwachting het energieverbruik van huishoudelijke apparaten die zij bezitten onderschatten. De verwachting is dan ook dat iemand die een bepaald apparaat bezit het energieverbruik ervan lager inschat dan iemand die dit apparaat niet bezit (hypothese 5).

Naar verwachting wordt dit effect nog sterker als mensen het bezit van een apparaat ook noodzakelijk vinden. Immers, mensen zijn dan nog minder snel geneigd om huishoudelijke apparaten weg te doen, en zijn nog sterker geneigd om hun perceptie van het energieverbruik aan te passen. Naar verwachting denkt iemand die het bezit van een apparaat noodzakelijk vindt dat dit apparaat minder energie verbruikt dan iemand die het bezit van dat apparaat niet noodzakelijk vindt (hypothese 6).

Het is mogelijk dat er ook cognitieve dissonantie optreedt voor apparaten waar mensen status aan toekennen. Over het algemeen vinden mensen het prettig om apparaten te bezitten die hun status verschaffen. Het is daarentegen vaak niet aantrekkelijk om hieraan een negatief aspect, zoals een hoog energieverbruik, te koppelen. We verwachten dan ook, om de ontstane cognitieve dissonantie te reduceren, dat mensen eerder hun perceptie van het energieverbruik van statusvolle apparaten aanpassen dan dat zij statusvolle apparaten minder aantrekkelijk vinden om te hebben. We verwachten daarom dat mensen die een apparaat statusvol vinden het energieverbruik van dit apparaat lager inschatten dan iemand die dit apparaat niet statusvol vindt (hypothese 7).

## 2.2. Samenhang tussen perceptie van energieverbruik en milieubesef en sociaal-demografische kenmerken

Een attitude is de neiging om in gunstige of ongunstige zin te reageren op een object, persoon, instantie of gebeurtenis (Ajzen, 1988). Doordat een attitude tot een negatief dan wel positief oordeel leidt ten opzichte van een object, hebben attitudes invloed op

de perceptie. Een attitude die voor de perceptie van het energieverbruik van huishoudelijke apparaten relevant zou kunnen zijn is *milieubesef*. Mensen met een hoog milieubesef vinden onder andere eenheid met de natuur, bescherming van het milieu, respect voor de aarde, (sociale) gelijkheid en een gevoel van thuishoren belangrijker dan mensen met een laag milieubesef.

Een veelgebruikte schaal die milieubesef meet is de *New Environmental Paradigm* (NEP) van Dunlap, Van Liere en hun collega's (Dunlap & Van Liere, 1978; Dunlap, Van Liere, Mertig, & Jones, 2000; Van Liere & Dunlap, 1981; Dunlap, Van Liere, Mertig, Catton, & Howell, 1992). Milieubesef zou invloed kunnen hebben op de perceptie van het energieverbruik van huishoudelijke apparaten omdat mensen met een hoog milieubesef wellicht meer kennis over en betrokkenheid bij energieverbruik hebben dan mensen met een laag milieubesef. In dit onderzoek wordt nagegaan of de perceptie van het energieverbruik van huishoudelijke apparaten samenhangt met milieubesef (onderzoeksvraag 1).

In dit onderzoek wordt ook voor een aantal sociaaldemografische kenmerken nagegaan of zij samenhangen met de perceptie van het energieverbruik van huishoudelijke apparaten. Dit is vooral van belang om doelgroepen te onderscheiden voor beleid. Er zal worden nagegaan wat de invloed is van geslacht, inkomen, woonomgeving (stedelijk, landelijk) en huishoudtype op de perceptie van het energieverbruik van huishoudelijke apparaten (onderzoeksvraag 2). Er wordt alleen gekeken naar het effect van inkomen en niet naar de invloed van sociaal-economische status (inkomen en opleidingsniveau), omdat dit onderzoek werd afgenomen onder studenten (die allen hetzelfde opleidingsniveau hebben). Overigens blijken effecten van sociaal-economische status op elektriciteitsverbruik vooral aan inkomen toegeschreven te moeten worden en niet aan opleidingsniveau (Steg, 1999).

### 3. METHODE

#### 3.1. Respondenten

In maart 2002 hebben 60 studenten, voornamelijk van de faculteit der Psychologische, Pedagogische en Sociologische Wetenschappen (PPSW) van de Rijksuniversiteit Groningen, een vragenlijst ingevuld. Zij werden in de kantine en de bibliotheek van deze faculteit benaderd. In totaal hebben 22 mannen en 38 vrouwen de vragenlijst ingevuld. Het inkomen van 28 respondenten lag onder de 500,-; het inkomen van 32 respondenten was hoger dan 500,-. Achtendertig respondenten woonden in een gedeeld huishouden (studentenhuis, bij ouders, woongroep) en 22 respondenten voerden een eigen huishouden (alleen, met partner en/ of met kinderen). Elf respondenten woonden in een combinatie van een landelijke en stedelijke omgeving, 49 respondenten woonden in een stedelijke omgeving.

#### 3.2. Materiaal

De vragenlijst bestond uit vier onderdelen. In het eerste deel werd de respondenten gevraagd 26 huishoudelijke apparaten te beoordelen op 7 eigenschappen. In het tweede deel schatten respondenten het energieverbruik van deze 26 apparaten in op een ratio-schaal. In het derde deel werd, met behulp van 8 stellingen, het milieubesef gemeten.

Tenslotte werd naar een aantal sociaal-demografische kenmerken gevraagd.

In de vragenlijst werd, zoals gezegd, het energieverbruik van huishoudelijke apparaten gedefinieerd als de hoeveelheid energie die een apparaat verbruikt gedurende een jaar in een gemiddeld huishouden waar dit apparaat aanwezig is. Dit betekent dat men zowel rekening moest houden met het *vermogen* (de hoeveelheid energie die een apparaat verbruikt als het apparaat aan of stand-by staat) als met de *tijd* dat het apparaat aan of stand-by staat.

Voor de vragenlijst zijn 26 huishoudelijke apparaten zorgvuldig geselecteerd: ze varieerden systematisch op energieverbruik, grootte, zichtbaarheid, en geluid. Deze selectie was gebaseerd op een database van het energieverbruik van huishoudelijke apparaten (BEK, 1997). Een aantal apparaten waar mensen wellicht status aan toekennen is ook toegevoegd aan de lijst (zie Tabel 1).

Tabel 1. 26 huishoudelijke apparaten en hun energieverbruik.

	Huishoudelijk apparaat	Verbruik per toestel (kWh/ jaar)
1	waterbed	734.3
2	boiler (<20 liter)	719.0
3	wasdroger	541.5
4	tweedeurskoelkast	461.5
5	diepvriezer	397.6
6	koelkast met vriesvak	285.9
7	televisie	137.9
8	satellietontvanger	108.0
9	videorecorder	107.9
10	koffiezetapparaat	79.9
11	stofzuiger	54.0
12	stereo-installatie	51.9
13	elektrische vloerverwarming (in badkamer)	47.4
14	zonnebank	32.0
15	antwoordapparaat	26.3
16	radio	22.6
17	whirlpool (bubbelbad)	20.8
18	centrifuge (losstaand)	15.0
19	föhn	10.8
20	elektrische deurbel	8.8
21	espressomachine	7.2
22	elektrische tandenborstel	5.1
23	tosti-ijzer	3.0
24	elektrische boormachine	1.8
25	citruspers	1.8
26	batterij-oplader	1.7

Noot: het verbruik per toestel is het verbruik in kWh per jaar van een 'gemiddeld' geïnstalleerd toestel, gemiddeld wat betreft het vermogen en het aantal toestellen met een stand-by functie (BEK, 1997).

### 3.2.1. Beoordeling van huishoudelijke apparaten op 7 kenmerken

In het eerste deel van de vragenlijst werd aan de respondenten gevraagd of het apparaat zich in hun huishouden bevond (ja/ nee). Vervolgens werd gevraagd om de apparaten op 6 kenmerken te beoordelen op 5-puntsschalen: energieverbruik (1: het apparaat ver-

bruikt heel veel energie – 5: het apparaat verbruikt heel weinig energie), grootte (1: het apparaat is groot – 5: het apparaat is klein), zichtbaarheid (1: het apparaat is goed zichtbaar – 5: het apparaat is niet zichtbaar), geluid (1: het apparaat maakt heel veel geluid – 5: het apparaat maakt geen geluid), noodzakelijkheid (1: het is heel noodzakelijk om het apparaat te bezitten – 5: het is totaal niet noodzakelijk om het apparaat te bezitten) en status (1: het apparaat geeft de eigenaar heel veel status – 5: het apparaat geeft de eigenaar geen status).

### 3.2.2. Perceptie van het energieverbruik van huishoudelijke apparaten

In het tweede deel van de vragenlijst werd dieper ingegaan op het energieverbruik van de 26 huishoudelijke apparaten. Respondenten vulden in hoeveel energie zij dachten dat de apparaten per jaar verbruikten in een gemiddeld huishouden. Om het schatten van het energieverbruik te vergemakkelijken was één apparaat gekozen als referentiepunt (een stereo-installatie). Het energieverbruik van de stereo-installatie werd van tevoren vastgesteld op 100 eenheden. Respondenten gaven voor alle overige apparaten aan hoeveel energie zij dachten dat deze apparaten verbruikten in vergelijking tot de stereo-installatie.

### 3.2.3. Milieubesef en achtergrondkenmerken

In het derde deel van de vragenlijst werd met behulp van een verkorte versie van de *New Environmental Paradigm* schaal (NEP) het milieubesef gemeten. Respondenten gaven op een 5-puntsschaal aan in hoeverre zij het eens waren met de stellingen (1: het helemaal eens – 5: helemaal mee oneens). De 8 geselecteerde stellingen waren voldoende betrouwbaar om ze alle op te nemen in een schaal voor milieubesef (Conbach's  $\alpha = .78$ ). Het gemiddelde milieubesef van de respondenten was 2.1, met een standaarddeviatie van 0.59. Dit houdt in dat over het algemeen het milieubesef van de respondenten vrij hoog was.

Tenslotte werd naar vier achtergrondkenmerken gevraagd: geslacht (man/ vrouw), soort huishouden (eigen huishouden (alleen, samen met partner en/ of kinderen, samen met kinderen) of gedeeld huishouden (bij ouders/ studentenhuis)), woonomgeving (landelijk, tussen landelijk en stedelijk in, stedelijk) en inkomen (minder dan € 500 en meer dan € 500).

## 3.3. Analyses

Alle 60 respondenten beoordeelden 26 apparaten op 7 kenmerken. Om na te gaan of er in het algemeen verbanden zijn tussen de perceptie van het energieverbruik van huishoudelijke apparaten en de zichtbaarheid, het geluid, de grootte, het bezit, de noodzaak en de status ervan werd eerst over alle respondenten en apparaten correlaties tussen de kenmerken berekend. Hierbij werd gedaan alsof 1 respondent 1 apparaat beoordeelde op de 7 kenmerken, in plaats van dat 1 respondent 26 apparaten heeft beoordeeld. Dat betekent dat voor deze analyses ervan werd uitgegaan dat er ( $26 \times 60 =$ ) 1560 respondenten waren. Deze aanname levert geen problemen op bij het berekenen van correlaties. Bij andere analyses, zoals regressieanalyses, is het niet plausibel om deze aanname te doen. Door de onderlinge afhankelijkheid van de data kon niet uitgegaan worden van gelijke varianties binnen de groepen, een voorwaarde voor het doen van regressieanalyses. Daarom is er besloten om analyses *per respondent* en analyses *per apparaat* uit te voeren.



De verwachtingen over het verband tussen de perceptie van het energieverbruik en de zichtbaarheid, geluid, grootte, bezit, noodzaak en status zijn op twee manieren gesteld. Ten eerste wilden we voor een aantal kenmerken (zichtbaarheid, grootte, geluid, status) weten hoe één respondent oordeelde over verschillende apparaten en hun kenmerken. Dit houdt in dat de *apparaten* worden beoordeeld *in vergelijking tot andere apparaten*, op basis van hun kenmerken. Hier worden, met andere woorden, oordelen over de apparaten onderling met elkaar vergeleken. We wilden bijvoorbeeld weten of respondenten het energieverbruik van grote apparaten hoger inschatten dan het energieverbruik van kleine apparaten. Om deze hypothesen te toetsen werden analyses per respondent uitgevoerd.

Ten tweede wilden we voor twee kenmerken (bezit en noodzaak) weten hoe *respondenten* de apparaten beoordelen *in vergelijking tot andere respondenten*. Bij deze analyses worden dus de respondenten onderling met elkaar vergeleken. We wilden bijvoorbeeld weten of mensen die een apparaat bezitten het energieverbruik van dit apparaat lager inschatten dan mensen die dit apparaat niet bezitten. Om deze hypothesen te toetsen werden analyses per apparaat uitgevoerd.

Om na te gaan of individuele kenmerken invloed hebben op de perceptie van het energieverbruik van huishoudelijke apparaten zijn ook analyses per apparaat uitgevoerd (variantieanalyses, Mann-Whitney toetsen).

## 4. RESULTATEN

### 4.1. Invloed van kenmerken van huishoudelijke apparaten op de perceptie van energieverbruik

Alle 60 respondenten hebben 26 apparaten beoordeeld op 7 kenmerken. Allereerst werd nagegaan in welke mate deze kenmerken (energieverbruik, zichtbaarheid, grootte, geluid, bezit, noodzaak, status) aan elkaar gerelateerd zijn door het berekenen van correlaties tussen de kenmerken over alle respondenten en apparaten heen.

Er werden drie behoorlijk hoge correlaties gevonden: tussen energieverbruik en grootte ( $r = .67$ ), tussen grootte en zichtbaarheid ( $r = .56$ ) en tussen aanwezigheid en noodzaak ( $r = .56$ ). Daarnaast was ook het verband tussen energieverbruik en zichtbaarheid tamelijk sterk ( $r = .38$ ) (zie Tabel 2).

Tabel 2. Correlaties tussen oordelen op 7 kenmerken van huishoudelijke apparaten.

	energie- verbruik	zichtbaarheid	grootte	geluid	bezit	noodzaak
zichtbaarheid	<b>.38</b>					
grootte	<b>.67</b>	<b>.56</b>				
geluid	.05	.16	.06			
bezit	-.18	.11	.16	.19		
noodzaak	-.04	.15	.02	.25	<b>.56</b>	
status	.26	-.02	.27	.08	-.21	-.17

Noot: de vier hoogste correlaties zijn vetgedrukt.

De hoge correlaties tussen het energieverbruik en de grootte ( $r = .67$ ) en de zichtbaarheid ( $r = .38$ ) van huishoudelijke apparaten wijzen erop dat respondenten het energieverbruik van zowel grote als zichtbare apparaten hoog inschatten en het energieverbruik van kleine en minder zichtbare apparaten laag inschatten. Grootte en zichtbaarheid hangen sterk samen ( $r = .56$ ); waarschijnlijk is één van beide de belangrijkste voorspeller van het energieverbruik van huishoudelijke apparaten. Een regressieanalyse kan hierover uitsluitsel geven (zie paragraaf 4.1.1).

Het verband tussen status en energieverbruik is niet erg sterk ( $r = .26$ ): respondenten schatten het energieverbruik van huishoudelijke apparaten waar zij status aan verbinden iets hoger in dan het energieverbruik van apparaten waar zij geen status aan verbinden. Dit effect is tegenovergesteld aan de verwachting. Overigens werd er ook een verband gevonden tussen de grootte en de status van de apparaten ( $r = .27$ ). Door het verband tussen de grootte en de status van apparaten is het goed mogelijk dat het energieverbruik door één van beide het beste voorspeld kan worden. Regressieanalyses zullen hierover meer duidelijkheid verschaffen (zie paragraaf 4.1.1).

Tussen de perceptie van het energieverbruik en de hoeveelheid geluid die de apparaten maakten werd geen verband gevonden ( $r = .05$ ). Ook tussen de perceptie van het energieverbruik en de aanwezigheid ( $r = -.18$ ) en de noodzaak van apparaten ( $r = -.04$ ) bestond geen sterk verband. Dit wijst erop dat de kenmerken geluid, aanwezigheid en noodzaak nauwelijks of geen relatie hebben met de perceptie van het energieverbruik van huishoudelijke apparaten.

Het sterke verband tussen de aanwezigheid van huishoudelijke apparaten en de noodzakelijkheid van deze apparaten ( $r = .56$ ) duidt erop dat respondenten de apparaten die zij bezitten ook noodzakelijk vinden. De causaliteit van het verband tussen aanwezigheid en noodzaak wordt door deze correlatie niet duidelijk: zijn de apparaten in een huishouden aanwezig omdat ze noodzakelijk zijn, of zijn de apparaten noodzakelijk geworden omdat ze aanwezig zijn?

#### 4.1.1. Invloed zichtbaarheid, grootte, geluid en status op perceptie van energieverbruik

De verwachting was dat mensen het energieverbruik hoger inschatten als een apparaat zichtbaar, groot of statusvol is en als het veel geluid maakt. Deze hypothesen werden getoetst met behulp van een stepwise regressieanalyse die voor elke respondent afzonderlijk werd uitgevoerd. Hieruit bleek dat voor één respondent geen van de vier kenmerken de perceptie van het energieverbruik van de 26 apparaten kon verklaren. Voor 51 van de 60 respondenten was de grootte van de apparaten de beste voorspeller van het geschatte energieverbruik. Bij 38 respondenten was grootte zelfs de enige voorspeller. De  $\beta$ -coëfficiënten toonden aan dat er in alle gevallen een positief verband was tussen de grootte en het geschatte energieverbruik van de 26 huishoudelijke apparaten. Met andere woorden: respondenten denken dat grootte apparaten veel energie verbruiken en dat kleine apparaten weinig energie verbruiken. Gemiddeld over alle respondenten verklaarde de grootte van de apparaten 46.7% van de variantie van het geschatte energieverbruik, een verassend hoog percentage.

Voor tien respondenten voorspelde de zichtbaarheid van de huishoudelijke apparaten een extra deel van de variantie van het geschatte energieverbruik van huishoudelijke apparaten. Voor slechts vijf van deze respondenten was de zichtbaarheid de beste voorspeller van het geschatte energieverbruik. Gemiddeld over alle respondenten verklaarde de zichtbaarheid van de huishoudelijke apparaten 3.5% van de variantie van het geschatte energieverbruik. De richting van het verband tussen de zichtbaarheid en het

geschatte energieverbruik is niet eenduidig: de  $\beta$ -coëfficiënten waren zowel positief als negatief. Dit betekent dat sommige mensen dachten dat zichtbare apparaten veel energie verbruikten en anderen dachten dat zichtbare apparaten weinig energie verbruikten. Het is zeer waarschijnlijk dat het gevonden verband door toeval werd veroorzaakt. Ten eerste omdat er geen systematisch verband tussen zichtbaarheid en energieverbruik werd gevonden en ten tweede omdat er 60 analyses (per respondent) zijn uitgevoerd. Hierdoor is de kans op het vinden van toevallig significante verschillen aanwezig.

Voor negen respondenten verklaarde de hoeveelheid geluid die de apparaten maken variantie in het geschatte energieverbruik van de huishoudelijke apparaten. Gemiddeld over alle respondenten was deze verklaarde variantie slechts 2.4%. Voor slechts twee respondenten was het geluid dat de apparaten maken de beste voorspeller van het geschatte energieverbruik. Voor vijf respondenten werd een negatieve  $\beta$ -coëfficiënt gevonden, voor vier een positieve  $\beta$ -coëfficiënt. Hierin is geen lijn te ontdekken. Daarom, en ook gezien het kleine aantal respondenten waarvoor een verband werd gevonden, kan worden geconcludeerd dat geluid geen belangrijke invloed heeft op de perceptie van het energieverbruik van huishoudelijke apparaten.

Voor vijf respondenten kon de status niet worden opgenomen in de regressieanalyse, omdat er geen variantie in de oordelen over de status van de apparaten bestond: zij gaven voor alle 26 apparaten aan dat zij geen status hadden. Voor slechts drie respondenten bleek status bij te dragen aan de voorspelling van het geschatte energieverbruik. Het percentage variantie in geschatte energieverbruik dat status verklaarde is gemiddeld over de zestig respondenten erg laag: 0.7%.

Uit deze regressieanalyses kan de conclusie worden getrokken dat de grootte van het apparaten het belangrijkste kenmerk is dat wordt gebruikt bij het schatten van het energieverbruik van huishoudelijke apparaten is. De eerder gevonden relatie tussen de zichtbaarheid en status met het geschatte energieverbruik (zie paragraaf 4.1) lijkt grotendeels toegeschreven te kunnen worden aan de grootte van de huishoudelijke apparaten.

#### 4.1.2. Invloed bezit en noodzaak op perceptie van energieverbruik

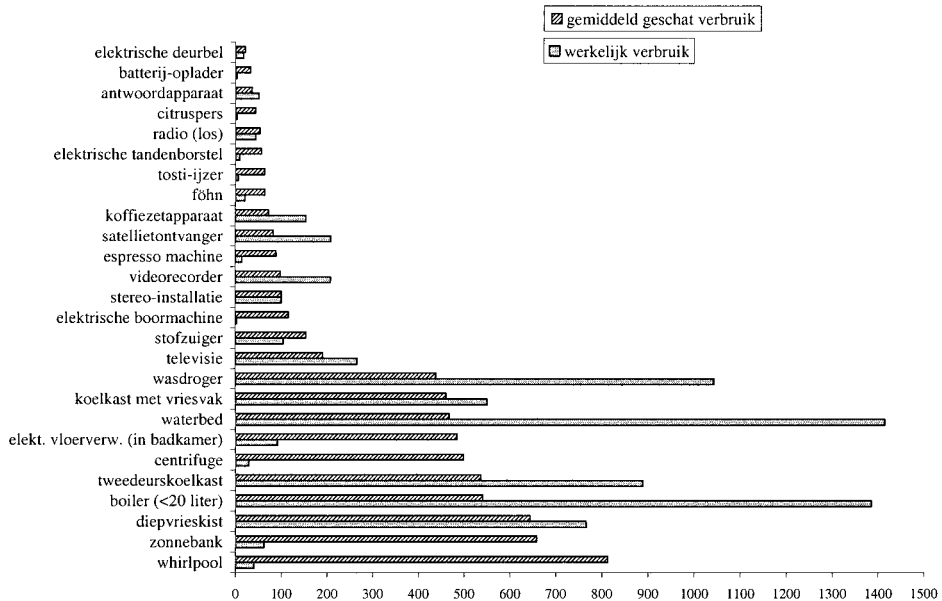
Om na te gaan of de aanwezigheid van huishoudelijke apparaten in een huishouden en de noodzakelijkheid van de apparaten gerelateerd zijn aan de perceptie van het energieverbruik werd voor elk apparaat afzonderlijk de correlaties tussen deze kenmerken berekend. Verwacht werd dat het energieverbruik van een apparaat lager wordt ingeschat door iemand die het apparaat bezit en/ of noodzakelijk vindt dan door iemand die het apparaat niet bezit en/ of niet noodzakelijk vindt. Analyses per apparaat maakten het mogelijk om het oordeel van respondenten onderling met elkaar te vergelijken.

Uit de correlaties bleek dat er geen verband is tussen het geschatte energieverbruik van huishoudelijke apparaten en de aanwezigheid van deze apparaten in een huishouden. Daarnaast bleek dat slechts voor 2 van de 26 apparaten (elektrische deurbel en elektrische vloerverwarming) een negatief verband werd gevonden tussen het oordeel over hoe noodzakelijk het bezit van dit apparaten werd gevonden en de perceptie van het energieverbruik hiervan. Dit betekent dat respondenten die een elektrische deurbel en elektrische vloerverwarming noodzakelijk vinden het energieverbruik van deze apparaten later inschatten dan respondenten die deze apparaten niet noodzakelijk vinden. Dat slechts voor twee apparaten een verband werd gevonden duidt een zwak verband tussen perceptie van noodzaak en perceptie van energieverbruik van apparaten.

#### 4.2. Werkelijk energieverbruik en perceptie van energieverbruik

Respondenten werd gevraagd op een ratioschaal aan te geven hoeveel energie zij dachten dat de 26 huishoudelijke apparaten verbruikten (zie paragraaf 3.3.2). Hiervoor moesten zij het energieverbruik van de huishoudelijke apparaten schatten ten opzichte van een referentiepunt: een stereo-installatie werd op 100 eenheden gesteld. Om de antwoorden van de respondenten te vergelijken met het werkelijke energieverbruik van de overige huishoudelijke apparaten, werd het werkelijke energieverbruik berekend ten opzichte van het verbruik van een stereo-installatie. In Figuur 1 is zowel het geschatte energieverbruik als het werkelijke energieverbruik van de 26 huishoudelijke apparaten weergegeven. Bij het bestuderen van Figuur 1 moet rekening worden gehouden met de verhoudingen van de weergegeven waarden. Het absolute verschil tussen het geschatte en het werkelijke energieverbruik is voor een apparaat dat in werkelijkheid veel energie verbruikt groter dan voor een apparaat dat in werkelijkheid weinig energie verbruikt. Het gaat ons vooral om de relatieve verschillen.

Over het algemeen blijkt dat het energieverbruik van apparaten die veel energie verbruiken werd onderschat. Dit neemt niet weg dat respondenten, terecht, dachten dat deze apparaten relatief de meeste energie verbruiken. Dit geldt bijvoorbeeld voor een televisie, koelkast met vriesvak, diepvrieskist, tweedeurskoelkast, wasdroger, boiler en waterbed. Deze apparaten zijn allemaal behoorlijk groot. Dit bevestigt dat grootte een belangrijke voorspeller is van de perceptie van het energieverbruik van huishoudelijke apparaten. Daarentegen werd het energieverbruik van een aantal apparaten met een laag



Figuur 1. Vergelijking werkelijke energieverbruik met waargenomen energieverbruik van 26 huishoudelijke apparaten. Voor deze vergelijking werd het energieverbruik van een stereo-installatie gesteld op 100 eenheden, het energieverbruik van de overige apparaten werd hieraan gereleerd.

energieverbruik consequent overschat door de respondenten. Dit zijn vooral de batterij-oplader, citruspers, elektrische boormachine, tosti-ijzer, elektrische tandenborstel, espressomachine, elektrische deurbel en föhn. Voor deze, kleine, apparaten geldt dat respondenten de apparaten, hoewel ze het energieverbruik overschatten, tot de apparaten rekenden die het minste energie verbruikten. Blijkbaar was het voor de respondenten moeilijk de *range* vast te stellen van het energieverbruik van de 26 huishoudelijke apparaten. De *range* van de het geschatte energieverbruik is veel kleiner dan die van het werkelijke energieverbruik.

Hoewel een deel van de misperceptie van het energieverbruik verklaard kan worden door de moeite die mensen hebben met het schatten van de *range* van apparaten, was er een aantal apparaten waarvoor deze verklaring niet opgaat. Ten eerste werd het energieverbruik van een waterbed en een boiler behoorlijk onderschat, terwijl het energieverbruik van een koelkast met vriesvak en een diepvrieskist toch nog redelijk accuraat was. Een boiler en waterbed verbruiken veel energie omdat ze grote hoeveelheden water verwarmen. Het lijkt erop dat veel mensen de hoeveelheid energie die nodig is om water te verwarmen onderschatten. Daarnaast hebben een boiler en een waterbed continu energie nodig, waardoor het energieverbruik hoog is. Dit kan ook een oorzaak zijn voor het onderschatten van het energieverbruik van deze apparaten.

Andere apparaten waarvan het energieverbruik relatief sterk werd onderschat waren een koffiezetapparaat, videorecorder en een satellietontvanger. Dat het energieverbruik van het koffiezetapparaat werd onderschat kan komen doordat mensen de energie die wordt verbruikt voor het warmhoudplaatje wellicht onderschatten. Daarnaast zijn een koffiezetapparaat, videorecorder en satellietontvanger tamelijk klein en niet goed zichtbaar.

Huishoudelijke apparaten waarvan het energieverbruik stelselmatig te hoog werd geschat waren een whirlpool, zonnebank, elektrische vloerverwarming en een centrifuge. Deze vier apparaten werden als grote apparaten gezien. Een zonnebank, centrifuge en whirlpool vindt men goed zichtbaar en aan een zonnebank, whirlpool en elektrische vloerverwarming werd redelijk wat status verbonden.

### 4.3. Invloed van individuele factoren op perceptie van energieverbruik

De individuele kenmerken die in dit onderzoek zijn meegenomen zijn milieubesef en de sociaal-demografische kenmerken geslacht, inkomen, soort huishouden en woonomgeving. Met behulp van variantieanalyses en Mann-Whitney toetsen werd nagegaan of deze factoren samenhangen met de perceptie van het energieverbruik (gemeten op een 5-puntsschaal).

#### 4.3.1. Samenhang tussen milieubesef en perceptie van energieverbruik

De verwachting dat mensen met een hoog milieubesef het energieverbruik van huishoudelijke apparaten anders in zouden schatten dan mensen met een laag milieubesef kon niet worden bevestigd. Uit variantieanalyses bleek dat milieubesef niet gerelateerd was aan de perceptie van het energieverbruik van huishoudelijke apparaten.

#### 4.3.2. Samenhang tussen geslacht, inkomen, soort huishouden en woonomgeving en perceptie energieverbruik

Om na te gaan of geslacht, inkomen, soort huishouden en woonomgeving samenhangen met de perceptie van het energieverbruik van huishoudelijke apparaten, werd per appa-

raat een Mann-Whitney toets gedaan. Variantieanalyses waren niet mogelijk, omdat de varianties in de groepen niet gelijk waren.

Voor één apparaat bleek een verschil in perceptie tussen mannen en vrouwen te bestaan: een satellietontvanger ( $Z = 2.50$ ,  $p < .05$ ). Omdat er voor elk apparaat afzonderlijk analyses zijn uitgevoerd is de kans op het vinden van toevallige significante verschillen aanwezig. Omdat slechts voor één van de 26 apparaten samenhang tussen geslacht en perceptie van het energieverbruik werd gevonden is het zeer waarschijnlijk dat dit resultaat aan toeval toegeschreven moet worden. Er is dan ook niet voldoende aanleiding om aan te nemen mannen en vrouwen verschillende percepties van energieverbruik van huishoudelijke apparaten hebben.

De twee onderscheiden inkomensgroepen (inkomen  $< € 500$  en inkomen  $> € 500$ ) bleken te verschillen in de perceptie van het energieverbruik van een boiler ( $Z = -2.24$ ,  $p < .05$ ). Dit resultaat geeft geen aanleiding om te veronderstellen dat inkomen samenhangt met perceptie van het energieverbruik van huishoudelijke apparaten.

Respondenten in een gedeeld huishouden versus eigen huishouden bleken helemaal niet te verschillen in hun perceptie van het energieverbruik van huishoudelijke apparaten. Het soort huishouden heeft dus geen invloed op de perceptie van het energieverbruik van huishoudelijke apparaten.

Tenslotte werd nagegaan of de woonomgeving van invloed was op de perceptie van het energieverbruik van huishoudelijke apparaten. Uit de Mann-Whitney toetsen bleek dat voor geen 26 van de huishoudelijke apparaten een verschil in perceptie bestond voor respondenten uit verschillende woonomgevingen (stad, tussen stad en platteland in of platteland). Met andere woorden, woonomgeving heeft geen invloed op perceptie van het energieverbruik van huishoudelijke apparaten.

## 5. CONCLUSIE EN DISCUSSIE

Het doel van dit onderzoek was na te gaan welke factoren van invloed zijn op de perceptie van het energieverbruik van huishoudelijke apparaten. Ten eerste was de verwachting dat *kenmerken van individuen* invloed hebben op de perceptie van het energieverbruik van huishoudelijke apparaten. In tegenstelling tot de verwachting bleken individuele factoren, in dit geval milieubesef en sociaal-demografische kenmerken, geen invloed te hebben op de perceptie van het energieverbruik van huishoudelijke apparaten. Op basis van dit onderzoek is het moeilijk een verklaring voor deze resultaten te geven, hiervoor zou nader onderzoek gedaan moeten worden. De resultaten geven echter wel aanleiding om te veronderstellen dat het irrelevant is om verschillende doelgroepen te onderscheiden bij de ontwikkeling van beleid op het gebied van perceptie van energieverbruik.

Ten tweede was de verwachting dat de perceptie van het energieverbruik van huishoudelijke apparaten afhankelijk is van kenmerken van deze apparaten. Op basis van het CBE model werd verwacht dat mensen apparaten categoriseren op basis van een beperkt aantal kenmerken, en dat men aan apparaten uit één categorie hetzelfde energieverbruik toekent. In dit onderzoek werd voor zes kenmerken van huishoudelijke apparaten nagegaan of zij invloed hebben op de perceptie van het energieverbruik van huishoudelijke apparaten: zichtbaarheid, grootte, geluid, noodzaak, bezit en status. Gebleken is dat het de grootte van het apparaat het meest belangrijke kenmerk is dat mensen gebruiken om het energieverbruik van huishoudelijke apparaten in te schatten: mensen denken dat grote

apparaten veel energie verbruiken en dat kleine apparaten weinig energie verbruiken. Hoewel er geen sprake is van een representatieve steekproef van de Nederlandse bevolking, is het gevonden verband tussen de grootte en de perceptie van het energieverbruik van huishoudelijke apparaten dusdanig sterk dat het zeer waarschijnlijk is dat hetzelfde verband wordt gevonden onder de Nederlandse bevolking.

De overige kenmerken van huishoudelijke apparaten bleken weinig invloed te hebben op de perceptie van het energieverbruik. De gevonden verbanden tussen zichtbaarheid en status van huishoudelijke apparaten en het geschatte energieverbruik moet grotendeels worden toegeschreven aan het verband tussen grootte en zichtbaarheid respectievelijk status van huishoudelijke apparaten. Dat slechts één kenmerk van huishoudelijke apparaten wordt gebruikt om het energieverbruik van huishoudelijke apparaten te schatten is in overeenstemming met de verwachting. Het CBE gaat er immers vanuit dat mensen zo weinig mogelijk kenmerken gebruiken om tot een oordeel te komen. Dit is snel en efficiënt. Daarnaast zijn er aanwijzingen gevonden dat mensen alleen kenmerken gebruiken die gemakkelijk waar te nemen zijn: de grootte van een apparaat is één van de eerste kenmerken die opvalt van huishoudelijke apparaten. Verder kan worden geconcludeerd dat cognitieve dissonantie geen belangrijke rol speelt bij de perceptie van het energieverbruik van huishoudelijke apparaten.

In overeenstemming met de resultaten van het onderzoek van Baird and Brier (1981) blijkt uit dit onderzoek dat mensen een simpele heuristiek gebruiken om het energieverbruik van huishoudelijke apparaten in te schatten: *grote apparaten verbruiken veel energie en kleine apparaten verbruiken weinig energie*. Deze heuristiek is over het algemeen bijzonder accuraat. In veel gevallen klopt het dat grote apparaten veel en kleine apparaten weinig energie verbruiken. Er zijn echter ook huishoudelijke apparaten waar deze heuristiek leidt tot mispercepties. Zo zijn er grote apparaten die relatief weinig energie verbruiken (bijvoorbeeld een centrifuge, zonnebank, whirlpool en elektrische vloerverwarming) en kleine apparaten die relatief veel energie verbruiken (bijvoorbeeld een koffiezetapparaat, videorecorder en satellietontvanger).

Hoewel uit dit onderzoek blijkt dat mensen vaak de heuristiek *grote apparaten verbruiken veel energie en kleine apparaten verbruiken weinig energie* gebruiken om het energieverbruik van huishoudelijke apparaten te schatten, hoeft dit niet de enige verklaring voor mispercepties van het energieverbruik van huishoudelijke apparaten te zijn. Andere factoren zouden hierin ook een rol kunnen spelen. Dit onderzoek geeft aanwijzingen dat het energieverbruik van apparaten die de functie 'water verwarmen' hebben (bijvoorbeeld waterbed) wordt onderschat. Ook lijken mensen het energieverbruik van apparaten die continu energie verbruiken te onderschatten (bijvoorbeeld een videorecorder). De resultaten uit dit onderzoek maken duidelijk dat er rekening moet worden gehouden met het verband dat mensen zien tussen de grootte en het energieverbruik van huishoudelijke apparaten in bijvoorbeeld informatiecampagnes. Ook bij het ontwerpen van informatielabels die aan huishoudelijke apparaten hangen zou hiermee rekening gehouden moeten worden.

## LITERATUUR

- Ajzen, I., 1988. Attitudes, personality and behavior. Milton Keynes: Open University Press.  
 Baird, J. C. and J.M. Brier, 1981. Perceptual awareness of energy requirements of familiar objects. *Journal of Applied Psychology*, 66, 90-96.

- BEK, 1997. Basisonderzoek Elektriciteitsverbruik Kleinverbruikers Arnhem: Energienet, Vereniging van Energiedistributiebedrijven in Nederland.
- Berretty, P. M., Todd, P. M. and L. Martignon, 1999. 'Categorization by Elimination'. In G. Gigerenzer and P. M. Todd (Eds.), *Simple heuristics that make us smart*. New York: Oxford University Press, pp. 235-254.
- CBS, 2004. Nederland langs de Europese meetlat. [www.cbs.nl](http://www.cbs.nl).
- Dunlap, R. M. and K.D. Van Liere, 1978. The new environmental paradigm: a proposed measuring instrument and preliminary results. *Journal of Environmental Education*, 9, 10-19.
- Dunlap, R. M., Van Liere, K. D., Mertig, A. D., Catton, W. R. and R.E. Howell, 1992. Measuring endorsement of an ecological worldview: a revised NEP-scale. Paper presented at the 1992 meeting of the Rural Sociological Society, State College, Pennsylvania.
- Dunlap, R. M., Van Liere, K. D., Mertig, A. D. and R.E. Jones 2000. Measuring endorsement of the new ecological paradigm: a revisited NEP-scale. *Journal of Social Issues*, 56, 425-442.
- Energie in cijfers, 2004a. CO<sub>2</sub>-emissies naar IPCC-sector in 2000. [www.energie.nl](http://www.energie.nl).
- Energie in cijfers, 2004b. Gemiddeld energiegebruik per huishouden. [www.energie.nl](http://www.energie.nl).
- Festinger, L., 1954. A theory of social comparison processes. *Human Relations*, 7, 117-140.
- Fiske, S. T. and S.E. Taylor, 1984. *Social cognition*. Reading, Ma: Addison-Wesley Publishing Company.
- Gatersleben, B. and C.A.J. Vlek, 1998. 'Household consumption, quality of life and environmental impacts: a psychological perspective and empirical study'. In: K.J. Noorman & T. Schoot Uiterkamp (Eds.), *Green households? Domestic consumers, environment and sustainability*. London: Earthscan Publications, pp. 141-183.
- O'Keefe, D., 1990. *Persuasion, theory and research*. Newbury Park: Sage Publications.
- Rosch, E., Mervis, C. B., Gray, W. D., Johnson, D. M., and P. Boyes-Bream, P., 1976. Basic objects in natural categories. *Cognitive Psychology*, 8, 382-439.
- Steg, L., 1999. *Verspilde energie? Wat doen en laten Nederlanders voor het milieu*. Den Haag: Sociaal en Cultureel Planbureau.
- Van Liere, K. D. and R.M. Dunlap, 1981. Environmental concern: does it make a difference how it's measured? *Environment and Behavior*, 13, 651-676.
- Vringer, K. & K. Blok, 2000. Long-term trends in direct and indirect household energy intensities: a factor in dematerialisation? *Energy Policy*, 8, 713-727.
- Wesp, R., 2000. Object recognition may distort size perception. *Perceptual and Motor Skills*, 90, 803-809.



## *12. Leveren publieke prestaties vertrouwen op? Een internationale vergelijking*

B. KUHRy en E.J.POMMER\*

### SAMENVATTING

Op verzoek van het ministerie van Binnenlandse Zaken en Koninkrijksrelaties heeft het Sociaal en Cultureel Planbureau in samenwerking met onderzoekers van andere instellingen recentelijk een vergelijkend rapport uitgebracht over de prestaties van de publieke sector in 29 landen ('Prestaties van de publieke sector', ook vertaald in het Engels en Frans). Naast een beschrijving van de verschillen tussen landen is ook gepoogd om deze te verklaren uit institutionele en andere achtergrondkenmerken en om na te gaan in hoeverre landen kunnen leren van 'best practices' elders. Een van de bevindingen van het rapport betreft de geringe samenhang tussen het niveau van de prestaties en de daarmee corresponderende uitgaven. In deze verslaglegging wordt de nadruk gelegd op een ander aspect, namelijk de relatie tussen de geleverde prestaties en het vertrouwen van de bevolking in de betrokken instellingen. Deze analyse lijkt uit te wijzen dat er ook maar een matig verband bestaat tussen werkelijkheid en perceptie.

### 1. INLEIDING

In het jaar 2000 hebben de lidstaten van de Europese Unie (EU) de Lissabon Agenda aanvaard. Die agenda legt lidstaten de ontzagwekkende taak op de Unie binnen tien jaar te maken tot de meest concurrerende economische zone in de wereld. Vier jaar later lijkt het in toenemende mate twijfelachtig of deze ambitieuze doelstelling zal worden gehaald. Hoe dan ook, prestaties van de publieke sector van nationale economieën vormen een doorslaggevende factor bij de wedren om de doelstellingen van de Lissabon Agenda volgens plan te verwezenlijken. Landen kunnen proberen het functioneren van hun publieke sector te verbeteren door succesvolle praktijken uit andere landen over te nemen. Met dit doel in het achterhoofd is het rapport "Public Sector Performance" opgesteld.<sup>1</sup>

De primaire onderzoeksvraag heeft betrekking op het in kaart brengen van verschillen in prestaties. Daarbij gaat het om verschillen in productiviteit, kwaliteit en doeltreffendheid van door de overheid georganiseerde dienstverlening. Vervolgens proberen wij na te gaan in hoeverre gevonden verschillen in prestaties samenhangen met

\* De auteurs zijn werkzaam op het Sociaal en Cultureel Planbureau. Aan de hier gepresenteerde resultaten hebben ook andere onderzoekers bijgedragen, te weten: C.A. de Kam (Rijksuniversiteit Groningen), A.G.J. van der Torre, R. Heesakker en L. Herweijer (allen SCP) en P. Smit (Wetenschappelijk Onderzoeks- en Documentatiecentrum).

ationale instituties van de in het onderzoek betrokken landen en met de ingezette middelen. Bij het laatstgenoemde aspect moet niet alleen rekening worden gehouden met publieke uitgaven, maar ook met private uitgaven.

Dit artikel geeft een samenvatting van de uitkomsten op een drietal beleidsterreinen: (1) onderwijs, (2) gezondheidszorg, (3) criminaliteitsbestrijding. Het achterliggende rapport gaat daarnaast ook in op economische prestaties en op het openbaar bestuur. Het vergelijkt prestaties van de publieke sector in de 15 oorspronkelijke EU-lidstaten, de 10 nieuwe lidstaten van de EU en in vier landen die geen lid van de Unie zijn: Australië, Canada, Nieuw-Zeeland en de Verenigde Staten (VS).<sup>2</sup> Bij de intrede van 10 nieuwe landen in de EU is het interessant om na te gaan in hoeverre deze op allerlei kenmerken verschillen van de oorspronkelijke 15 EU-landen en in hoeverre reeds sprake is van een zekere convergentie met betrekking tot deze verschillen. De Angelsaksische landen zijn toegevoegd omdat het hoogontwikkelde landen betreft met een verhoudingsgewijs geringe collectieve lastendruk. Dit heeft enerzijds gevolgen voor het ondernemingsklimaat, maar anderzijds uiteraard ook voor het niveau van collectieve voorzieningen.

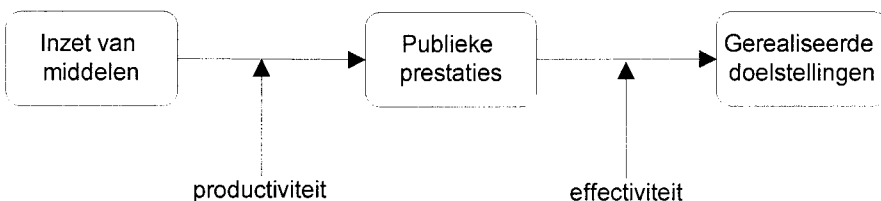
Landen zetten bepaalde middelen (personeel, materieel, kapitaal) in om publieke diensten te leveren. Deze activiteiten dragen mede bij aan het realiseren van gewenste doelen. Dit is nader toegelicht in Figuur 1.

De productiviteit heeft betrekking op de relatie tussen prestaties en middelen, de effectiviteit op de mate waarin de gestelde doelen worden gerealiseerd. Het SCP-rapport gaat enerzijds in op de meting van de productie dan wel het gebruik van diensten (aantallen leerlingen, aantal behaalde diploma's, aantal behandelde patiënten, aantal veroordeelde misdadigers) en anderzijds op de daaraan gelieerde effecten (aangeleerde vaardigheden, opleidingsniveau van de bevolking, levensverwachting, gezondheidstoestand, criminaliteit).

Aan dergelijke analyses kleef een aantal problemen:

- 1) vaak is alleen de kwantiteit van de productie (het 'productievolume') direct meetbaar,
- 2) de kwaliteit moet meestal in kaart worden gebracht via een subjectieve meting: de oordelen van burgers of andere betrokkenen,
- 3) de bereikte effecten zijn veelal maar ten dele toe te schrijven aan de geproduceerde dienst. Zo wordt de gezondheidstoestand van de bevolking ook in sterke mate beïnvloed door leefstijl en algemene voorzieningen (drinkwater, sanitair en dergelijke).

In de praktijk richt onze analyse zich vooral op effecten en ingezette middelen. Daarnaast wordt ook aandacht gegeven aan productie en productiviteit en wordt een relatie



Figuur 1. Input en output van publieke dienstverlening.

gelegd met een subjectieve maat voor de kwaliteit van de dienstverlening: het vertrouwen van burgers.

Dit artikel gaat in het bijzonder in op de relatie tussen de prestaties van publieke voorzieningen en het vertrouwen van burgers in de betrokken sectoren. De relatie tussen deze grootheden is niet vanzelfsprekend. Zo is niet altijd duidelijk wat men bij een bepaalde voorziening voor ogen heeft, welke kenmerken de respondent in de oordeelsvorming betreft, welke verwachtingen hij koestert en welke ervaring hij met die voorziening heeft. Betere prestaties van de overheid hoeven daardoor niet meteen te leiden tot meer vertrouwen. Voorts kan het vertrouwen betrekking hebben op verschillende aspecten van een voorziening, zoals de inzet van personele en materiële middelen (proces), de geleverde producten (output) en de resulterende effecten (outcome). Vertrouwen kan ook betrekking hebben op de veronderstelde integriteit en competentie van de betreffende beroepsgroep. Bij gebrek aan specifieke kennis en ervaringen met afzonderlijke voorzieningen zal vertrouwen soms niet meer zijn dan een afgeleide van algemene houdingen. Ondanks deze kanttekeningen is vertrouwen de best beschikbare landelijke meting van tevredenheid.

## 2. ONDERWIJS

### 2.1. Inleiding

Overheidsbemoediging met de productie en consumptie van onderwijs is gerechtvaardigd door positieve externe effecten (omdat ook anderen dan ouders en leerlingen er bij zijn gebaat) en om gelijke toegang te bewerkstelligen, een zaak van sociale rechtvaardigheid. Wij onderscheiden vier onderwijssystemen afhankelijk van de differentiatie van het voortgezet onderwijs. Nederland hoort bij de landen met de sterkste differentiatie. De Scandinavische landen worden gekenmerkt door een uniforme 'grundskola' voor alle kinderen van 6-16 jaar en Angelsaksische landen als de VS door een uniforme junior en senior 'highschool'.

### 2.2. Deelname

Geheel volgens verwachting is de onderwijsdeelname onder leerlingen van de leerplichtige leeftijd nagenoeg 100%. Voor jongeren uit de leeftijdsgroep 15-19 jaar ligt de onderwijsdeelname tussen 72% (Italië) en 91% (België). Nederland, Frankrijk en Duitsland kennen eveneens hoge deelnamepercentages. Meer dan de helft van alle jongeren begint in Nederland aan een hogere opleiding (op universitair niveau of hbo-niveau). Daarmee bevindt Nederland zich hoog in een middengroep van landen. De lijst wordt aangevoerd door Australië, Finland, Nieuw-Zeeland, Polen en Zweden met een score van 65% of meer. Wanneer ook rekening wordt gehouden met kortere, beroepsmatig georiënteerde hogere opleidingen die in sommige andere landen een prominente plaats innemen, zakt Nederland af naar een plaats in de middenmoot.

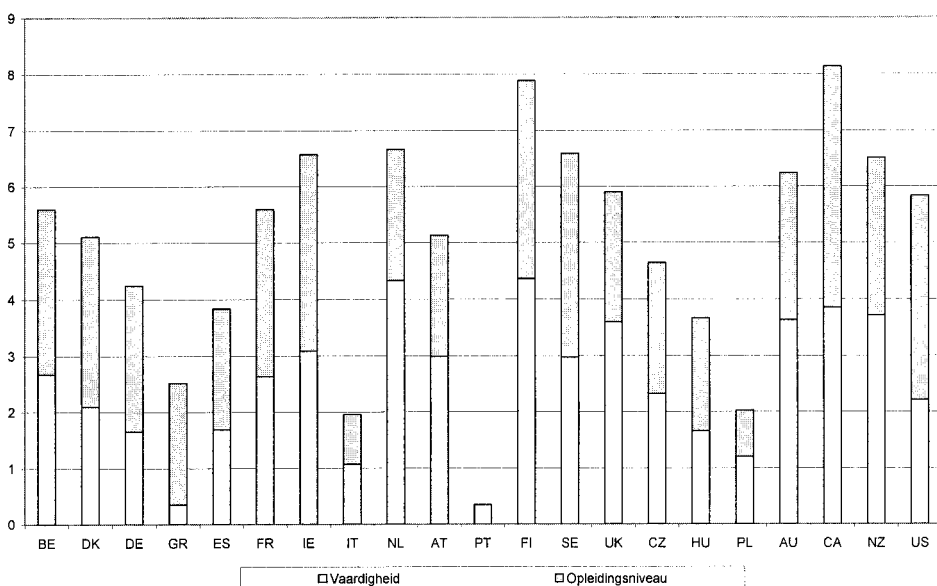
### 2.3. Effectiviteit

Een van de maten voor de effectiviteit van het onderwijs zijn internationale prestatie-

toetsen die leesvaardigheid, wiskundig inzicht en begrip voor de wetenschappelijke aanpak van 15-jarigen meten. Nederlandse leerlingen scoren hoog bij deze toetsen, net als Finse en Canadese leerlingen. De VS nemen een middenpositie in, terwijl Luxemburg, een aantal Mediterrane landen en de nieuwe lidstaten met lage testcores uit de bus komen.

Naast het niveau van de uiteindelijk verworven kennis en vaardigheden ('educational achievement') is ook het formeel bereikte opleidingsniveau ('educational attainment') een belangrijke maatstaf om de kwaliteit van een onderwijsstelsel te bepalen. Een indicator daarvoor is het deel van de bevolkingsgroep van 25-34 jaar dat tenminste een hogere voortgezette opleiding heeft voltooid. Het landengemiddelde bedraagt circa 62%. Nederland scoort op dit criterium slechts iets hoger (65%) en moet elf landen laten voorgaan. Portugal is hekkensluiter met een score van niet meer dan 20%. Nederland neemt eveneens een middenpositie in ten aanzien van een tweede indicator, het aantal afgestudeerden van hogere opleidingen.

De afzonderlijke indicatoren zijn omgerekend tot 'rapportcijfers' tussen 0 en 10. Vervolgens is een gemiddeld rapportcijfer voor effectiviteit berekend, dat rekening houdt met zowel indicatoren voor 'achievement' als voor 'attainment' (zie figuur 2). Wanneer landen op basis van deze maatstaf worden gerangschikt, komen Canada en Finland als lijststaanvoerders naar voren. Deze landen scoren hoog in termen van zowel 'achievement' als 'attainment'. Zij worden gevolgd door de andere Angelsaksische landen, Zweden en Nederland, dat een hoge score voor 'achievement' met een tamelijk middelmatige score voor 'attainment' combineert. Polen en de meeste Zuid-Europese landen presteren slecht.



Bron: OESO; SCP-bewerking.

Figuur 2. Effectiviteit van het onderwijs (rapportcijfer).

#### 2.4. *Gelijkheid van kansen*

Het bevorderen van gelijke onderwijskansen voor kinderen met een verschillende sociaal-economische achtergrond is een belangrijke doelstelling van de onderwijspolitiek in Nederland. Dit zal in andere landen vermoedelijk niet anders zijn. Idealiter vallen er weinig kinderen uit de boot (hoge 'achievement' score) en is er weinig verschil in prestaties van kinderen uit gezinnen met een hoge en lage sociaal-economische status. Scores op tests die leesvaardigheid meten, wekken de indruk dat beide oogmerken elkaar niet uitsluiten. Er hoeft dus geen onoplosbaar conflict te bestaan tussen gelijkheid van kansen en kwaliteit van het onderwijs. Een sterk punt van goed presterende onderwijsstelsels is dat zij ook beter zijn in het verminderen van verschillen die verband houden met de uiteenlopende sociale achtergrond van leerlingen.

Bestaat er verband tussen de kenmerken van onderwijsstelsels, de gemiddelde slaagkans ('achievement') en de gelijkheid van onderwijskansen? Over het geheel genomen, tekent zich een patroon af. De meeste landen met een gedifferentieerd onderwijsstelsel doen het wat minder goed bij het verzekeren van gelijke toegang tot het onderwijs en boeken middelmatige tot zwakke scores als het gaat om de resultaten van hun onderwijsstelsel. Nederland lijkt eerder een uitzondering die de regel bevestigt. Landen met een meer uniform stelsel (van het Scandinavische of Angelsaksische type) scoren in het algemeen heel behoorlijk in termen van gelijke kansen en resultaten, terwijl landen uit de tussengroep in beide opzichten een grote spreiding rondom het gemiddelde kennen. Hoewel nader onderzoek zeker nodig is, voordat definitieve conclusies kunnen worden getrokken, wijzen onze voorlopige resultaten er op dat brede, weinig gedifferentieerde onderwijsstelsels in het algemeen beter presteren.

#### 2.5. *Uitgaven*

Gemiddeld besteden landen 5,5% van hun BBP aan onderwijs. De lijst wordt aangevoerd door Denemarken, Zweden en Cyprus, waar de onderwijsuitgaven ongeveer 8% BBP beslaan. In de staart vinden we Griekenland, Luxemburg en Slowakije, met een onderwijsuitgavenquote van circa 4%. Nederland neemt met 5,3% een middenpositie in en is recent een aantal plaatsen op de ranglijst gestegen.

De effectiviteit van onderwijsstelsels kan ook worden afgezet tegen de uitgaven voor onderwijs. Nederland weet een redelijk evenwicht te vinden tussen uitgaven en prestaties. Nieuw-Zeeland levert echter vergelijkbare prestaties tegen geringere kosten en Finland en Canada leveren betere prestaties tegen hogere kosten. Landen als de VS en Zweden boeken redelijke resultaten, ten koste van in verhouding hoge uitgaven. Veel Zuid- en Centraal-Europese landen kennen een betrekkelijk laag uitgavniveau, maar boeken ook tamelijk zwakke resultaten. Per saldo lijkt er nauwelijks een verband te bestaan tussen prestaties en uitgaven.

#### 2.6. *Lissabon Agenda*

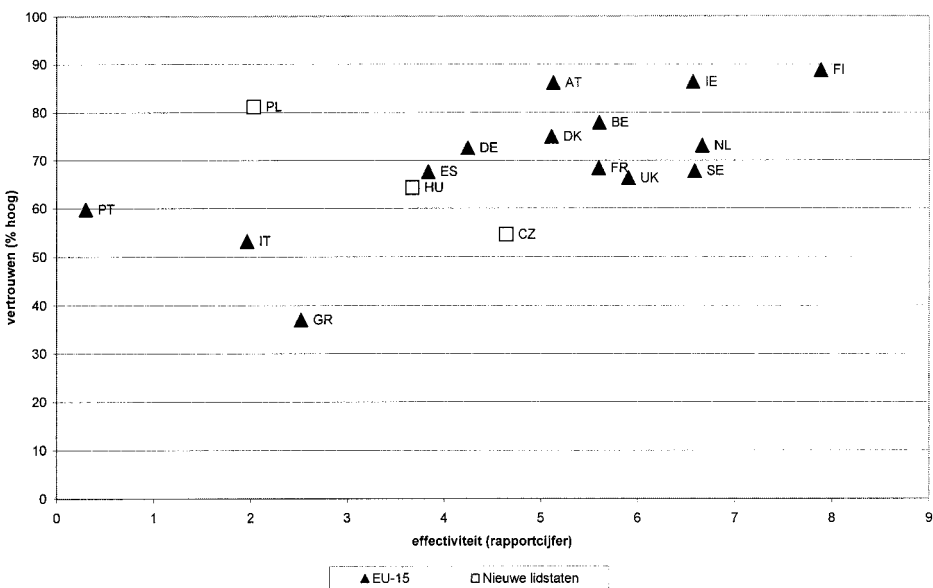
In de Lissabon Agenda, waarin de toenmalige EU-lidstaten zich de uitdagende taak hebben gesteld om de EU binnen tien jaar om te vormen tot het meest concurrerende economische gebied ter wereld, is vastgelegd dat de schooluitval in 2010 moet zijn teruggebracht tot 10%. In dat zelfde jaar zou 80% van alle 25-64-jarigen ten minste hoger voortgezet onderwijs moeten hebben genoten. Het is twijfelachtig, of deze beide

doelstellingen in Nederland kunnen worden gehaald tenzij door verlaging van het eisenpakket of door aanpassingen in het onderwijsstelsel. De vroege en radicale differentiatie van leerwegen, draagt mogelijk bij aan de stigmatisering van leerlingen. Specifieke kenmerken van het Nederlands lager voortgezet onderwijs kunnen voorts bijdragen aan gebroken schoolloopbanen en de arbeidsmarktpositie van voortijdige schoolverlaters verzwakken. Zou het lager voortgezet onderwijs bijvoorbeeld slechts drie jaar duren (in plaats van vier jaar) of – zoals in sommige landen – beginnen op de leeftijd van 11 jaar, dan zouden leerlingen meer mogelijkheden hebben om eerder over te stappen naar een voortgezette beroepsopleiding, en zouden zij sterker geprikkeld worden om hun opleiding af te maken.

### 2.7. Vertrouwen

Figuur 3 gaat in op een subjectieve kwaliteitsmaat, het vertrouwen dat burgers hebben in het onderwijssysteem en brengt dit in verband met de eerder besproken prestaties. Hierbij ontbreken gegevens voor de niet-Europese landen. Het percentage in de figuur heeft betrekking op de som van de rubrieken ‘tamelijk hoog’ en ‘zeer hoog’. Het vertrouwen is hoog in Finland, Malta, Ierland, Oostenrijk, Polen en Slovenië. Een lage score is kenmerkend voor Griekenland, Italië, Tsjechië en Portugal. Nederland behoort in dit opzicht tot de middenmoot.

Het zal geen verbazing wekken dat prestaties en vertrouwen een positieve correlatie vertonen ( $r = 0,57$ ), met Finland en Ierland (en in iets mindere mate Nederland) aan de positieve kant van het spectrum en een groep Zuid-Europese landen aan de andere kant. De positie van Tsjechië en Polen is zeer opvallend: het Tsjechische onderwijssysteem



Bron: OESO, European Values Survey; SCP-bewerking.

Figuur 3. Relatie tussen effectiviteit en vertrouwen in het onderwijs.

is redelijk effectief, maar het vertrouwen daarin van de burgers zeer gering, en voor Polen geldt juist het omgekeerde.

Er lijkt nauwelijks een relatie te bestaan tussen stelselkenmerken en vertrouwen: De groep landen met een gedifferentieerd onderwijssysteem omvat zowel landen met een hoog (Oostenrijk) als met een gering vertrouwen (Tsjechië). Voor de tussengroep (Ierland versus Italië) en de groep landen met een uniform onderwijssysteem (Finland versus Portugal) geldt hetzelfde. Ook lijkt er nauwelijks een verband te bestaan tussen het vertrouwen en de onderwijsuitgaven.

### 3. GEZONDHEIDSZORG

#### 3.1. Inleiding

Onbelemmerde toegang tot zorgvoorzieningen voor wie zorg nodig hebben, hoge kwaliteit van zorgvoorzieningen en financiële houdbaarheid van het zorgstelsel zijn drie kerndoelstellingen van het overheidsbeleid inzake de gezondheidszorg in nagenoeg alle landen. De beheersing van de uitgaven voor gezondheidszorg is één van de grootste uitdagingen waarmee overheden op dit moment worden geconfronteerd. De zorguitgaven lopen op, om een aantal redenen. Een belangrijke oorzaak is dat de kosten van zorgproducten veel sneller stijgen dan de kosten van andere goederen en diensten. Velen verwachten dat ook de vergrijzing van de bevolking de zorguitgaven in de toekomst verder zal opstuwten.

De meeste landen proberen de financiële houdbaarheid van hun zorgstelsel te handhaven door in te grijpen in het zorgaanbod en door in sterkere mate op marktkrachten te vertrouwen. Tot maatregelen om het aanbod aan te passen horen: (a) verruiming van het aanbod buiten de muren van dure instellingen, (b) aanscherping van de budgetten van zorgaanbieders, en (c) beperking van het pakket van verplicht verzekerde zorgvoorzieningen. De werking van prijssignalen kan worden versterkt door invoering of verhoging van eigen betalingen van zorggebruikers om zo de zorgvraag in te tomen. Verder hopen beleidsmakers de bevolking over te halen gezonder te gaan leven. Bestrijding van ernstige vormen van milieuverontreiniging kan eveneens aan de volksgezondheid ten goede komen.

#### 3.2. Stelselkenmerken

Hoewel de doelstellingen van beleidsmakers in alle landen ruwweg identiek zijn, hebben landen gekozen voor heel verschillend vormgegeven stelsels om die doelen te bereiken. Dit geldt zowel voor het aanbod als voor de financiering van de gezondheidszorg. In sommige landen ligt het zorgaanbod in private handen: dokters zijn gevestigd als zelfstandig ondernemer, de ziekenhuizen worden particulier geëxploiteerd. In andere landen staat de grote meerderheid van de artsen op de loonlijst van een staatsziekenhuis. Wat de financiering betreft kan een onderscheid worden gemaakt tussen stelsels van het Beveridge-type (belastinggefinancierd) en Bismarck-type (premiegefinancierd). Het eerste type komt voor in de Angelsaksische, de Scandinavische en de Zuid-Europese landen, het tweede in West- en Centraal-Europa. Daarbij ligt de nadruk óf op particuliere betalingen (premies voor de particuliere ziektekostenverzekering daarbij inbegrepen), óf op heffingen van de overheid (belastingen, premies voor een verplicht

gestelde sociale ziektekostenverzekering). In feite kennen de meeste landen een gemengd systeem. Aan beide uiteinden van het spectrum vinden we de VS, met een hoofdzakelijk privaat stelsel, en Zweden, met een nagenoeg volledig publiek stelsel. Nederland is een middenmoter; de particuliere uitgaven hebben hoofdzakelijk betrekking op premies van particuliere verzekeraars en niet op eigen bijdragen. In Portugal, Finland, Spanje, Griekenland, Polen en Nieuw-Zeeland daarentegen, hebben private betalingen hoofdzakelijk de vorm van directe betalingen van zorggebruikers.

### *3.3. Gebruik van zorg*

Intramurale zorg wordt verleend in algemene ziekenhuizen, psychiatrische ziekenhuizen en verpleeghuizen. Oostenrijk springt er uit met een relatief hoog aantal ziekenhuisopnamen: 270 per duizend inwoners. Ook andere landen uit de groep corporatistische landen (in het bijzonder Frankrijk en Duitsland) kennen een hoog aantal opnamen per duizend inwoners, evenals Finland en de Oost-Europese landen. Daarentegen is het aantal opnamen betrekkelijk laag (ongeveer 100 per duizend inwoners) in Canada, Nederland, Portugal, Spanje en de VS. Het aantal bezoeken aan de huisarts en aan de medisch specialist loopt tussen landen sterk uiteen. In twee landen – Hongarijë en Tsjechië – is het doktersbezoek extreem hoog, met gemiddeld 12 artsbezoeken per persoon per jaar. In de meeste landen bezoeken mensen gemiddeld tussen 4 en 8 keer per jaar een dokter. Nederland (met 6 artsbezoeken per persoon per jaar) zit iets beneden het gemiddelde van de EU-15.

### *3.4. Wachtlijsten*

Op basis van de beschikbare gegevens kunnen wij alleen cijfers presenteren over wachttijden voor niet-spoedeisende zorg (niet-urgente ingrepen). Negen van de EU-15 landen rapporteren het bestaan van aanzienlijke tot lange wachttijden voor niet-urgente ingrepen. Vijf EU-landen (België, Oostenrijk, Frankrijk, Duitsland en Luxemburg) geven aan geen wachttijden van betekenis te kennen. Nederland neemt in dit opzicht een middenpositie in.

In het algemeen gesproken neemt de kans op het bestaan van wachtlijsten af naarmate meer capaciteit beschikbaar is, ruimere financiële middelen voor handen zijn en hogere eigen bijdragen van zorggebruikers worden gevraagd. Wachtlijsten komen dus hoofdzakelijk voor in zorgstelsels met een beperkte rol voor prijssignalen. Sinds er een duidelijk verband bestaat tussen (grotere) capaciteit en (kortere) wachtlijsten, is het verleidelijk om te concluderen dat extra financiële middelen wachtlijsten korter maken. Dit hoeft echter niet noodzakelijkerwijs het geval te zijn. In de praktijk kan een ongeclausuleerde groei van het aanbod leiden tot extra vraag. Beleidsmakers kunnen daarom overwegen gerichte financiële prikkels in het leven te roepen, om te voorkomen dat extra aanbod zijn eigen vraag schept. Verbetering van het management van ziekenhuizen, zoals een efficiëntere planning en routing van patiënten, en het tegengaan van ‘no shows’ kan ook helpen bij het verkorten van wachttijden.

### *3.5. Effectiviteit*

Het voornaamste doel van het zorgstelsel is uiteraard de gezondheidstoestand van de bevolking te verbeteren. Een goede gezondheid – te beschouwen als een objectieve



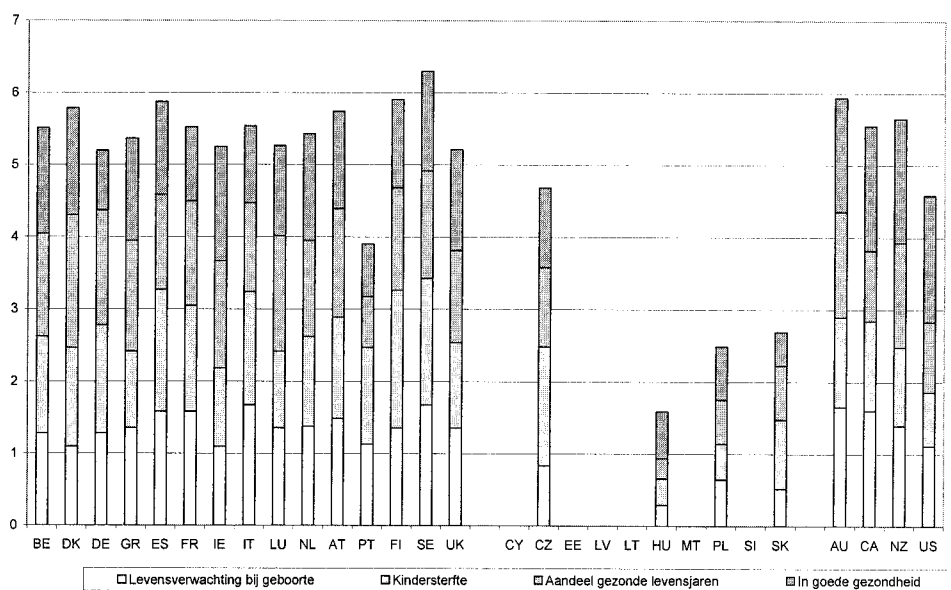
kwantiteitsmaatstaf – komt tot uiting in een gemiddeld lange levensverwachting, lage kindersterfte, een groot aantal in gezondheid doorgebrachte levensjaren en een algemeen gevoel van in gezondheid te leven. Deze vier indicatoren vormen samen een graadmeter voor de gezondheidstoestand van de bevolking die kan worden gebruikt om de effectiviteit van nationale zorgstelsels te meten. Figuur 4 brengt de samengestelde score voor effectiviteit in beeld.

De waarde van deze graadmeter ligt tussen 1,6 in Hongarije en 6,3 in Zweden. De meeste landen scoren tussen een 5 en een 6. Portugal, de VS en de nieuw toegetreden EU-lidstaten halen echter geen 5.

### 3.6. Uitgaven

In het jaar 2000 gaven de EU-15 landen meer dan 8% van hun BBP uit aan gezondheidszorg. In de nieuw toegetreden landen was dat 5% tot 7% van het BBP. Met 13% van het BBP hebben de VS veruit de hoogste uitgavenquote. Deze extreem hoge uitgaven zijn deels het gevolg van de hoge inkomens die in de VS in de zorgsector worden verdiend en het juridisch claimgedrag van patiënten bij vermeende onjuiste behandelingen.

De grote verschillen in zorguitgaven per hoofd van de bevolking hangen nauw samen met het BBP per hoofd. Hoe hoger het inkomen per hoofd, hoe groter de vraag naar gezondheidszorg en hoe hoger de kostprijs van de zorg. Er lijkt geen verband te bestaan tussen type zorgstelsel en hoogte van de zorguitgaven als aandeel van het BBP. Verder valt op dat verschillen in de zorguitgavenquote nauwelijks verband houden met de mate waarin de bevolking is vergrijsd.



Bron: OESO, World Health Organisation; SCP-bewerking.

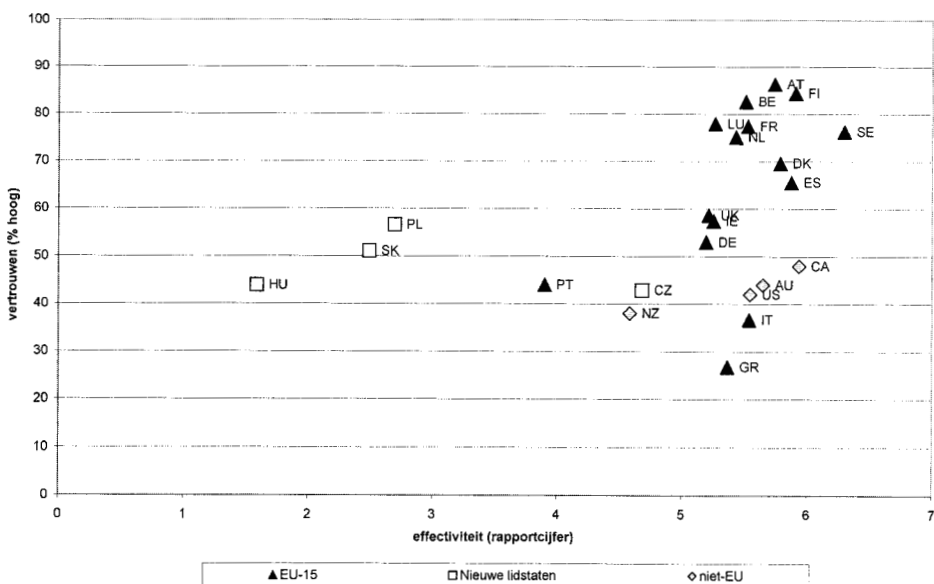
Figuur 4. Samengestelde index voor de gezondheidstoestand, 2001 (rapportcijfer).

Er bestaat evenmin een direct verband tussen de zorguitgaven in een land (als aandeel van het BBP) en de kwaliteit/effectiviteit van zijn zorgstelsel. De VS combineren een extreem hoge uitgavenquote met een beperkte effectiviteit. Daartegenover is het stelsel in Tsjechië redelijk effectief, bij een in verhouding beperkte inzet van middelen. De lage uitgavenquote in sommige nieuw toegetreden landen wordt mede verklaard, doordat de salarissen in de gezondheidszorg daar laag zijn, wat tot onvrede onder de beroepsgroep leidt. Tamelijk verrassend is dat de samenhang tussen de kwaliteit van het stelsel en de uitgavenquote nauwelijks wordt beïnvloed door verschillen in de leeftijdsopbouw en (on)gezonde levensstijl van de bevolking.

### 3.7. Vertrouwen

Het vertrouwen van de bevolking in het zorgstelsel is weergegeven langs de verticale as van figuur 5. Hier zijn Malta, Oostenrijk en Finland koploper. De rij wordt gesloten door enkele mediterrane landen als Griekenland en Italië. Ook de zorgstelsels in de nieuwe Oost-Europese lidstaten genieten doorgaans geen groot vertrouwen van de bevolking. Nederland behoorde in 2000 tot de subtop, met de kanttekening dat het vertrouwen in het zorgstelsel daar de afgelopen jaren – mede door een publieke discussie over wachtlijsten in de zorg – een geduchte knauw heeft gekregen: in 2002 vond ruim 30% van de volwassen Nederlandse bevolking dat de kwaliteit van de medische zorg in Nederland er de afgelopen vijf jaar op achteruit is gegaan, terwijl nog geen 20% vond dat de kwaliteit is toegenomen (Sociaal en Cultureel Rapport 2002, p. 294).

Uit figuur 5 blijkt dat voorts dat er een zwakke, maar significante correlatie ( $r = 0,36$ ) bestaat tussen de effectiviteit van het zorgstelsel en het vertrouwen in het zorgstelsel.



Bron: European Values Survey, niet-Europese landen bijgeschat.

Figuur 5. Vertrouwen in het gezondheidszorgstelsel naar effectiviteit, 2000.

Er lijkt ook enig verband te zijn tussen de institutionele inrichting van het zorgsysteem en het vertrouwen dat de bevolking heeft in het zorgstelsel. Met name de West-Europese landen met een Bismarck-stelsel genieten, op Duitsland na, een hoog vertrouwen van de bevolking. Het vertrouwen van de bevolking in de landen met een Beveridge-stelsel en lage eigen bijdragen is redelijk, met Canada als negatieve uitschieter. De Centraal-Europese landen en de landen met een Beveridge-stelsel en hoge eigen bijdragen krijgen, op Finland en Zweden na, nauwelijks een voldoende van hun ingezetenen. De ontevredenheid over het publieke deel van de gezondheidszorg is in Griekenland, Italië en Portugal altijd al betrekkelijk groot geweest. Vooral de plattelandsgebieden komen er in deze landen betrekkelijk slecht af. In Italië geven met name de bewoners van de zuidelijke provincies blijk van ongenoegen over het zorgaanbod. In Griekenland en Portugal staat het publieke stelsel onder druk, hetgeen gevolgen heeft voor de kwaliteit van de zorg en de hoogte van de eigen bijdragen. In beide landen moet de particuliere sector het verlies aan kwaliteit van het publieke stelsel compenseren, hetgeen voor de lagere inkomensgroepen problematisch is. Deze blijven aangewezen op een publiek bestel met lagere kwaliteit: langere wachtlijsten en een lager niveau van medische faciliteiten en personele diensten. De landen met een sterk geprivatiseerd stelsel, de VS en Griekenland krijgen een ruime onvoldoende van de bevolking. Het vertrouwen in het zorgstelsel blijkt maar zwak samen te hangen met de zorguitgaven per hoofd van de bevolking.

## 4. POLITIE EN JUSTITIE

### 4.1. Inleiding

De ontwikkeling van de criminaliteit in een land kan specifieke, nationale oorzaken hebben. Ook kan sprake zijn van een internationale trend. Zo steeg bijvoorbeeld in de afgelopen tien jaar de geweldscriminaliteit in Nederland en de meeste omringende landen met ruwweg hetzelfde percentage. Dan ligt het voor de hand naar oorzaken te zoeken die niet specifiek Nederlands zijn.

### 4.2. Stelselkenmerken

Bij de classificatie van rechtssystemen kan men aansluiten bij het traditionele onderscheid tussen de Angelsaksische ‘common law’ traditie en de continentaal-Europese ‘civil law’ traditie. Overigens zijn ook tal van andere stelselkenmerken van belang alsmede de omvang van particuliere veiligheidsdiensten. Daarnaast is ook de repressiviteit van het strafrechtstelsel (zwaarte van straffen en inzet van personeel per inwoner) van betekenis. Straffen zijn zwaar en de personeelsinzet is hoog in met name Zuid- en Centraal-Europa en de VS. In de VS, bijvoorbeeld, worden meer dan 50 gevangenisdagen per geregistreerd misdrijf opgelegd tegen 5 in Nederland.

### 4.3. Geregistreerde misdrijven

Zweden, Nieuw Zeeland en het Verenigd Koninkrijk kennen het hoogste aantal geregistreerde misdrijven met meer dan 10.000 geregistreerde delicten per honderdduizend inwoners. Met minder dan 2.000 misdrijven doen Ierland, Cyprus en Slowakije het zeer

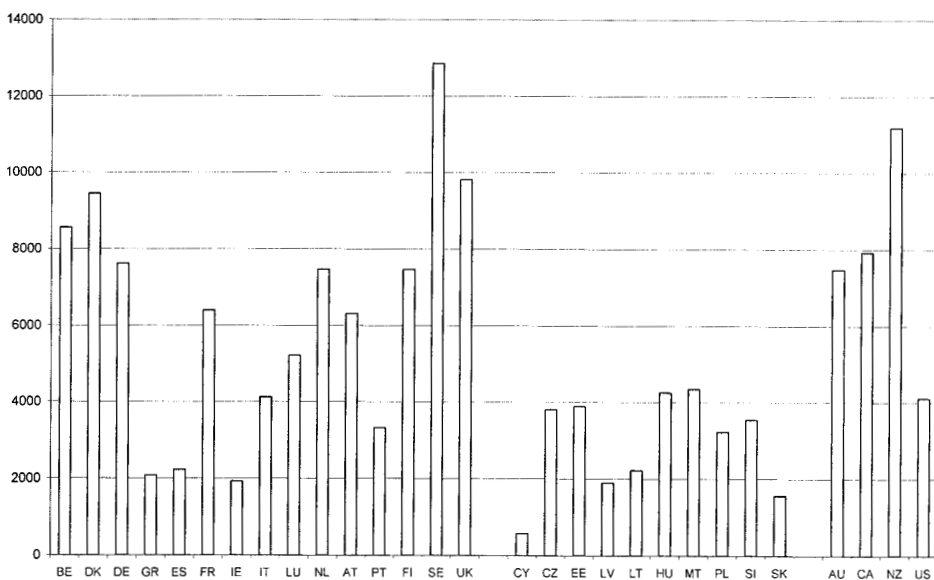
goed. Verhoudingsgewijs goed komen ook de Zuid-Europese landen en de nieuwkomers uit de verf. Dat geldt ook voor de VS, die toch in de volksmond een veel slechtere reputatie genieten. Nederland behoort met een score van 8.000 tot de subtop (zie figuur 6). De hoge positie van Nederland is vooral een gevolg van het grote aantal fietssendiefstallen.

Het meest dramatische delict is ongetwijfeld moord of doodslag. Merkwaardig is dat deze delicten relatief frequent voorkomen in een aantal landen die verder in het algemeen niet worden gekenmerkt door hoge criminaliteitscijfers. Dat geldt voor de Baltische staten, maar ook voor de VS.

Het aantal drugs misdrijven is in Nederland uitzonderlijk laag. Dit valt te verklaren door het Nederlandse gedoogbeleid, waarbij niet het bezit van drugs voor eigen gebruik, maar alleen de handel in verdovende middelen wordt vervolgd. Maar ook het problematisch drugsgebruik en het percentage jongeren dat ooit marihuana of cannabis heeft gebruikt, is in Nederland betrekkelijk laag. Daarentegen is Nederland onmiskenbaar een internationaal knooppunt voor drugshandel.

#### 4.4. Functioneren strafrechtelijke keten

Een integraal beeld voor het functioneren van de gehele strafrechtelijke keten ontstaat door het aantal veroordeelde personen uit te drukken als percentage van het aantal geregistreerde misdrijven. Op dat criterium scoort Griekenland met meer dan 40% het hoogst, gevolgd door een aantal nieuwe lidstaten met scores tussen de 15% en 25%. Van de EU-landen hebben Finland en Engeland/Wales het hoogste percentage (15% à 20%). Nederland is hekkensluiter met 7%. Echter, dit laatste cijfer zou aanzienlijk toe-



Bron: European Sourcebook 2003.

Figuur 6. Aantal geregistreerde misdrijven per 100.000 inwoners (exclusief verkeersmisdrijven), 2000.

nemen indien transacties van de openbare aanklager als veroordelingen worden meegeteld.

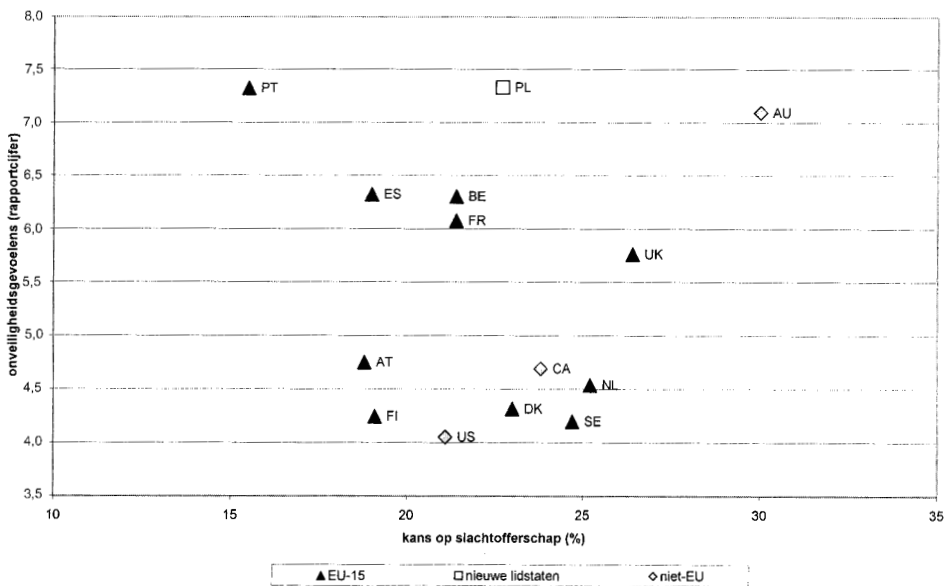
#### 4.5. Inzet van personeel

De voor politie en justitie ingezette personeelssterkte varieert tussen 270 (Finland) en 830 (Italië) per honderdduizend inwoners. Met 400 scoort Nederland iets onder het gemiddelde. Een betrekkelijk hoge inzet is kenmerkend voor de Centraal- en Zuid-Europese landen, een lage inzet in de Scandinavische landen, het Verenigd Koninkrijk, Australië en Canada. De laatstgenoemde drie landen en de VS worden echter gekenmerkt door een omvangrijke particuliere beveiligingssector. De arbeidsproductiviteit van de strafrechtelijke keten kan worden uitgedrukt als quotiënt van het aantal veroordeelde verdachten en het ingezette personeel. De zo gedefinieerde productiviteit is zeer hoog (5 tot 8 veroordeelden per fte) in Finland, Zweden, en Engeland/Wales. Bij Zweden past hier overigens de kanttekening dat bij samengestelde misdrijven meerdere veroordelingen worden geteld. De productiviteit is laag (minder dan 1 veroordeelde per personeelslid) in Ierland, de Zuid-Europese landen en bij een aantal nieuwe lidstaten. Nederland hoort met 1,7 veroordeelden per fte tot de middenmoot. Ook hier geldt de kanttekening dat dit aantal aanzienlijk toeneemt als rekening wordt gehouden met transacties van de openbare aanklager.

#### 4.6. Het perspectief van de burger

Slechts een betrekkelijk klein deel van de misdrijven (10% á 20%) wordt aangegeven bij en geregistreerd door de politie. Daarom geven bevolkingsenquêtes een zuiverder beeld van de criminaliteit dan politieregistraties. De kernvragen in dit type enquêtes hebben betrekking op de mate waarin de ondervraagde burgers zelf slachtoffer zijn geworden van criminaliteit. Opmerkelijk is dat Nederland uit deze enquêtes te voorschijn komt als het land met een hoge kans op slachtofferschap. Dit blijkt hoofdzakelijk het gevolg van het veelvuldig voorkomen van twee soorten van delicten, namelijk fietsendiefstal en vandalisme aan auto's. Bij specifieke andere delicten is het beeld duidelijk anders.

Figuur 7 gaat in op de relatie tussen het niveau van criminaliteit volgens de slachtofferenquêtes en de veiligheidsgevoelens van de bevolking. Deze laatste zijn gebaseerd op de opgetelde gestandaardiseerde scores van drie vragen in de slachtofferenquêtes: geschatte kans op inbraak en gevoelens van onveiligheid in huis en op straat. Het percentage burgers dat de kans op inbraak als waarschijnlijk tot zeer waarschijnlijk inschat varieert bijvoorbeeld tussen 58% in Portugal en 13% in Oostenrijk (in Nederland 18%). Onveiligheidsgevoelens bij het alleen wandelen in het donker variëren tussen 35% in Australië, Polen en Spanje en 14% in Zweden en de VS (wederom 18% in Nederland). De figuur illustreert dat er geen verband blijkt te bestaan tussen gevoelens van onveiligheid en de feitelijke kans op slachtofferschap ( $r = -0,03$ ). Zowel in Portugal, Polen als Australië voelen mensen zich onveilig, terwijl de kans op het ondervinden van een misdrijf in deze landen respectievelijk zeer laag, gemiddeld en hoog is. Nederland is een land waar de burgers zich betrekkelijk veilig voelen, maar waar ze toch een verhoudingsgewijs hoge kans hebben om een misdrijf te ondervinden.



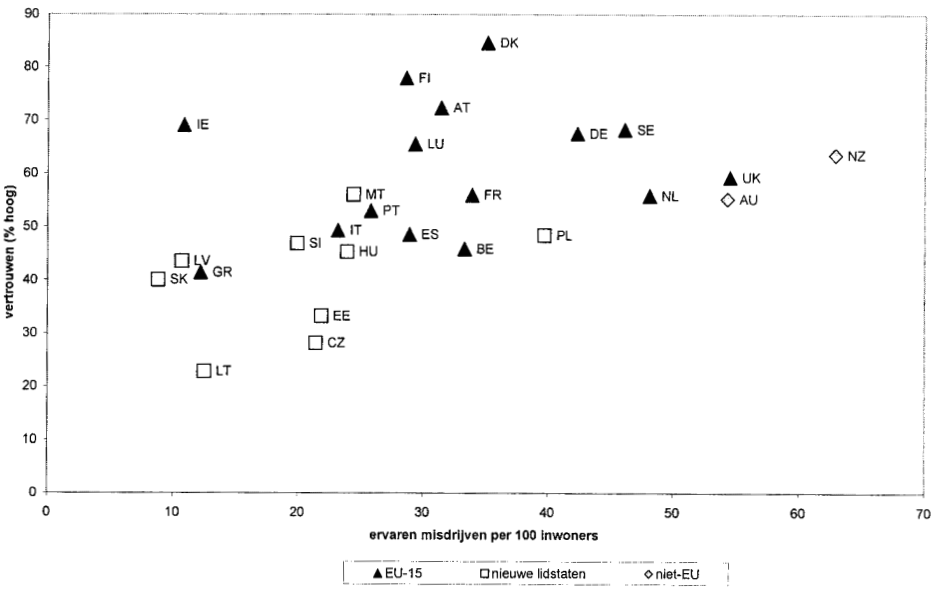
Bron: International Crime Victimization Survey. SCP-bewerking.

Figuur 7. Angst voor criminaliteit en kans op slachtofferschap, 2000.

#### 4.7. Vertrouwen

Figuur 8 geeft (langs de verticale as) informatie over het vertrouwen in, respectievelijk, politie en justitie. In Denemarken heeft 92% van de bevolking vertrouwen in de politie, tegen 28% in Litouwen. De meeste landen die behoren tot de EU-15 scoren evenals de niet-Europese Angelsaksische landen hoog tot gemiddeld. Nederland en België nemen een middenpositie in met respectievelijk 66% en 54%. Griekenland doet het met 38% beduidend slechter. Ook de nieuwkomers scoren laag, met Malta als positieve uitzondering. Het beeld voor justitie is vergelijkbaar, met dien verstande dat het vertrouwen in justitie over het algemeen wat lager is dan dat in de politie. Hier voert Denemarken de lijst aan met 79% en is Litouwen hekkensluiter met 19%. Nederland doet het met 50% wat beter dan gemiddeld Griekenland, Letland en Hongarije schuiven naar het midden op. Nieuw-Zeeland doet het matig en Australië slecht. Bij België, Australië en Italië is het oordeel over justitie (met respectievelijk 39%, 35% en 31% vertrouwen) beduidend slechter dan dat over de politie.

De criminaliteitscijfers in figuur 8 hebben geen betrekking op de geregistreerde criminaliteit, maar op de volgens slachtofferenquêtes door burgers ervaren criminaliteit. Voor landen waarvoor geen gegevens uit slachtofferenquêtes beschikbaar zijn is een proportioneel aangepast cijfer voor de geregistreerde criminaliteit gebruikt. In tegenstelling tot de verwachtingen lijkt er een zwak positief verband te bestaan tussen criminaliteit en vertrouwen in politie en justitie (figuur 8). De Noord- en West-Europese landen, alsmede Australië en Nieuw-Zeeland worden gekenmerkt door een gemiddelde tot hoge criminaliteit, die gepaard gaat met een eveneens groot tot middelmatig vertrouwen in politie en justitie. De Zuid-Europese en vooral de nieuwe lidstaten paren daarentegen een laag vertrouwen in politie en justitie aan een laag niveau van criminaliteit.



Bron: European Sourcebook, European Values Survey. SCP-bewerking.

Figuur 8. Politie en justitie: criminaliteit en vertrouwen.

Deze verrassende uitkomsten worden in tabel 1 in verband gebracht met enkele andere bevindingen.

Tabel 1. Enkele dwarsverbanden bij politie en justitie (geschematiseerd).

	crimi- nali- liteit	repressi- viteit (straf- maat)	repressi- viteit (omvang perso- neel)	bestraffings- kans	produc- tiviteit	vertrou- wen
Noord- en West-Europa						
Australië, Nieuw-Zeeland	hoog	laag	laag	laag m.u.v. Finland en Engeland	hoog	hoog
Zuid- en Centraal-Europa	laag	hoog	hoog	hoog m.u.v. Portugal en Spanje	laag	laag

Bron: SCP.

De genoemde landengroepen onderscheiden zich ook met betrekking tot repressiviteit (strafmaat en inzet van middelen), bestraffingskans en productiviteit. De lage productiviteit in landen met een geringe criminaliteit heeft te maken met de hoge inzet van personeel, zowel per hoofd van de bevolking als ten opzichte van de omvang van de crimi-

naliteit. Het lage vertrouwen in de politie en justitie in die landen heeft wellicht te maken met een ander aspect dat aan repressiviteit kan kleven: de sterke gerichtheid op het aanpakken van criminaliteit en het opsporen van daders zou er toe kunnen leiden dat politie en justitie het minder nauw nemen met de rechten van de verdachte en met de gegedeheid van de bewijslast. Verder kan een samenhang bestaan tussen de lage salariering van het overheids personeel en de matige moraal met betrekking tot corruptie in een aantal van die landen. Een alternatieve verklaring kan zijn dat een gebrek aan vertrouwen in politie en justitie burgers beweegt om geen aangifte te doen. En inderdaad is de discrepantie tussen het aantal in slachtofferenquêtes gerapporteerde misdrijven en het aantal geregistreerde misdrijven verhoudingsgewijs hoog in landen als Polen, Portugal en Spanje. Dit registratieprobleem lijkt echter maar een deel van het verschil te kunnen verklaren. Alleen op basis van nadere enquêtes, met name in die landen waar momenteel geen slachtofferenquête wordt gehouden, kan dit onderwerp verder worden uitgediept.

## 5. TOT BESLUIT

### 5.1. Totale prestaties

Om de totale prestaties op de genoemde terreinen te meten zijn de prestatiescores voor onderwijs (paragraaf 2), gezondheid (paragraaf 3) en criminaliteitsbestrijding (paragraaf 4) tot één index samengevoegd. Deze index, weergegeven in figuur 9, geeft de allocatiefunctie van de overheid weer.

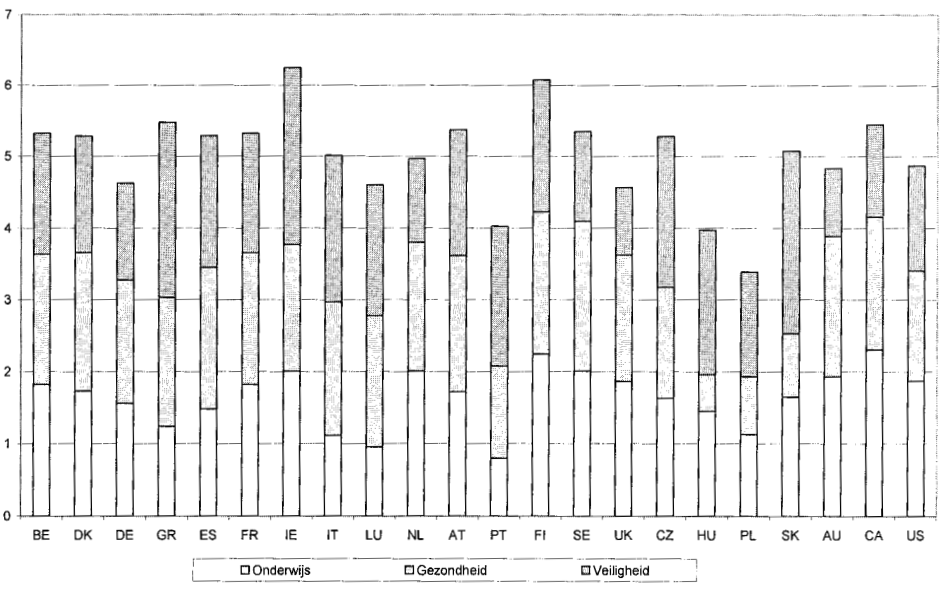
Opmerkelijk is dat veel landen een score hebben die weinig afwijkt van het gemiddelde. Ierland en Finland komen duidelijk boven dit gemiddelde uit, gevolgd door een groep EU-15 landen, Tsjechië, Slowakije en Canada. Nederland is te rekenen tot de middenmoot. Polen, Hongarije en Portugal komen duidelijk onder het gemiddelde uit. Geen enkel land scoort uitstekend op alle drie sectoren. Ierland heeft de topositie verworven door een uitstekende score op criminaliteit, een goede score op onderwijs en een gemiddelde score op gezondheid. Finland heeft een zeer goede score op onderwijs en gezondheid, maar moet de topositie aan Ierland overlaten door een matige score op veiligheid.

### 5.2. Ingezette middelen

Uitgedrukt als percentage van het BBP bedragen de opgetelde private en particuliere uitgaven voor onderwijs, gezondheidszorg en politie en justitie tussen de 12% en 22% van het BBP. Het laagst zijn ze in Tsjechië, Slowakije, Hongarije, Polen en Ierland. De subtop wordt gevormd door Zweden, Denemarken en Duitsland met 17% à 18%. Eenzaam aan de top staat de VS met 22%. Nederland is, zoals gewoonlijk, een middenmoter met 15%.

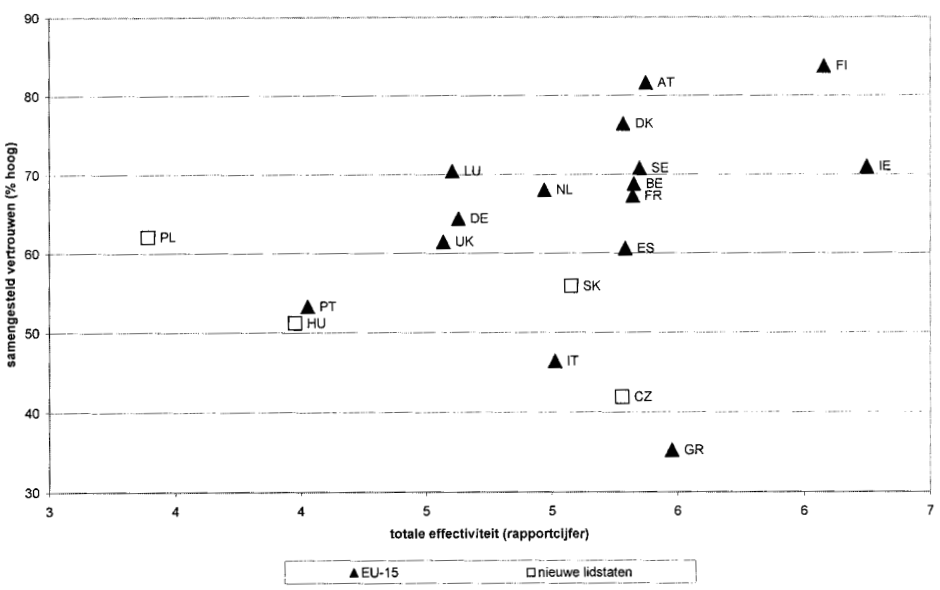
Het meeste waar voor hun geld (dwz. de hoogste kosteneffectiviteit) krijgen Ierland en Finland. Hongarije en Polen hebben lage prestaties en lage kosten; de VS hoge kosten en matige prestaties. Nederland bevindt zich in een omvangrijke middengroep.





Bron: SCP.

Figuur 9. Totale prestatiescore voor onderwijs, gezondheidszorg en veiligheid (rapportcijfer).



Bron: European Values Survey, SCP-bewerking.

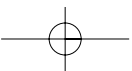
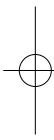
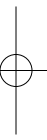
Figuur 10. Totale effectiviteit en samengesteld vertrouwen in onderwijs, gezondheidszorg en politieljustitie.

### 5.3. Vertrouwen

Bij onderwijs is de correlatie tussen effectiviteit en vertrouwen tamelijk hoog, bij zorg is deze zwak en bij politie/justitie negatief. De geaggregeerde uitkomsten voor deze sectoren (zie figuur 10) laten dan ook slechts een zwak positief verband zien ( $r = 0,30$ ). Rechtsboven treffen we de Scandinavische landen, Ierland en Oostenrijk met hoge prestaties en hoog vertrouwen. Links onder landen als Portugal en Hongarije met een lage score voor beide items. Nederland is een middenmoter. Polen valt uit de toon door lage prestaties, gecombineerd met een hoog vertrouwen. Voor Tsjechië, Italië en Griekenland geldt het omgekeerde: redelijke prestaties, gecombineerd met een laag vertrouwen.

#### NOTEN

1. Het volledige rapport kan worden aangetroffen op de website [www.scp.nl](http://www.scp.nl). Zie dit rapport voor literatuurverwijzingen.
2. In de figuren zijn de landen afgekort tot twee letters conform internationale afspraken. Enkele minder bekende afkortingen zijn AT (Oostenrijk), ES (Spanje), Estland (EE), Letland (LV), Litouwen (LT), Slowakije (SK) en Slovenië (SI).



## *13. Beter een goed verhaal dan de hele waarheid\**

H. HOIJTINK

### SAMENVATTING

In dit essay zal worden betoogd dat verhalen, inzichten, ideeën en theorieën van onderzoekers expliciet moeten worden verdisconteerd bij de analyse van empirische data, om op die manier de kans op het achterhalen van de waarheid te vergroten. Door middel van voorbeelden zal worden geïllustreerd dat deze verhalen en theorieën als een soort mal over de data kunnen worden gelegd door ongelijkheidsrestricties op de parameters van statistische modellen te leggen. Met behulp van Bayesiaanse model selectie kan worden bepaald welke mal het beste past. Het belangrijkste resultaat is de posterior kans (dat wil zeggen de kans na observatie van de data) dat een mal (en dus het bijbehorende verhaal of theorie) de beste is van de malen (competitieve verhalen en theorieën) die onderzocht worden.

### 1. INLEIDING

De meeste wetenschappers zijn geboeid door hun vakgebied. Naast het bestuderen en toepassen van de aanwezige kennis, zoeken ze naar nieuwe inzichten en ontwikkelen ze nieuwe theorieën. Met theorieën bedoel ik ideeën die niet empirisch geverifieerd zijn, ideeën die nog aan de werkelijkheid getoetst moeten worden. Vaak worden nieuwe ideeën besproken met vakgenoten. Soms blijkt dat de gesprekspartner er andere ideeën op nahoudt. In dat geval is er sprake van competitieve ideeën dan wel competitieve theorieën. Soms is er overeenstemming tussen de gesprekspartners en vindt er alleen een aanscherping van de ideeën plaats. In dat geval is er sprake van competitie met jezelf. Het klopt niet wat ik denk, ook wel de niets of nul-hypothese genoemd, versus, het klopt wel wat ik denk ook wel de niet-niets of alternatieve hypothese genoemd. Door het gebruik van de begrippen niets en niet-niets hypothese wil ik uitdrukken dat de traditionele nul en alternatieve hypothese mijns inziens van beperkte waarde zijn bij het zoeken naar nieuwe inzichten en het ontwikkelen van nieuwe theorieën (zie ook, Berger en Sellke (1987) en Cohen (1994)). Hierop zal in het navolgende verder worden ingegaan.

\* Bewerking van de rede uitgesproken bij het aanvaarden van het ambt van hoogleraar ‘toegepaste Bayesiaanse statistiek’ aan de faculteit Sociale Wetenschappen van de Universiteit Utrecht op vrijdag 16 januari 2004.

Ideeën en theorieën zou ik goede verhalen willen noemen, waarmee een deel van de titel van dit essay van een context is voorzien. Goed, want geformuleerd door experts op hun vakgebied. Verhalen, omdat de empirische verificatie nog ontbreekt. De Bayesiaanse terminologie voor goede verhalen is voorkennis, en omdat u dat waarschijnlijk bekender in de oren klinkt, geef ik u ook de vertaling in het Engels: *prior knowledge* (Congdon, 2001:30-34; Howson, 2002; Sober, 2002).

Onderzoekers laten het meestal niet bij theoretiseren. De volgende stap is het toetsen van de theorieën aan de werkelijkheid. Om dit te kunnen doen, moet de werkelijkheid eerst gekwantificeerd worden. Dit gebeurt door middel van het verzamelen van empirische gegevens, dan wel het zoeken van bestaande databestanden. Impliciet wordt vaak aangenomen dat deze gegevens de hele werkelijkheid, of, om het in termen van de titel van dit essay te formuleren, de hele waarheid representeren. Zoals u uit de titel op kunt maken, heb ik enige bedenkingen bij deze hele waarheid. Ik kan me namelijk moeilijk voorstellen dat de hele waarheid te vangen is in een hoeveelheid getallen, of, om het netter te formuleren de data matrix. Het lijkt mij veel redelijker een concrete datamatrix te zien als een venster of een raam dat een blik geeft op de hele waarheid. En, als er vitrage voor het venster hangt of als het raam beslagen is, dan kan de blik op de waarheid er vertekend uitzien. Hierbij valt te denken aan ontbrekende waarnemingen, onbetrouwbaar gemeten variabelen, en het gebruik van niet geheel vergelijkbare groepen in een pseudo-experimenteel design. Merk op, dat er voor al deze problemen moderne statistische oplossingen zijn. Ontbrekende waarnemingen kunnen multiple geïmputeerd worden (Rubin, 1987; Schafer, 1997), latente regressie modellen en *errors in variables* modellen maken onbetrouwbare variabelen hanteerbaar (Muthen and Muthen, 1998; Glöckner-Rist and Hoijtink, 2003), en *propensity scores* zijn een prachtig hulpmiddel om onvergelijkbare groepen vergelijkbaar te maken (Rosenbaum and Rubin, 1983; Holland, 1986). Dit verandert echter niets aan het feit dat een datamatrix slechts een indirecte kijk op de waarheid biedt.

Nadat de theorieën geformuleerd zijn, en de empirische gegevens verzameld zijn, volgt als derde stap de analyse van deze gegevens. Dit gebeurt vaak op basis van een statistisch model, of, anders geformuleerd, een mathematische beschrijving van de populatie waaruit de empirische gegevens afkomstig zijn. Alhoewel een statistisch model een fantastisch hulpmiddel is, geeft het geen beschrijving van de hele waarheid. Statistische modellen zijn altijd gebaseerd op aannames -zoals lineariteit, onafhankelijkheid, en normaliteit (Tabachnick and Fidell, 1996:58-126)- waarvan het aannemelijk is dat ze een redelijke beschrijving van de werkelijkheid geven, maar die zeker niet de hele waarheid vormen. Vaak kunnen verschillende, elk op zich redelijke, statistische modellen geformuleerd worden. Dat laatste is problematisch omdat antwoorden op onderzoeksvragen afhangen van het statistische model dat gebruikt is. Net als voor de data geldt dus dat het statistische model slechts een indirecte kijk op de waarheid biedt.

Na het formuleren van theorieën, het verzamelen van data, en het kiezen van een statistisch model is het tijd voor een confrontatie tussen theorie en data. Hier zal in de volgende secties verder op in worden gegaan.

## 2. CONFIRMATIEVE ANALYSES MET BEHULP VAN HYPOTHESETOETSING EN BAYESIAANSE MODELSELECTIE

Een populaire benadering voor het onderzoeken van theorieën is het toetsen van hypothesen. Hierbij worden een aantal niets of nul hypothesen tegen de bijbehorende niet-niets of alternatieve hypothesen afgezet. Het is niet altijd eenvoudig om uit de resulterende significanties en p-waarden af te leiden welke ideeën of theorieën het meest plausibel zijn. Een verbetering is het vervangen van de niet-niets hypothese door een op grond van theoretische overwegingen geconstrueerde alternatieve hypothese. Nog beter is het Bayesiaans selecteren van de beste van een aantal op grond van theoretische overwegingen geconstrueerde alternatieve hypothesen.

Laat ik u een eenvoudig voorbeeld geven aan de hand van een dataset gepubliceerd in Neter et al. (1996:677) welke is weergegeven in Tabel 1.

Tabel 1. 'Breakfast Cereal' Verkoop Data.

Winkel	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19
Verkoop	11	17	16	14	15	12	10	15	19	11	23	20	18	17	27	33	22	26	28
Ontwerp	1	1	1	1	1	2	2	2	2	2	3	3	3	3	4	4	4	4	4

De data betreffen de verkoopcijfers voor 'breakfast cereal' in 19 vergelijkbare winkels. Het enige verschil is het ontwerp van de productverpakking. Ontwerp 1 was een kort dik doosje voorzien van een stripfiguur; ontwerp 2 een lang slank doosje voorzien van een stripfiguur; ontwerp 3 een kort dik doosje voorzien van een afbeelding van een atleet; en, ontwerp 4 een lang slank doosje voorzien van een afbeelding van een atleet. Er zijn drie theorieën oftewel verhalen in omloop over de wijze waarop de verkoopcijfers gerelateerd zijn aan het ontwerp van het doosje. Theorie 1 stelt dat stripfiguren 'minder goed verkopen' dan atleten, en, dat daarna slanke doosjes 'het beter doen' dan dikke doosjes. Oftewel,  $\mu_1 < \mu_2 < \mu_3 < \mu_4$ , waarbij  $\mu_2$  het gemiddelde verkoopcijfer voor het geïndexeerde ontwerp weergeeft. Theorie 2 stelt dat slanke doosjes beter verkopen dan dikke, en, dat daarna atleten het beter doen dan stripfiguren. Oftewel,  $\mu_1 < \mu_3 < \mu_2 < \mu_4$ . Theorie 3 stelt dat stripfiguren het beter doen dan atleten, en, daarna dat slank het beter doet dan dik. Oftewel,  $\mu_3 < \mu_4 < \mu_1 < \mu_2$ .

We kunnen de theorieën op verschillende manieren evalueren. Ik zal drie mogelijke scenario's schetsen en toepassen.

### Scenario 1: Standaard Hypothesetoetsing

Na toetsing met behulp van een F-toets (Grimm, 1993:269-275) blijkt dat de nulhypothese  $H_0: \mu_1 = \mu_2 = \mu_3 = \mu_4$  een significantie heeft van .00, oftewel, de nulhypothese dient te worden verworpen. Dan is dat aardig, want significant, maar het helpt niet bij het kiezen tussen de drie theorieën. In de praktijk wordt dat opgelost door naar de vier gemiddelden te kijken, en vervolgens te beslissen bij welke theorie ze het beste passen. Voor ontwerp 1 tot en met 4 zijn de gemiddelden achtereenvolgens 14.6, 13.4, 19.5 en 27.2. Het lijkt erop dat deze gemiddelden slecht bij theorie 3 passen, maar ze wijzen niet eenduidig naar theorie 1 of 2. Dat laatste geldt zeker als de standaard deviaties van

ongeveer 3.0 per ontwerp in ogenschouw worden genomen. Verder is er door op deze manier te werken niet langer sprake van een formele toetsing omdat de keuze tussen de theorieën gebaseerd is op een subjectieve evaluatie van de vier gemiddelden.

### *Scenario 2: Theoriegestuurde Hypothesetoetsing*

Een betere benadering is drie hypothesen te toetsen. Ten eerste de nulhypothese tegen een alternatief waarin de gemiddelden geordend zijn aan de hand van de theorie 1. En vervolgens hetzelfde met de ordeningen gegeven door theorie 2 en theorie 3. Dit geeft een enorme *power*-winst, en, bij verwerping van de nulhypothese weet je waar je aan toe bent zonder dat daar nog een subjectieve evaluatie van de gemiddelden voor nodig is. Alhoewel deze benadering niet in veelgebruikte software zoals SPSS geïmplementeerd is, is ze prima uitvoerbaar (Barlow et al., 1972:116-153; Robertson et al., 1988: 59-86). Een nadeel van deze benadering is dat alle nulhypotesen verworpen kunnen worden. Het is dan nog steeds niet duidelijk welke van de drie theorieën de beste is.

*Tabel 2. Theoriegestuurde Hypothesetoetsing en Bayesiaanse Modelselectie.*

Theorie	Significantie	Posterior Kans
1	.000	.985
2	.000	.015
3	.779	.000

De resultaten van de theoriegestuurde hypothesetoetsing staan in de tweede kolom van Tabel 2. Het blijkt dat theorie 3 afvalt (de nulhypothese kan niet worden verworpen), maar dat de toetsingsresultaten geen informatie leveren over de relatieve kwaliteit van theorieën 1 en 2.

### *Scenario 3: Bayesiaanse Model Selectie*

Het is dan ook nog beter om de drie theorieën direct met elkaar te vergelijken. De klassieke statistiek heeft, wanneer het aankomt op het vergelijken van dit soort niet geneste hypothesen weinig te bieden. Ten eerste is het niet duidelijk welke van de drie theorieën de rol van nulhypothese aan moet nemen. En mocht hier al een keuze in gemaakt kunnen worden, dan is het zo lastig een significantie of p-waarde uit te rekenen dat de vraag zich opdringt of dit überhaupt wel mogelijk is. De Bayesiaanse statistiek heeft geen moeite met het vergelijken van niet geneste hypothesen zoals de drie theorieën. De Bayesiaanse statistiek kwantificeert de ondersteuning die elke theorie in de data vindt, en presenteert deze als de a posteriori kans (Kass and Raftery, 1995; Hoijtink and Klugkist, 2004) dat elk van de modellen waar is. Deze a posteriori kans is een centraal begrip uit de Bayesiaanse statistiek, en is letterlijk de kans dat een theorie waar is na observatie van de data. Als deze kansen .9 en .1 zijn voor respectievelijk theorie A en B, dan vindt A veel meer ondersteuning in de data dan B. Zijn deze kansen .5 en .5, dan hebben de data geen voorkeur voor A of B. Gebruikmakend van de methode beschreven in Klugkist, Laudy and Hoijtink (2004), zijn a posteriori kansen uitgerekend voor elk van de drie theorieën. Zoals kan worden gezien in de derde kolom van Tabel 2, blijkt bij een *directe* vergelijking van de drie theorieën dat de eerste theorie superieur is aan de andere twee. Na observatie van de data is de kans dat de eerste theorie de beste is .985.

Deze schets van drie scenario's geeft aan dat het expliciet verdisconteren van theorieën in statistische modellen grote voordelen kan hebben. Dit geldt in toenemende mate als de data een mistige of vertroebelde blik op de waarheid geven, en als het statistische model slechts een benadering is van de populatie waar uitspraken over gedaan worden. Conclusies geheel door de data laten bepalen, ontaardt vaak in een spelletje "ik zie, ik zie wat jij niet ziet". En zeker als er meerdere niets tegen niet-niets hypothesen getoetst worden, is de platitude "wie zoekt zal vinden" op zijn plaats. Het expliciet meenemen van een theorie, betekent dat het statistische model wordt uitgebreid met een mal voor dat het over de data gelegd wordt. Vervolgens wordt gekeken of de mal past of niet. Dit zal zelfs bij mistige data en een imperfect statistisch model goed werken. Als er consensus is dat een van een set theorieën, een goede beschrijving van de waarheid geeft, dan zijn er meerdere malen die over de data gelegd kunnen worden. Ook met mistige data en een imperfect statistisch model zou het dan mogelijk moeten zijn te bepalen welke mal het beste past.

### 3. HET COMBINEREN VAN THEORIEËN EN STATISTISCHE MODELLEN

De zojuist geschetste benadering waarin theoretische malen over de data worden gelegd, is mijn invulling van de Bayesiaanse benadering. Kenmerkend voor deze benadering is het expliciete gebruik van voorkennis bij de data analyse. In het voorbeeld gegeven in de vorige sectie is de voorkennis hypothetisch, namelijk drie verschillende theorieën met betrekking tot het effect van designs op verkoopcijfers. In de Bayesiaanse benadering worden twee imperfecte visies op de waarheid gecombineerd en met elkaar geconfronteerd: het beeld dat via een statistisch model geschetst wordt door de data, en, de theorieën van de onderzoekers. Alhoewel data, statistisch model en theorieën elk op zich hun zwakke kanten hebben, zou de combinatie van deze drie ingrediënten het zicht op de hele waarheid wel eens sterk kunnen verbeteren.

Binnen de statistiek is er ook een niet-Bayesiaanse benadering die onderzoekers in staat stelt theorieën in modellen te vertalen, om vervolgens met behulp van empirische data het beste model te selecteren. Dit zijn de lineair structurele modellen zoals geïmplementeerd in softwarepakketten zoals AMOS, Mplus en LISREL. Deze modellen bieden u de mogelijkheid uw theorieën te schetsen in termen van geobserveerde variabelen (weergegeven door vierkantjes), latente variabelen (weergegeven door rondjes) en af dan wel aanwezige relaties tussen deze variabelen (weergegeven door pijlen). De steun van de data voor elke theorie wordt gekwantificeerd door middel van zogenaamde informatie-indices. Deze hebben welluidende namen zoals AIC, CAIC en BIC. De laatste afkorting staat voor *Bayesian Information Criterion* en kan gezien worden als de klassieke tegenhanger van de Bayesiaanse a posteriori kans. Een mogelijk nadeel van deze benadering is het rigide karakter van de specificatie van de relaties tussen variabelen: een relatie bestaat, of een relatie bestaat niet. Het is vaak realistischer aan te geven welke relaties zwakker en welke sterker zijn. Ik zal het laatste illustreren aan de hand van een factor-analyse zoals gerapporteerd door Orth en Holancova (2004).

Genoemde auteurs onderzoeken *Men's and women's responses to sex role portrayals in advertisements*. Hiertoe laten zij 320 consumenten foto's zien waarop een man en



een vrouw zijn afgebeeld. Twee voorbeelden zijn een foto waarop een man in overall huishoudelijke schoonmaakwerkzaamheden verricht terwijl een goedgeklede vrouw er al telefonerend naar staat te kijken, en een foto waarop deze rollen zijn omgedraaid. In totaal zijn er 16 condities met 20 consumenten in elke conditie. Naast sekse van de consument kan de getoonde foto al dan niet een beroepsmatige setting hebben, en zijn er vier verschillende sekserol weergaven (waarvan er hierboven twee geschetst zijn). Om de emotionele response op de foto's te meten, moest elke consument de foto in de conditie waaraan deze consument is toegewezen scoren met 21 items uit de *disapproval-approval-surprise* schaal (Bhat et al., 1998). Het doel is te onderzoeken of en hoe de emotionele response van de conditie afhangt. Als onderdeel van hun onderzoek rapporteren Orth en Holancova (2004) het resultaat van een confirmatieve factor analyse na uitsluiting van de items *happy* en *loving*. Bij deze confirmatieve analyse zijn een aantal factor ladingen op nul gezet (relatie bestaat niet) en zijn andere ladingen vrijgelaten (relatie bestaat). Een schets van hun resultaten kan worden gevonden in Tabel 3.

Tabel 3. Factor Ladingen Confirmatieve Analyse Orth en Holancova (2004).

Item	Disapproval	Approval	Surprise
Angry	.6		
Contemptuous	.7		
Disgusted	.6		
Distrustful	.6		
Irritated	.6		
Revolted	.5		
Scornful	.7		
Skeptical	.6		
Uneasy	.7		
Worried	.6		
Curious		.6	
Envious		.7	
Excited		.6	
Interested		.6	
Involved		.5	
Stimulated		.7	
Wishful		.7	
Surprised			.7
Astonished			.6

Verder is nog vermeldenswaardig dat de factoren *disapproval* en *approval* een correlatie van  $-.66$  hebben, en dat Orth en Holancova (2004) rapporteren dat het model naar tevredenheid past (een significantie voor de chi-square goodness of fit test van  $.057$ ).

Een vraag die bij dit soort confirmatieve factor analyses gesteld kan worden is in hoeverre het rigide op nul stellen van een factor lading realistisch is, en, in hoeverre het de resultaten beïnvloedt. Hieronder staan drie alternatieve en meer realistische confirmatieve specificaties die aansluiten bij de hierboven gegeven specificatie. Merk op dat dit geen kritiek is op de door Orth en Holancova (2004) uitgevoerde analyses, voor zover ik weet is er namelijk nog geen software waarmee de hieronder volgende specificaties geëvalueerd (schatten factor ladingen en bepalen van de fit van het model) kunnen wor-

den. De letters 'D', 'A' en 'S' zullen worden gebruikt om de ladingen behorend bij de factoren *disapproval*, *approval* en *surprise* aan te duiden. De cijfers 1, 2 en 3 zullen worden gebruikt om te refereren aan de item sets die bij Orth en Holancova(2004) laden op respectievelijk de factoren *disapproval*, *approval* en *surprise*.

*Specificatie 1:*  $D1 > .3$ ,  $|D2| < .3$ ,  $|D3| < .3$ ,  $|A1| < .3$ ,  $A2 > .3$ ,  $|A3| < .3$ ,  $|S1| < .3$ ,  $|S2| < .3$ ,  $S3 > .3$ . Deze specificatie lijkt op de specificatie van Orth en Holancova (2004), maar voorkomt dat factorladingen op nul worden gefixeerd. Ladingen zijn ofwel groter dan .3 (item is indicatief voor de bijbehorende factor), ofwel de absolute waarde van ladingen zijn kleiner dan .3 (item is niet of heel zwak indicatief voor de bijbehorende factor).

*Specificatie 2:*  $D1 > .3$ ,  $D2 < -.3$ ,  $|D3| < .3$ ,  $A1 < -.3$ ,  $A2 > .3$ ,  $|A3| < .3$ ,  $|S1| < .3$ ,  $|S2| < .3$ ,  $S3 > .3$ . Deze specificatie geeft weer dat het best eens zo zou kunnen zijn dat de items uit set 2 (de *approval* items) negatief laden op de factor *disapproval* en, dat de items uit set 1 (de *disapproval* items) negatief laden op de factor *approval*.

*Specificatie 3:*  $DA1 > .3$ ,  $DA2 < -.3$ ,  $|DA3| < .3$ ,  $|S1| < .3$ ,  $|S2| < .3$ ,  $S3 > .3$ . Deze specificatie geeft weer dat er een *disapproval-approval* factor is en daarnaast nog een *surprise* factor. De items uit set 1 en 2 hebben een tegengestelde lading op de *disapproval-approval* factor.

Deze specificaties zijn niet alleen realistischer dan specificaties waarbij factorladingen op nul worden gefixeerd, ze zijn ook informatiever. Van de eerste naar de derde specificatie wordt het onderscheid tussen de factoren *disapproval* en *approval* steeds kleiner. Op deze manier worden de ideeën van onderzoekers die verwachten dat het twee te onderscheiden factoren zijn, onderzoekers die twijfelen, en, onderzoekers die verwachten dat het een factor is gevangen in een mal. Net als voor het eerder gegeven voorbeeld met geordende gemiddelden, zal in de toekomst deze mal gecombineerd kunnen worden met het bijbehorende statistische model, waarna de ondersteuning van de data voor elke mal door middel van posterior kansen kan worden bepaald. Of, om het in termen van het motto van deze rede te formuleren, door een of meer goede verhalen aan het onderzoek toe te voegen, verheldert de blik op de hele waarheid die door empirische data en statistische modellen geboden wordt.

Wat mij is opgevallen, is dat de onderzoekers goede verhalen vaak formuleren in termen van 'groter dan', 'kleiner dan' en 'gelijk aan', of, 'beter dan', 'slechter dan' en 'ongeveer even goed'. Ik heb u daar tot nu toe twee voorbeelden van gegeven, een met geordende gemiddelden, en een met gerestricterde factorladingen. Hiermee is mijn voorliefde om theorieën door middel van ongelijkheidsrestricties toe te voegen aan statistische modellen gemotiveerd en verklaard: onderzoekers kunnen er mee uit de voeten. Helaas voor de onderzoekers, is de statistische theorie en de statistische software nog niet zover dat een en ander makkelijk en voor een brede familie van statistische modellen gebruiksklaar is. Voor informatie met betrekking tot de laatste stand van zaken kunnen onderzoekers bij mij terecht.

Tot nu toe heeft de nadruk gelegen op het confirmeren van theorieën. Echter, ook als onderzoekers geen theorieën hebben, en voornamelijk exploratief bezig willen zijn, kan

de Bayesiaanse benadering worden gebruikt. In de volgende sectie zal daar kort een voorbeeld van worden gegeven, waarna dit essay met een discussie zal worden afgesloten.

#### 4. EXPLORATIEVE ANALYSES MET BEHULP VAN HYPOTHESE-TOETSING EN BAYESIAANSE MODELSELECTIE

In het voorgaande zijn de traditionele nul en alternatieve hypothese omschreven als de niets en de niet-niets hypothese. Zoals eerder geschetst, heb ik een voorkeur voor alternatieve hypothesen die kleur bekennen. Dat wil zeggen hypothesen waarmee onderzoekers aangeven waarnaar ze op zoek zijn, oftewel, hypothetische modellen voor de waarheid. Soms echter, is onderzoek voornamelijk explorerend, en zijn er geen kleur-bekennende alternatieve hypothesen. In dat geval zou onze aandacht zich op de nul-hypothese kunnen richten. Laat ik twee eenvoudige voorbeelden geven. In de praktijk willen onderzoekers niet alleen weten of twee gemiddelden significant van elkaar verschillen, maar ook of de *effect-size* (Cohen, 1992) zo groot is, dat het verschil inhoudelijk gesproken relevant is. Ook willen onderzoekers meestal niet alleen weten of predictoren significant samenhangen met de afhankelijke variabele, maar ook hoe groot het percentage verklaarde variantie is. Er ontstaat dan een soort tweetrapsraket: bij een significant toetsingsresultaat wordt ook gekeken of de grootte van het gevonden effect interessant is.

Ik ondersteun deze benadering omdat statistische software momenteel niet veel meer dan dit te bieden heeft. Ik zie echter ook mogelijkheden voor verbetering. De effectgrootte kan in de nul-hypothese ingebouwd worden. Wat denkt u van H0: de twee gemiddelden verschillen minder dan .5 van elkaar versus H1: de twee gemiddelden verschillen meer dan .5 van elkaar. Of H0: het percentage verklaarde variatie is kleiner dan 10 procent versus H1: het percentage verklaarde variatie is groter dan 10 procent. Het is een directe manier om de onderzoeksvraag te beantwoorden. Het is geen tweetrapsraket. Zowel de klassieke als de Bayesiaanse statistiek bieden aanknopingspunten om tot evaluatie van dit soort hypothesen te komen. In de Bayesiaanse benadering worden twee groepen van populaties onderscheiden. In het voorbeeld met de gemiddelden, ten eerste populaties waarin de gemiddelden minder dan .5 van elkaar verschillen, ten tweede populaties waarin de gemiddelden meer dan .5 van elkaar verschillen. Vervolgens wordt voor beide groepen van populaties de kans uitgerekend dat ze de datamatrix van de onderzoeker voortbrengen. Voor de liefhebbers, deze kans wordt *marginal likelihood* (Kass en Raftery, 1995) genoemd. Voor elke populatie wordt de *marginal likelihood* vervolgens omgerekend tot een posterior kans, oftewel de kans na observatie van de data dat de nul- en alternatieve hypothese waar zijn.

Laat ik ter afsluiting aan de hand van gesimuleerde data een voorbeeld van het laatste geven. In een populatie heeft de criterium variabele  $y$  gemiddelden van respectievelijk 0 en .2 en een variantie van 1 voor twee groepen personen. In termen van Cohen (1992) correspondeert dit met een kleine effectgrootte ter grootte van .2. Na het trekken van een steekproef van 500 personen per groep uit de voor hem onbekende populatie, vindt een denkbeeldige onderzoeker gemiddelden en standaard deviaties van respectievelijk .02 en .97, en, .22 en .95. Toetsing van de nul-hypothese dat beide gemiddelden aan

elkaar gelijk zijn geeft een significantie van .001. Alhoewel dit mooi is, is het resultaat pas relevant voor de denkbeeldige onderzoeker als de effectgrootte minimaal .30 is. De gevonden effectgrootte in de steekproef is  $(.22-.02)/.96=.21$ . Wat moet de onderzoeker besluiten? Weliswaar is .21 minder dan .30, maar op het oog ook weer niet zoveel dat een effectgrootte van .30 in de populatie geheel uitgesloten is. Hoe het besluit ook uit mag vallen, het blijft een subjectieve evaluatie van de in de steekproef gevonden effectgrootte.

Met behulp van de benadering zoals beschreven in Klugkist, Laudy en Hoijtink (2004) en Hoijtink en Klugkist (2004), zijn de posterior kansen van  $H_0$ : de effectgrootte is kleiner dan .30 en  $H_1$ : de effectgrootte is groter dan .30 uitgerekend. Deze kansen waren .83 voor  $H_0$  en .17 voor  $H_1$ , oftewel, na observatie van de data is de kans op een relevante effectgrootte redelijk klein t.o.v. de kans op een irrelevante effectgrootte. Dit laat zien dat ook in explorerend onderzoek Bayesiaanse modelselectie van nut kan zijn. In plaats van competitieve verhalen te vertalen in competitieve alternatieve hypothesen, is er nu gekozen voor het inbouwen (wederom gebruikmakend van ongelijkheidsrestricties) van de gewenste effectgrootte ('het goede verhaal') in de nul en alternatieve hypothese. Het resultaat is een formele procedure waarmee de vraag van de onderzoeker in een stap beantwoord kan worden.

## 5. DISCUSSIE

In dit essay is een betoog gehouden om goede verhalen, dat wil zeggen (competitieve) ideeën en theorieën expliciet te verdisconteren bij de analyse van empirische data. Door middel van een drietal voorbeelden is geïllustreerd dat Bayesiaanse modelselectie met behulp van posterior kansen een geëigende procedure kan zijn voor het evalueren van theorieën die door het gebruik van ongelijkheidsrestricties als een mal over het te gebruiken statistische model zijn gelegd. In deze context lijkt Bayesiaanse modelselectie ruimere mogelijkheden te bieden dan meer traditionele op hypothesetoetsing gebaseerde procedures. Dit wil niet zeggen dat Bayesiaanse modelselectie in de context zoals geschetst in dit essay probleemloos is. De problemen vallen uiteen in twee groepen. Ten eerste zijn de procedures zoals voorgesteld in dit essay op dit moment alleen ontwikkeld voor toepassingen in de context van de variantieanalyse zoals in het voorbeeld met geordende gemiddelden, en het voorbeeld met twee gemiddelden die al dan niet relevant van elkaar verschillen (zie ook, Klugkist, Laudy en Hoijtink, 2004, en, Hoijtink en Klugkist, 2004). De ontwikkeling van analoge procedures voor andere statistische modellen (lineair structurele modellen, modellen voor kruistabellen) is een taak voor de komende jaren. Ten tweede moet er in de Bayesiaanse benadering een prior verdeling worden gespecificeerd voor elk van de competitieve modellen. Posterior kansen kunnen in meer of mindere mate afhangen van de keuzes die met betrekking tot deze prior verdelingen worden gemaakt. Klugkist en Hoijtink (2004) presenteren een zogenaamde *encompassing* prior verdeling die de eigenschap heeft dat voor bepaalde klassen van restricties (bijvoorbeeld de restricties gebruikt in het voorbeeld met geordende gemiddelden) de posterior kansen vrijwel onafhankelijk zijn van de gekozen *encompassing* prior. Gebaseerd op dezelfde prior verdeling laten zij ook zien dat voor andere klassen van restricties (bijvoorbeeld of twee gemiddelden relevant van elkaar verschillen) de posterior kansen sterk afhangen van de gekozen *encompassing* prior. De

belangrijkste kritiek op de Bayesiaanse benadering is immers dat resultaten soms sterk afhangen van de gekozen prior verdelingen. Een punt van verdere studie zal dan ook zijn het kiezen van prior verdelingen zodanig dat resultaten er niet of nauwelijks van afhangen. In situaties waarin dit niet mogelijk is, lijkt het mogelijk prior verdelingen te baseren op een klein deel van de data waarna het overblijvende deel van de data voor model selectie gebruikt wordt (O'Hagan, 1995, Berger en Perricchi, 1996). In situaties waarin geen van beide mogelijk is, kunnen zogenaamde subjectieve prior verdelingen gebruikt worden. Omdat resultaten dan af gaan hangen van persoonlijke subjectieve keuzes van de onderzoeker, zal deze dan in staat moeten zijn de gemaakte keuzes dusdanig goed te motiveren dat ze voor het forum van collegae acceptabel zijn.

## LITERATUUR

- Barlow, R.E., D.J. Bartholomew, H.M. Bremner and H.D. Brunk, 1972. *Statistical Inference Under Order Restrictions*. New York: Wiley.
- Berger, J.O. and T. Sellke, 1987. Testing a point null hypothesis: the irreconcilability of p-values and evidence. *Journal of the American Statistical Association*, 82, 112-122.
- Berger, J.O. and Perricchi, L., 1996. The intrinsic Bayes factor for model selection and prediction. *Journal of the American Statistical Association*, 91, 109-122.
- Bhat, S., T.W. Leigh and D.L. Wardlow, 1998. The effect of consumer prejudices on ad processing: Heterosexual consumers' responses to homosexual imagery ads. *Journal of Advertising*, 27, 9-25.
- Cohen, J., 1992. A power primer. *Psychological Bulletin*, 112, 155-159.
- Cohen, J., 1994. The earth is round ( $p < .05$ ). *American Psychologist*, 12, 997-1003.
- Congdon, P., 2001. *Bayesian Statistical Modeling*. New York: John Wiley and Sons.
- Glöckner-Rist, A. and H. Hoijtink, 2003. The Best of Both Worlds: Factor Analysis of Dichotomous Data Using Item Response Theory and Structural Equation Modeling. *Structural Equation Modelling*, 10, 544-565.
- Grimm, L.G., 1993. *Statistical Applications for the Behavioral Sciences*. New York: John Wiley and Sons.
- Hoijtink, H. and I. Klugkist, 2004. Using the Bayesian approach in psychological research. Unpublished Manuscript.
- Holland, P.W., 1986. Statistics and causal inference. *Journal of the American Statistical Association*, 81, 945-960.
- Howson, C., 2002. Bayesianism in statistics. In: R. Swinburne, *Bayes Theorem*, pp. 39-70. Oxford: University Press.
- Kass, R.E. and A. Raftery, A. 1995. Bayes factors. *Journal of the American Statistical Association*, 90, 773-795.
- Klugkist, I., O. Laudy and H. Hoijtink, 2004. Inequality constrained analysis of variance: a Bayesian approach. Unpublished Manuscript.
- Klugkist, I. and H. Hoijtink, 2004. Encompassing Priors. Unpublished Manuscript.
- Muthen, B.O. and L. Muthen, 1998. *Mplus User's Guide*. Los Angeles: Muthen and Muthen.
- Neter, J., M.H. Kutner, C.J. Nachtsheim, and W. Wasserman, 1996. *Applied Linear Statistical Models*. Chicago: McGrawhill/Irwin.
- O'Hagan, A., 1995. Fractional Bayes factors for model comparisons. *Journal of the Royal Statistical Society, Series B*, 57, 99-138.
- Orth, U.R. and D. Holancova, 2004. Men's and women's responses to sex role portrayals in advertisements. *International Journal of Research in Marketing*, 21, 77-88.
- Robertson, T., F.T. Wright, and R.L. Dykstra, 1988. *Order Restricted Statistical Inference*. New York: Wiley

- Rosenbaum, P. and D. Rubin, 1983. The central role of the propensity score in observational studies for causal effects. *Biometrika*, 70, 41-55.
- Rubin, D.B., 1987. *Multiple Imputation for Nonresponse in Surveys*. New York: John Wiley.
- Schafer, J.L., 1997. *Analysis of Incomplete Multivariate Data*. London: Chapman and Hall.
- Sober, E., 2002. Bayesianism, its scope and limits. In: R. Swinburne, *Bayes Theorem*, pp. 21-38. Oxford: University Press.
- Tabachnick, B.G. and L.S. Fidell, 1996. *Using Multivariate Statistics*. New York: Harper-Collins.

