

1. De functie van reclame: Informatie of Behang?

E.G. SMIT*

SAMENVATTING

Hoe gebruiken mensen reclame in massamedia? In deze bijdrage wordt het onderzoek besproken dat antwoord geeft op deze vraag. Dit onderzoek bestaat uit een vragenlijstonderzoek bij een representatief deel van de Nederlandse bevolking (N=1065). In dit onderzoek wordt het onderzoeksmodel getest voor vier verschillende media: dagbladen, tijdschriften, radio en televisie. Het blijkt dat zowel het gebruik als de evaluatie van reclame verschillend is bij print en audiovisuele media. Printreclame wordt positiever geëvalueerd in termen van informatie en men claimt meer aandacht aan advertenties te schenken. Met name televisiereclame wordt minder positief geëvalueerd, meer geassocieerd met irritatie en vaker vermeden. Radiocommercials worden met name genegeerd en beleefd als onderdeel van de achtergrondfunctie van het medium ("behang").

1. INLEIDING

Reclame is iets dat we allemaal kennen er herkennen als we het tegenkomen, op straat, in de kranten die we lezen, voor en tijdens televisieprogramma's of tijdens het websurfen. Reclame omgeeft ons waar we ook zijn. Hierbij hebben we niet te maken met neutrale informatie. Reclame heeft een doel en is ontworpen om ons iets te vertellen; bijvoorbeeld een product te kopen, een naam van een bedrijf betrouwbaar te vinden of ons prettig te voelen bij de aankoop van een bepaald merk. Een belangrijk aspect van reclame is het massamediale karakter, oftewel reclame is onderdeel van ons mediagebruik. Het onderzoek dat in dit artikel centraal staat, heeft het doel te beschrijven wat mensen doen met reclame ('advertising use' of reclamegebruik) en om dit gebruik te verklaren. Het concept reclamegebruik is gedefinieerd als gedrag met betrekking tot reclame in de diverse media (zie Smit, 1999). Dit gedrag bestaat uit het selecteren of vermijden van reclame en het geven van aandacht aan reclame als deel van de media-inhoud. Dit laatste betekent dat de nadruk ligt op de manier waarop mensen reclame beleven in media die meestal worden gebruikt voor andere redenen dan reclame. Deze conceptualisatie is met name gebaseerd op de Uses and Gratifications Approach (U&G); een benadering die als een van de weinige theoretische benaderingen – binnen communicatiewetenschap en aanverwante gebieden – de gebruikerskant centraal stelt binnen het communi-

* De auteur zou graag drs. Marjolein Moorman, prof. dr. Peter Neijens en dr. Lisette Pondman willen bedanken voor hun bijdrage aan het artikel en het promotie-onderzoek.

catieproces (McQuail, 1997). Een andere vergelijkbare benadering die in Nederland wordt gebruikt, is het Handelingstheoretisch perspectief (Renckstorf, 1989).

Centraal binnen de U&G is de veronderstelling dat mediagebruik is gerelateerd aan behoeften van het publiek (Palmgreen, Wenner & Rosengren, 1985). Effecten van mediagebruik – de Gratifications Obtained – zijn enerzijds gebaseerd op de inhoud van het gebruikte medium en anderzijds op het gebruik zelf of de sociale situatie. Hiernaast spelen mediakenmerken een centrale rol in het structureren van deze effecten. Critici van de U&G benadering stellen dat het concept gebruik onvoldoende is uitgewerkt en daardoor vaag (Elliot, 1977; Palmgreen et al., 1985). Sommigen proberen het concept helderder te krijgen door de meer uitgewerkte Expectancy Value theorie te integreren in de benadering (Palmgreen & Rayburn, 1982). Deze attitude theorie veronderstelt dat mensen handelen op basis van positieve verwachtingen ten aanzien van hun gedrag. Toegepast op mediagebruik betekent dit dat gebruik kan worden verklaard door de opinies die mensen hebben met betrekking tot kenmerken van dat medium of mediuminhoud en de verwachting dat het mediagebruik effect zal hebben ('beliefs'). Deze opinies worden positief of negatief geëvalueerd ('evaluations') en kunnen resulteren in motieven voor het gebruik van het medium ('gratifications sought') en tenslotte in mediagebruik of mediaconsumptie (Palmgreen & Rayburn, 1985). Het resultaat van dit mediagebruik wordt 'gratifications obtained' genoemd en kan gebaseerd zijn op de gezochte gratificaties (de verwachte gratificaties) of kan onverwacht worden verkregen. In het onderzoek dat hier zal worden besproken, wordt dit theoretische raamwerk toegepast op reclame. Reclame vormt een specifiek soort mediuminhoud, omdat de meeste mensen een medium niet zullen gebruiken voor dit soort informatie. Hiernaast zijn reclame-uitingen geen onderdeel van de redactionele inhoud, maar betaalde boodschappen van adverteerders met een specifiek doel. Er zijn weinig studies voorhanden die ook de U&G benadering hebben toegepast op reclame. Een uitzondering zijn Royne Stafford en Stafford (1996) die de U&G benadering toepasten in hun onderzoek naar motieven voor het vermijden van televisiecommercials.

De centrale vraag "Hoe en waarom gebruiken mensen reclame in massamedia?" wordt beantwoord door een literatuurstudie en een survey. Op basis van de literatuur is een conceptueel model ontwikkeld dat vervolgens is getest voor vier media (kranten, tijdschriften, radio en televisie) door middel van een survey onder een steekproef van de Nederlandse bevolking (N=1065).

2. RELEVANTE LITERATUUR

Op basis van een kwalitatieve meta-analyse (Bouwman & Neijens, 1991; Van den Hooff, 1994) zijn 50 empirische studies systematisch vergeleken en geïntegreerd (zie voor een gedetailleerde beschrijving: Smit, 1999, hoofdstuk 2). Deze analyse liet onder andere zien dat met name televisiereclame centraal stond in onderzoek naar reclamegebruik en dat maar een paar studies probeerden *verschillende* media te vergelijken. Ten behoeve van de literatuurstudie werden de relaties tussen reclamegebruik (oftewel het selecteren, vermijden of aandacht schenken aan reclame) en andere (verklarende) variabelen eruit gelicht.

Uit de analyse bleek dat *het vermijden* van reclame in de meeste studies werd gerelateerd aan de situatie van het gebruik (zoals de aanwezigheid van andere kijkers of de dag van de week), aan medium gerelateerde variabelen (zoals het type reclameblok en

de programma's rondom deze reclameblokken) en aan demografische variabelen (zoals sekse, opleiding en leeftijd) (zie bijvoorbeeld de studies van Abernethy, 1991; Bunn, 1982; Capocasa & Lucchi, 1986; Danaher, 1995; Heeter & Greenberg, 1995; Horsley, 1986; Kitchen, 1985; Sapolsky & Forrest, 1989; Zufryden, Pedrick & Sankaralingam, 1993). Maar enkele studies stelden de meer psychologische gebruikers georiënteerde variabelen centraal, zoals houding of opinies ten aanzien van reclame (e.g. Lanigan, 1997; Mittal, 1994; Royne Stafford & Stafford, 1996). Deze studies lieten in het algemeen zien dat opvattingen over irritatie en een negatieve evaluatie van met name televisiereclame leidden tot meer zapedrag.

In studies naar *aandacht* werd daarentegen de nadruk gelegd op de inhoud van de uitingen en de psychologische gebruikerskenmerken, zoals iemands houding ten aanzien van reclame (e.g. Bogart & Tolley, 1988; Burns & Foxman, 1989; Celci & Olson, 1988; Goodstein, 1993; James & Kover, 1992; Olney, Holbrook & Batra, 1991; Walker & Dubitsky, 1994). Het bleek dat een positieve evaluatie van reclame resulteerde in meer aandacht voor de reclame-inhoud. Tevens bleek dat hoe meer mensen van mening waren dat reclame voordelen heeft en mooi is uitgevoerd, hoe meer zij de reclame waardeerden (positieve attitude). Omgekeerd gold dat hoe meer men in de veronderstelling was dat er te veel reclame is en dat reclame kwetsend of schijnheilig is, hoe minder leuk reclame werd gevonden (negatieve attitude). Hiernaast bleek dat reclame verschillende functies kan hebben voor individuen, zoals een bron van vermaak of informatie over producten en diensten. Deze functies hadden een positieve relatie met de attitude ten aanzien van reclame.

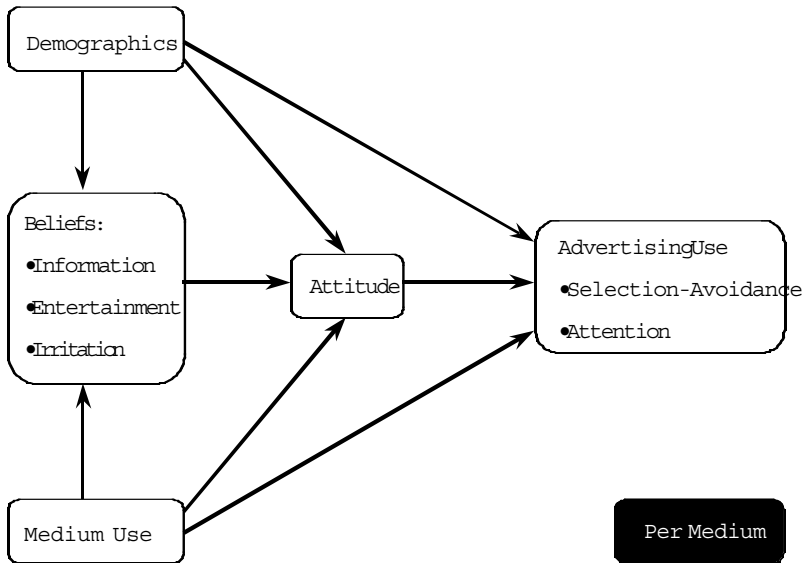
Op het eerste gezicht passen genoemde studies goed binnen het U&G raamwerk. Verschillen ontstaan met name op het gebied van terminologie: verwachtingen (U&G) zijn niet bestudeerd, maar beliefs en functies. Hiernaast laten de meeste studies zien dat maar een paar variabelen kunnen worden beïnvloed door de adverteerder, namelijk scheduling en reclame-inhoud. De meeste variabelen zijn onder controle van de gebruikers zelf, wat precies de basisveronderstelling is van de U&G benadering.

Kortom, gebaseerd op de literatuur zal het antwoord op de onderzoeksvraag "Waarom gebruiken mensen reclame?" moeten worden gevonden in kenmerken van de gebruikers zelf, hun evaluatie van reclame (attitudes en opvattingen), de situatie van het gebruik en kenmerken van het medium waarin de reclame verschijnt.

3. CONCEPTUEEL MODEL

Binnen het onderzoek ligt de nadruk op de relatie tussen reclamegebruik en de volgende gebruikerskenmerken: opvattingen (beliefs) over irritatie, informatie en vermaak, attitude ten aanzien van reclame, mediagebruik en demografische variabelen (zie Figuur 1). Hiernaast is het mediumkenmerk *pacing* (zie verder) opgenomen in het model. Omdat het mogelijk moest blijven meerdere media te vergelijken in één studie, moest het aantal verklarende variabelen beperkt blijven. Variabelen met betrekking tot situatie en andere mediakenmerken dan *pacing* zijn daarom niet opgenomen in deze studie.

Reclamegebruik is gedefinieerd als het gedrag van mediagebruikers ten aanzien van reclame in deze media. Onderscheid is gemaakt tussen gedragingen voor en tijdens blootstelling (vergelijk Levy en Windahl's (1984, 1985) conceptualisatie van 'audience activity'). Vóór blootstelling kan iemand kiezen voor een bepaald type media-inhoud,



Figuur 1. Conceptueel model.

oftewel het selecteren of juist vermijden van reclameboodschappen. Tijdens blootstelling kan het reclamegebruik bestaan uit verschillende gradaties van aandacht voor de reclame-inhoud. Reclamegebruik bestaat in deze studie dus uit het *selecteren of vermijden* van reclame-inhoud (voor blootstelling) en het schenken van *aandacht* aan reclame (tijdens blootstelling).

Reclamegebruik wordt verondersteld direct te worden beïnvloed door de *houding (attitude)* die mensen hebben ten aanzien van reclame in het betreffende medium. Een centraal kenmerk van verschillende attitude definities is “the idea of readiness of response” (Oskamp, 1991, p.7). Oftewel, een attitude is geen gedrag, niet iets dat iemand doet, maar wordt in het algemeen gezien als “a disposition to respond in a favorable or unfavorable manner to given objects” (ibid., 1991, p.8). In dit onderzoeksproject is attitude gedefinieerd als de mate waarin een persoon positief of negatief reageert op reclame in een specifiek medium (zie voor een uitgebreide definitie hoofdstuk 3 in Smit, 1999). De veronderstelling in Figuur 1 is verder dat de attitude wordt beïnvloed door de *opvattingen (beliefs)* die mensen hebben ten aanzien van het attitude object, oftewel reclame in een bepaald medium. Fishbein en Ajzen (1975; Oskamp, 1991, p.11) definiëren een ‘*belief*’ als “statements indicating a person’s subjective probability that an object has a particular characteristic”. Beliefs zijn cognitief (gedachten, ideeën), terwijl attitudes affectief zijn (gevoelens, emoties). Beliefs met een oordeel over een object worden “*evaluative beliefs*” genoemd (Oskamp, 1991, p.12). Dit type beliefs lijkt erg op attitudes en komt veel voor in reclameonderzoek, bijvoorbeeld “I like advertising” (attitude) en “I think that newspaper advertising provides me with useful information about bargains” (belief). Gebaseerd op de literatuurstudie worden de volgende soorten beliefs onderscheiden: opvattingen over de negatieve aspecten van reclame (irritatie) en opvattingen over de positieve aspecten van reclame. Binnen de positieve opvattingen is een tweedeling gemaakt tussen de functies van reclame met betrekking tot de informatie die reclame verschaft en de vermaakaspecten van reclame. De verwachting is dat

beliefs met betrekking tot informatie en vermaak een positieve relatie zullen hebben met attitude, terwijl irritatie negatief gerelateerd zal zijn aan attitude.

De volgende *demografische variabelen* worden verwacht van invloed te zijn op reclamegebruik, houding en opvattingen: leeftijd, sekse, opleiding, inkomen en het aantal personen in een huishouden. De verwachting is dat jongere mensen, vrouwen, mensen met minder inkomen en mensen die deel uitmaken van grotere huishoudens positiever zijn over reclame (houding en opvattingen) en reclame meer zullen gebruiken (in termen van minder vermijding en meer aandacht) dan andere mensen. Door tegenstrijdige resultaten kon op basis van de literatuur geen richting worden verondersteld met betrekking tot opleiding.

Ook kon op basis van de literatuur geen richting worden geformuleerd met betrekking tot de relatie tussen *mediagebruik* en reclamegebruik. Aan de ene kant zouden heavy (TV) users reclame vaker vermijden, terwijl aan de andere kant vaker aandacht wordt geschonken aan televisiereclame wanneer mensen meer televisiekijken (duur plus frequentie). Mediagebruik is in onze studie gemeten met twee variabelen, namelijk zowel de tijd die mensen besteden aan het medium (*duur*) en de *frequentie* waarmee ze dit doen.

Media verschillen op meerdere aspecten. Eén van deze aspecten is *pacing*, oftewel de waargenomen controle over het moment van gebruik en het tempo van het gebruik (Neijens & Smit, 2000; Van Cuilenburg, 1998; Van Raaij, 1991). Print media (ook wel “search media” genoemd) zijn media met interne pacing: het zijn de gebruikers die bepalen op welk moment en met welke snelheid zij deze media gebruiken. De audiovisuele media (of “display media”) daarentegen zijn media met externe pacing: het is het medium dat het moment en de snelheid van informatieoverdracht bepaalt. In dit onderzoek worden de beschreven variabelen gemeten voor zowel media met interne pacing (kranten en tijdschriften) als media met externe pacing (televisie en radio).

4. METHODE

4.1. Pilot

Om uitspraken te verzamelen voor het survey onderzoek werden, gedurende één tot anderhalf uur, zeven mannen en elf vrouwen in de leeftijd van 12 tot 74 jaar geïnterviewd. Deze pilot liet over het algemeen zien dat de verklaringen, die de geïnterviewden gaven voor hun media- en reclamegebruik, consistent waren met het conceptueel model. Ook bleek dat het voor de respondenten moeilijk was om een mening te formuleren over reclame in het algemeen. Gedrag en meningen ten aanzien van reclame zijn duidelijk gekoppeld aan het medium waarbinnen advertenties en commercials verschijnen. Deze ‘medium afhankelijkheid’ is een belangrijk onderdeel van het survey onderzoek waarin dezelfde variabelen voor vier media werden gemeten.

4.2. Survey

In het vragenlijstonderzoek werd het reclamegebruik met betrekking tot vier verschillende media gemeten, namelijk televisie, radio, kranten en tijdschriften. De respondenten werden benaderd en geïnterviewd met behulp van Computer Assisted Telephone Interviewing (CATI). Om het vraagesprek niet te lang te laten worden, werd maar één

medium meegenomen in het telefonische interview. Hierbij werden de vier media random toegewezen aan de respondenten. Aan het eind van ieder telefonisch interview werd de respondent gevraagd of hij/zij ook wilde meedoen aan het schriftelijke vervolgonderzoek. In dit vervolgonderzoek werden de vragen gesteld over de drie media die nog niet aan bod waren gekomen tijdens het telefoongesprek.

4.2.1. Het meetinstrument

De vragenlijst bestond uit 30 vragen die hetzelfde waren in de CATI en de schriftelijke versie. Hiernaast had de CATI vragenlijst 21 extra vragen met betrekking tot: de attitude ten aanzien van reclame, algemene gebruikerskenmerken (sekse, leeftijd, opleiding, grootte huishouden, het gemiddeld aantal werkuren per week) en enkele vragen over een op dat moment recente campagne voor bier.

De uitspraken met betrekking tot reclamegebruik werden gemeten op een 4-puntsschaal (van “nooit” tot “altijd”). Voor de printmedia werden de volgende uitspraken voorgesteld met betrekking tot reclamegebruik na de vraag wat de respondent zou doen wanneer hij/zij reclame zou tegenkomen (reclamegebruik): het overslaan van de advertentie (vermijding), het opzoeken van specifieke advertenties (selectie), aandacht schenken aan de advertentie, en kijken naar de advertentie (aandacht). Het vermijden van televisie- en radioreclame werd gemeten door middel de uitspraken: overschakelen naar een andere zender, het uitzetten van het geluid en het verlaten van de kamer en iets anders doen. Selectie en aandacht van televisie- en radioreclame werd gemeten met dezelfde uitspraken als bij print.

De beliefs werden gemeten door tien uitspraken per medium op een 5-puntsschaal. Deze statements waren (hier weergegeven voor de krant): (1) “Ik vind dat advertenties in de krant nuttige informatie geven over aanbiedingen”; (2) “Ik vind advertenties in de krant leuk om te lezen”; (3) “Ik vind dat advertenties in de krant zinvolle informatie geven over het productgebruik van andere mensen”; (4) “Ik vind dat advertenties in de krant nuttige informatie geven over nieuwe producten”; (5) “Ik vind advertenties in de krant amusant”; (6) “Ik vind dat advertenties in de krant op een ongewenst moment komen”; (7) “Ik vind advertenties in de krant te schreeuwerig”; (8) “Ik vind advertenties in de krant ongeloofwaardig”; (9) “Ik vind dat advertenties in de krant veel te vaak worden herhaald”; en (10) “Ik vind advertenties in de krant te veel op elkaar lijken”.

Attitude werd gemeten op een 5-puntsschaal (van “zeer negatief” tot “zeer positief”) door respondenten te vragen of zij in het algemeen positief of negatief waren ten aanzien van reclame op televisie (op de radio, in kranten en in tijdschriften).

Mediagebruik, ten slotte, werd gemeten door respondenten te vragen hoeveel dagen per week (frequentie) zij gemiddeld hun krant lezen (hun tijdschrift lezen, hoe vaak zij naar de radio luisterden en televisie keken). Na deze vraag werden zij gevraagd: “en wanneer u leest (luistert, kijkt) hoeveel minuten doet u dit, gemiddeld?” (duur).

4.2.2. Steekproef

Gedurende een periode van 27 avonden (tussen 18.00 en 22.00 uur) werden in totaal 4716 door de computer gegenereerde telefoonnummers gedraaid. Meer dan eenderde van deze nummers (37.7%) bleek niet te bestaan of was een fax- of bedrijfsnummer. Van de resterende 2936 nummers had 3.3% een probleem met de verbinding en was 23.1% onbereikbaar (oftewel de telefoon werd niet beantwoord na het vijf keer te hebben geprobeerd). Naast deze resultaten, weigerden 1057 mensen om mee te doen (36%) en deden 1105 personen mee aan het onderzoek (37.6%). Van deze 1105 interviews

werden 40 verwijderd omdat er sprake was van een item non-response van meer dan 75% van de antwoorden. In totaal bestond de steekproef dus uit 1065 respondenten. Van deze respondenten deed 50.2% ook mee aan het schriftelijke deel van het onderzoek (n=535).

Een vergelijking van de steekproef met de samenstelling van de Nederlandse populatie laat zien dat de steekproef representatief is. In Tabel 1 worden de netto steekproeven per medium en methode gepresenteerd. Chi² testen laten zien dat deze sub steekproeven niet verschillen wat betreft sekse, leeftijd of opleiding ($p > .90$). Een consequentie van de combinatie van methoden van dataverzameling is dat het mogelijk kan zijn dat de twee methoden hebben gezorgd voor verschillende resultaten (mode effect). T testen laten echter zien dat dit niet het geval is voor de gemiddelde scores op de eerder beschreven variabelen.

Tabel 1. Aantal Respondenten per Medium en Methode.

Medium	Methode		Totaal
	CATI	Mail	
Televisie	272	392	664
Radio	259	404	663
Dagbladen	267	411	678
Tijdschriften	267	398	665

5. RESULTATEN

In deze paragraaf zullen de belangrijkste resultaten van het survey onderzoek worden gepresenteerd, in sommige gevallen aangevuld met informatie uit de diepte-interviews (pilot). De resultaten zijn gecategoriseerd naar de centrale thema's: (a) reclamegebruik en (b) evaluatie van reclame (attitude en beliefs). Ook wordt een mediumtype vergelijking gemaakt.

5.1. Hoe gebruiken mensen reclame?

De diepte-interviews laten in het algemeen zien dat er weinig aandacht werd geschonken aan print en radioreclame, terwijl televisiereclame meestal werd vermeden. De consequentie hiervan was dat de geïnterviewden aangaven dat zij het moeilijk vonden om te beschrijven of, hoe en wanneer ze aandacht schonken aan reclameboodschappen. Om het reclamegebruik meer saillant te maken in het hoofdonderzoek (survey), werden de respondenten eerst gevraagd om terug te halen wanneer zij voor het laatst het medium gebruikten (wanneer, met wie, waar, hoe lang) en daarna om een schatting te maken van het aantal reclameboodschappen dat zij toen zagen. De respondenten die aangaven dit te kunnen doen (maar eenderde van de steekproef, de rest zei "weet niet"), claimden gemiddeld twee keer zoveel print als audiovisuele reclame-uitingen te hebben gezien. Ook gaven de respondenten aan gemiddeld meer aandacht aan printreclame te besteden (zie Smit, 1999, voor meer details).

Alle respondenten werden gevraagd in welke mate zij aandacht schenken aan reclame en in welke mate zij reclame vermijden (in het algemeen). Gemiddeld werd televisie-

reclame vaker vermeden door het verlaten van de kamer ($M = 2.59$, $SD = .71$) dan door zappen ($M = 2.22$, $SD = .82$) of iets anders doen ($M = 2.08$, $SD = .71$). De vermijdingsstrategie ‘muten’ (het uitzetten van het geluid) werd zelden gebruikt ($M = 1.41$, $SD = .78$). Hiernaast werden alle vermijdingsstrategieën (zappen, muten, iets anders doen en het verlaten van de kamer) nauwelijks toegepast bij radio commercials ($M = 1.35$, $SD = .66$; $M = 1.25$, $SD = .61$; $M = 1.65$, $SD = .71$; $M = 2.26$, $SD = 1.00$, respectievelijk). Televisie commercials werden gemiddeld “soms” bekeken ($M = 2.15$, $SD = .72$). Naar radiocommercials werd iets vaker geluisterd ($M = 2.51$, $SD = 1.02$). Hiernaast werd aan beide typen reclame maar weinig aandacht besteed (TV: $M = 1.80$, $SD = .59$; radio: $M = 1.55$, $SD = .59$).

In tegenstelling tot radio en televisie, werden kranten- en tijdschriftadvertenties soms van te voren geselecteerd ($M = 1.45$, $SD = .64$; $M = 1.31$, $SD = .55$, respectievelijk) en vaker aandacht aan gegeven ($M = 2.10$, $SD = .72$; $M = 1.95$, $SD = .73$, respectievelijk). Als vermijdingsstrategie sloegen respondenten advertenties over wanneer zij werden geconfronteerd met reclame bij het lezen of doorbladeren van hun krant of tijdschrift ($M = 2.76$, $SD = .72$; $M = 2.77$, $SD = .80$, respectievelijk).

De moeilijkheid om blootstelling aan reclame voor de geest te halen en het verschil tussen het gebruik van print en audiovisuele reclame zou kunnen worden verklaard door de behangfunctie van reclame. Audiovisuele reclame (met name radio) wordt beleefd als onderdeel van de continue stroom aan beelden en geluid waaraan men nauwelijks aandacht hoeft te besteden. Deze behangfunctie kwam verschillende keren terug tijdens de diepte-interviews. Een extreem voorbeeld hiervan was het echtpaar dat elke dag naar één van de Nederlandse commerciële zenders luisterde. Het echtpaar ontkende dat er commercials werden uitgezonden: “Er is nauwelijks reclame op de radio, nee, speciaal niet op deze zender. Hier hoor je nooit commercials, er is het nieuws, en ... nee, geen reclame. Dat is juist waarom we hiernaar luisteren. We luisteren naar de muziek, het nieuws ..., totaal geen reclame!”. Opvallend detail is dat “deze zender” aan stond tijdens het interview. Terwijl we het opgenomen gesprek uitwerkten, hoorden we meerdere commercials. De behangfunctie van audiovisuele media blijkt tevens uit het onderzoek naar mediabeleving van Platform’95 (1998) en het tijdbestedingsonderzoek (Faasse & Van Meerem, 1997).

5.2. Hoe evalueren mensen reclame?

Op basis van factoranalyse werden de tien belief statements per medium gereduceerd tot drie evaluatie dimensies: irritatie (items 6-10), vermaak (items 2 en 5) en informatie (items 1, 3 en 4). Nadat bleek dat ook de Cronbach’s alfa voldoende was per dimensie ($> .65$) werden drie nieuwe variabelen samengesteld (zie voor details: Smit, 1999). De gemiddelde scores voor deze variabelen staan samengevat in Tabel 2.

Naast reclamegebruik verschilt ook de evaluatie van reclame voor met name televisie en dagbladen. Krantreclame werd, gemiddeld, het meest positief geëvalueerd. Het scoorde het hoogst op de belief dimensie informatie en het laagst op irritatie, terwijl televisie werd gezien als meest negatief. Televisiereclame scoorde het hoogst op irritatie en het laagst op informatie.

Ook de gemiddelde scores op de attitudevraag laten het verschil tussen print en audiovisuele media zien (zie tabel 3). Zowel tabel 3 als tabel 2 laten zien dat radio- en tijdschriftreclame minder extreem werden beoordeeld; hun gemiddelden vallen als het ware telkens tussen televisie en krant in. Tijdschriftreclame werd over het algemeen

Tabel 2. Gemiddelde Scores op de Belief Dimensies per Medium.

Belief dimensies	Medium			
	Dagbladen (n=485)	Tijdschriften (n=482)	Radio (n=471)	Televisie (n=529)
Irritatie (n=284)	2.66 (.61) ^{a x}	2.85 (.68) ^{b y}	3.22 (.73) ^{c y}	3.65 (.74) ^{d y}
Informatie (n =322)	3.07 (.71) ^{c y}	2.97 (.78) ^{b z}	2.43 (.84) ^{a x}	2.41 (.80) ^{a x}
Vermaak (n =340)	2.63 (.81) ^{d x}	2.59 (.84) ^{c x}	2.40 (.93) ^{a x}	2.46 (.92) ^{b x}

Noot. De schalen variëren van 1 (geheel mee oneens) tot 5 (geheel mee eens). De scores zijn gemiddelde scores met de standaard afwijkingen tussen haakjes. ^{a,b,c,d} vergelijking per rij en ^{x, y, z} vergelijking per kolom waarbij verschillende subscripts significant verschillen ($p < .05$) (Paired Samples T-test, missing cases are excluded list wise).

vergelijkbaar met krantreclame beoordeeld; als informatief en het bieden van vermaak. Radioreclame werd beoordeeld als neutraal wat betreft gemiddelde scores op attitude en de drie belief dimensies.

Tabel 3. Gemiddelde Scores Attitude ten aanzien van Reclame.

Attitude	Medium			
	Dagbladen	Tijdschriften	Radio	Televisie
Positief/negatief ten aanzien van ... reclame (n=934)	3.56 (.93) ^d	3.35 (1.04) ^c	2.84 (1.12) ^b	2.42 (1.19) ^a

Noot. De schalen variëren van 1 (zeer negatief) tot 5 (zeer positief). De scores zijn de gemiddelde scores met standaard afwijkingen tussen haakjes. ^{a,b,c,d} vergelijkingen met verschillende subscripts verschillen significant ($p < .05$). (Paired Samples T-test, missing cases are excluded list wise).

Zowel de pilot als het hoofdonderzoek laten zien dat kranten – inclusief reclame-inhoud – werden beleefd als een medium dat informatie levert. Het medium televisie werd geëvalueerd als het hebben van een vermaak en ontspanningsfunctie. De commercials rondom of binnen televisieprogramma's werden meestal beleefd als ongewenste confrontatie (resultierend in irritatie), omdat zij af doen aan eerder genoemde vermaakfunctie. Tijdschrift en radiogebruik werden gezien als meer divers dan televisie en krant. Tijdschriften dienen als vermaak, het bieden van ontspanning alsmede informatie over nieuwe producten en geven ideeën. Deze diversiteit werd ook gereflecteerd in de beleving van reclame in tijdschriften: als informatief en vermaak. Radio biedt zowel informatie als vermaak en kan worden gebruikt tijdens het doen van andere dingen. Deze achtergrondfunctie wordt weerspiegeld in de beleving van radioreclame.

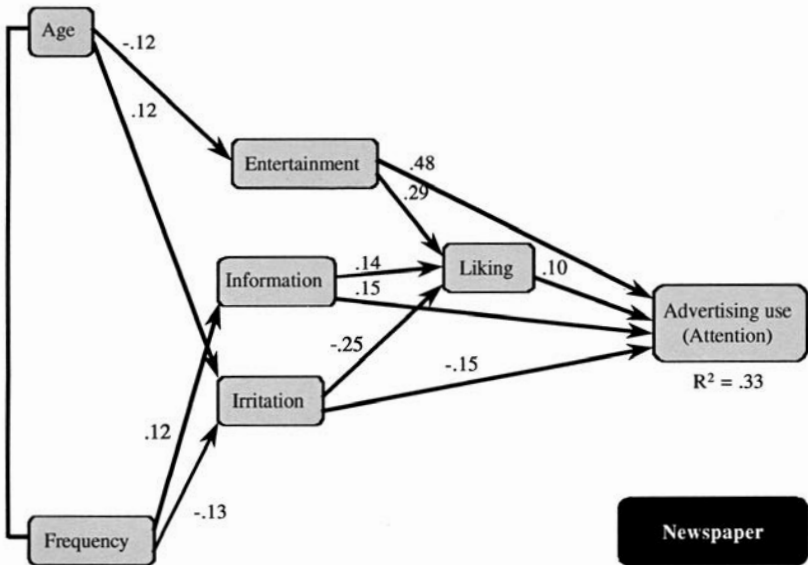
5.3. Waarom gebruiken mensen reclame?

Voor het construeren van de te verklaren variabele reclamegebruik gebruikten we de gedragsstatements uit het survey onderzoek. Deze uitspraken werden gereduceerd tot één (of meerdere) factoren per medium op basis van factor analyse (PCA, zie voor details Smit, 1999). De gebruiksfactor voor de printmedia bestond uit één factor, name-

lijk het wel of niet geven van aandacht aan reclame in kranten en tijdschriften. Voor de audiovisuele media levert PCA drie factoren op: aandacht, vermijding en fysieke vermijding. Het conceptuele model uit figuur 1 werd vervolgens getest in twee stappen. Eerst werden regressieanalyses uitgevoerd per afhankelijke variabele met het doel het aantal te testen significante relaties te beperken. Dit betekent dat reclamegebruik werd verklaard door alle andere variabelen in het model, gevolgd door een analyse ter verklaring van attitude (door beliefs, mediagebruik en demografische variabelen) en analyses voor de drie belief dimensies (verklaard door mediagebruik en demografische variabelen). Als tweede stap werd een LISREL analyse uitgevoerd per medium om de relaties in één keer te toetsen en om zo een indicatie te geven van de fit tussen het model en de data. De resultaten van deze tweede stap worden in deze paragraaf gepresenteerd.

5.3.1. Reclame in kranten

Het te toetsen causale model bestaat uit dertien relaties. Uit de LISREL analyse bleek dat dit model past ($\chi^2(12) = 6.49, p = .89, AGFI = .99; R^2_{\text{gebruik (aandacht)}} = .33, R^2_{\text{attitude}} = .17$), alhoewel het niet al te zuinig is ($PGFI = .33$). De LISREL schattingen (gebaseerd op Maximum Likelihood) laten zien dat twee relaties niet significant zijn, namelijk de relatie tussen opleiding en gebruik, en de relatie tussen opleiding en informatie¹. Deze relaties zijn niet weergegeven in onderstaand model (figuur 2). De weergegeven relaties zijn significant ($p < .05$) met gestandaardiseerde padcoëfficiënten als waarden. Figuur 2 laat zien dat met name de belief dimensie vermaak (entertainment) direct is gerelateerd aan reclamegebruik. Informatie, de afwezigheid van irritatie en een positieve attitude hebben een directe relatie met deze variabele, alhoewel minder sterk. De drie belief dimensies zijn ook indirect gerelateerd aan gebruik, via attitude. Het model laat zien dat respondenten aangeven meer aandacht te schenken aan krantreclame



Figuur 2. Verklaring gebruik dagbladreclame.

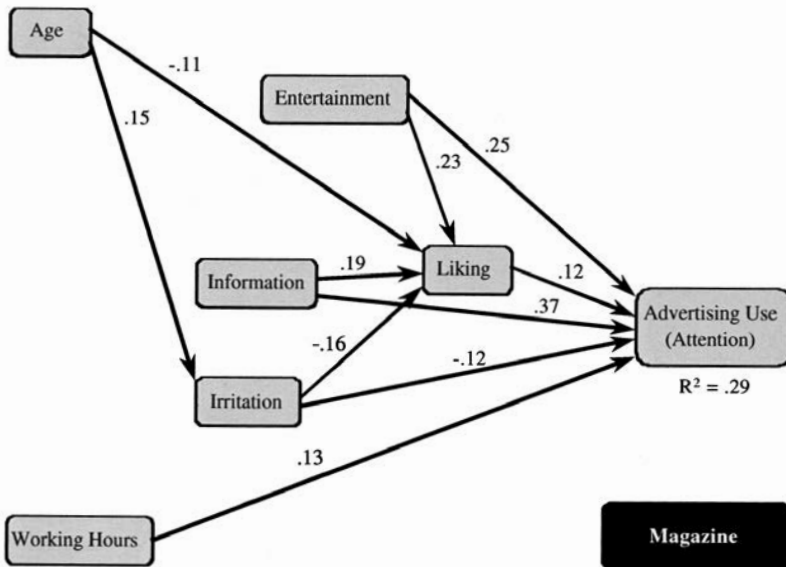
(gebruik), en positiever te staan ten aanzien van reclame (attitude), wanneer zij van mening zijn dat krantreclame een informatie en vermaak functie heeft, en geen irriterende elementen. Ook blijkt dat reclamegebruik niet direct is gerelateerd aan de frequentie van het mediumgebruik en leeftijd. Gebruik is alleen indirect gerelateerd aan deze variabelen via de beliefs. Zo blijkt dat jongere respondenten positiever zijn over krantreclame (meer vermaak, minder irritatie). Ook blijkt dat hoe vaker respondenten een krant lezen, hoe positiever zij zijn over de informatiefunctie van reclame en hoe minder zij van mening zijn dat reclame in kranten kan irriteren. Oftewel, frequente krantlezers zijn positiever over reclame in kranten. Wellicht leren frequente gebruikers door het gebruik dat zij de reclame-inhoud kunnen gebruiken; een resultaat dat in lijn is met de eerder besproken Expectancy Value theorie.

5.3.2. Reclame in tijdschriften

LISREL resultaten laten zien dat het model in figuur 3 passend is ($\chi^2(10) = 5.70, p = .84$; AGFI = .99, PGFI = .28; $R^2_{\text{gebruik (aandacht)}} = .29, R^2_{\text{attitude}} = .13$). De relaties tussen opleiding en informatie, en opleiding en aandacht blijken niet significant te zijn².

Het belangrijkste verschil tussen dit model voor tijdschriftreclame en het vorige model voor krantreclame is de sterkere rol van informatie. Dit betekent dat de respondenten die verschillen in hun mening over de informatieve functie van tijdschriftreclame verschillen in hun bereidheid om aandacht te besteden aan advertenties in tijdschriften. Net als bij krantreclame dragen demografische kenmerken nauwelijks bij aan de verklaring van reclamegebruik in tijdschriften. Alleen de variabele werkuren blijkt direct gerelateerd te zijn aan gebruik; Hoe meer uren respondenten gemiddeld werken in een week, hoe meer zij aangeven aandacht te hebben voor tijdschriftreclame.

Kortom, figuur 3 laat zien dat mensen meer aandacht schenken aan advertenties in tijdschriften wanneer zij van mening zijn dat tijdschriftreclame relevant is voor hen in ter-



Figuur 3. Verklaring gebruik tijdschriftreclame.

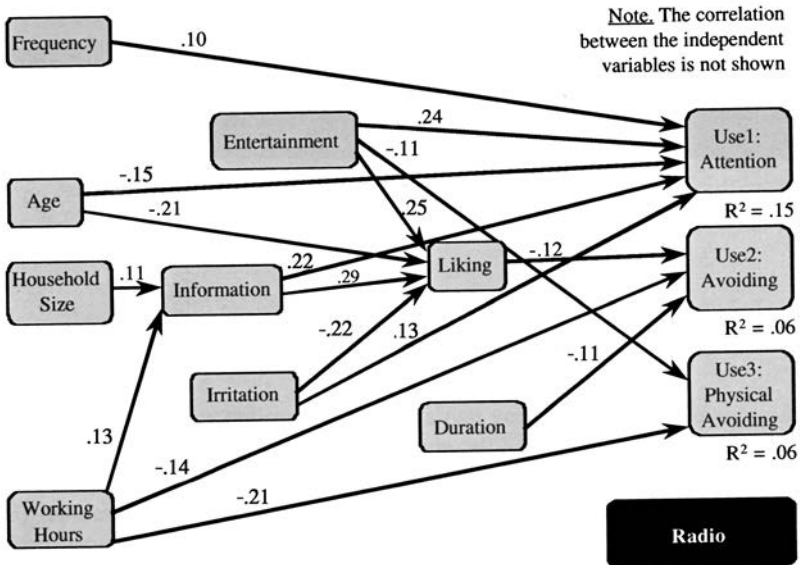
men van verwachte informatie en vermaak. Bovendien geeft men aan tijdschriftreclame vaker te gebruiken wanneer zij deze reclamevorm positiever beoordelen.

5.3.3. Reclame op de radio

Gebaseerd op de regressieanalyses zijn 22 relaties getest door middel van LISREL. Deze analyse laat zien dat het model voor het verklaren van radiogebruik net past ($\chi^2(48) = 65.11, p = .05; R^2_{\text{gebruik1 (aandacht)}} = .15; R^2_{\text{gebruik2 (vermijding)}} = .06; R^2_{\text{gebruik3 (fysieke vermijding)}} = .06; R^2_{\text{attitude}} = .25$). Hoewel de fit goed genoeg is (AGFI = .96), is het model niet erg zuinig (PGFI = .45). Drie relaties blijken niet significant te zijn³. Deze relaties en relaties met een pad coëfficiënt lager dan .10 zijn niet opgenomen in figuur 4.

Omdat drie typen gebruik zijn opgenomen in één model, is de grafische weergave lastig. In het kort laat figuur 4 zien dat opvattingen over de vermaak functie en het bieden van informatie direct de aandacht voor reclame beïnvloeden (gebruik1). Naast deze relatief sterke relaties is aandacht negatief gerelateerd aan leeftijd en irritatie, en positief aan frequentie van radiogebruik. Oftewel, hoe vaker mensen naar de radio luisteren, hoe minder zij ervan uitgaan dat radiocommercials irriteren, en hoe jonger zij waren, hoe meer aandacht zij schonken aan deze commercials. Naast leeftijd, zijn ook het gemiddeld aantal werkuren per week en de grootte van het huishouden indirect gerelateerd aan aandacht (via informatie). Dit betekent dat mensen die meer uren per week maken, meer geloven in de informatiefunctie van radioreclame en dus meer bereid zijn aandacht te schenken aan de commercials. Dit is ook het geval voor respondenten die leven in huishoudens van meerdere personen.

De houding ten aanzien van radioreclame wordt direct beïnvloed door de aanwezigheid van informatie en vermaak in radiocommercials en de afwezigheid van irriterende elementen. Naast beliefs, wordt houding ook direct beïnvloed door leeftijd: hoe jonger



Figuur 4. Verklaring gebruik radioreclame.

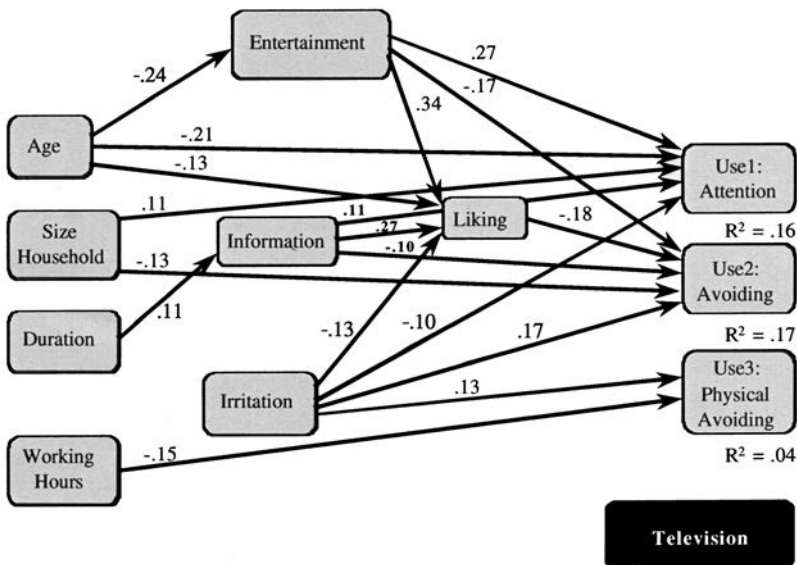
men is, hoe positiever over radioreclame. Ten slotte verklaart ook het gemiddeld aantal werkuren per week enige variantie in de attitude, alhoewel in beperkte mate en via de variabele informatie. De andere afhankelijke variabelen in het model (gebruik2 en gebruik3) konden nauwelijks worden verklaard (beide maar 6% verklaarde variantie). Gevonden relaties zijn tussen vermijding en (negatievere) attitude, (minder) vermaak, (minder) werkuren per week en per dag gemiddeld (korter) luisteren naar de radio.

5.3.4. Reclame op televisie

De LISREL resultaten laten zien dat het model bestaande uit 21 relaties past ($\chi^2(30) = 44.97$, $p = .24$, $AGFI = .97$, $PGFI = .38$; $R^2_{\text{gebruik1 (aandacht)}} = .16$, $R^2_{\text{gebruik2 (vermijding)}} = .17$, $R^2_{\text{gebruik3 (fysieke vermijding)}} = .04$, $R^2_{\text{houding}} = .26$). Figuur 5 laat alle significante relaties zien met een gestandaardiseerd pad coëfficiënt van minimaal .10.

Aandacht voor televisiereclame (gebruik1) wordt met name verklaard door zowel opvattingen over het vermaak in televisiereclame als de leeftijd van de respondenten. In figuur 5 is te zien dat jongere respondenten positiever zijn over televisiereclame (wat betreft vermaak en attitude) en meer bereid zijn aandacht te schenken aan televisiereclame. Aandacht was ook, alhoewel in mindere mate, direct gerelateerd aan de grootte van het huishouden (hoe meer personen, hoe meer aandacht) en aan de andere twee belief dimensies (informatie en afwezigheid van irritatie).

In tegenstelling tot radioreclame kan vermijding bij televisie wel worden verklaard (gebruik2). Het model in figuur 6 laat zien dat vermijding een negatieve relatie heeft met vermaak, attitude en informatie, en een positieve relatie met irritatie. Oftewel, hoe negatiever respondenten denken over televisiereclame, hoe meer zij deze reclamevorm proberen te vermijden door zapping, muten en het gewoon niet kijken. Naast een negatieve evaluatie, is vermijding ook gerelateerd aan de grootte van het huishouden: respondenten uit grotere gezinnen blijken reclame minder te vermijden. Dit effect



Figuur 5. Verklaring gebruik televisie reclame.

wordt ook gevonden in andere studies naar zappedrag die aantonen dat dit vermijddrag minder voorkomt wanneer televisie wordt gekeken met meerdere mensen (zie bijvoorbeeld Zufryden et al., 1993).

5.4. Intermedia vergelijking

De modellen uit de vorige paragraaf verklaren gemiddeld 23.3% van de variantie in aandacht voor reclame. Met name aandacht voor reclame in kranten en tijdschriften wordt relatief goed verklaard (respectievelijk 33% en 29%). Aandacht voor reclame in audiovisuele media kan minder goed worden verklaard (16% en 15% voor respectievelijk televisie en radio). Om de vergelijking van de vier media met betrekking tot aandacht te vergemakkelijken, worden in deze paragraaf de gestandaardiseerde totale effecten van alle variabelen op aandacht vergeleken (zie tabel 4).

Tabel 4. Het vergelijken van vier de modellen (totale effect scores).

Variabelen	Aandacht voor reclame			
	Dagbladen	Tijdschriften	Radio	Televisie
Attitude	.10	.12	—	—
Irritatie	-.17	-.14	-.13	-.10
Informatie	.16	.39	.22	.11
Vermaak	.51	.28	.24	.27
Frequentie	.04	—	.10	—
Duur	—	—	—	.01
Leeftijd	-.08	-.03	-.15	-.28
Sekse ¹	—	—	—	—
Opleiding	—	—	—	—
Werkuren	—	.13	.03	—
Grootte huishouden	—	—	.02	.11
R ²	.33	.29	.15	.16

Noot. Scores zijn gebaseerd op de gestandaardiseerde totale effecten van de LISREL modellen, alleen significant effecten zijn opgenomen ($p < .05$). ¹ Dummy variabele (1 = mannen, 0 = vrouwen). — betekent 'excluded' of niet significant.

Over het algemeen is vermaak, en in mindere mate informatie, het meest belangrijk voor het verklaren van aandacht voor reclame in alle media. Zoals voorspeld op basis van de Expectancy Value theorie, beïnvloeden de functies van de mediuminhoud (in dit onderzoek dus de vermaak- en informatiefunctie van reclame) het gebruik van deze inhoud. Opvallend is hiernaast de invloed van leeftijd op aandacht voor televisiereclame. Zoals voorspeld, zijn jongere respondenten positiever over televisiereclame dan ouderen en claimen vaker aandacht te besteden aan commercials.

6. DISCUSSIE

Claude, een personage van Zola in 'Une Victime de la réclame' (1866/1997), volgde elke aanbeveling uit de advertenties op en leefde in een ware hel. Iedere dag, 's morgens en 's avonds, las Claude de kranten op zoek naar advertenties (zijn 'raadsheren')

die hem adviseerden wat te kopen, waar te leven, welke kleren te dragen, welke boeken te lezen, wat te denken, en, uiteindelijk ... hoe te sterven. Claude werd een kale idioot (van alle haarmiddeltjes) en stierf uiteindelijk in een bad dat hem een zestien jaar jonger uiterlijk had moeten geven. Niemand gebruikt reclame zoals Claude dat deed: het zoeken van advertenties die hem levensregels zouden verschaffen. Maar, hoe gebruiken mensen dan wel de reclame waarmee ze worden geconfronteerd tijdens de media die ze gebruiken? De beschreven studie probeert antwoord te geven op deze vraag. Kiezen mensen voor reclame? En als ze met reclame worden geconfronteerd, zijn ze dan bereid om ernaar te kijken of zullen ze reclame zoveel mogelijk proberen te ontlopen? Bovendien: waarom gebruiken mensen wel of geen reclame? Is men op zoek naar informatie of kijkt men naar de commercials omdat ze leuk zijn?

Het onderzoek laat zien dat respondenten hun gebruik en evaluatie van reclame anders beleven voor print en audiovisuele media. Print reclame wordt positiever geëvalueerd wat betreft informatie en er wordt meer aandacht aan geschonken. Televisiereclame wordt minder positief geëvalueerd, meer geassocieerd met irritatie, en vaker vermeden. Radiocommercials worden vooral genegeerd en gezien als onderdeel van de achtergrondfunctie van het medium zelf.

Twee verklaringen zijn gegeven voor de mediaverschillen. Ten eerste is gesteld dat aan audiovisuele commercials minder aandacht wordt geschonken omdat zij worden beleefd als onderdeel van de behangfunctie van het medium. In de tweede plaats wordt televisiereclame negatief geëvalueerd omdat, in tegenstelling tot printreclame, commercials worden beleefd als een interruptie van de vermaakfunctie van het medium televisie. Voor krantreclame geldt dat zowel het medium als de advertenties worden geëvalueerd en gebruikt voor hun informatie.

Ter verklaring van reclamegebruik zijn vier modellen getoetst. In het algemeen blijkt dat de evaluatie van de vermaakfunctie belangrijk is voor het verklaren van aandacht voor reclame in alle media. Verschillen zijn vooral gevonden voor de variabele leeftijd: jongeren – gewend aan televisiereclame – evalueren audiovisuele reclame positiever en besteden meer aandacht aan reclame dan ouderen doen. Evaluatie van de informatiefunctie blijkt het meest belangrijk voor printreclame. In het algemeen kan worden gesteld dat een informatiefunctie alleen relevant is wanneer het medium de gebruiker een bepaalde mate van controle biedt, zoals bij de “search media” (oftewel print media of media met interne pacing) het geval is. Wanneer reclame alleen maar wordt getoond aan de gebruiker – zoals bij de audiovisuele media, oftewel de media met externe pacing – is een algemene beoordeling in termen van “leuk – niet leuk” en “wel of geen irritatie” een geschiktere strategie.

De behangfunctie is echter niet uniek voor reclame; commerciële boodschappen worden beleefd als alle soorten informatie in onze “information affluent society”. De snelle ontwikkeling van nieuwe communicatietechnologieën en de verbetering van de infrastructuur hebben geleid tot een overvloed aan dagelijkse informatie die wordt verstuurd en ontvangen via massamedia. Omdat de mentale capaciteit om met het aanbod van informatie om te gaan beperkt is (“information overload”), passen mensen zich aan door de omgeving snel te scannen en steeds minder aandacht te schenken aan media-inhoud (Van Cuilenburg, 1998). De mate waarin een boodschap succes heeft om door te dringen, zal in toenemende mate afhangen van de relevantie voor de consument die zich afvraagt “What is in it for me?”. Deze ontwikkeling betekent niet dat al het reclamegeld weggegooid is of dat reclame “falls on deaf ears” (Ducoffe, 1995). Het betekent echter wel dat het belangrijk is om feeling te houden met de gebruikers, om bood-

schappen te maken die leuk en relevant zijn, en te realiseren dat reclame vaak maar ‘de zwakkere kracht’ is. Volgens Ehrenberg (1998), werkt reclame “in a lower-key way”, door het verschaffen van brand awareness, het bevestigen van bestaande attitudes en door het verdedigen van de merkpositie in de hoofden van de mensen (“Here I am”). Zo stelt hij: “What is required of an ad is not heavy persuasion of either the already experienced customer or of the unconcerned newcomer, but just to ‘Tell a good story well’. This is to get and maintain attention, memorability, and salience for the brand” (ibid., 1998, p.10).

De relevantie van onderzoek naar reclame is de dominante aanwezigheid van reclame in ons dagelijks leven. Zowel wat betreft de media- en reclamebestedingen als de manier waarop mensen deze aanwezigheid beleven op straat en op televisie. Deze aanwezigheid wordt echter niet gereflecteerd in de literatuur. Onze literatuurstudie laat onder andere zien dat de meeste studies naar aandacht voor reclame of het vermijden van reclame zijn gebaseerd op *televisiereclame*. Bovendien worden zelden studies verricht waarbij meerdere mediumtypen centraal staan (NB: In Nederland is een paar jaar geleden een groot onderzoek gedaan naar mediabeleving waarbij wel meerdere media werden vergeleken, zie: Platform’95 1998; Bronner & Neijens, 1999).

Naast de roep om meer onderzoek naar mediumtype vergelijking, zouden ook andere factoren aandacht moeten krijgen in onderzoek. In dit onderzoek is de keuze gemaakt om minder aandacht te besteden aan de invloed van specifieke reclame-inhoud en bepaalde medium gerelateerde variabelen (vergelijk het SPOT onderzoek naar zapgedrag, zie: Van Meurs, 1998). Zo verklaarden de modellen uit de vorige paragraaf gemiddeld 23.3% van de variantie in aandacht voor reclame. Een verklaarde variantie van bijna 25% is redelijk goed, althans in vergelijking met andere sociaal wetenschappelijke studies. Toch betekent dit dat zo’n 75% van de variantie niet kan worden verklaard in het model. Een deel van deze onverklaarde variantie kan worden toegeschreven aan het uitsluiten van genoemde reclame- en mediavariabelen.

NOTEN

1. Ook na het weglaten van deze twee relaties (en dus het weglaten van de variabele opleiding) is het model passend ($\chi^2(9) = 6.27, p = .71$; AGFI = .99, PGFI = .32).
2. Ook na het weglaten van deze twee relaties (en dus het weglaten van de variabele opleiding) is het model passend ($\chi^2(5) = 3.25, p = .66$; AGFI = .99, PGFI = .18).
3. Ook na het weglaten van deze drie relaties (en dus het weglaten van de variabelen sekse en opleiding) is het model passend ($\chi^2(35) = 32.48, p = .59$; AGFI = .98, PGFI = .44).

LITERATUUR

- Abernethy, A.M. (1991). Differences between advertising and program exposure for car radio listening. *Journal of Advertising Research*, 31(2), 33–42.
- Bogart, L., & Tolley, B.S. (1988). The search for information in newspaper advertising. *Journal of Advertising Research*, 28(2), 9–19.
- Bronner, F. & Neijens P. (1999). *Hoe beleven mensen hun media?* In J. van Cuilenburg, P. Neijens & O. Scholten (Eds.), *Media in overvloed* (pp. 118-133). Amsterdam: University Press.
- Bunn, D. (1982). Audience presence during breaks in television programs. *Journal of Advertising Research*, 22(5), 35–39.
- Burns, A.C., & Foxman, E.R. (1989). Some determinants of the use of advertising by married women. *Journal of Advertising Research*, 29(5), 57–63.

- Capocasa, A., & Lucchi, R. (1986). *Understanding audiences of TV commercial breaks: What people do, how they react, how much they recall?* Paper presented at Esomar seminar on new developments in media research, Helsinki (Finland), 9th–12th April.
- Celsi, R.L., & Olson, J.C. (1988). The role of involvement in attention and comprehension processes. *Journal of Consumer Research*, 15(2), 210–224.
- Danaher, P.J. (1995). What happens to television ratings during commercial breaks? *Journal of Advertising Research*, 35(1), 37–47.
- Ducoffe, R.H. (1995). How consumers assess the value of advertising. *Journal of Current Issues and Research in Advertising*, 17(1), 1–18.
- Ehrenberg, A. (1998). *Justifying advertising budgets: The weak and strong theories*. Paper presented at the Admap conference on monitoring advertising performance, Londen (UK), January 22.
- Elliot, P. (1977). Media organizations and occupations: An overview. In J. Curran, M. Gurevitch & J. Woollacott (Eds.), *Mass communication and society* (pp. 142–173). London: Edward Arnold.
- Faasse, J., & van Meerem, L. (1997). *Reading in competition: The time budget survey*. Paper presented at the Worldwide Readership Symposium, Vancouver (Canada), no date.
- Fishbein M., & Ajzen, I. (1975). *Belief, attitude, intention, and behavior: An introduction to theory and research*. London: Addison–Wesley.
- Goodstein, R.C. (1993). Category–based applications and extensions in advertising: Motivating more extensive ad processing. *Journal of Consumer Research*, 20(1), 87–99.
- Heeter, C., & Greenberg, B. (1985). Profiling the zappers. *Journal of Advertising Research*, 25(2), 15–19.
- Horsley, C. (1986). *A quantified study on what people do during commercial breaks*. Paper presented at the Esomar seminar on new developments in media research, Helsinki (Finland), 9th–12th April.
- James, W.L., & Kover, A.J. (1992). Observations: Do overall attitudes towards advertising affect involvement with specific advertisements? *Journal of Advertising Research*, 32(5), 78–83.
- Kitchen, P.J. (1985). The effects of VCRs and remote control on behavior during commercial breaks. *Admap (January)*, 28–34.
- Lanigan, D. (1997). The logic of likeable ads. *Campaign*, 74(1), 32–33.
- Levy, M.R., & Windahl, S. (1984). Audience activity and gratifications: A conceptual clarification and exploration. *Communication Research*, 11(1), 51–78.
- Levy, M.R., & Windahl, S. (1985). The concept of audience activity. In K.E. Rosengren (Ed.), *Media gratifications research* (pp. 109–122). California: Sage.
- McQuail, D. (1997). *Audience analysis*. London: Sage.
- Mittal, B. (1994). Public assessment of TV advertising: Faint praise and harsh criticism. *Journal of Advertising Research*, 34(1), 35–53.
- Neijens, P., & Smit, E.G. (2000). Onderzoek naar bereikskwaliteit. In A. den Boon & P. Neijens (Eds.), *Media & Reclame: Mediaplanning en bereiksonderzoek* (pp. 413–435). Groningen: Wolters–Noordhoff.
- Olney, T.J., Holbrook, M.B, & Batra, R. (1991). Consumer responses to advertising: The effects of ad content, emotions, and attitude towards the ad on viewing time. *Journal of Consumer Research*, 17(4), 440–453.
- Oskamp, S. (1991). *Attitudes and opinions* (2nd ed.). New Jersey: Prentice–Hall.
- Palmgreen, P., & Rayburn, J.D. (1982). Gratifications sought and media exposure: An expectancy value model. *Communication Research*, 9(4), 561–580.
- Palmgreen, P., & Rayburn, J.D. (1985). An expectancy–value approach to media gratifications. In K.E. Rosengren, P. Palmgreen & L. Wenner (Eds.), *Media Gratification Research* (pp. 61–73). London: Sage.
- Palmgreen, P., Wenner, L.A, & Rosengren, K.E. (1985). Uses and gratifications research: The past ten years. In K.E. Rosengren, P. Palmgreen & L. Wenner (Eds.), *Media Gratification Research* (pp. 61–72). London: Sage.

- Platform'95 (1998). *Themadossier mediabeleving*. Amsterdam: author.
- Pollay, R.W. (1986). The distorted mirror: Reflections on the unintended consequences of advertising. *Journal of Marketing*, 50(2), 18–36.
- Renckstorf, K. (1989). Mediennutzung als sociales Handeln: Zur Entwicklung einer handlungstheoretischen Perspektive der empirischen (Massen)-Kommunikationsforschung. *Kölner Zeitung für Soziologie und Sozialpsychologie, Sonderheft*, 120-145.
- Royme Stafford, M., & Stafford, T.F. (1996). Mechanical commercial avoidance: A uses and gratifications perspective. *Journal of Current Issues and Research in Advertising*, 18(2), 27–38.
- Sapolsky, B.S., & Forrest, E. (1989). Measuring VCR 'ad-avoidance'. In M.R. Levy (Ed.), *The VCR Age: Home video and mass communication* (pp. 148–167). Newbury Park: Sage.
- Smit, E.G. (1999). *Mass media advertising: Information or Wallpaper?* Amsterdam: Spinhuis Uitgeverij.
- van Cuilenburg, J.J. (1998). *Het spel der krachten en machten*. Paper presented at the Media-Werkgroep conference "De kracht van de onmacht". Amsterdam, 19–20 October.
- van den Putte, S.J.H.M. (1993). *On the theory of reasoned action*. Unpublished doctoral dissertation, University of Amsterdam, The Netherlands.
- van Meurs, A. (1998). Zapp!: A study on switching behaviour during commercial breaks. *Journal of Advertising Research*, 38(1), 43-53.
- van Raaij, F. (1991). Scannen en focussen. *Blad*, 1, 6-25.
- Walker, D., & Dubitsky, T.M. (1994). Why liking matters. *Journal of Advertising Research*, 34(3), 9–17.
- Zola, E. (1866/1997). *Une victime de la réclame*. (In 1997 translated in Dutch by L. van Nes). Amsterdam: Perdu.
- Zufryden, F.S., Pedrick, J.H., & Sankaralingam, A. (1993). Zapping and its impact on brand purchase behavior. *Journal of Advertising Research*, 33(1), 58–66.

2. *Het SCORE-model*

Een methode voor het bepalen van de betrouwbaarheidsmarges van kijkcijfers)*

A. DEN BOON

SAMENVATTING

Het voordeel van steekproeven is dat met beperkte middelen iets gezegd kan worden over een grote populatie. Maar steekproeven geven geen zekerheid over uitkomsten, er is altijd een steekproefmarge rond de resultaten. Bij de meeste ad hoc steekproeven kunnen de steekproefmarges eenvoudig berekend worden. Maar bij een panelonderzoek is dat lastiger. Bij het CKO, een panel respondenten dat de bron vormt van het kijkonderzoek in Nederland, is dat nog veel complexer. De redenen zijn de gestratificeerde en dubbel geclusterde steekproef en de steeds andere resultaten die moeten worden gerapporteerd, zoals de kijkdichtheden van tijdvakken, programma's, reclameblokken, spots en reclamecampagnes over steeds andere doelgroepen. In dit artikel wordt het SCORE-model gepresenteerd: een eenvoudige manier voor het berekenen van de steekproefmarges in steeds wisselende omstandigheden, rekening houdend met het complexe steekproefdesign van het CKO. Uit de resultaten blijkt dat het SCORE-model goed werkt. Uit het onderzoek blijkt tevens hoe belangrijk het steekproef design voor de steekproefmarges is: na aanpassing van enkele designkenmerken verdubbelde de effectieve steekproefomvang van het CKO.

1. WAT IS EEN KIJKCIJFER WAARD?

De kwaliteit van bereiksonderzoek staat in Nederland sterk in de belangstelling. Dat geldt ook voor kijkcijfers. Aanvankelijk was er een bijna algemeen gedeeld geloof in de kwaliteit van de kijkcijfers. Maar na vijftientig jaar intensief gebruik en vertrouwen sloegen rond 1995 twijfel en scepsis toe en verschenen kritische publicaties over de validiteit van de kijkcijfers. Dankzij de inspanningen van de BVA, SPOT en Intomart is daarover inmiddels een meer evenwichtig oordeel ontstaan, waarin genuanceerd naar betekenis en waarde van het televisie bereiksonderzoek wordt gekeken. Wel bleven twijfels over de betrouwbaarheid van kijkcijfers, vooral bij televisieplannen voor reclamecampagnes. Hoe groot zijn de steekproefmarges van individuele spots en hoe groot zijn de steekproefmarges van reclamecampagnes op televisie met soms honderden spots met verschillende kijkdichtheden, verdeeld over allerlei zenders en tijdstippen?

* Dit artikel is gebaseerd op het rapport Vuistregels voor kijkcijfers, het SCORE-model van Arie den Boon en Michel Wedel

1.1. Het meten van kijkcijfers

Kijkcijfers meten is niet eenvoudig. Iedere dag zijn er andere televisieprogramma's en die programma's trekken een steeds wisselend publiek. Reclameblokken worden tussen en binnen die programma's uitgezonden. Ze verschillen alleen door hun inhoud: de reclamespots. Sommige spots worden in een burst uitgezonden, heel frequent in een grote reeks reclameblokken en op een groot aantal zenders. Andere volgen een dripschema, af en toe een spot om de reclamebekendheid niet te laten wegzakken. Voor bereiksonderzoekers is het ondoenlijk zonder elektronische instrumenten het juiste aantal kijkers vast te stellen: geen respondent kan onthouden wat hij of zij gisteren allemaal gezien heeft. Daarom worden kijkcijfers gemeten met behulp van peplemeters in een steekproef huishoudens. Een peplemeter bestaat uit twee elementen, een setmeter en de eigenlijke peplemeter. De setmeter registreert op welke zender de televisie staat afgestemd (elke toestel in een uitverkoren huishouden heeft een setmeter). Zodra een toestel wordt aangezet, uitgezet of er van kanaal wordt gewisseld wordt dit geregistreerd. Daar hoeft de respondent niets aan te doen. De peplemeter is een soort afstandbediening waarmee respondenten aangeven wanneer zij beginnen te kijken en wanneer ze stoppen met kijken. Iedere persoon in het huishouden heeft een eigen nummer en meldt zich aan en af. Op dit moment zijn er in het Continu Kijkonderzoek, dat in opdracht van SPOT wordt uitgevoerd door Intomart, ongeveer 1250 huishoudens met peplemeters. De gegevens van de setmeter en de peplemeter worden gecombineerd en iedere nacht haalt de Intomart computer de kijkgegevens op bij de huishoudens. De gegevens van ieder van de respondenten wordt verwerkt en de volgende ochtend produceert Intomart de ruwe, gewogen data en een grote reeks rapporten voor de gebruikers van de kijkcijfers.

Omdat van de respondenten ook allerlei andere gegevens bekend zijn, kunnen allerlei doelgroepen worden gevormd. Van de doelgroepen is per zender en per tijdstip per dag precies bekend wat de kijkdichtheid van een programma of reclameblok is. De kijkdichtheid is het gemiddeld aantal kijkers, uitgedrukt in een percentage. Zo betekent een kijkdichtheid van 2% in een doelgroep dat van de doelgroep gedurende het programma gemiddeld 2% heeft gekeken. Of beter geformuleerd: dat gemiddeld 2% van de doelgroep tijdens dat programma was aangemeld.

1.2. Televisieplanning

De belangrijkste gebruikers van kijkcijfers zijn de publieke en commerciële omroepen en de media- en reclamebureaus. Exploitanten en televisieplanners gebruiken deze cijfers ieder voor hun eigen doeleinden. Exploitanten bepalen er hun uitzendbeleid mee, tv-planners beoordelen de aantrekkelijkheid van programma's en blokken voor specifieke doelgroepen en stellen op basis hiervan een televisieplan samen voor adverteerders. Een televisieplan bestaat uit een reeks spots (reclameruimte voor een tv-commercial in een reclameblok), verdeeld over een groot aantal reclameblokken. Met behulp van historische analyses van de prestaties van programma's en blokken en met een dosis gezond verstand en ervaring wordt bepaald hoe groot de kijkdichtheid zal zijn van een reclameblok dat in de komende maanden zal worden uitgezonden. Gedurende de campagne wordt de prestatie van de reclamecampagne gevolgd en worden slecht presterende blokken vervangen door andere. Na uitzending van de campagne wordt de

campagne geëvalueerd. Dat wil zeggen dat aan de hand van de werkelijk behaalde kijkcijfers bepaald wordt hoeveel de adverteerder voor zijn geld heeft gekregen. De criteria daarbij zijn het aantal GRP's (het bruto aantal contacten met de doelgroep, uitgedrukt in procentpunten), de kosten per GRP, het aantal bereikte personen in de doelgroep, de contactfrequentie, enzovoort. De dagelijkse stroom kijkcijfers uit het CKO worden zeer intensief gebruikt.

1.3. Het verschil met gedrukte media

Kijkcijfers verschillen sterk van de bereikscijfers van de gedrukte media, die eens per half jaar worden opgeleverd. Bij televisie ligt de nadruk op de dagelijkse veranderingen in bereik. Daarom wordt gebruik gemaakt van een in omvang beperkt panel met een peoplemetersysteem. Bij print wordt gebruik gemaakt van telefonische vragenlijsten en zeer grote ad hoc steekproeven (n = 18.000 per jaar). Bij print worden zeer veel doelgroepkenmerken verzameld, zoals interesses, productgebruik, lifestyle, etc. De sterk gesegmenteerde tijdschriften eisen van het onderzoek dat de steekproef vergaand kan worden uitgesplitst. Dat kan bij televisie niet en is ook minder noodzakelijk omdat televisie nog steeds een massamedium is. Bij televisie ligt de nadruk op veranderingen in kijken naar een grote verscheidenheid van programma's, bij de sterk gesegmenteerde gedrukte media ligt de nadruk op het bereik van individuele titels door een grote verscheidenheid consumenten.

1.4. Veranderingen in het televisielandschap

In de afgelopen vijf jaar hebben zich enkele veranderingen voltrokken in het televisielandschap die consequenties hebben voor het kijkgedrag en voor de betrouwbaarheid van kijkcijfers. De belangrijkste veranderingen zijn het grotere aantal televisiezenders en het grotere aantal zenders dat zich op specifieke doelgroepen richt. Bovendien wordt er gedurende een langere periode per dag uitgezonden. Toch is de gemiddelde kijktijd is echter maar licht gestegen. De consequentie is dat de kijktijd over meer zenders en over een langere periode wordt verdeeld. De gemiddelde kijkdichtheid van de publieke zenders en de grote commerciële zenders is daardoor tot minder dan de helft afgenomen. Was de gemiddelde kijkdichtheid van een publieke zender enkele jaren geleden nog 4%, inmiddels is die gedaald tot minder dan 2%. Door een verdere sterke toename van het aantal zenders zal die kijkdichtheid de komende jaren ongetwijfeld nog verder dalen. Dat betekent dat er meer reclamespots moeten worden ingekocht om dezelfde reclaimedruk te kunnen opbouwen. De consequentie van die lage kijkdichtheden is dat de kijkcijfers van een programma of een reclameblok gebaseerd worden op een steeds kleiner aantal personen in de steekproef van het CKO. Dat heeft gevolgen voor de betrouwbaarheidsmarges van de resultaten. Het is al langer bekend dat de relatieve marges van een reclamespot soms wel 20%, 50% of nog veel hoger kunnen zijn. Tot nu toe was het echter niet duidelijk hoe die marges op eenvoudige wijze berekend en toegepast zouden kunnen worden in de dagelijkse praktijk van TV-planning. Ten behoeve van SPOT zijn in de vorm van het SCORE-model vuistregels ontwikkeld voor de bepaling van de steekproefmarges van kijkcijfers.

2. ONTWIKKELING VAN HET SCORE-MODEL

De kijkdichtheden van het CKO worden op een vrijwel eindeloos aantal manieren uitgesplitst en geaggregeerd op basis van een aantal kenmerken: zenders, programma's, spots, tijdvakken, (campagne)perioden en doelgroepen. Voor de planningspraktijk is echter een beperkte set van ongeveer 40 verschillende doelgroepen. Met het beschikbaar komen van respondentgegevens komen ook andere doelgroepen voor de planner binnen bereik. In dit rapport beperken we ons tot de standaarddoelgroepen. De CKO-gegevens worden door Intomart gerapporteerd met uitsplitsingen naar zenders, programma's, spots en daarnaast naar perioden en tijdvakken (zoals jaren, seizoenen, maanden, weken, dagen, minuten). Daarnaast levert Intomart gegevens op campagneniveau, bijvoorbeeld voor een doelgroep over zenders en tijdvakken heen.

2.1. Factoren die invloed hebben op steekproefmarges

Het CKO is een panelonderzoek en gebaseerd op een steekproef. Steekproefresultaten kunnen variëren als gevolg van toevallige steekproefafwijkingen. Voor eerdere analyses van de factoren die invloed hebben op de steekproefmarges zijn verschillende buitenlandse studies beschikbaar, vooral van Twyman, Wilcox en Reeve (1992 en 1996). Voor publicaties over de marges van het CKO wordt verwezen naar het artikel "De betrouwbaarheid van kijkcijfers" (Den Boon, 1993) en naar het rapport "Validiteit en betrouwbaarheid van kijkcijfers" (Den Boon en Wedel, 1997).

2.1.1. Aantal kijkers

In dit rapport wordt uitgegaan van relatieve steekproefmarges of relatieve betrouwbaarheidscijfers. Dat betekent dat niet de absolute waarde van de marges wordt bekeken, maar de marge ten opzichte van de gevonden waarde van de kdh. De formule voor het 95% betrouwbaarheidsinterval van de kijkdichtheid in een aselechte steekproef is de volgende:

$$b_a = \pm 1,96 \sqrt{\frac{p(1-p)}{n}}$$

Waarin:

b_a = betrouwbaarheidsmarge bij 95% zekerheid

p = kijkdichtheid (variërend van 0,00-1,00)

n = steekproefomvang

Van alle te berekenen steekproefmarges zal 95% het werkelijke aantal kijkers bevatten. Behalve de betrouwbaarheidsintervallen van een zender in een tijdvak kunnen we ook de relatieve steekproefmarges berekenen. We delen dan nog eens door p en houden bij kleine waarden van p na vereenvoudiging tot onze verrassing alleen nog het aantal kijkers in de formule over:

$$b_a' \cup \pm 2 \sqrt{\frac{1}{np}} = \pm 2 \sqrt{\frac{1}{\text{kijkers}}}$$

De steekproefmarges zijn dus alleen afhankelijk van een voldoende groot aantal kijkers. Dat maakt het berekenen van marges een stuk eenvoudiger! Stel dat de kijkdichtheid van een spot in de doelgroep huisvrouwen gelijk aan 5%. Het aantal kijkers

bedraagt dan 50 en de relatieve marge is 28%. Met 95% zekerheid ligt de werkelijke waarde van de gevonden kdh van 5% tussen 3,6% en 6,4%. Bij 200 kijkers is de relatieve marge gehalveerd tot 14%. Maar bij een vijfmaal kleinere kijkdichtheid van 1% moet de steekproef vijfmaal zo groot worden om dezelfde steekproefmarges te behalen! Helaas zijn er daarvoor onvoldoende huisvrouwen in de steekproef.

De steekproefmarges kunnen ook berekend worden door middel van de standaarddeviatie van kijkgedrag. Deze levert precies dezelfde resultaten op. Het voordeel van de standaarddeviatieformule is dat deze ook toegepast worden op allerlei andere kengetallen die TV-planners gebruiken, zoals de gemiddelde kijktijd, zenderaandeel, zenderbereik, campagnebereik en frequentie. Wilcox en Reeve, 1992; Twyman en Wilcox, 1996 geven voorbeelden van toepassing. Het nadeel is dat de standaarddeviaties bekend moeten zijn, zowel van de individuele spots en programma, als voor combinaties van spots in campagnes en combinaties van programma's. Die gegevens zijn echter niet altijd voorhanden. Als vuistregelmethode zijn ze daarom niet erg geschikt.

De zojuist gegeven methoden hebben wel een beperking. In de praktijk worden kijkdichtheden van spots namelijk vaak opgeteld of gemiddeld, bijvoorbeeld tot campagne-resultaten. De steekproefmarges nemen door optellen en middelen af, zodat de resultaten betrouwbaarder worden (Muilwijk, Snijders en Moors, 1992). Hoe groter het aantal spots, hoe kleiner de steekproefmarges. Hoe groot de marges zijn hangt af van:

1. de omvang van de steekproef (het aantal waarnemingen);
2. de overlap (O) tussen respondenten met de kijkdichtheden;
3. de correlatie (C) tussen het kijken naar de verschillende spots.

Stel dat we de marge van twee spots willen berekenen. De spots bevinden zich in twee reclameblokken die kort na elkaar worden uitgezonden. Van ieder van die spots is de marge bepaald. De overlap is gelijk aan 1,00 (alle panelleden doen elke keer weer mee met het onderzoek) en de correlatie tussen het kijkgedrag naar de spots is 0,40. Dan kunnen we de effectieve steekproefomvang berekenen:

$$n_{\text{eff}} = \frac{2 \cdot n}{1 + 0.C} = \frac{2 \cdot n}{1 + 1 \cdot 0,4} = 1,42 \cdot n$$

In dit geval is n_{eff} gelijk aan 1,42. We zien dus dat door het optellen van de twee metingen een toename van de betrouwbaarheid geeft, maar niet tweemaal zo groot. De toename wordt beperkt door de correlatie tussen het kijken van de respondenten.

Stel dat de kijkdichtheden van twee programma's op twee verschillende dagen voor de doelgroep huisvrouwen beide gelijk zijn aan 5%. De overlap is gelijk aan 100% en de correlatie is 0. De marge die dan resulteert daalt van 28% naar 20%, omdat er tweemaal zoveel waarnemingen zijn als in het eerste geval.

Stel dat er twee metingen van kijkdichtheden zijn van de doelgroep huisvrouwen ($n=1000$). We nemen aan dat de overlap in het voorbeeld van de huisvrouwen gelijk is aan 0,90 (10% is door omstandigheden niet op beide dagen in de steekproef betrokken) en dat de correlatie in kijkgedrag van de doelgroep huisvrouwen tussen beide dagen gelijk is aan 0,2. Dan kunnen we daarmee de betrouwbaarheidsmarge bepalen. De mar-

ge is nu 22%, een vrijwel verwaarloosbaar verschil met de eerder gevonden 20%. Als de correlatie in kijkgedrag gelijk is aan 1,00, wordt de marge weer vrijwel gelijk aan 28%.

De berekeningswijze werkt in aangepaste vorm ook bij meer inschakelingen. Deze formule is precies als de varianties van de metingen aan de S spots hetzelfde zijn. De correlaties tussen de spots mogen wel verschillen. Stel dat in de campagne gericht op huisvrouwen 100 keer een kdh van 5% wordt gerealiseerd. De overlap is 0,96, de correlatie is 0,2. De marge is dan gelijk aan 13%. De marge is door de 100 inschakelingen sterk afgenomen, maar niet met een factor 100, omdat de metingen zijn gecorreleerd. Bij een kijkdichtheid van slechts 1% is in het geval van 100 spots de marge teruggelopen van 63% tot 28%.

We kunnen nu enkele vuistregels formuleren. De marges nemen af naarmate:

1. het aantal waarnemingen ($n \times p$) groter is;
2. de overlap tussen de steekproeven kleiner is;
3. de waarnemingen minder met elkaar correleren.

De conclusie is dat het sommeren (en middelen) van kijkdichtheden, zoals dat gebeurt in campagnes, tot drastische verbetering van de betrouwbaarheid van de marges kan leiden. Omdat er sprake is van panelonderzoek en campagnes veelal betrekking hebben op een bepaalde doelgroep, is de overlap in de metingen groot. Hoe hoger de correlatie, des te minder zal de steekproefmarge afnemen. In de praktijk blijkt dat de invloed van de correlatie echter tamelijk beperkt is. In het beste geval is er geen correlatie en kan de steekproefmarge van een campagneresultaat bepaald worden met de wortel uit het aantal waarnemingen, wanneer er wel een hoge correlatie is, zal de marge echter groter zijn.

2.2. Afhankelijke waarnemingen

De bovenstaande formule geldt alleen bij een aselechte steekproef. Dat betekent dat de steekproefelementen (personen) aselekt getrokken en onafhankelijk van elkaar moeten zijn. De essentie van de aanname is dat het kijkgedrag onafhankelijk moet zijn tussen individuen in de steekproef. Deze voorwaarde is niet altijd vervuld. Zo zijn de elementen binnen een huishouden niet onafhankelijk omdat hun kijkgedrag niet onafhankelijk is (we spreken dan van clustering). Ook in andere gevallen zijn de elementen waarover de kdh wordt berekend niet altijd onafhankelijk, bijvoorbeeld als gevolg van programmatrouw. Als kijkers steeds naar dezelfde programma's kijken worden de omringende spots voor een relatief groot deel door dezelfde mensen bekeken. Wanneer in die gevallen de kijkdichtheden van een reeks spots in een campagne worden gemiddeld of opgeteld tot GRP's zijn de waarnemingen afhankelijk.

Het kan voorkomen dat een steekproef een betere afspiegeling is van de populatie dan wanneer een aselechte steekproef zou zijn getrokken. Dat is bijvoorbeeld het geval als de steekproef met behulp van bekende populatiegegevens wordt gestratificeerd. De steekproefmarges zijn dan kleiner dan bij een aselechte steekproef, of zijn gelijk aan die van een grotere aselechte steekproef. In de praktijk blijkt de stratificatie van de CKO-steekproef echter maar weinig kwaliteitsverbetering op te leveren (zie Den Boon en Wedel,

1997). Ook Twyman en Wilcox (1996) vinden in het Verenigd Koninkrijk dat stratificatie een verwaarloosbare afname van de betrouwbaarheidsmarges geeft. We zullen de geringe verbetering van de steekproef door stratificatie buiten beschouwing laten.

2.3. Vuistregels voor kijkcijfers: het SCORE-model

Voor het berekenen van de marges van kijkcijfers van campagnes moet nog met enkele andere zaken rekening worden gehouden: weging van de respondenten in de steekproef en de clustering van respondenten. Weging is het toekennen van wegingsfactoren aan de respondenten, zodat sommigen zwaarder meetellen dan anderen en de totale steekproef een betere afspiegeling is van de populatie. De wegingsfactoren variëren sterk van doelgroep tot doelgroep. Het wegen van steekproeven is een gebruikelijke techniek, die in vrijwel elk marktonderzoek wordt toegepast. Met wegen verkrijgt een steekproef de aantrekkelijke eigenschap dat de representativiteit op de weegvariabelen wordt gegarandeerd. Maar wegen heeft zijn prijs. Door de weging neemt de effectieve steekproefomvang af.

Het tweede steekproefeffect is clustering. Clustering treedt op als de steekproefelementen niet onafhankelijk zijn, maar op bepaalde variabelen afhankelijk van elkaar zijn. Clustering maakt het trekken van een steekproef vaak een stuk goedkoper, maar ook dit heeft een prijs. Wanneer gezinsleden gezamenlijk televisiekijken, is hun kijkgedrag geclusterd, waardoor het aantal onafhankelijke waarnemingen afneemt. In sommige doelgroepen is de afname van de effectieve steekproef groter dan in andere.

Het wegingseffect en het clustereffect zijn beide doelgroepafhankelijk en kunnen bij elkaar worden genomen tot een steekproefeffectfactor E . Die factor kan variëren van 0,00 tot meer dan 1,00. Dat de effectieve steekproef groter kan zijn dan de werkelijke steekproef komt echter niet veel voor, in de meeste gevallen is de effectieve steekproefomvang veel kleiner dan de getrokken steekproef.

We gebruiken voor de schatting van de marges het SCORE-model met de volgende parameters:

S: Spots: het aantal spots in een campagne.

C: Correlatie tussen het kijken op tijdstip van meting 1 en 2 (indien er meer metingen zijn, de gemiddelde correlatie).

O: Overlap tussen de metingen. Kies als waarde 1,00 en trek voor iedere maand waarop de analyse betrekking heeft 0,02 af.

R: Rating: de gemiddelde kijkdichtheid als fractie.

E: Effectieve steekproefomvang. Deze waarde is doelgroepspecifiek en corrigeert zowel voor de wegingsfactoren en voor de gemiddelde clustering van de doelgroep in de steekproef. Wanneer een campagne over een lange periode wordt geanalyseerd en de kijkdichtheden berekend zijn op bijvoorbeeld de helft van het totale aantal in de betreffende doelgroep, moet dat ook in dit cijfer worden verwerkt.

Voor het bepalen van de relatieve steekproefmarge voegen we de parameters toe aan de volgende formule:

$$\text{Marge}_{\text{som}} = \pm 2 \sqrt{\frac{1 + C \times O \times (S - 1)}{S \times R \times E}}$$

Bij een campagne van 100 spots is de gemiddelde correlatie tussen alle spots (C) vrijwel altijd kleiner dan $C = 0,30$, die stellen we hier op $0,05$. Merk op dat hier de gemiddelde correlatie tussen alle $S(S-1)/2$ paarsgewijze combinaties van spots in de formule voorkomt. Bij dezelfde gemiddelde kijkdichtheid van $0,10$ is de relatieve marge nu nog maar $0,15$. Wanneer er onzekerheid is over bepaalde parameters, is het aan te raden een sensitiviteitsanalyse te doen, om het effect van een parametervariatie te onderzoeken. Bijvoorbeeld, de gekozen waarde van de correlatie in het bovenstaande voorbeeld is een ruwe schatting. Die correlatie kan liggen tussen $0,0$ en $0,3$. Een nieuwe berekening van het SCORE-model geeft marges van $\pm 0,06$ ($C = 0,0$) en $\pm 0,34$ ($C = 0,3$). Bij een dergelijke onzekerheid over parameterwaarden zou men de meest prudente marges (in dit geval $0,34$) kunnen nemen.

Ook de marges van kijkdichtheden van tijdvakken kunnen met de SCORE-formule worden berekend, mits de tijdvakken een niet te lange periode omvatten. De marges van meer tijdvakken kunnen worden berekend met:

S: Het aantal tijdvakken.

C: De correlatie tussen het kijken in tijdvak van meting 1 en 2. In het algemeen is de correlatie van de kijkdichtheden van verschillende tijdvakken groter dan bij spots.

O: De overlap tussen de metingen. Net als bij spots is de overlap groot, zodat deze veelal op $1,00$ kan worden gesteld.

R: De rating, of gemiddelde kijkdichtheid in het tijdvak.

E: Kies de waarde van E uit de tabel met de effectieve steekproefwaarde (zie Den Boon & Wedel) van de betreffende doelgroep.

Stel dat de marge van het brutobereik van een enkel tijdvak moet worden berekend voor de doelgroep Totaal 6 plus. We gaan uit van de volgende waarden:

S: Het aantal inschakelingen is 1 .

C: De correlatie bedraagt $0,0$.

O: De overlap is $0,0$.

R: De kijkdichtheid van het tijdvak is 10% .

E: De effectieve steekproefgrootte is 777 .

De relatieve steekproefmarge van dat tijdvak is dan gelijk aan 23% .

Wanneer we de marge van de gemiddelde kijkdichtheid van dat tijdvak gedurende een week willen weten, vullen we het SCORE-model als volgt in:

S: Het aantal inschakelingen is 7 .

C: De gemiddelde correlatie bedraagt $0,20$.

O: De overlap is $1,0$.

R: De kijkdichtheid van het tijdvak is 10% .

E: De effectieve steekproefgrootte is 777 .

Dat betekent dat de marge 13% bedraagt. Bij een maand (30 inschakelingen van het tijdvak en een lagere gemiddelde correlatie van $0,05$) is de marge gelijk aan 6% . De gevoeligheid van parameterwaarden kan in voorkomende gevallen worden onderzocht.

2.3.1. Lange tijdvakken

Bij tijdvakken die een lange periode omvatten is de correlatie binnen het tijdvak onbekend. Dan wordt de marge eerst met behulp van de standaarddeviatie methode berekend. Na berekening van deze marges kan bepaald worden hoe sterk die afnemen als meer van dergelijke lange tijdvakken worden ingeschakeld. Het SCORE-model wordt gebruikt om de afname van de marges bij S inschakelingen te bepalen (zie Den Boon & Wedel). Stel dat een tijdvak een marge heeft van 20% en het betreffende tijdvak wordt 9 keer ingeschakeld. De correlatie tussen de tijdvakken is $C = 0,30$ en de overlap is gelijk aan $O = 1,00$. Dan neemt de marge af tot 12%. Als de correlatie gelijk aan nul zou zijn geweest, zou de marge zijn teruggelopen tot eenderde van 20% of 6,7%.

3. EMPIRISCHE TOETSING VAN HET SCORE-MODEL

Het SCORE-model is een theoretisch model. Er worden weliswaar empirische gegevens gebruikt, maar het model vindt zijn oorsprong in theoretische noties over steekproefdesign en steekproefmarges. Daarom is een reeks van empirische analyses uitgevoerd voor het toetsen (valideren) van het model. Het blijkt dat de resultaten van het model en van de empirische analyse dicht bij elkaar liggen. De conclusie uit de empirische toetsing is dat het SCORE-model een goed beeld geeft van de empirisch bepaalde marges. Dat is voor media-inkoop van groot belang. De empirisch bepaalde marges kunnen nog fractioneel afwijken van de werkelijke marges, maar dat is voor veel toepassingen niet relevant. Zo geeft het SCORE-model soms conservatieve schattingen van de marges. Bedacht moet worden dat de empirische marges zelf ook weer schattingen zijn van de werkelijke marges. De werkelijke marges blijven voor ons verborgen, al kunnen we door de toegepaste statistische methoden daarvan wel een goede indicatie krijgen.

Om de toepasbaarheid van het model is zo groot mogelijk te maken zijn alleen gegevens in de formule opgenomen die voor TV-planners steeds voorhanden zijn, namelijk de gemiddelde kijkdichtheid en het aantal spots. Met de kennis van de doelgroep kan eenvoudig de effectieve steekproefomvang worden verkregen. Er zijn twee parameters in het model die extra aandacht behoeven, namelijk de correlatie C tussen de spots en de waarde van E , de effectieve steekproefomvang.

3.1. De correlatie van kijkgedrag

De correlatie tussen kijkdichtheden kan in het SCORE-model een belangrijke rol spelen. Hoe sterk correleren de kijkdichtheden in de verschillende omstandigheden? We weten dat de correlatie tussen twee kijkdichtheden toeneemt naarmate:

1. de tijd tussen de gemeten spots of tijdvakken korter is (dan kijken immers dezelfde mensen nog);
2. de duur van de gemeten kijkeenheid van spots, reclameblokken, tijdvakken of programma's langer is (hoe langer de periode, des te groter is overeenstemming tussen publieken);
3. een reclamecampagne specifiek op een doelgroep is gericht en meer gestreefd wordt naar het vergroten van de contactfrequentie in plaats van bereik (dan kijken meer dezelfde mensen);

4. de kijkdichtheden hoger zijn;
5. de gemeten spots of tijdvakken sterker het kijkpatroon van de doelgroep volgen. Zo is er een hogere correlatie als de ingeschakelde spots zich op dezelfde tijdstippen van verschillende dagen bevinden.

Opeenvolgende minuut kijkdichtheden hebben een zeer hoge correlatie, maar deze correlatie neemt snel af naarmate de afstand tussen beide groter is. De kijkdichtheden van spots die zijn verdeeld over de dag zijn meestal niet of nauwelijks gecorreleerd. Naarmate de periode waarop de kijkdichtheid betrekking heeft langer is, nemen de onderlinge correlaties tussen de kijkdichtheden juist toe. Bij tijdvakken langer dan twee uur beginnen de onderlinge correlaties al betekenisvolle waarden aan te nemen. Bij lange tijdvakken of dag-, week-, maand of jaarkijkdichtheden nemen de correlaties sterk toe. Ook geldt dat de gemiddelde correlatie bij een groot aantal spots of tijdvakken een grotere invloed heeft op de marges dan bij inschakeling van slechts twee of drie spots.

In de praktijk is de correlatie C tussen het kijken naar spots door een doelgroep op twee tijdstippen vrijwel altijd zeer laag en verwaarloosbaar. Uit een reeks van 96.000 correlaties tussen minuut kijkdichtheden verdeeld over de dag blijkt dat de gemiddelde waarde gelijk is aan 0,017. Slechts 2% heeft een waarde groter dan $C = 0,20$. Uitzonderingen zijn er echter wel, want televisiekijken is vrij sterk gebonden aan tijdbesteding. Zo zijn er soms hoge ($C = 0,35$) correlaties tussen het kijken op twee opeenvolgende dagen naar dezelfde zender en op hetzelfde tijdstip. Een enkele keer is er zelfs een uitschieter van $C = 0,60$. Ook zijn er relatief sterke correlaties tussen opeenvolgende momenten in hetzelfde blok en in enkele gevallen ook tussen opeenvolgende blokken. Dat laatste geldt vooral als het publiek is blijven kijken, zoals bij programma-onderbrekende blokken. Bij de inschakeling van veel spots nemen de correlaties gemiddeld weer sterk af, omdat de spots dan noodzakelijkerwijs over allerlei programma's en soms ook zenders zijn verdeeld. Ook komen er nogal eens licht negatieve correlaties voor, bijvoorbeeld als de doelgroepen sterk uiteenlopen. Gemiddeld kan met een correlatie van $C = 0,05$ rekening gehouden worden.

De correlaties tussen de kijkdichtheden van tijdvakken zijn hoger dan die tussen spots. Dat blijkt uit de correlaties van een groot aantal tijdvakken, dagen, doelgroepen, zenders en data. Opvallend is dat deze correlatie in veel gevallen dicht bij $C = 0,30$ liggen. In enkele gevallen zijn de correlaties hoger, tot een maximum van $C = 0,45$, maar er zijn ook veel lagere correlaties ($C = 0,10$). In gevallen waarin de kijkdichtheden van een doelgroep bij een tijdvak en zender laag zijn, nemen de gemiddelde correlaties af tot $C = 0,10$ of lager. De correlaties tussen de kijkdichtheden op werkdagen en weekenddagen verschillen weinig, al zijn die soms enigszins lager. Voor de berekeningen van de steekproefmarges is het in de meeste gevallen verstandig met een correlatie van 0,30 te rekenen.

3.2. De grote invloed van het steekproefdesign

In het SCORE-model staat de effectieve steekproefomvang E in de noemer. De effectieve steekproefomvang van een doelgroep wordt bepaald door de weging en de clustering van de steekproef. Dit is het steekproef design effect. E wordt berekend door:

$$E = E1 \times E2 \times N_{\text{doelgroep}}$$

E1 is het effect van de steekproefweging. E1 is vrijwel altijd kleiner dan 1, want in de meeste gevallen is het effect van de respondenten in de doelgroep die een weefactor hebben kleiner dan 1 groter dan de invloed van de respondenten met een weefactor groter dan 1. In een heel enkel geval is E1 groter dan 1,00, bijvoorbeeld bij kinderen. Dit betekent dat de betreffende doelgroep in de steekproef is oververtegenwoordigd en terug wordt gewogen. De statistische betekenis daarvan is dat de steekproef voor die groep beter is dan een aselechte steekproef.

Het effect van de clustering wordt gegeven door E2. E2 heeft vrijwel steeds een belangrijke negatieve bijdrage aan de effectieve steekproefomvang: clustering door gezamenlijk kijken in huis heeft een sterk negatief effect op de effectieve steekproef. Gelukkig is dat vooral bij grote doelgroepen, want bij kleine doelgroepen zijn er vaak maar één of gemiddeld iets meer dan één persoon per huishouden. Het effect van de clustering is geen doelgroepconstante en deels afhankelijk van de kijkdichtheid van het specifieke programma. Het clustereffect is sterker (en de waarde van E2 is kleiner) naarmate:

1. de doelgroep meer personen per huishouden bevat;
2. de kijkdichtheid hoger is omdat het programma beter aansluit bij die doelgroep;

Als veel mensen kijken en de kijkdichtheid hoog is, is de kans dat er meer mensen per huishouden kijken groter. Bij kleine zenders en weinig bekeken programma's is het effect van E2 dus kleiner. In het algemeen zal bij campagnes gericht op een doelgroep met meer personen per huishouden de kans groot zijn dat er meer huisgenoten naar dat programma kijken.

Het product van E1 en E2 is vrijwel altijd kleiner dan 1. De betekenis van E is in het algemeen veel groter dan van correlatie C. Zou in het bovenstaande rekenvoorbeeld het product van de waarden van E1 en E2 niet 0,55 zijn, maar 1,00 dan zou de effectieve steekproefgrootte gelijk zijn aan 234 en de marge niet 53% zijn, maar 35%.

3.3. Consequenties voor het steekproefdesign

Dit alles geeft nog eens aan dat de wijze waarop de steekproef is getrokken en gewogen van grote betekenis is voor de betrouwbaarheid van de resultaten. Na het verschijnen van het validatierapport kijkcijfers heeft Intomart een reeks van maatregelen genomen. De steekproef ging omhoog van 1000 naar 1250 huishoudens. De nogal ineffectieve wegingprocedure werd aangepast. Er werd voornamelijk op huishoudkenmerken gewogen en niet op individuele kenmerken, waardoor het effect beperkt was. Ook de clustering van respondenten in regio en steden werd verminderd. Als gevolg daarvan is de effectieve steekproefomvang gemiddeld verdubbeld! Het is dus veel effectiever de steekproeftrekking verder te verbeteren en daarmee de effectieve steekproefomvang te vergroten.

Voor kijkonderzoek zal er altijd een ruime marge bij kleine zenders en kleine doelgroepen blijven bestaan. In veel gevallen is het daarom noodzakelijk grotere hoeveelheden spots gezamenlijk te beoordelen, of over langere tijdvakken te rapporteren in plaats van individuele spots of tijdvakken.

4. TOT SLOT

Een artikel van Tony Twyman en Steve Wilcox (1996) over betrouwbaarheidsmarges van de kijkcijfers in het Verenigd Koninkrijk heeft de veelzeggende toevoeging: "... and how to live with it." De problematiek van de steekproefmarges is namelijk fundamenteel onoplosbaar zolang kijkcijfers worden gebaseerd op een steekproef kijkers. Natuurlijk kan er veel worden gedaan door de steekproef te vergroten, de steekproef-trekking aan te passen en de weging te beperken. Deze maatregelen kosten echter veel geld en steeds zal de wens bestaan om resultaten verder uit te splitsen en te verbijzonderen. Daarom is het zaak de gebruikers te informeren welke wijzen van gebruik wel betrouwbare resultaten opleveren en welke niet. Met dit rapport hopen wij een bijdrage geleverd te hebben aan de mogelijkheden gebruikers te informeren over de waarde van kijkcijfers.

LITERATUUR

- Boon, A.K. den. (1993). De betrouwbaarheid van kijkcijfers. Een onderzoek naar de factoren die de betrouwbaarheid van kijkcijfers beïnvloeden. *Massacommunicatie 21*, (4), 262-288.
- Boon, A.K. den, en M. Wedel. (1997). *Validiteit en betrouwbaarheid van kijkcijfers*. BVA / Associatie Nederlandse Adverteerders.
- Boon, A.K. den, en M. Wedel. (1999). *Vuistregels voor kijkcijfers, het SCORE-model*. SPOT: Amsterdam.
- Fisher, R.A. en Yates, F. (1948). *Statistical tables for biological, agricultural and medical research*. London: Oliver and Boyd.
- Kish, L. (1965). *Survey Sampling*. John Wiley: New York.
- Muilwijk, J., Snijders, T.A.B. en Moors, J.J.A. (1992). *Kanssteekproeven*. Leiden: Stenfert Kroese.
- Twyman T., and Wilcox, S. (1996). *The variability of audience measurement data and how to live with it*. (Paper presented to the ARF/ESOMAR Broadcast Audience research Symposium in San Francisco).
- Wilcox, S. en Reeve, B. (1992). *Statistical efficiencies in the New UK television audience measurement panels*. Paper presented at the ARF/ESOMAR Worldwide Broadcast Audience Research Symposium in Toronto, Canada.

3. De invloed van Internet op extern informatiezoekgedrag bij de aankoop van financiële diensten¹

P.W.J. PLASMEIJER

SAMENVATTING

Dit artikel gaat in op de invloed van Internet op het zoekgedrag van consumenten voorafgaand aan de aankoop van financiële diensten en de implicaties hiervan voor marketing (Plasmeijer, 1999). Een instrument is ontwikkeld dat kan worden gebruikt om extern informatiezoekgedrag ten behoeve van een koopbeslissing voor financiële diensten te karakteriseren aan de hand van twee dimensies van informatiezoekgedrag, resulterend in vier informatiezoekgedrag segmenten. Het onderzoek wijst uit dat Internetgebruik leidt tot een verhoogde zoekintensiteit en een verandering van zoekrichting bij consumenten. Financiële dienstverleners moeten anticiperen op dit veranderend informatiezoekgedrag teneinde de koopbeslissing van consumenten positief te beïnvloeden. Hiertoe is een aanpassing van de interne organisatie, vaardigheden en activiteiten van financiële dienstverleners vereist.

1. INLEIDING

Internet is een interactief medium dat in hoog tempo wordt omarmd door consumenten en leveranciers. Het is een informatiebron, een marketing kanaal en een wereldwijd communicatiemedium dat gebruikt kan worden voor het zoeken naar informatie, amusement, interactie, educatie en hyperrealiteit. Interactiviteit staat voor een kort feedback interval. Het leidt tot controle over het communicatieproces door zowel aanbieders als consumenten (Williams et al, 1988; Van Raaij, 1998). Interactie op Internet omvat zowel machine-interactiviteit als persoonsinteractiviteit (Hoffman en Novak, 1996). Machine-interactiviteit betreft interactie met databases, zoals het plaatsen van informatie op Internet of het opvragen van informatie uit databases van bijvoorbeeld commerciële bedrijven. Persoonsinteractiviteit heeft betrekking op communicatie tussen individuen onderling, bijvoorbeeld in chatboxen en nieuwsgroepen. De toename van deze directe interactiemogelijkheden op Internet zal, naar verwachting, leiden tot veranderingen in consumentengedrag en resulteren in relaties tussen aanbieders en afnemers, waarbij de consument meer controle heeft over het communicatieproces. Het is voor

¹ Dit artikel is gebaseerd op het proefschrift "The Influence of the Internet on Prepurchase External Search for Financial Services" (1999) van Pauline Plasmeijer, promotores: J.C. Hoekstra en W.F. van Raaij.

aanbieders van belang inzicht te verkrijgen in veranderingen in extern informatiezoekgedrag van consumenten als gevolg van Internet. Het externe informatiezoekgedrag biedt immers binnen het aankoopproces de eerste mogelijkheid voor bedrijven om informatie te verstrekken aan consumenten en de koopbeslissing positief te beïnvloeden (Wilkie en Dickson, 1985).

Dit artikel beschrijft een studie naar de invloed van Internet op extern informatiezoekgedrag. Vanuit een aantal overwegingen is deze uitgevoerd voor de financiële dienstverlening. Ten eerste vereist de hevige concurrentie binnen de financiële dienstverlening een efficiënte en effectieve distributie, alsmede het aangaan van langdurige, directe relaties met loyale klanten (Reichheld, 1996; Grönroos, 1994; Fain and Roberts, 1997). Internet kan hierbij een belangrijke rol spelen, aangezien het beide doelstellingen kan ondersteunen. Ten tweede is de financiële dienstverlening van nature een bedrijfstak waarbij nieuwe technologieën snel worden aanvaard en geïmplementeerd. Dit heeft te maken met het niet-fysieke en elektronische karakter van financiële diensten. Deze sector leent zich derhalve zeer goed voor distributie via Internet. Ten derde biedt de hoge penetratie van financiële diensten (o.a. de betaalrekening) reële kansen voor een brede adoptie door consumenten van elektronische distributie van financiële diensten via Internet.

Dit artikel behandelt allereerst de gehanteerde onderzoeksmethode (paragraaf 2) en het domein van extern informatiezoekgedrag (paragraaf 3). Vervolgens wordt de ontwikkeling van een model van extern informatiezoekgedrag besproken (paragraaf 4), gevolgd door een analyse van de invloed van Internet op extern informatiezoekgedrag (paragraaf 5). Het artikel sluit af met de belangrijkste conclusies (paragraaf 6) en met aanbevelingen voor verder onderzoek (paragraaf 7).

2. METHODE VAN ONDERZOEK

In deze studie staan vijf onderzoeksvragen centraal. Deze vragen worden in het onderstaande weergegeven, en bij elke vraag wordt kort de gehanteerde methode vermeld. In het vervolg van dit artikel wordt op de methoden nader ingegaan.

1. Wat is het domein van extern zoekgedrag ten behoeve van een koopbeslissing voor financiële diensten?

Het domein van extern informatiezoekgedrag wordt gedefinieerd door middel van literatuuronderzoek. Het doel van het literatuuronderzoek is een zo compleet mogelijk beeld te verkrijgen van de beschikbare wetenschappelijke kennis binnen dit domein. De belangrijkste determinanten van extern informatiezoekgedrag worden geïdentificeerd.

2. Hoe kan het huidige extern zoekgedrag van consumenten naar financiële diensten worden omschreven?

Met behulp van exploratief onderzoek (deskundigen interviews en kwalitatief consumenten onderzoek) wordt het domein huidig extern informatiezoekgedrag naar financiële diensten verder gedefinieerd en de set determinanten van extern informatiezoekgedrag aangevuld. Protocolen van de consumenten interviews worden gemaakt ten behoeve van analyse en interpretatie.

3. Is het mogelijk om segmenten te identificeren op basis van extern zoekgedrag ten behoeve van een koopbeslissing voor financiële diensten?

De determinanten van extern informatiezoekgedrag worden gevalideerd met behulp van een kwantitatief consumentenonderzoek. Hiertoe worden schaalitems ontwikkeld voor de afzonderlijke determinanten. Met behulp van factoranalyse en Cronbach's Alpha worden betrouwbare constructen (factoren) van extern zoekgedrag vastgesteld. De factorscores vormen de basis voor clusteranalyse, welke moet uitwijzen of clusters voor zoekgedrag kunnen worden gevonden. Vervolgens wordt een discriminantanalyse uitgevoerd ter validatie van de gevonden clusters en om deze clusters in andere datasets te kunnen vinden, hetgeen noodzakelijk is bij de volgende deelvraag. Dit resulteert in een gevalideerd instrument ter karakterisering van extern informatiezoekgedrag.

4. Wat is de invloed van Internet op het domein van extern zoekgedrag ten behoeve van een koopbeslissing voor financiële diensten in het algemeen en voor de segmenten in het bijzonder?

Empirisch onderzoek wordt uitgevoerd om inzicht te verkrijgen in de invloed van Internet op extern informatiezoekgedrag voor financiële diensten in het algemeen en voor de gevonden clusters in het bijzonder. Dit onderzoek wordt uitgevoerd op de database van een financiële dienstverlener, en richt zich op verschillen die bestaan in extern informatiezoekgedrag naar financiële diensten tussen klanten die Internet gebruiken en klanten die geen gebruik maken van Internet.

5. Hoe kan de financiële dienstverlening Internet integreren in haar marketing communicatie?

De inzichten verkregen in de voorgaande onderzoeksvragen worden gebruikt om antwoord te geven op de vraag hoe de financiële dienstverlening Internet kan integreren in haar marketingcommunicatie.

3. DOMEIN VAN EXTERN INFORMATIEZOEKGEDRAG

Het domein van extern zoekgedrag is gedefinieerd vanuit marketingperspectief door middel van een uitgebreide literatuurstudie. Extern informatiezoekgedrag is een onderdeel van de oriëntatiefase van de consumptiecyclus en de tweede fase in het koopproces (Kotler, 1994; Kotler en Armstrong, 1998; Antonides en Van Raaij, 1998). Inzicht in extern informatiezoekgedrag van consumenten is van belang voor aanbieders omdat dit de eerste fase in het koopproces is, waarbij een aanbieder informatie kan verschaffen en de koopbeslissing positief kan beïnvloeden (Wilkie en Dickson, 1985).

Extern zoekgedrag bestaat uit twee elementen, te weten het extern *zoekgedrag* en het extern *zoekproces* (Engel, Blackwell en Miniard, 1990; Antonides en Van Raaij, 1998). Ten aanzien van het extern zoekgedrag worden twee hoofdvormen onderkend: doelgericht informatie zoeken en incidenteel leren van informatie die men ongevraagd ontvangt. Het extern informatiezoekproces van consumenten bestaat uit drie onderdelen, te weten: zoekintensiteit, zoekrichting en zoekvolgorde (Engel, Blackwell, en Miniard, 1990).

Zoekintensiteit heeft betrekking op de mate waarin naar informatie wordt gezocht. De

hoeveelheid informatie die consumenten opvragen, wordt beïnvloed door economische factoren (kosten en baten) en psychologische factoren, te weten bekwaamheid en motivatie (Schmidt en Spreng, 1996; Beatty en Smith, 1987). Specifiek voor financiële diensten geldt dat gepercipieerd risico, product complexiteit, en 'satisficing' determinanten lijken te zijn voor de intensiteit van extern informatiezoekgedrag (Maute en Forrester, 1991; Simon, 1976). Satisficing betekent dat consumenten niet op zoek zijn naar de optimale beslissing, maar dat zij genoeg nemen met de eerste oplossing die aan bepaalde eisen voldoet (Simon, 1976). Hieruit blijkt het belang voor aanbieders inzicht te hebben in het extern zoekgedrag waardoor de effectiviteit van commerciële inspanningen kan worden vergroot.

Zoekrichting betreft zowel het type informatie en de informatiebronnen die door consumenten gebruiken, als ook de mediakeuze van consumenten (Engel, Blackwell en Miniard, 1990). Informatiebronnen bestaan uit vier categorieën, te weten: publieke bronnen, commerciële bronnen, sociale bronnen, en directe inspectie (Engel, Blackwell en Miniard, 1990; Peter en Olson, 1990; Kotler, 1994; Schmidt en Spreng, 1996; Antonides en Van Raaij, 1998). Media kunnen worden ingedeeld in de categorieën massamedia, interactieve media, interpersoonlijke communicatie, en 'computer-mediated environments' (Hoffman en Novak, 1996). Op Internet komen alle informatiebronnen samen waardoor Internet een complete informatiebron kan zijn.

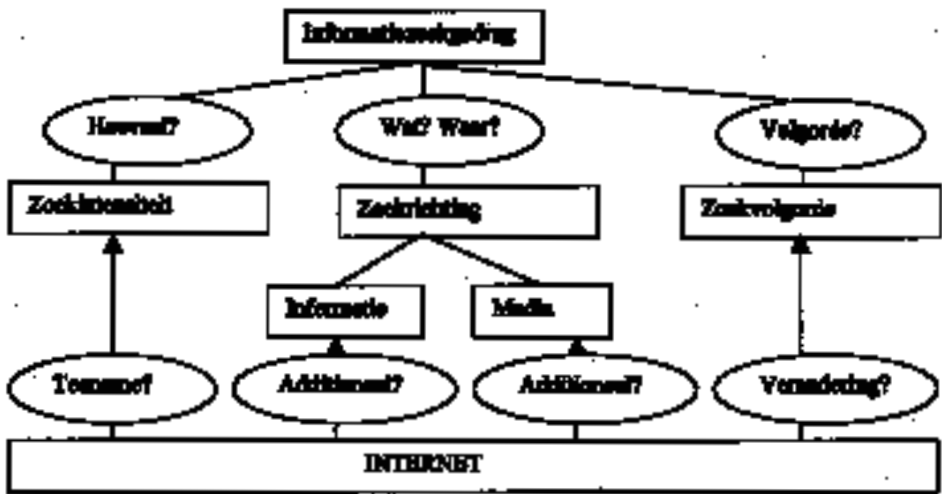
Zoekvolgorde heeft betrekking op beslissingsregels voor de volgorde waarin informatie van informatiebronnen via media wordt verkregen en verwerkt (Bettman, 1979; Van Raaij, 1988). Lazarsfeld, Berelson en Gaudet (1948) hebben de theorie van de 'tweetrapscommunicatie' ontwikkeld. De eerste trap bestaat uit het verkrijgen van informatie via massa media, de tweede trap uit het verkrijgen van informatie door sociale communicatie tussen consumenten onderling. Hierbinnen zijn meerdere varianten te onderkennen (Antonides en Van Raaij, 1998).

Aanbieders met kennis van het extern informatiezoekgedrag van consumenten op de drie genoemde elementen van het zoekproces (zoekintensiteit, zoekrichting, zoekvolgorde) kunnen koopbeslissingen van consumenten beïnvloeden door de juiste informatie op het juiste tijdstip via het juiste medium aan te bieden (figuur 1).

4. MODEL VAN EXTERN INFORMATIEZOEKGEDRAG

4.1. *Kwalitatief consumentenonderzoek*

De inzichten in extern informatiezoekgedrag en in het informatiezoekproces ten behoeve van een koopbeslissing inzake financiële diensten verkregen uit literatuurstudie zijn aangevuld met veldonderzoek. Het veldonderzoek bestaat uit een deskundigen interview en een kwalitatief consumentenonderzoek. De doelstelling van het deskundigen interview is 'kennis en ervaring van diegenen die bekend zijn met het algemene onderwerp boven water te krijgen' (Churchill, 1995). Aan het deskundigen interview hebben vijf deskundigen op het gebied van direct marketing van financiële diensten deelgenomen. Het deskundigen interview tezamen met het literatuuronderzoek heeft geresulteerd in een gesprekspuntenschema voor het kwalitatief consumentenonderzoek. Aan



Figuur 1. Onderzoeksdomein.

dit onderzoek hebben in totaal 40 personen deelgenomen. De respondenten zijn geselecteerd op basis van het criterium dat zij in het afgelopen half jaar een autoverzekering en/of een lijfrentekoopsomverzekering hebben afgesloten en/of hebben deelgenomen aan een beleggingsfonds. De geselecteerde respondenten zijn diegenen die bij deze aankoop het zoekproces daadwerkelijk hebben uitgevoerd. Daarnaast is gestreefd naar een spreiding in socio-demografische kenmerken en naar distributiekanaal (direct writer versus intermediair). Het gesprekspuntenschema bestond globaal uit de volgende onderwerpen:

1. achtergrond van de respondenten;
2. daadwerkelijk extern informatiezoekgedrag:
 - gebruik van informatiebronnen (zoekintensiteit, zoekrichting, zoekvolgorde);
 - gebruik van media (zoekintensiteit, zoekrichting, zoekvolgorde);
 - gehanteerde criteria ten aanzien van informatiebronnen en media;
 - type informatie dat gezocht werd.

Aan de hand van het gesprekspuntenschema is inzicht verkregen in het extern *zoekgedrag* van de respondenten en in de wijze waarop de respondenten het *zoekproces* hebben vormgegeven. In het algemeen kan worden gesteld dat de *zoekintensiteit* bij financiële diensten gering is. Dit heeft te maken met het eerder genoemde ‘satisficing’, maar ook met een algehele geringe betrokkenheid met financiële diensten. De eerder genoemde hoofdindeling van zoekgedrag (doelgericht zoeken versus incidenteel leren) is teruggevonden. Ten aanzien van de *zoekrichting* zijn twee distributiekanaalen relevant, te weten het directe distributiekanaal en het intermediairkanaal. Een deel van de consumenten heeft vaker een voorkeur voor publieke en commerciële informatiebronnen. Een ander deel volgt vaak het advies op van sociale informatiebronnen en zij hebben veelal behoefte aan persoonlijk contact met de financiële dienstverlener om de onzekerheid weg te nemen. Consumenten verschillen ook in hun voorkeur voor media met interne pacing (bijvoorbeeld: krant, brochure) of voor media met externe pacing. Media met interne pacing stellen consumenten in staat informatie in hun eigen tempo

tot zich te nemen, waardoor zij als het ware zelf het medium beheersen in tegenstelling tot media met externe pacing (Antonides en Van Raaij, 1998). Ten aanzien van de *zoekvolgorde* van het extern informatiezoekproces zijn verschillen vastgesteld in volgorde van gebruikte informatiebronnen. Deze verschillen hebben betrekking op de gebruikte categorieën van informatiebronnen zoals beschreven in hoofdstuk 3, maar ook op de gebruikte informatiebronnen binnen een categorie.

Het literatuuronderzoek, de deskundigen interviews en het kwalitatief consumenten onderzoek hebben een set van determinanten voor extern informatiezoekgedrag opgeleverd zoals weergegeven in tabel 1.

Tabel 1. Determinanten van extern informatiezoekgedrag.

Nr	Determinanten	Literatuur onderzoek	Deskundigen interviews	Consumenten interviews
1	Doelgericht zoeken / incidenteel leren	X	X	X
2	Pacing (intern / extern)	X	X	X
3	Actief / passief zoekgedrag	X	X	X
4	Situationele betrokkenheid	X	X	—
5	Beheersing cq controle door de consument	X	X	X
6	Retrieval / delivery media	X	X	X
7	Systematisch / onsystematisch zoekgedrag	—	—	X
8	Niveau van satisficing	X	X	X
9	Gevoelig voor tijdsdruk	X	X	X
10	Gepercipieerd risico	X	X	—
11	Gepercipieerde financiële positie	—	—	X
12	Subjectieve kennis	X	—	—
13	Gepercipieerde ervaring met het product	—	X	X
14	Gepercipieerde kennis van rechten en verplichtingen	—	—	X
15	Gepercipieerde kennis van terminologie	—	—	X
16	Product complexiteit	X	X	X
17	Gepercipieerde productverschillen	X	X	—
18	Gebruik maken van sociale bronnen	X	—	X
19	Behoeft aan persoonlijk contact	X	—	X
20	Interactiviteit (persoons- / machine-)	X	—	X

X : geïdentificeerde determinant

— : niet geïdentificeerde determinant

4.2. Kwantitatief consumentenonderzoek

4.2.1. Inleiding

Het exploratief onderzoek is opgevolgd door een kwantitatief, validerend consumentenonderzoek op basis van de determinanten zoals vermeld in tabel 1. De determinanten situationele betrokkenheid, gepercipieerd risico, subjectieve kennis, en gepercipieerde productverschillen zijn niet meegenomen in het kwantitatief consumentenonderzoek, aangezien deze niet naar voren zijn gekomen uit het kwalitatief consumentenonderzoek. De overige 16 determinanten zijn geoperationaliseerd in de vorm van stellingen, 28 in totaal, ten behoeve van het kwantitatief telefonisch consumenten

onderzoek. De stellingen zijn gemeten met behulp van een Likert schaal, waarbij respondenten wordt gevraagd aan te geven in hoeverre zij het eens of oneens zijn met elke stelling in een serie vragen (Slater en Narver, 1996; Churchill, 1995; Zwart, 1996). In het onderzoek is gebruik gemaakt van een vijf-punts Likert schaal, te weten: helemaal mee eens, mee eens, noch mee eens / noch mee oneens, oneens, helemaal mee oneens.

Een pre-test van de 28 stellingen is uitgevoerd in de vorm van een kwalitatieve toets door de deskundigen die aan het deskundigen interview hebben deelgenomen. Een gestructureerde vragenlijst is ontwikkeld, welke bestaat uit een selectie criterium (ouder dan 15 jaar), de 28 stellingen en socio-demografische variabelen. Het onderzoek is uitgevoerd onder 402 consumenten die in het afgelopen jaar een autoverzekering (50 procent) of een beleggingsfonds (50 procent) hebben afgesloten. Hierbij is een verdeling aangehouden van 50 procent afgesloten via het directe distributiekanaal en 50 procent via het intermediairkanaal. Analyse van de verzamelde data bestond uit 3 stappen, zoals weergegeven in tabel 2.

Tabel 2. Analyse methode.

Stap	Doelstelling	Analyse methode
1	Identificatie en validatie van constructen van informatiezoekgedrag	Principale componentenanalyse Cronbach alpha
2	Identificatie en validatie van informatiezoekgedrag segmenten	Clusteranalyse Discriminantanalyse
3	Beschrijving van informatiezoekgedrag segmenten	Chi-kwadraat analyse t-test

4.2.2. Identificatie en validatie van zoekgedrag constructen

Principale componentanalyse met Varimax rotatie is uitgevoerd om het aantal factoren (constructen) vast te stellen op basis van drie criteria, te weten: het 'latent roots' criterium, de scree test en de factorladingen. De Varimax rotatie is gehanteerd teneinde de verklaring van de factoren te faciliteren, doordat de Varimax rotatie de factorladingen van respondenten op de stellingen zoveel mogelijk richting 0 of 1 stuurt (SAS/STAT, 1990, p.40, Churchill, 1995; Van Raaij, 1977; Zwart, 1996).

Uit de eerste factoranalyse op de 28 statements bleek dat negen factoren een eigenwaarde groter dan 1 hebben, welke gezamenlijk 57.06% van de totale variantie verklaren. De scree plot van de eigenwaarden is gemaakt. Volgens Churchill (1995) is de laatste echte factor het punt in de scree plot waarvoor de 'scree' begint. In deze scree plot is factor 2 de laatste echte factor. Ten aanzien van de factorladingen is vastgesteld dat elke factor door tenminste drie stellingen moet worden verklaard (waarde kleiner dan -0.4 of groter dan 0.4). De factoren 1, 2 en 6 worden door tenminste drie stellingen verklaard. Op basis van consistentie van de uitkomsten van het 'latent roots' criterium (eigenwaarde >1), de scree test, en de factorladingen (verklaard door > 2 stellingen) is het aantal factoren vastgesteld op twee.

Vervolgens is een tweede principale componentanalyse met Varimax rotatie uitgevoerd op de 28 stellingen met een beperking van twee factoren. Hierbij zijn 15 stellingen

geïdentificeerd met hoge factorladingen (kleiner dan -0.4 of groter dan 0.4). Een derde principale componentanalyse op de 15 hoogladende stellingen beperkt tot twee factoren is vervolgens uitgevoerd. Het doel hiervan was het beschrijven van factor 1 en factor 2 aan de hand van de stellingen met hoge factorladingen op deze factoren. Daarnaast zijn hiermee de factorscores voor de respondenten berekend, welke nodig waren voor de volgende analysefase, clusteranalyse. Factor 1 en factor 2 verklaren gezamenlijk 23.92% van de totale variantie.

De interne consistentie van de 2 factoren is vastgesteld met behulp van een betrouwbaarheidsanalyse (Anderson, Gerbing en Hunter, 1987). De betrouwbaarheid van de factoren of schalen heeft betrekking op de afwezigheid van meetfouten in de response patronen van respondenten en kan worden geanalyseerd met behulp van een 'item-to-total' correlatie en met behulp van de Cronbach Alpha coëfficiënt (Cronbach, 1951; Nunnally, 1978). Bij exploratief onderzoek moet de Cronbach Alpha coëfficiënt een waarde hebben van tenminste 0.7 (Nunnally, 1978). Voor de 'item-to-total' correlatie geldt dat stellingen met een waarde kleiner dan 0.25 moeten worden verwijderd (Kerlinger, 1973). Tabel 3 geeft de resultaten van deze analyses weer.

Tabel 3. Interne schaal consistentie.

	Aantal stellingen	Minimum 'item-to-total' correlatie	Cronbach Alpha
Factor 1	8	0.30	0.74
Factor 2	7	0.31	0.72

De schaalitems zijn weergegeven in tabel 4 en tabel 5. Aan de hand van de schaalitems zijn de twee constructen omschreven. Factor 1 heeft betrekking op de mate van gestructureerd informatiezoekgedrag, en factor 2 heeft betrekking op de mate van gepercipieerde zekerheid.

Tabel 4. Schaalitems voor factor 1 – Mate van gestructureerd informatiezoekgedrag.

- S101	Ik lees eigenlijk alleen informatie over < > als ik die toevallig tegenkom.
S102	Als ik informatie zoek over < > zoek ik vergelijkbare aanbiedingen van verschillende maatschappijen.
S104	Ik vraag zelf brochures op over < >.
S106	Ik zoek zo goed mogelijk naar informatie die ik nodig heb over < >
- S107	Ik heb het liefst dat iemand anders informatie over < > voor mij opzoekt en het een en ander uitzoekt.
S108	Ik zorg er zelf voor dat ik de informatie die ik nodig heb over < > ook krijg.
S109	Als ik informatie zoek over < > vraag ik veel informatie op bij verschillende instanties.
S114	Voor het opvragen van informatie over < > trek ik veel tijd uit.

< > : afhankelijk van respondent, betreft het hier autoverzekeringen of beleggingsfondsen.

- : negatieve correlatie met de factor.

Tabel 5. *Schaalitems voor factor 2 – Gepercipieerde zekerheid.*

- S117	Ik heb zelf nog niet zo veel met <> te maken gehad.
- S119	In brochures over <> worden vaak moeilijke woorden gebruikt.
S120	De meeste vaktermen over <> die ik tegenkom ken ik.
- S125	Verzekeraars doen meestal zo ingewikkeld over <> die ze verkopen.
S126	Ik vind <> makkelijk te begrijpen.
S127	Ik heb vaak het gevoel dat aanbieders mij niet alles vertellen over <>.
S128	Ik weet meestal precies wat <> inhouden.

<> : afhankelijk van respondent, betreft het hier autoverzekeringen of beleggingsfondsen.

- : negatieve correlatie met de factor.

4.3.2. *Identificatie en validatie van informatiezoekgedrag segmenten*

De factorscores van de respondenten die zijn gegenereerd uit de derde principale componentanalyse met Varimax rotatie (twee factoren, 15 stellingen) vormden de basis voor het identificeren en valideren van informatiezoekgedrag segmenten. Hiertoe zijn drie stappen uitgevoerd.

1. Clusteranalyse, gericht op het vaststellen van het aantal homogene clusters waarbij respondenten worden toegewezen aan clusters met zo min mogelijk verschillen binnen een cluster en zoveel mogelijk verschillen tussen clusters onderling (Churchill, 1995).
2. Vergelijken van gemiddelde waarden, waarmee wordt getest in hoeverre de stellingen gebruikt voor de clusteranalyse leiden tot significante verschillen tussen de clusters onderling.
3. Validatie van de homogene clusters, door het testen van de clusteranalyse. Hiertoe is een discriminantanalyse uitgevoerd, welke zich richt op het discriminerende karakter van de stellingen voor de identificatie van de clusters.

Ad 1. Clusteranalyse

Ward's Minimum-Variantie clustermethode is gebruikt omdat deze de beste algehele performance heeft (Milligan, 1981). Het is een hiërarchische clustermethode die kleine clusters samenvoegt en streeft naar het produceren van clusters van ongeveer dezelfde grootte met een maximale homogeniteit binnen de clusters (Ward, 1963; SAS, 1990). Bevredigende methoden voor het vaststellen van het aantal clusters ontbreken (Everitt, 1979; Hartigan, 1985; Bock, 1985). Voor hiërarchische clustermethoden bestaan echter drie statistische criteria die zeer goed presteren, te weten het Cubic Cluster Criterion (CCC), Pseudo F en Pseudo t2 (Milligan en Cooper, 1985; SAS, 1990). Consensus tussen deze drie criteria vereist een lokaal maximum van CCC en Pseudo F, dat samenvalt met een lage Pseudo t2 die wordt opgevolgd door een hogere Pseudo t2 bij de volgende clustersamenvoeging (SAS, 1990). Uit tabel 6 blijkt dat op basis van consensus tussen de drie criteria, het aantal clusters kan worden vastgesteld op vier.

Tabel 6. CCC, Pseudo F en Pseudo t^2

Aantal clusters	CCC	Pseudo F	Pseudo t^2
2	-3.79	191.70	211.38
3	-11.13	228.90	126.76
4	-11.06	230.75	69.05
5	-12.35	215.88	91.86
6	-12.70	212.97	70.82

Ad 2. Vergelijken van gemiddelde waarden

Vergelijking van de gemiddelde waarden (gebaseerd op de Likert schaal) van de clusters op de 15 stellingen wijst uit dat bij alle 15 stellingen significante verschillen tussen de vier clusters zijn vastgesteld. Dit is een indicatie voor een goede clusteroplossing (Churchill, 1995).

Ad 3. Validatie van de clusters

Twee doelstellingen liggen ten grondslag aan de uitgevoerde discriminantanalyse:

1. Ontwikkeling van een discriminantfunctie waarmee respondenten worden toegewezen aan één van de vier clusters;
2. Evaluatie van de discriminantfunctie door middel van geschatte foutpercentages, die refereren aan misclassificatie van respondenten;

De discriminantanalyse wijst de 402 respondenten toe aan de vier clusters op basis van hun scores op de 15 stellingen. Het aantal misclassificaties is 64 (15.83 procent). Dit is een bevredigende uitkomst. Derhalve kan geconcludeerd worden dat vier betrouwbare clusters zijn gevonden.

Het afgeleide discriminant criterium uit deze dataset kan worden toegepast op andere datasets. Dit is relevant voor de volgende fase van het onderzoek: de invloed van Internet op het zoekgedrag van de vier verschillende clusters.

4.2.4. Beschrijving van de clusters

De vier clusters worden gevormd op basis van de dimensies ‘mate van gestructureerd informatiezoekgedrag’ en ‘gepercipieerde zekerheid’ zoals weergegeven in tabel 7.

Tabel 7. Informatiezoekgedrag segmenten.

	Gepercipieerde zekerheid	Gepercipieerde onzekerheid
Gestructureerd zoekgedrag	Gefundeerde beslissers	Eeuwig zoekende beslissers
Ongestructureerd zoekgedrag	Ongefundeerde beslissers	Impulsieve beslissers

De clusters zijn getypeerd op basis van de score op de 15 stellingen (t-toets, Chi-kwadraattoets) en de socio-demografische variabelen.

Gefundeerde beslissers

Gefundeerde beslissers zoeken op een gestructureerde wijze naar informatie en voelen zich zeker op het terrein van financiële diensten. Gefundeerde beslissers besteden behoorlijk wat tijd aan het actief verkrijgen van gedetailleerde informatie over financië-

le diensten en raadplegen een verscheidenheid aan informatiebronnen. Zij zorgen ervoor dat ze ook daadwerkelijk de benodigde informatie verkrijgen. Deze groep voelt zich zeker omdat zij ervaring heeft met financiële diensten en kennis heeft van producten en terminologie. Zij vinden financiële diensten dan ook gemakkelijk te begrijpen. Dit segment bestaat voornamelijk uit hoog opgeleide consumenten en mannen. Ruim 50 procent van de respondenten bestaat uit gefundeerde beslissers.

Eeuwig zoekende beslissers

Eeuwig zoekende beslissers zoeken op een gestructureerde wijze naar informatie en voelen zich onzeker op het terrein van financiële diensten. Ook eeuwig zoekende beslissers besteden behoorlijk wat tijd aan het verkrijgen van gedetailleerde informatie door een verscheidenheid aan informatiebronnen te raadplegen. Ook zij zorgen ervoor dat ze daadwerkelijk de door hen gewenste informatie ontvangen. De gevoelens van onzekerheid bij deze groep vloeien voort uit hun geringe ervaring met financiële diensten en hun beperkte kennis van terminologie en productinhoud van financiële diensten. Eeuwig zoekende beslissers hebben het idee dat financiële diensten complex zijn, en dat de financiële dienstverlening het hen niet gemakkelijker maakt. Ondanks het uitgebreide informatiezoekgedrag van deze groep, blijven de eeuwig zoekende beslissers twifelen of ze wel de juiste keuze hebben gemaakt. Persoonlijk contact is belangrijk om de onzekerheid te reduceren. Dit segment bestaat voornamelijk uit lager opgeleide consumenten en vrouwen. In totaal behoort 14% van de respondenten tot het segment 'eeuwig zoekende beslissers'.

Impulsieve beslissers

Impulsieve beslissers worden gekarakteriseerd door ongestructureerd zoekgedrag en een gevoel van onzekerheid met betrekking tot financiële diensten. Het ongestructureerde zoekgedrag van deze groep wordt gekarakteriseerd door in het algemeen uitsluitend informatie te lezen waar zij onverwachts tegenaan lopen. Zij houden er niet van zelf actief informatie te zoeken, en hebben liever dat iemand anders dat voor hen doet. Impulsieve beslissers lijken op eeuwig zoekende beslissers met betrekking tot hun onzekerheid. Ook zij hebben beperkte kennis en ervaring met financiële diensten en hebben het gevoel dat financiële dienstverleners hen niet alles vertellen. Evenals bij de eeuwig zoekende beslissers bestaat dit segment uit meer vrouwen dan mannen en uit meer lager opgeleide consumenten. Ruim 10% van de respondenten behoort tot het segment 'impulsieve beslissers'.

Ongefundeerde beslissers

Ongefundeerde beslissers worden gekarakteriseerd door een beperkt en ongestructureerd informatiezoekgedrag alsmede door een gevoel van zekerheid op het gebied van financiële diensten. Het zoekproces van ongefundeerde beslissers komt overeen met het zoekproces van impulsieve beslissers. Dit wil zeggen dat ongefundeerde beslissers in het algemeen slechts reageren op informatie die zij ongevraagd ontvangen. Door hun kennis van en ervaring met financiële diensten vinden zij deze producten eenvoudig te begrijpen, hetgeen hen een gevoel van zekerheid geeft. Evenals bij de gefundeerde beslissers bestaat dit segment uit relatief veel hoog opgeleide consumenten en mannen. Bijna een kwart van de respondenten behoort tot het segment 'ongefundeerde beslissers'.

Op basis van literatuuronderzoek, kwalitatief onderzoek en kwantitatief onderzoek is in het bovenstaande een model ontwikkeld, gericht op de financiële dienstverlening, voor de karakterisering van consumenten op basis van hun zoekgedrag in termen van zoekintensiteit, zoekrichting en zoekvolgorde. Dit model vormt het uitgangspunt voor het onderzoek naar de invloed van Internet op het zoekgedrag van consumenten bij de aankoop van financiële diensten.

5. INVLOED VAN INTERNET OP INFORMATIEZOEKGEDRAG

5.1. *Onderzoeksmethode*

Om inzicht te verkrijgen in veranderingen die optreden in het extern informatiezoekgedrag van consumenten als gevolg van het gebruik van Internet, zijn twee groepen vergeleken door middel van een database analyse. Dit omvat het analyseren van consumenten gegevens met betrekking tot informatie aanvragen, die zijn vastgelegd in een prospect- en klantendatabase van een direct writing financiële dienstverlener. Hiertoe zijn twee groepen via een random selectieprocedure uit de database geselecteerd. Ten eerste een Internetgroep, oftewel consumenten die ten minste één maal informatie voor financiële diensten bij de betreffende onderneming via Internet heeft aangevraagd. Ten tweede een Controle groep, bestaand uit consumenten die Internet nog nooit hebben gebruikt voor dergelijke doeleinden bij de betreffende financiële instelling.

Deze twee groepen zijn benaderd voor een telefonisch onderzoek. Een vragenlijst is ontwikkeld die betrekking heeft op de twee dimensies van het model van extern informatiezoekgedrag: de mate waarin men gestructureerd dan wel ongestructureerd zoekt, en de mate waarin iemand zich zeker voelt op het betreffende onderwerp. Met behulp van hun antwoorden op de 15 stellingen en de ontwikkelde discriminantfunctie zijn respondenten toegewezen aan een van de vier informatiezoekgedrag segmenten. Daarnaast bevatte de vragenlijst drie vragen omtrent Internetgebruik en socio-demografische variabelen. Het telefonisch onderzoek heeft geresulteerd in twee bestanden, een Internetgroep (n=923) en een controle groep (n=719). Aan deze datafiles zijn gegevens uit de database van de financiële dienstverlener toegevoegd. Deze additionele gegevens hebben betrekking op elke informatieaanvraag ten behoeve van een potentiële aankoop van de respondent bij de financiële dienstverlener. De gegevens die per informatieaanvraag zijn toegevoegd aan de datafile hebben betrekking op zoekintensiteit (aantal informatieaanvragen) en op zoekrichting (type informatie aanvraag en mediumgebruik). Bij het type informatieaanvraag is een onderscheid gemaakt naar brochures en persoonlijke offertes. Mediumgebruik is onderverdeeld naar coupon, telefoon en Internet. Verschillen in zoekintensiteit en zoekrichting tussen de Internetgroep en de controle groep zijn gedefinieerd als het 'Internet effect'.

5.2. *Resultaten*

Het onderzoek heeft een aantal inzichten opgeleverd, welke achtereenvolgens worden besproken.

5.2.1. *Informatiezoekgedrag segmenten en Internet*

Internet wordt gebruikt door alle vier typen informatiezoekers. Echter, niet elke groep

maakt in dezelfde mate gebruik van Internet bij hun zoekproces naar financiële diensten (zie tabel 8). Uit tabel 8 blijkt tevens dat als Internet een rol speelt, het segment gefundeerde beslissers groter is (57%), het aandeel eeuwig zoekende beslissers gelijk blijft (14%), en het aandeel impulsieve beslissers (9%) en ongefundeerde beslissers (20%) kleiner is.

Tabel 8. Verdeling informatiezoekgedrag segmenten.

	N	Gefundeerde beslissers	Eeuwig zoekende beslissers	Impulsieve beslissers	Ongefundeerde beslissers
Internet groep	923	57 % (523)	14 % (127)	9 % (85)	20 % (188)
Controle groep	719	45 % (323)	14 % (104)	15 % (106)	26 % (186)

In het algemeen blijkt dat consumenten die Internet gebruiken in het zoekproces naar financiële diensten ervaren Internetgebruikers zijn. Deze consumenten zijn gemiddeld twee jaar actief op Internet en circa vijf uur per week online.

5.2.2. Internet en zoekintensiteit

Consumenten die Internet gebruiken bij het zoekproces naar financiële diensten, vragen gemiddeld 80 procent meer informatie aan dan consumenten die geen gebruik maken van Internet (tabel 9). Deze toename is het sterkst zichtbaar bij de impulsieve beslissers (114%) en het laagst bij ongefundeerde beslissers (69%). Dit betekent dat Internet wordt gebruikt voor zowel doelgericht zoeken (gefundeerde beslissers) als voor incidenteel leren (impulsieve beslissers).

Tabel 9. Zoekintensiteit: aantal informatie aanvragen.

Informatiezoekgedrag segmenten	Internet groep		Controle groep	
	Gemiddeld	Standaard deviatie	Gemiddeld	Standaard deviatie
Gefundeerde beslissers	3.30	2.58	1.87	1.63
Eeuwig zoekende beslissers	2.93	2.58	1.64	0.98
Impulsieve beslissers	3.34	2.85	1.56	1.04
Ongefundeerde beslissers	3.11	2.52	1.84	1.74
Totaal	3.22	2.57	1.78	1.51

Significant $p < 0.001$

5.2.3. Internet en zoekrichting

De zoekrichting (welke informatie, welke media) van consumenten die gebruik maken van Internet, is anders dan die van de controle groep. Met betrekking tot het type informatieaanvragen blijkt dat persoonlijke informatie in de vorm van offertes het meest voorkomt. Toch worden via Internet meer algemene brochures aangevraagd over financiële diensten dan via andere media (34 procent in plaats van 14 procent), zie tabel 10.

Tabel 10. Zoekrichting: type informatieaanvragen.

	Algemene brochures		Persoonlijke offertes	
	Internetgroep	Controlegroep	Internetgroep	Controlegroep
Totaal	34 %	14%	66%	86%

Significant $p < 0.001$

Internet is een additioneel medium, hetgeen blijkt uit het feit dat toename van de zoekintensiteit zo goed als volledig via Internet wordt gerealiseerd. Consumenten die gebruik maken van Internet geven de voorkeur aan informatieaanvragen via Internet, gevolgd door telefonische aanvragen. Dit gaat ten koste van het aantal couponaanvragen (tabel 11). Snelle responstijden en interactiviteit leiden kennelijk tot een voorkeur voor media met vergelijkbare karakteristieken.

Tabel 11. Zoekrichting: mediumgebruik.

	Internet		Telefoon		Coupon	
	Internet groep	Controle groep	Internet groep	Controle groep	Internet groep	Controle groep
Gefundeerde beslissers	55%	0%	24%	39%	22%	61%
Eeuwig zoekende beslissers	58%	0%	24%	38%	18%	62%
Impulsieve beslissers	54%	0%	24%	35%	21%	64%
Ongefundeerde beslissers	56%	0%	24%	37%	21%	63%
Totaal	55%	0%	24%	38%	21%	62%

Significant $p < 0.001$

6. CONCLUSIE

Het is van belang voor aanbieders te anticiperen op veranderend informatiezoekgedrag van consumenten als gevolg van Internet, met als doel consumenten adequaat van informatie te blijven voorzien en hun koopbeslissing te beïnvloeden. Consumenten die op zoek zijn naar financiële diensten en gebruik maken van Internet, vragen meer informatie op dan van oudsher het geval was. Deze toename aan informatieaanvragen wordt met name gerealiseerd via Internet. Daarnaast neemt het aantal telefonische informatieaanvragen ook toe, echter in mindere mate dan via Internet. Een multi-channel benadering van aanbieders is van strategisch belang. De coördinatie en integratie van klantcommunicatie over alle media heen dient gewaarborgd te zijn.

Alle vier typen informatiezoekers maken gebruik van Internet, maar zij verschillen van elkaar met betrekking tot mediagebruik en informatiebehoefte. Aanbieders kunnen inspelen op de vier afzonderlijke informatiezoekgedrag-segmenten door elk type met de voor hen geschikte media en informatie te bedienen. In het algemeen geldt dat het gericht toepassen van interactieve media (o.a. het World Wide Web) en 'computer-mediated' communicatie (e-mail, e-zines) de focus van aanbieders verdient.

Consumenten die op een gestructureerde manier naar informatie zoeken, hebben

behoefte het zoekproces zelf te beheersen. Hiertoe zijn media met interne pacing waarbij consumenten 'on demand' informatie kunnen opzoeken, zoals de Gouden Gids en het World Wide Web, van belang. Via deze media kan globale en gedetailleerde informatie, op toegankelijke wijze, beschikbaar worden gesteld aan deze gefundeerde beslissers en eeuwig zoekende beslissers. Consumenten die op een ongestructureerde wijze informatie verzamelen, zoals de ongefundeerde beslissers en de impulsieve beslissers, moeten met name worden bediend met media met externe pacing zoals televisie, radio en 'webvertising'. Via deze media kan op initiatief van de aanbieder algemene informatie ongevraagd worden verstrekt aan deze ongestructureerde informatiezoekers.

Daarnaast is het van belang om de interactieve dialoog met consumenten goed in te vullen. Dit is met name relevant voor consumenten die worden gekarakteriseerd door onzekerheid (eeuwig zoekende beslissers en impulsieve beslissers). Hiertoe kunnen naast de traditionele interpersoonlijke communicatievormen zoals telefoon en persoonlijk gesprek, nieuwe vormen van 'one-to-one' communicatie worden ingezet. Voorbeelden zijn: e-mail, chat, en virtuele adviseurs en 'live agents'. Deze inzichten gecombineerd met een gesignaleerde toename van de zoekintensiteit naar financiële diensten met gemiddeld 80 procent zal een grote druk leggen op het aanpassen van de interne organisatie, vaardigheden, en activiteiten van aanbieders (Day, 1998). Geconcludeerd kan worden dat Internet een belangrijk additioneel informatiekanaal is. Een adequate inzet van Internet binnen de gehele media mix is dan ook van strategisch belang voor het realiseren van structurele directe relaties met klanten.

7. AANBEVELINGEN VOOR VERDER ONDERZOEK

De uitgevoerde studie is een eerste stap in het verkrijgen van inzicht in de impact van Internet op het koopproces van consumenten. Uit deze studie blijkt dat Internet leidt tot wijzigingen in zoekintensiteit en zoekrichting van consumenten bij de aankoop van financiële diensten. Het onderzoek is uitgevoerd terwijl de Internetpenetratie nog relatief beperkt was en het gebruikersprofiel nauwe verwantschappen vertoonde met innovators en early adopters. De penetratie van Internet neemt in hoog tempo toe, waardoor het gebruik van Internet door een bredere groep mensen wordt geadopteerd, en de early majority inmiddels op Internet aanwezig is. De bredere gebruikersgroep van Internet kan ertoe leiden dat de gevonden verdeling over de informatiezoekgedrag segmenten aan verandering onderhevig is. Het is dan ook van belang om longitudinaal onderzoek op te starten naar het zoekgedrag van consumenten bij de aankoop van financiële diensten. Hierbij zou de focus moeten liggen op de verdeling van de typen informatiezoekgedrag segmenten enerzijds en de zoekintensiteit en zoekrichting anderzijds.

Daarnaast vormen de inzichten in de vier informatiezoekgedrag segmenten een goed aanknopingspunt voor het onderzoeken van de wijze waarop aanbieders kunnen inspelen op de onderling verschillende behoeften van deze informatiezoekgedrag segmenten door de marketingcommunicatie hierop af te stemmen. Het uitgevoerde onderzoek heeft het gedrag van consumenten inzichtelijk gemaakt, het toepassen van deze kennis is nog een braakliggend terrein.

Het onderzoek naar extern informatiezoekgedrag van consumenten is uitgevoerd voor

financiële diensten. Toekomstig onderzoek naar een toepasbaarheid van dit model bij andere bedrijfstakken dient plaats te vinden. Op basis hiervan kan inzichtelijk worden in hoeverre het model generaliseerbaar is met alle marketingimplicaties van dien.

Het onderzoek is uitgevoerd op de database van slechts één financiële dienstverlener. Hierdoor is het derde element van informatiezoekgedrag – zoekvolgorde – niet meegenomen. Een onderzoeksmodel waarbij over meerdere aanbieders heen de volgorde van het zoekproces in kaart wordt gebracht is van groot belang. Internet biedt de nodige mogelijkheden hiertoe. Zo kan gedacht worden aan een E-panel waarbij het totale surf- en interactiegedrag wordt vastgelegd en geanalyseerd.

Kortom, een aanzet voor het verkrijgen van inzicht in de invloed van Internet op consumentengedrag is gegeven. De resultaten geven een indicatie van wijzigingen in consumentengedrag waar aanbieders op moeten anticiperen teneinde de koopbeslissing te kunnen blijven beïnvloeden. Uit het onderzoek blijkt naast een grote voorkeur voor het medium Internet, een toename van het telefoongebruik. De hoge penetratie van mobiele telefoons, de integratie van telefoon en Internet zullen naar verwachting deze mediumvoorkeur versterken. De opkomst van nieuwe media welke plaatsafhankelijk zijn, kunnen veranderingen in consumentengedrag een extra impuls geven. Een zeer breed onderzoeksdomein ligt open.

LITERATUUR

- Anderson, J.C., D.W. Gerbing, and J.E. Hunter. 1987. "On the Assessment of Unidimensional Measurement: Internal and External Consistency, and Overall Consistency Criteria". *Journal of Marketing Research* 24 (November): 432-437.
- Antonides, G, and W.F. van Raaij. 1998. *Consumer Behaviour. A European Perspective*. Chichester: John Wiley.
- Beatty, S.E. and S.M. Smith. 1987. "External Search Effort: an Investigation Across Several Product Categories". *Journal of Consumer Research* 14 (June): 83-95.
- Bettman, J.R. 1979. *An Information Processing Theory of Consumer Choice*. Menlo Park, California: Addison-Wesley Publishing Company.
- Bock, H.H. 1985. "On Some Significance Tests in Cluster Analysis". *Journal of Classification* 2: 77-108.
- Churchill, G.A. Jr. 1987. *Marketing Research. Methodological Foundations*, 4th edition, Fort Worth: the Dryden Press.
- Churchill, G.A. Jr. 1995. *Marketing Research. Methodological Foundations*, 6th edition, Fort Worth: the Dryden Press.
- Cronbach, L.J. 1951. "Coefficient Alpha and the Internal Structure of Tests" *Psychometrika* 16 (11): 2-49.
- Day, G.S. 1998. "Organizing for interactivity". *Journal of Interactive Marketing* 12 (1) (Winter): 47-53.
- Engel, J.F., R.D. Blackwell, and P.W. Miniard. 1990. *Consumer Behaviour*. Chicago: Dryden Press.
- Everitt, B.S. 1979. "Unresolved Problems in Cluster Analysis". *Biometrics* 35: 169-181.
- Fain, D. and M.L. Roberts. 1997. "Technology vs. Consumer Behaviour: the Battle for the Financial Services Customer". *Journal of Direct Marketing* (Winter) 11 (1): 44-54.
- Grönroos, C. 1994. "From Marketing Mix to Relationship Marketing: Towards a Paradigm Shift in Marketing". *Management Decision* 32 (2): 4-20.

- Hartigan, J.A. 1985. "Statistical Theory in Clustering". *Journal of Classification* 2: 63-76.
- Hoffman, D.L. and T.P. Novak. 1996. "Marketing in Hypermedia, Computer-Mediated Environments: Conceptual Foundations". *Journal of Marketing* 60 (July): 50-68.
- Kamakura, W.A., S.N. Ramaswani, and R.K. Srivastava. 1991. "Applying Latent Trait Analysis in the Evaluation of Prospects for Cross-selling of Financial Services". *International Journal of Research in Marketing* 8 : 329-349.
- Kerlinger, F.N. 1973. *Foundations of Behavioral Research*. New York: Holt, Rinehart and Winston Inc.
- Kotler, P. 1994. *Marketing Management: Analysis, Planning and Control*. Englewood Cliffs: Prentice Hall.
- Kotler, P. and G. Armstrong. 1998. *Principles of Marketing*. 8th edition. New Jersey: Prentice Hall.
- Kumar, A. and W.R. Dillon. 1987a. "The Interaction of Measurement and Structure in Simultaneous Equation Models with Unobservable Variables". *Journal of Marketing Research* 24 (February): 98-105.
- Kumar, A. and W.R. Dillon. 1987b. "Some Further Comments on Measurement/Structure Interaction and the Unidimensionality of Constructs". *Journal of Marketing Research* 24 (November): 438-444.
- Lazarsfeld, P.F., B.R. Berelson, and H. Gaudet. 1948. *The People's Choice*. New York: Columbia University Press.
- Malhotra, N.K. 1996. *Marketing Research, An Applied Orientation*. 2nd. Edition. New Jersey: Prentice Hall.
- Maute, M.F. and W.R. Forrester Jr. 1991. "The Effects of Attribute Qualities on Consumer Decision Making". *Journal of Economic Psychology* (December): 643-666.
- Milligan, G.W. 1981. "A Review of Monte Carlo Tests of Cluster Analysis". *Multivariate behavioral research* 16: 379-407.
- Milligan, G.W. and M.C. Cooper. 1985. "An Examination of Procedures for Determining the Number of Clusters in a Data Set". *Psychometrika* 50: 159-179.
- Nunnally, J.C. 1967. *Psychometric Theory*. New York: McGraw-Hill.
- Nunnally, J.C. 1978. *Psychometric Theory* 2nd edition. New York: McGraw-Hill.
- Peter, J.P. and J.C. Olson. 1990. *Consumer Behaviour and Marketing Strategy*. Boston: Irwin.
- Plasmeijer, P.W.J. 1999. The Influence of the Internet on Prepurchase External Search for Financial Services. *Academisch proefschrift*. Erasmus Universiteit Rotterdam. 156 p.
- Reichheld, F.F. 1996. *The Loyalty Effect. The Hidden Force behind Growth, Profits and Lasting Value*. Boston, Massachusetts: Harvard Business School Press.
- Sarle, W.S. 1983. "Cubic Clustering Criterion". *SAS Technical Report A-108*. Cary NC: SAS Institute Inc.
- SAS Institute Inc. 1990. *SAS/STAT User's Guide* version 6. 4th edition. Volume 1 and 2. Cary NC: SAS Institute Inc..
- Schmidt, J.B. and R.A. Spreng. 1996. "A Proposed Model of External Consumer Information Search". *Journal of the Academy of Marketing Science* 24 (3): 246-456.
- Siegel, Sidney. 1956. *Nonparametric Statistics for the Behavioral Sciences*. New York: McGraw-Hill Book Company Inc
- Sharma, S. 1996. *Applied Multivariate Techniques*. New York: John Wiley & Sons Inc.
- Simon, H.A. 1976. "Rational Choice and the Structure of the Environment". In *Method and Appraisal in Economics*. Ed. S.J. Latsis. Cambridge: Cambridge University Press, 129-148.
- Slater, S.F. and J.C. Narver. 1996. "Competitive Strategy in the Market Focused Business". *Journal of Marketing* 59 (3): 339-358.
- Van Raaij, W.F. 1977. *The Use of Principal Components and Factor Analysis*, Dept. Economic Psychology. Tilburg University (December)
- Van Raaij, W.F. 1988. "Information Processing and Decision Making Cognitive Aspects of Economic Behaviour". In *Handbook of Economic Psychology*. Van Raaij, W.F., van Veldhoven,

- G.M., and Wärneryd, K.E. (eds.). Dordrecht: Kluwer Academic Publishers, 74-10.
- Ward, J.H. 1963. "Hierarchical Grouping to Optimize an Objective function". *Journal of the American Statistical Association* 58: 236-244.
- Wilkie, W.L. and P.R. Dickson. 1985. *Shopping for Appliances: Consumers' Strategies and Patterns of Information Search*. Cambridge, MA: Marketing Science Institute.
- Williams, F., R.E. Rice, and E.M. Rogers. 1998. *Research Methods and the New Media*. New York: Free Press.
- Zwart, P.S. 1996. *Methoden van marktonderzoek*. Amsterdam/Brussel: Elsevier.

4. De invloed van tevredenheid op cross-buying

Koppeling vragenlijst data met de klantendatabase¹

P.C. VERHOEF, P.H. FRANSES en J.C. HOEKSTRA

SAMENVATTING

Bedrijven hebben tegenwoordig steeds vaker de beschikking over grote databases met klantgegevens. Daarnaast wordt er door veel bedrijven (continu) onderzoek gedaan naar de klanttevredenheid. In dit artikel ontwikkelen de auteurs een model waarin de klanttevredenheidsgegevens afkomstig uit een enquête gerelateerd worden aan het koopgedrag ('cross-buying') van individuele klanten verkregen uit de klantendatabase van een Nederlandse verzekeraar. Ze tonen aan, dat een relatief groot positief verschil tussen de tevredenheid over de onderneming en de tevredenheid over de belangrijkste concurrent een sterke positieve invloed heeft op cross-buying bij de verzekeraar. Het ontwikkelde model kan gebruikt worden om in te schatten of investeringen in tevredenheid bijdragen tot een hogere klantwaarde en dus een hogere winstgevendheid.

1. INLEIDING

Tevredenheid is een belangrijk begrip geworden binnen marketing en marktonderzoek (De Ruyter, 2000). Er wordt er hierbij van uitgegaan dat tevredenheid de loyaliteit van klanten positief beïnvloedt (Reichheld, 1996). Gegevens over de tevredenheid zijn dan ook belangrijke indicatoren voor de toekomstige loyaliteit van klanten. Als gevolg hiervan doen steeds meer bedrijven (continu) onderzoek naar de tevredenheid van hun klanten. Dit is vooral het geval in de dienstensector, waarbij prestaties op zogenaamde 'moments of truth' een zeer sterke invloed hebben op de loyaliteit van klanten (Zeithaml en Bitner, 1996). Opvallend is echter, dat er weinig empirisch bewijs is voor de relatie tussen tevredenheid en loyaliteit in termen van gedrag.

Het toegenomen belang van tevredenheid past binnen de trend binnen marketing, dat bedrijven steeds meer lange termijn relaties trachten te ontwikkelen met klanten. Bij de ontwikkeling van deze relaties is tevredenheid van groot belang. Daarnaast wordt er bij de ontwikkeling van lange termijn relaties steeds vaker gebruik gemaakt van klantendatabases, waarbij op basis van vooral voorgaand koopgedrag en socio-demografische gegevens, getracht wordt voor elke individuele klant een bij hem of haar passend individueel aanbod te doen (Blattberg et al., 1995; Hoekstra et al., 1999). Denk hierbij

¹ Dit artikel is gebaseerd op Verhoef P.C., Franses P.H. en Hoekstra J.C. (1999), The Impact of Satisfaction on the Breadth of a Relationship with a Multi-Service Provider, Ribes, Report 9955.

onder meer aan de recente introductie van klantenkaarten in de levensmiddelen-sector door bijvoorbeeld Albert Heijn en Edah. Gegevens uit de klanten databases worden veelal gebruikt voor de selectie van kansrijke adressen bij direct marketing acties (zie o.a. Bult, 1993; Van der Scheer, 1997). Deze data hebben echter ook andere toepassingsmogelijkheden (Glazer, 1991). Zo voorspellen Leeflang en Wittink (2000) in hun bijdrage aan het millennium NVMI Jaarboek, dat er in de toekomst modellen gebouwd gaan worden, die tevredenheid-, intentie- en preferentiedata zullen relateren aan daadwerkelijk gedrag van klanten verkregen uit een klantendatabase. Met deze modellen kan de marktonderzoeker inzicht krijgen in de daadwerkelijke determinanten van het koopgedrag en in de vraag of investeringen in bijvoorbeeld tevredenheid daadwerkelijk leiden tot loyalere klanten. In de internationale marketingliteratuur zijn hier al een aantal voorbeelden van bekend (Bolton, 1998; Bolton en Lemon, 1999; Bolton et al., 2000).

In deze bijdrage zullen wij zo'n model rapporteren, waarbij tevredenheidsdata van klanten van een verzekeraar worden gerelateerd aan het koopgedrag van deze klanten bij deze verzekeraar. Met dit model willen we ten eerste laten zien, dat bedrijven door het koppelen van verschillende databronnen, in dit geval een klantendatabase en een enquête, inzicht kunnen krijgen in de determinanten van het koopgedrag van hun klanten over de tijd. Ten tweede willen we met deze bijdrage inzicht verschaffen in de relatie tussen tevredenheid en gedragsloyaliteit.

Dit artikel is als volgt opgebouwd. In paragraaf 2 behandelen we kort de relatie tussen tevredenheid en loyaliteit. Vervolgens beschrijven we in paragraaf 3 het model, terwijl in paragraaf 4 de onderzoeksmethode en een beschrijving van de data gegeven wordt. In paragraaf 5 bespreken we de belangrijkste resultaten. We eindigen dit artikel met een conclusie, implicaties voor zowel het marktonderzoek & informatiemanagement als het management van klantrelaties en aanbevelingen voor toekomstig onderzoek.

2. RELATIE TEVREDENHEID EN LOYALITEIT

In de marketing literatuur is veel aandacht geweest voor de vraag wat tevreden klanten kunnen betekenen voor een onderneming. Bekend zijn de artikelen en boeken van Reichheld en zijn mede auteurs (Reichheld, 1996; Reichheld en Sasser, 1990). Zij stellen nadrukkelijk, dat tevredenheid leidt tot een hogere loyaliteit en dus winstgevendheid. Anderson et al. (1994) tonen inderdaad op geaggregeerd niveau een positief verband aan tussen de gemiddelde tevredenheid van klanten van een bedrijf en de winstgevendheid. Op individueel klantniveau zijn er verschillende onderzoeken die aantonen, dat tevreden klanten loyaler zeggen te zijn (Anderson en Sullivan, 1993; Bloemer, 1994). Veelal wordt op individueel klantniveau in dit empirisch onderzoek de tevredenheid van klanten gekoppeld aan de koopintentie van klanten (Anderson and Sullivan, 1993; Bloemer, 1994; Boulding et al., 1993). Hierbij kunnen echter twee belangrijke kanttekeningen geplaatst worden. Ten eerste is het algemeen bekend dat de voorspellende waarde van koopintentie nogal beperkt kan zijn (Morwitz et al., 1998). Immers, er is een verschil tussen zeggen en uiteindelijk doen. Ten tweede kan er in vragenlijsten door zogenaamde 'carry-over' of 'back-fire' effecten een overschatting van de gevonden samenhangen ontstaan (Bickart, 1993). Kortweg houden deze effecten in

dat antwoorden op een vraag in een vragenlijst consistent met eerdere vragen in de vragenlijst worden beantwoord.

Recent is echter een aantal studies verschenen, die op individueel klantniveau dynamische modellen schatten, waarin een verband wordt gelegd tussen tevredenheid en het daadwerkelijke koopgedrag bij een aanbieder. Zo stelt Bolton (1998) een positief verband vast tussen tevredenheid en relatieduur, terwijl Bolton en Lemon (1999) de relatie tussen tevredenheid en het gebruik van een service bestuderen. Tenslotte bestuderen Bolton et al. (2000) zowel de relatie tussen tevredenheid, relatieduur en het gebruik van een service. Doordat deze studies allemaal een positief effect van tevredenheid op gedragsloyaliteit aantonen, gaan zij in tegen de vanuit de praktijk gerezen twijfels over de relatie tussen tevredenheid en loyaliteit (Gale, 1997).

Naast de bestudeerde dimensies van gedragsloyaliteit (relatieduur en het gebruik (consumptieniveau) van een service of product), is er echter nog een andere dimensie waarop tevredenheid invloed heeft: 'cross-buying', ofwel het kopen van verschillende diensten of producten bij dezelfde aanbieder. Cross-buying kan verstrekkende consequenties hebben voor de totale financiële waarde van een klant. Cross-buying is van belang bij bedrijven die meerdere producten of diensten aanbieden. Denk hierbij bijvoorbeeld aan een verzekeringsmaatschappij, die naast autoverzekeringen ook ziektekosten –en inboedelverzekeringen aanbiedt. Maar ook voor telecommunicatie-bedrijven is cross-buying van belang; zo zal men tegenwoordig naast een vast abonnement ook trachten een mobiele aansluiting aan klanten te verkopen. Het is dan ook niet voor niets, dat veel bedrijven instrumenten, zoals loyaliteitsprogramma's en direct mailings inzetten om dit gedrag te stimuleren. Bekend zijn bijvoorbeeld de relatievoordeelprogramma's in de verzekeringsmarkt, waarbij bij aankoop van meerdere verzekeringen pakketkorting wordt gegeven. In de volgende paragraaf leiden we het model af, waarin we tevredenheid relateren aan cross-buying, waarbij we ook controleren voor de invloed op het koopgedrag van een relatieprogramma, reeds afgesloten verzekeringen en eigenschappen van klanten, zoals leeftijd en huishoudgrootte.

3. HET MODEL

3.1. Theoretische achtergrond

Zoals beschreven bestuderen we in het model de cross-buying van klanten. Deze cross-buying operationaliseren we door te kijken naar het verschil in het aantal afgenomen verschillende diensten (Serv) tussen twee tijdsperioden. De afhankelijke variabele in ons model in periode T_1 wordt dus als volgt gedefinieerd:

$$\Delta \text{Serv}_i = \text{Serv}_i - \text{Serv}_0 \quad (1)$$

De onderliggende gedachte van het model is dat bij de keuze om minder, evenveel of meer diensten af te nemen, de klant van de onderneming streeft naar een zo hoog mogelijke waarde of nut, die hij kan ontlenen aan de afgenomen diensten (Oliver en Winer, 1987; Bolton, 1998). Het vaststellen van deze waarde is echter lastig en zal voornamelijk gebaseerd zijn op eerdere ervaringen met de aanbieder. We veronderstel-

len daarom, dat deze waarde afhangt van de huidige tevredenheid. Daarnaast hebben klanten vaak ervaring met meerdere aanbieders. We veronderstellen dan ook dat de klant ook deze ervaringen zal meenemen in zijn beslissing wel of niet meer diensten af te nemen. Tenslotte zullen we in ons model ook rekening houden met andere belangrijke variabelen, zoals koopgedrag uit het verleden, ingezette marketinginstrumenten en socio-demografische kenmerken van de klant (Engel et al., 1996; Kamakura et al., 1991; Winer, 1985).

3.2. Model Specificatie

De specificatie van het model is vooral gebaseerd op literatuur op het gebied van klanttevredenheid. In de onderstaande toelichting gaan we vooral in op de invloed van tevredenheid, het verschil in tevredenheid met de concurrent en de invloed van relatieduur.

3.2.1 Tevredenheid

Zoals reeds eerder beschreven is er in de literatuur bewijs, dat tevredenheid een positief effect heeft op koopintentie, relatieduur of het gebruik van een service (Bolton, 1998; Bolton en Lemon, 1999; Zeithaml et al., 1996). In ons model veronderstellen we dan ook dat een hogere tevredenheid met de onderneming (OTevr) leidt tot een hogere kans op cross-buying bij deze onderneming.

3.2.2. Vergelijking met concurrent

Klanten nemen hun producten of diensten vaak van verschillende aanbieders af. In de literatuur wordt dan ook gesteld, dat klanten de prestaties van verschillende aanbieders vergelijken (Inman et al., 1997). Bolton et al. (2000) tonen aan dat een verschil in prestatie tussen de twee bedrijven invloed heeft op zowel het gebruik van een dienst als het doen van herhalingsaankopen. Het beslissingsproces bij cross-buying verschilt van dat van bijvoorbeeld het beslissingsproces bij herhalingsaankopen, omdat het bij cross-buying verschillende typen producten, met verschillende (nieuwe (onbekende)) karakteristieken betreft. Men mag ervan uitgaan dat hierdoor het beslissingsproces bij cross-buying uitgebreider zal zijn dan bij herhalingsaankopen, met als gevolg dat de invloed van concurrentie belangrijker wordt. Indien er dus bij herhalingsaankopen een invloed van het verschil in prestatie met de concurrent wordt gevonden, des te meer zal er ook een effect gevonden moeten worden bij cross-buying. In ons model nemen we het verschil in tevredenheidsscores mee tussen de onderneming en een concurrent van de onderneming ($OTevr - CTevr$). Hierbij veronderstellen we bovendien, dat in overeenstemming met de prospect theorie, een positief verschil in tevredenheid een ander effect heeft, dan een negatief verschil in tevredenheid. Daarnaast veronderstellen we dat er pas een effect zal optreden indien het absolute verschil ($\tau_{p/n}$) in tevredenheidsscores tussen de onderneming en de concurrent groot genoeg is (Oliver, 1996; Zeithaml et al., 1993).

3.2.3 Invloed relatieduur

Relatieduur heeft in ons model twee effecten. Ten eerste veronderstellen we dat klanten met meer ervaring hun tevredenheid sterker laten wegen, omdat zij meer waarde hechten aan hun eigen opinies (Bolton, 1998). Deze klanten hebben immers meer meegeemaakt met de onderneming en weten beter of de prestaties van een onderneming goed zijn. In ons model nemen we als gevolg hiervan een interactieterm van relatieduur en

tevredenheid met de onderneming op. Ten tweede suggereert de relatiemarketing-literatuur, dat een relatie zich pas ontwikkelt na een verkenningsperiode (Dwyer et al., 1987). Na deze ontwikkelingsperiode wordt de relatie stabiel. Op basis van deze kennis mogen we dus een positief effect van relatieduur verwachten. Het gegeven dat er een verkenningsperiode is in een relatie, zou er op kunnen wijzen dat we zelfs te maken hebben met een niet-lineair verband. In ons model zullen we daarom zowel een lineaire invloed van relatieduur (Relduur), als een kwadratische invloed van relatieduur opnemen.

3.3 Model

Op basis van bovenstaande specificatie kan uiteindelijk het volgende mathematische model gespecificeerd worden:

$$\begin{aligned}
 Dserv_{i,1-0} = & \beta_1 + \beta_2 OTevr_{i,0} + \beta_3 (OTevr_{i,0} - CTevr_{i,0}) (als\ OTevr_{i,0} - CTevr_{i,0} \geq \tau_{p,0}) + \\
 & \beta_4 (lOTevr_{i,0} + CTevr_{i,0}) (als\ OTevr_{i,0} - CTevr_{i,0} \leq \tau_{n,0}) \quad (2) \\
 & + \beta_5 Relduur_{i,0} * OTevr_{i,0} + \sum_{j=1}^2 \beta_{5+j} (Relduur_{i,0})^j + \beta_n X_{i,0} + \varepsilon_{i,0},
 \end{aligned}$$

Hierbij staat X voor variabelen zoals marketinginstrumenten en socio-demografische klantkenmerken. In de volgende paragraaf gaan we in op het testen van het model.

4. METHODOLOGIE

4.1. Context

In dit artikel bestuderen we het gedrag van klanten van een Nederlandse verzekeringsmaatschappij, die gebruik maakt van het directe kanaal (direct writer). In de verzekeringsmarkt trachten bedrijven lange-termijn relaties aan te gaan met hun klanten. Klanten worden dan ook als een belangrijk ‘bezit’ beschouwd. De verzekeringsmarkt is bovendien zeer geschikt, omdat verzekeringsmaatschappijen verschillende typen verzekeringen aanbieden en er dus sprake kan zijn van cross-buying. Daarnaast zijn verzekeraars, en in het bijzonder die verzekeraars die gebruik maken van het directe kanaal, aantrekkelijk, omdat deze vaak al gedurende langere tijd gebruik maken van grote klantendatabases in hun marketingstrategie (Spring et al., 2000).

4.2. Database

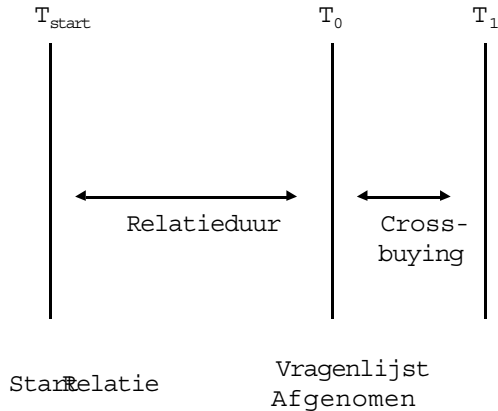
Zoals al eerder beschreven is een belangrijk vernieuwend element in dit artikel, dat we enquêtedata combineren met data uit de klantendatabase. De vragenlijstdata betreffen een onderzoek naar de tevredenheid van klanten die recent een schade afhandeling hebben gehad. In de verzekeringsmarkt kunnen dit soort contacten als ‘moments of truth’ worden bestempeld.

In deze totale (onderzoeks-)database beschouwen we twee tijdsperiodes. De eerste periode start bij het begin van de relatie met de klant (T_{start}). Deze start is voor iedere klant verschillend. Deze periode eindigt op het tijdstip waarop de vragenlijst wordt

afgenomen (T_0). Deze tijdsperiode is dus gelijk aan de relatieduur van de klant met het bedrijf. Op T_0 start de tweede periode. Deze periode eindigt na een jaar op T_1 . We hebben voor een periode van een jaar gekozen, omdat er bij verzekeringen vaak sprake is van jaarlijkse contracten of betaalmomenten. Een schematische weergave van de tijdsperiodes staat in figuur 1.

4.3. De vragenlijst

De vragenlijst is telefonisch afgenomen in november 1995 onder een gestratificeerde steekproef van 435 consumenten. Uiteindelijk hebben 309 respondenten meegewerkt aan het onderzoek (response 71%). Omdat één van de centrale variabelen in het model de vergelijking met de concurrent is, hebben we alleen de respondenten meegenomen die ervaring(en) hebben met de dienstverlening van een concurrent en de dienstverlening van deze concurrent ook kunnen beoordelen. Hierdoor blijven er 204 respondenten over.



Figuur 1. Toelichting meetprocedure.

4.4. Meting variabelen

De meting van de belangrijkste variabelen in ons model is weergegeven in tabel 1. Informatie over het aantal afgenomen verzekeringen komt uit de klantendatabase. De tevredenheidsscores zijn afkomstig uit de vragenlijst. Tevredenheid met het bedrijf is gemeten op een 10-punts-schaal (1= zeer ontevreden, 10 = zeer tevreden) met 1 item (Anderson en Sullivan, 1993). Tevredenheid met de belangrijkste concurrent is gemeten, door eerst aan respondenten te vragen of ze ervaring met een andere verzekeraar hebben. Vervolgens moesten ze aangeven welke verzekeraar dat is. Tenslotte is de tevredenheid over deze concurrent wederom gemeten met een 10-punts-schaal (Ping, 1993). De relatieduur met het bedrijf is gemeten als de tijdsperiode tussen de start van de relatie en het tijdstip waarop de vragenlijst is afgenomen. Zoals beschreven nemen we ook een aantal socio-demografische variabelen mee in het model, te weten deelname aan een loyaliteitsprogramma, het type reeds afgesloten verzekeringen en het totaal aantal afgesloten polissen. De opgenomen socio-demografische variabelen betreffen leeftijd, huishoudgrootte en geslacht. Deze variabelen zijn afkomstig uit de vragenlijst.

Het loyaliteitsprogramma gaf in 1995 klanten korting op premies, afhankelijk van het aantal afgesloten polissen en de lengte van de relatie. Het programma belooft dus loyaal gedrag.

Tabel 1. Meting variabelen in model.

Afhankelijke variabele	Meting
Cross-buying	Verschil in aantal afgesloten polissen tussen T_1 en T_0 .
Verklarende variabelen	Meting
Tevredenheid met onderneming	Hoe tevreden bent u in het algemeen over de dienstverlening van de verzekeringsmaatschappij? (1 = zeer ontevreden, 10 = zeer tevreden)
Tevredenheid concurrent	Hoe tevreden bent u in het algemeen over de dienstverlening van concurrent X? (1= zeer ontevreden, 10=zeer tevreden)
Relatieduur	Aantal jaren sinds de relatie begon.
Overige variabelen	Meting
Leeftijd	Wat is uw leeftijd?
Huishoudgrootte	Aantal personen in huishouden
Geslacht	0 = man, 1 = vrouw
Type polis	Aantal polissen per type polis
Loyaliteitsprogramma	Deelname aan loyaliteitsprogramma (0 = geen deelname, 1 wel deelname)

4.5. Analyse methode

De afhankelijke variabele in ons model (de verandering in het aantal afgenomen – verschillende – verzekeringen) is in wezen ordinaal geschaald, omdat de afgesloten verzekeringen verschillend van karakter zijn. Daarnaast werken we met een aantal categorieën: een verschil in aantal afgenomen polissen dat kleiner is dan of gelijk aan -1 ; een verschil van 0, een verschil van 1 en een verschil dat groter is dan of gelijk aan 2. Aangezien de afhankelijke variabele ordinaal geschaald is, mag statistisch gezien het standaard regressie model (kleinste kwadraten) niet gebruikt worden, maar moet er gebruik worden gemaakt van een geordend regressie model (Long, 1997). Dit model kan vergeleken worden met de bekendere logit en probit modellen. Evenals in deze modellen wordt in een geordend regressie model het geobserveerde gedrag (cross-buying) gemodelleerd met behulp van een latente variabele, die lineair afhangt van de onafhankelijke variabelen². Bij de schatting van het model nemen we van de controle variabelen (leeftijd, geslacht etc.) alleen die variabelen mee die een redelijk significantieniveau hebben (p-waarde <0.20).

Een belangrijk onderdeel in onze analyse is het vaststellen van de waarde waarbij het verschil tussen de tevredenheid over de onderneming en de tevredenheid over de concurrent invloed heeft op cross-buying. Hiervoor hebben wij de volgende procedure

² Voor een gedetailleerde beschrijving van deze techniek verwijzen wij naar Long (1997) hoofdstuk 5.

gevolgd. We hebben het model voor elk van de gemeten verschilwaarden geschat en vervolgens is de combinatie met de laagste waarde voor het Akaike Informatie Criterium (AIC) gekozen. Hierdoor wordt die verschilwaarde gevonden, waarbij het model, gegeven het aantal variabelen in het model, de hoogste fit heeft (Pindyck en Rubinfeld, 1998). Na vergelijking van de AIC-waarden, blijkt dat de AIC-waarde minimaal is bij een positief van verschil van 4 of meer, terwijl we geen eenduidige waarde voor een negatief verschil vinden. We hebben daarom besloten om deze waarde op -1 te zetten.

5. RESULTATEN

In deze paragraaf beschrijven we de resultaten van ons model. We starten echter met een korte beschrijving van de variabelen in de database.

5.1. Beschrijving variabelen

Wat betreft het aantal afgesloten polissen behoudt 74% van de klanten hetzelfde aantal in de door ons onderzochte tijdsperiode, terwijl in deze periode bij 17.7 % van de klanten het aantal afgesloten polissen toeneemt. Dit houdt dus in, dat bij 8.3% van de klanten het aantal afgesloten polissen vermindert in het bestudeerde jaar. De gemiddelde relatieduur is ruim 10 jaar. De klanten zijn gemiddeld zeer tevreden over de verzekeraar. De score voor de concurrent ligt beduidend lager. Procentueel is 41.6% van de respondenten meer tevreden over deze verzekeraar dan over de concurrent.

5.2. Resultaten model

In tabel 2 staan de schattingsresultaten van het model. Het model is zeer significant blijkende uit de Likelihood Ratio statistiek van 48.27 ($p=0.00$). Het model verklaart ongeveer 14% van de variantie. De resultaten laten verder zien, dat tevredenheid geen directe invloed heeft op cross-buying ($p=0.85$). Echter, er wordt wel een zeer significant effect gevonden van een positief verschil tussen tevredenheid en tevredenheid met de concurrent van 4 of hoger ($p=0.00$). Nadere analyse wijst uit dat voor een verschil van +4 er minimaal een 8 moet worden gescoord op de tevredenheid, terwijl in veel gevallen de concurrent een 5 of lager scoort. Opvallend is, dat er geen effect wordt gevonden van een negatief verschil ($p=0.34$). Daarnaast blijkt, dat in tegenstelling tot wat we verwachten, er geen interactie-effect is van tevredenheid en relatieduur ($p=0.53$). Ofwel, klanten die langer bij de verzekeraar zijn, hechten niet meer waarde aan hun tevredenheid. Wel vinden we een significant effect van de gekwadrateerde relatieduur ($p=0.02$) op cross-buying. Dit betekent dat klanten met een langere relatieduur een steeds grotere kans hebben om meer verzekeringen af te nemen. Deze relatie is ook onderzocht voor hogere orde effecten (derde-machts etc.). Deze bleken echter niet aanwezig. Leeftijd heeft een negatieve invloed ($p=0.01$), terwijl het bezit van een inboedelverzekering een positieve invloed heeft ($p=0.03$). Dit laatste effect kan verklaard worden door het feit dat bij deze verzekering in 1995 nog gewerkt werd met lange termijn contracten. Opvallend is tenslotte dat onder deze groep klanten het deelnemen aan het loyaliteitsprogramma geen invloed heeft op cross-buying.

Tabel 2. Resultaten schatting model.

Variabele	Ongestandaardiseerde coëfficiënt	Absolute z-waarde ¹	p-waarde
Tevredenheid	-0.02	0.17	0.85
Tevredenheid – Tevredenheid Concurrent ($\tau_p \geq 4$)	0.38	5.11	0.00
Tevredenheid – Tevredenheid Concurrent ($\tau_n \leq -1$)	0.17	0.95	0.34
Relatieduur*Tevredenheid	0.005	0.63	0.53
Relatieduur	-0.151	1.47	0.14
Relatieduur ²	0.005	2.28	0.02
Covariaten			
Leeftijd	-0.03	2.67	0.01
Inboedelverzekering	0.43	2.22	0.03
Rechtshulpverzekering	0.53	1.33	0.18
Fit-criteria model			
Likelihood ratio statistiek	48.27 (0.00)		
AIC	1.5886		
Pseudo R ²	0.14		

6. CONCLUSIE, IMPLICATIES, BEPERKINGEN EN TOEKOMSTIG ONDERZOEK

6.1. Conclusie

In dit artikel is een model gepresenteerd waarin tevredenheidsdata gekoppeld zijn aan gegevens over het koopgedrag uit een database. Hierbij hebben we specifiek gekeken naar het verschil in het aantal afgesloten polissen over de tijd bij een Nederlandse verzekeraar. In de Nederlandse marktonderzoek literatuur is een model waarin dit soort data aan elkaar gekoppeld worden nog niet eerder gepubliceerd³. De ontwikkeling van het model is echter in lijn met de ontwikkelingen die momenteel plaats vinden op het gebied van marketingmodellen (Leeftink en Wittink, 2000). Hier zien we door de toegenomen hoeveelheid data (Franses, 1998) een toename van het belang van (econometrische) modellen binnen marketing. Hierdoor worden dit soort modellen niet alleen meer toegepast bij frequent aangekochte goederen, zoals ketchup en wasmiddelen, maar ook in andere markten, zoals in dit geval de verzekeringsmarkt. Daarnaast wordt door de toename van continu onderzoek (Lehmann et al., 1998) naar bijvoorbeeld klanttevredenheid de opname van dit soort data in marketingmodellen bevorderd. Met behulp van dit soort modellen kan dan worden vastgesteld of investeringen in bijvoorbeeld de tevredenheid van klanten enig rendement zullen opleveren in termen van een hogere loyaliteit.

De resultaten van het model laten in ons geval zien, dat we cross-buying slechts gedeel-

³ De auteurs sluiten niet uit, dat bij marktonderzoeksbureaus of in bedrijven dit soort modellen wel al zijn ontwikkeld.

telijk kunnen verklaren met de door ons onderzochte variabelen. Opvallend is, dat tevredenheid pas een sterke invloed heeft als de score 4 of meer punten hoger is dan de tevredenheid met de concurrent. Er is dus pas echt een effect, wanneer de onderneming in de perceptie van de klant op individueel klantniveau een veel betere prestatie levert dan de concurrent. Opvallend is verder dat we geen effect vinden van een negatief verschil in de tevredenheidsscores. Dit gaat in tegen resultaten uit de prospect theorie, waarin in het algemeen gevonden wordt dat negatieve verschillen een grotere impact hebben dan positieve verschillen (Oliver, 1996). Echter, ook Bolton, et al. (2000) vinden geen significante invloed van een negatief verschil op het gebruiksniveau van een dienst. Wij verklaren ons resultaat uit het feit, dat het door veel klanten toch lastig gevonden wordt om een eenmaal afgesloten verzekering op te zeggen. Er is dus sprake van hoge switch kosten. Het opzeggen van polissen zal waarschijnlijk alleen gebeuren, wanneer het negatieve verschil een permanent karakter heeft. Daarnaast vinden we een positief effect van de gekwadrateerde relatieduur. Dit bevestigt het idee van Dwyer et al. (1987) dat een relatie zich pas goed ontwikkelt na een periode van verkenning. Tenslotte bewijzen onze resultaten dat het zinvol is om in dit soort modellen ook transactiegegevens mee te nemen.

6.2. Implicaties

Wij zien twee belangrijke implicatiegebieden voor ons onderzoek: marktonderzoek & informatie management en het management van relaties met klanten.

6.2.1 Marktonderzoek en informatiemanagement

Zoals al eerder gezegd laat dit artikel duidelijk zien dat attitudedata gekoppeld kunnen worden aan het daadwerkelijk koopgedrag. Voor marktonderzoekers en informatie-managers is het dan echter wel van belang dat zij deze data op individueel niveau bewaren. Database-technisch betekent dit dat de data uit verschillende bronnen op basis van een relatienummer gekoppeld moeten gaan worden. In ons geval had het marktonderzoeksbureau van de verzekeraar de onderzoeksgegevens op relatienummer-niveau bewaard. Vervolgens was de verzekeraar voor al deze relatienummers in staat de relevante historische data uit de klantendatabase te destilleren. Dit betekent dus wel dat voor het maken van dergelijke analyses een goed databasebeheer noodzakelijk is, waarbij historische data in de datawarehouse van de onderneming bewaard moeten worden. Door het schatten van dit soort modellen kan de onderzoeker antwoord geven op de vraag of tevredenheid het koopgedrag bij een onderneming beïnvloedt, en zo ja, hoe dat gebeurt. Gaat men een stap verder, dan moet het in principe mogelijk zijn in een soort beslissingsondersteunende setting, de marketing manager te adviseren over de mate waarin een investering in kwaliteit leidt tot een hogere tevredenheid, en hoe vervolgens deze tevredenheid weer leidt tot een hogere loyaliteit en hogere opbrengsten. Door een afweging te maken tussen deze hogere opbrengsten en de investering in tevredenheid kan bekeken worden of dit soort investeringen rendabel is (Blattberg en Deighton, 1996). Hierbij zou men ook voort kunnen bouwen op modellen die reeds gebruikt worden om de klantwaarde (lifetime value) te berekenen (Dwyer, 1989). Ook is het mogelijk door het toevoegen van meerdere marketinginstrumenten in het model aan te geven aan de manager wat nu de werkelijke determinanten zijn van het koopgedrag van de klant. Zo liet ons model zien, dat het ingezette loyaliteitsprogramma bij klanten die net geclaimd hebben geen invloed heeft op het afsluiten van verzekeringen. Daarnaast kun-

nen natuurlijk ook bijvoorbeeld prijs(percepties) en direct mailings meegenomen worden.

6.2.2. Management van relaties

Voor het management van relaties hebben de resultaten van ons onderzoek vooral betrekking op cross-buying. Wil een onderneming de afname van meer diensten bevorderen met het verhogen van de klanttevredenheid, dan moet de onderneming er rekening mee houden dat ze daadwerkelijk veel beter moeten presteren dan de concurrentie. De vraag is of de investeringen die hiervoor nodig zijn opwegen tegen de extra opbrengsten. Door het gedrag te koppelen aan de marges van de afgenomen diensten en dit voor verschillende scenario's te doen, kan hier enig inzicht in worden verkregen. We merken hier echter nog wel bij op dat tevredenheid niet alleen op cross-buying een positief effect kan hebben, maar ook op de relatieduur en op bijvoorbeeld mond-tot-mond reclame. Om daadwerkelijk het effect van een verhoging van tevredenheid op de winstgevendheid van klanten door te rekenen, zullen idealiter deze facetten van het gedrag van de klant ook meegenomen moeten worden.

6.3. Beperkingen en toekomstig onderzoek

Er zijn een aantal beperkingen van ons onderzoek. Ten eerste is het model alleen geschat in de verzekeringsmarkt. In toekomstig onderzoek zou het model ook geschat kunnen worden in andere markten. Daarbij komt dat we het model hebben geschat voor een verzekeraar die gebruik maakt van het directe kanaal. Klanten in dit kanaal zijn vaak behoorlijk prijsgevoelig. Dit zou kunnen leiden tot een overschatting van het effect van de concurrentie. Toekomstig onderzoek zou bijvoorbeeld het gedrag van klanten van verzekeringsagenten kunnen bestuderen. Ten tweede hebben we in ons model alleen het loyaliteitsprogramma als marketinginstrument meegenomen. Toekomstig onderzoek zou ook andere instrumenten, zoals prijs en de verstuurd direct mailings, in het model kunnen incorporeren. Ten derde hebben we te maken met een relatief kleine steekproef en weinig variatie in de afhankelijke variabele. Hierdoor is het lastig om significante effecten te vinden. Daarnaast is het hierdoor lastig om gesegmenteerde modellen te schatten (Wedel, 2000; Wedel en Kamakura, 1998). In toekomstig onderzoek zou dan ook een grotere steekproef bekeken kunnen worden. Ten vierde hebben we in dit onderzoek alleen gekeken naar het gedrag van claimers. De vraag is echter of dezelfde verbanden ook opgang doen bij klanten die niet recent een schade afhandeling hebben gehad. Ten vijfde hebben we in dit artikel alleen gekeken naar het korte-termijn effect van tevredenheid. Onderzoek toont echter aan dat er daadwerkelijk ook sprake kan zijn van een lange-termijn effect (Mazursky en Geva, 1989). In ons eigen werk (Verhoef et al., 1999) besteden we hier al aandacht aan. Hier is echter nog zeer veel ruimte voor verder onderzoek. Tenslotte kan men in toekomstig onderzoek beslissingsondersteunende systemen ontwikkelen (Wierenga en Van Bruggen, 2000). Deze systemen kunnen ingezet worden bij de bepaling van het budget voor eventuele investeringen in tevredenheid. Dit past binnen de trend om klanten als een 'asset' te beschouwen (Srivastava et al., 1998).

LITERATUUR

- Anderson, E.W., C. Fornell & D.R. Lehmann, 1994, Customer Satisfaction, Market Share en Profitability – Findings from Sweden, *Journal of Marketing*, Vol. 58 (3), blz. 53-66.
- Anderson, E.W. & M.W. Sullivan, 1993, The Antecedents and Consequences of Customer Satisfaction for Firms, *Marketing Science*, Vol. 12 (2), blz. 125-143.
- Bickart, B., 1993, Carry-over en Backfire Effects in Marketing Research, *Journal of Marketing Research*, Vol. 30 (1), blz. 52-62.
- Blattberg, R.C. & J. Deighton, 1996, Managing Marketing by the Customer Equity Test, *Harvard Business Review*, Vol. 74 (July-August), blz. 136-144.
- Blattberg, R.C., R. Glazer & J.D.C. Little 1994, *The Marketing Information Revolution*, Harvard Business School Press, Boston.
- Bloemer, J., 1994, *Loyaliteit en Tevredenheid*, Jaarboek Nederlandse Vereniging voor Marktonderzoek en Informatiemanagement.
- Bolton, R. N., 1998, A Dynamic Model of the Duration of the Customer's Relationship with a Continuous Service Provider: the Role of Satisfaction, *Marketing Science*, Vol. 17 (1), blz. 45-65.
- Bolton, R.N. & K.N. Lemon, 1999, A Dynamic Model of Customers' Usage of Services: Usage as an Antecedent and Consequence of Satisfaction, *Journal of Marketing Research*, Vol. 36 (2), blz. 171-186.
- Bolton, R. N., P.K. Kannan & M.D. Bramlett , 2000, Implications of Loyalty Program Membership and Service Experiences for Customer Retention and Value, *Journal of the Academy of Marketing Science*, Vol. 28 (1), blz. 95-108.
- Boulding, W., A. Kalra, R. Staelin & V.A. Zeithaml, 1993, A Dynamic Process Model of Service Quality: From Expectations to Behavioral Intentions, *Journal of Marketing Research*, Vol. 30 (1), blz. 7-27.
- Bult, J.R., 1993, *Target Selection for Direct Marketing*, Proefschrift, Rijksuniversiteit Groningen
- De Ruyter, K., 2000, *Tevredenheid: Een Uitgebalanceerd Begrip in Marketing en Marktonderzoek*, Jaarboek Nederlandse Vereniging voor Marktonderzoek en Informatiemanagement, blz. 35-56.
- Dwyer, R.F., P.H. Schurr & S. Oh, 1987, Developing Buyer-Seller Relationships, *Journal of Marketing*, Vol. 51 (April), blz. 11-27.
- Dwyer, R.F., 1989, Customer Lifetime Valuation to Support Marketing Decision Making, *Journal of Direct Marketing*, Vol. 3 (4), blz. 8-15.
- Engel, J.F., R.D. Blackwell & P.W. Miniard, 1993, *Consumer Behavior*, 7th edition, The Dryden Press, Fort Worth.
- Franses, P.H., 1998, *Na de Banvloek of hoe het de Econometrie verder Vergaat*, Eburon, Delft
- Gale, B., 1997, Satisfaction is not enough, *Marketing News*, October 27, blz. 18.
- Glazer, R., 1991, Marketing in an Information-Intensive Environment – Strategic Implications of Knowledge as an Asset, *Journal of Marketing*, 55 (4), blz. 1-19.
- Inman, J. J., J.S. Dyer & J. Jia, 1997, A Generalized Utility Model of Disappointment and Regret Effect on Post-Choice Evaluation, *Marketing Science*, Vol. 16 (2), blz. 97-111.
- Hoekstra, J.C., P.S.H. Leeftang & D.R. Wittink, 1999, The Customer Concept , *Journal of Market Focused Management*, Vol. 4, blz. 43-75.
- Kamakura, W.A., S.N. Ramaswami & R.K. Srivastava, 1991, Applying Latent Trait Analysis in the Evaluation of Prospects for Cross-Selling of Financial Services, *International Journal of Research in Marketing*, Vol. 8 (4), blz. 329-350.
- Leeftang, P.S.H. & D.R. Wittink, 2000, *Marketing Modellen: Een Terugblik en een Blik op de Toekomst*, Jaarboek Nederlandse Vereniging voor Marktonderzoek en Informatiemanagement, blz. 81-92.
- Lehmann, D.R., S. Gupta & J.H. Steckel, 1998, *Marketing Research*, Addison Wesley, Reading Massachusetts.

- Long, J.S., 1997, *Regression Models for Categorical and Limited Dependent Variables*, Advanced Quantitative Techniques in the Social Sciences Series 7, Sage Publication, Thousand Oaks.
- Mazursky, D. & A. Geva, 1989, Temporal Decay in Satisfaction-Purchase Intention Relationship, *Psychology and Marketing*, Vol. 6 (3), blz. 211-227.
- Morwitz, V.G., J.H. Steckel & A. Gupta, 1997, When do Purchase Intentions Predict Sales?, Marketing Science Institute Working Paper, June, Report No. 97-112.
- Oliver, R.L., 1996, *Satisfaction: A Behavioral Perspective on the Consumer*, McGraw-Hill, New York.
- Oliver, R.L. & R.S. Winer, 1987, A Framework for the Formation and Structure of Consumer Expectations: Review and Propositions, *Journal of Economic Psychology*, Vol. 8, blz. 469-499.
- Pindyck, R.S. & D.L. Rubinfeld, 1998, *Econometric Models and Economic Forecasts*, 4th edition, McGraw-Hill, Boston.
- Ping, R.A. Jr., 1993, The Effects of Satisfaction and Structural Constraints on Retailer Exiting, Voice, Loyalty, Opportunism and Neglect, *Journal of Retailing*, Vol 69 (3), blz. 320-352.
- Reichheld, F.F., 1996, *The Loyalty Effect*, Cambridge, MA, Harvard Business School Press.
- Reichheld, F.F. & W.E. Sasser, 1990, Zero Defections – Quality Comes to Service, *Harvard Business Review*, Vol. 68 (5), blz. 105-111.
- Spring, P.N, P.C. Verhoef, J.C. Hoekstra & P.S.H. Leeflang, 2000, The Commercial Use of Segmentation and Predictive Modeling Techniques for Database Marketing, Working Paper, Rijksuniversiteit Groningen.
- Srivastava, R.K., T.A. Shervani & L. Fahey, 1998, Market-Based Assets and Shareholder Value: A Framework for Analysis, *Journal of Marketing*, Vol. 62 (January), blz. 2-18.
- Van der Scheer, H.R., 1997, *Quantitative Approaches for Profit Maximization in Direct Marketing*, Proefschrift, Rijksuniversiteit Groningen.
- Verhoef, P.C., P.H. Franses & J.C. Hoekstra, 1999, The Impact of Satisfaction on the Breadth of the Relationship with a Multi-Service Provider, Ribes, Report 9955.
- Wedel, M., 2000, Vijfentwintig Jaar Marktonderzoekmethoden, *Jaarboek Nederlandse Vereniging voor Marktonderzoek en Informatiemanagement*, blz. 75-80.
- Wedel, M. & W.A. Kamakura, 1998, *Market Segmentation, Conceptual and Methodological Foundations*, Kluwer Academic Publishers, Massachusetts.
- Wierenga, B. en G.H. Van Bruggen, 2000, Marketing Management Support Systemen: Lessen uit de Afgelopen Vijfentwintig Jaar en Perspectieven voor het Nieuwe Millennium, *Jaarboek Nederlandse Vereniging voor Marktonderzoek en Informatiemanagement*, blz. 193-215.
- Winer, R.S., 1985, A Price Vector Model of Demand for Consumer Durables, Preliminary Developments, *Marketing Science*, Vol. 4 (1), blz. 74-90.
- Zeithaml, V.A., L.L. Berry & A. Parasuraman, 1993, The Nature and Determinants of Customer Expectations of Service, *Journal of the Academy of Marketing Science*, Vol. 21 (1), blz. 1-12.
- Zeithaml, V.A., L.L. Berry & A. Parasuraman, 1996, The Behavioral Consequences of Service Quality, *Journal of Marketing*; Vol. 60 (2), blz. 31-46.
- Zeithaml, V.A. & M.J. Bitner, 1996, *Services Marketing*, 1st edition, McGraw-Hill, New York.

5. Winkeltrouw in de kledingbranche

Een empirische vergelijking tussen Nederlandse en Belgische consumenten

K. DE WULF en G. ODEKERKEN-SCHRÖDER

SAMENVATTING

Het belangrijkste doel van dit onderzoek is om expliciet het gezichtspunt van de consument te belichten door na te gaan wat de invloed is van twee kernvariabelen op relatiemarketing succes: de relatie-oriëntatie van de winkel zoals die wordt gepercipieerd door de consument en de relatie-geneigdheid van deze consument. Bovendien willen we onderzoeken welke antecedenten ten grondslag liggen aan beide factoren. Persoonlijke interviews met 338 en 302 bezoekers van respectievelijk een groot Nederlands en Belgisch winkelcentrum tonen aan dat het ontwikkelen van sterke relaties met consumenten afhangt van de gepercipieerde relatie-oriëntatie van de winkel door de consument evenals van de intrinsieke neiging van consumenten om relaties aan te gaan. De resultaten bieden duidelijke richtlijnen voor de kleding detailhandel om relatiemarketing inspanningen te optimaliseren en om doelgroepen beter te segmenteren.

1. INLEIDING

Fournier, Dobscha en Mick (1998, p. 42) stelden dat “relationship marketing is in vogue. Managers talk it up. Companies profess to do it in new and better ways every day. Academics extol its merits.” Het toenemend aantal ‘special issues’ over relatiemarketing van gerenommeerde marketing tijdschriften ondersteunt deze stelling (bijv., *European Journal of Marketing* 1996; *Industrial Marketing Management* 1997; *International Journal of Research in Marketing* 1997; *Journal of Business Research* 1999; *Journal of Marketing Management* 1997; *Journal of the Academy of Marketing Science* 1995). Ondanks de algemeen erkende waarde van het opbouwen van lange termijn relaties vroegen Fournier, Dobscha en Mick (1998, p. 42) zich af of “caught up in our enthusiasm for our information-gathering capabilities and for the potential opportunities that long-term engagements with customers hold, is it possible that we have forgotten that relationships take two?”

De meeste definities van relatiemarketing benadrukken het bestaan van voordelen voor beide partijen (bijv., Grönroos 1990). Deze veronderstelling impliceert dat zowel de aanbieder als de consument interesse moeten hebben om een langdurige relatie aan te gaan. Aangezien herhalingsaankopen de aanbieder aantrekkelijkere financiële resultaten, marktkennis, stabielere marktomstandigheden, toegenomen verkoopmogelijkheden en een flexibere benadering van de markt bieden (Reichheld en Sasser 1990), is het logisch dat de meeste aanbieders de voordelen inzien van lange termijn relaties. Der-

halve is het niet verbazingwekkend dat de waarde van relatiemarketing vooral benaderd is vanuit het oogpunt van de aanbieder, dit ten koste van het oogpunt van de consument (Sheth en Parvatiyar 1995). Toch onderkennen academici meer en meer het belang van een klantvisie in het onderzoeken van langdurige relaties (Gruen 1995). Barnes (1997) stelde bijvoorbeeld dat een relatie niet bestaat tenzij de consument expliciet deze relatie ervaart. De veronderstelling dat een relatie gevormd kan worden met iedere consument leidt vaak tot een verspilling van kostbare middelen, simpelweg omdat de consument geen relatie wil. Als reactie hierop neemt onze studie expliciet een klantenbenadering door het introduceren van twee nieuwe variabelen (relatie-oriëntatie van de winkel en relatiegeneigdheid van de consument), onderzoekt ze hun antecedenten en evalueert ze hun invloed op relatiemarketing succes (tevredenheid met de relatie, vertrouwen, commitment ten opzichte van de relatie en koopgedrag).

De rest van dit artikel is als volgt gestructureerd. Ten eerste introduceren we ons conceptuele model en presenteren we de gerelateerde onderzoekshypothesen. Vervolgens gaan we in op de toegepaste onderzoeksmethodologie en bespreken we de resultaten. In het laatste deel lichten we de implicaties voor de marketing praktijk toe en verschaffen we suggesties voor vervolgonderzoek.

2. CONCEPTUEEL MODEL EN HYPOTHESEN

2.1. Relatiegeneigdheid van de consument en haar bepalende factoren

Relatiegeneigdheid van de consument. Gwinner, Gremler en Bitner (1998) beweerden dat het succes van relatiemarketing niet alleen afhankelijk is van de strategie of implementatie ervan, maar ook van de voorkeuren van de individuele consument. Meer specifiek gebruikten Christy, Oliver en Penn (1996) de term “psychologically predisposed” om het idee uit te drukken dat sommige consumenten intrinsiek geneigd zijn om langdurige relaties aan te gaan. In dit onderzoek introduceren we het begrip “relatiegeneigdheid van de consument” en definiëren we dit als de relatief stabiele en bewuste neiging van een consument om relaties aan te gaan met aanbieders van een bepaalde productcategorie. Het doel van onze studie is dan ook om te onderzoeken in hoeverre de *algemene* neiging van een consument om langdurige relaties te hebben verband houdt met de sterkte van de relatie die deze consument met een *specifieke* aanbieder heeft. Bovendien is deze studie een eerste aanzet om de vraag te beantwoorden waarom mensen relaties aangaan. In het kader hiervan meten we de invloed van vier antecedenten (sociale binding, sociale waardering, winkelplezier en betrokkenheid bij de productcategorie) op de mate waarin een consument relatiegeneigd is.

Sociale binding. In overeenstemming met Cheek en Buss (1981), definiëren we sociale binding als een individueel kenmerk van een consument dat de neiging weergeeft om contact te zoeken met anderen en liever samen te zijn dan alleen. Het aangaan van lange termijn relaties met aanbieders kan een mogelijke manier zijn om deze behoefte aan sociale binding te bevredigen (Forman en Sriram 1991). Shim en Eastlick (1998) stelden dat consumenten niet puur op zoek hoeven te zijn naar producten en diensten van een aanbieder, maar ook naar sociale contacten. Derhalve formuleren we de volgende hypothese:

H₁: Een grotere behoefte aan sociale binding door een consument leidt tot een hogere mate van relatiegeneigdheid van deze consument.

Sociale waardering. In overeenstemming met Brock et al. (1998), definiëren we sociale waardering als een individueel kenmerk van een consument dat de wens weergeeft om door anderen gerespecteerd te worden. Er wordt verondersteld dat sociale waardering het soort relatie bepaalt en de ontwikkeling ervan beïnvloedt (Kirkpatrick en Davis 1994). Forman en Sriram (1991) stelden dat sommige mensen een relatie met een aanbieder aangaan omdat ze op zoek zijn naar sociale waardering. Daarom kunnen we verwachten dat consumenten met een grote behoefte aan sociale waardering meer relatiegeneigd zullen zijn dan consumenten met een beperkte behoefte aan sociale waardering. Dit leidt tot de volgende hypothese:

H₂: Een grotere behoefte aan sociale waardering door een consument leidt tot een hogere mate van relatiegeneigdheid van deze consument.

Winkelplezier. Op basis van Bellenger en Korgaonkar (1980), definiëren we winkelplezier als een individueel kenmerk van een consument dat de neiging weergeeft om winkelen leuker te vinden dan anderen. Het construct winkelplezier duidt tevens op het verschil tussen hedonistisch en utilitaristisch ingestelde consumenten. Terwijl utilitaristische consumenten zich vooral richten op het vervullen van hun winkeltaak streven hedonistische consumenten naar plezier en entertainment tijdens het winkelen (Hirschman en Holbrook 1982). Bellenger en Korgaonkar (1980) toonden aan dat mensen die winkelen leuk vinden, zelden vantevoren weten wat ze willen kopen, waardoor ook de kans verkleint dat ze bij één bepaalde winkel hun aankopen zullen verrichten. Als een gevolg hiervan verwachten we dat consumenten die geen plezier hebben in winkelen eerder geneigd zijn om een relatie aan te gaan met eenzelfde winkel omdat ze wellicht verwachten dat deze relatie hun onplezierige winkeltaak wat kan verzachten. De volgende hypothese wordt voorgesteld:

H₃: Een lagere mate van winkelplezier door een consument leidt tot een hogere mate van relatiegeneigdheid van deze consument.

Betrokkenheid bij de productcategorie. In overeenstemming met Mittal (1995), definiëren we betrokkenheid bij een productcategorie als het langdurige gepercipieerde belang van een productcategorie gebaseerd op de onderliggende behoeften en waarden van een consument. Diverse onderzoekers hebben gesuggereerd dat individuen die zeer betrokken zijn bij een productcategorie ook een grotere neiging vertonen om trouw te zijn (King en Ring 1980). Christy, Oliver en Penn (1996) benadrukten dat betrokken consumenten een goede basis vormen voor het opbouwen van langdurige relaties. Benaderingen door de aanbieder, hoe goed bedoeld dan ook, zouden als een gevolg hiervan door de consument als onwenselijk gezien kunnen worden indien deze consument nauwelijks betrokken is bij de betreffende productcategorie. Derhalve veronderstellen we dat:

H₄: Een grotere betrokkenheid bij een productcategorie door een consument leidt tot een hogere mate van relatiegeneigdheid van deze consument.

2.2. Relatie-oriëntatie van de winkel en haar bepalende factoren

Relatie-oriëntatie van de winkel. We definiëren de relatie-oriëntatie van een winkel als de mate waarin een consument percipieert dat de winkel actief inspanningen doet die erop gericht zijn om vaste klanten te behouden. Deze inspanningen kunnen betrekking hebben op het product of de dienst, maar ook op aspecten van de relatie zelf. We gaan ervan uit dat dit construct voortbouwt op gerelateerde concepten zoals bijvoorbeeld “relational selling behavior” in een context van klant-verkoper relaties (Crosby, Evans en Cowles 1990). Ondanks overeenkomsten tussen deze concepten bestaan er ook sub-

stantiële verschillen. Zo is ons construct relatie-oriëntatie van de winkel ruimer dan alleen inspanningen van een persoonlijke verkoper, wat benadrukt wordt in onderzoeken die “relational selling behavior” bestuderen.

Percepties van relatie-oriëntatie zouden versterkt kunnen worden door de relatiegeïnteresse van de consument om relaties met aanbieders in het algemeen aan te gaan. Relatiegeïnteresseerde consumenten zouden de relatie- inspanningen van enige aanbieder door een soort roze bril kunnen bekijken. Zo toonden Simpson, Gangestad en Lerma (1990) aan dat mensen die op zoek zijn naar een romantische relatie, potentiële partners aantrekkelijker vinden dan mensen die al een relatie hebben. In lijn hiermee verwachten we dat consumenten die relatiegeïnteresseerd zijn een aanbieder ook eerder relatie-geïnteresseerd zullen vinden.

H₅: Een hogere mate van relatiegeïnteresse van de consument leidt tot een hogere mate van gepercipieerde relatie-oriëntatie van de winkel door deze consument.

Velen kampen nog met de vraag welke specifieke strategieën succesvol toegepast kunnen worden om klantentrouw te bevorderen, vaak voortvloeiend uit onjuiste veronderstellingen en onzekerheid over wat belangrijk is voor consumenten (Sirohi, McLaughlin en Wittink 1998). Dit onderzoek gaat na wat de invloed is van vier soorten relatiemarketing inspanningen (direct mail, gedifferentieerde behandeling, interpersoonlijke communicatie en tastbare beloningen) op de mate waarin een winkel als relatie-geïnteresseerd gepercipieerd wordt.

Direct mail. We definiëren direct mail als de perceptie van een consument van de mate waarin de aanbieder vaste klanten op de hoogte houdt via direct mail. Doordat directe communicatie interesse kan opwekken bij de consument, kunnen we direct mail beschouwen als een stimulerende factor voor de groei van een relatie (Duncan en Moriarty 1998). Als een gevolg hiervan willen we onderzoeken of direct mail een antecedent is van relatie-oriëntatie. Onze hypothese luidt dan ook:

H₆: Een hoger mate van gepercipieerd direct mail gebruik leidt tot een hogere mate van gepercipieerde relatie-oriëntatie van de aanbieder door de consument.

Gedifferentieerde behandeling. We definiëren gedifferentieerde behandeling als de perceptie van een consument van de mate waarin een aanbieder vaste klanten anders behandelt en bedient dan niet-vaste klanten (Gwinner, Gremler en Bitner 1998). Sheth en Parvatiyar (1995, p. 264) onderkennen dat focus en selectiviteit ten grondslag liggen aan relatiemarketing, wat impliceert dat niet alle consumenten op dezelfde wijze behandeld dienen te worden. Peterson (1995) stelde dat zo'n afwijkende behandeling de aanbieder in staat stelt om tegemoet te komen aan de menselijke basisbehoefte om belangrijk gevonden te worden. Derhalve verwachten we te kunnen aantonen dat consumenten gedifferentieerde behandeling als een relatie-inspanning zullen ervaren. In overeenstemming hiermee formuleren we de volgende hypothese:

H₇: Een hogere mate van gepercipieerde gedifferentieerde behandeling leidt tot een hogere mate van gepercipieerde relatie-oriëntatie van de aanbieder door de consument.

Interpersoonlijke communicatie. We definiëren interpersoonlijke communicatie als de perceptie van een consument van de mate waarin een aanbieder op een interpersoonlijke en warme manier omgaat met vaste klanten (Metcalf, Frear en Krishnan 1992). Het belang van persoonlijke uitwisselingen tussen consumenten en aanbieders in het beïnvloeden van relatiemarketing uitkomsten is niet verbazingwekkend gezien het feit dat relaties inherent sociale processen zijn (Beatty et al. 1996). Evans, Christiansen en Gill (1996) stelden dat de sociale interactie die mensen ervaren wanneer ze gaan winkelen voor velen de primaire reden is om bepaalde winkels te bezoeken. Voorbeelden van

sociale relatievoordelen zijn gevoelens van bekendheid, vriendschap en sociale steun (Berry 1995), persoonlijke herkenning en het gebruik van de naam van de consument (Howard, Gengler en Jain 1995), de consument als persoon kennen, vriendelijke gesprekken aangaan en persoonlijke warmte tonen (Crosby, Evans en Cowles 1990). Deze redenering is samengevat in de volgende hypothese:

H₈: Een hogere mate van gepercipieerde interpersoonlijke communicatie leidt tot een hogere mate van gepercipieerde relatie-oriëntatie van de aanbieder.

Tastbare beloningen. We definiëren tastbare beloningen als de perceptie van een consument van de mate waarin de aanbieder tastbare beloningen zoals kortingen of cadeautjes geeft aan vaste klanten in ruil voor hun loyaliteit. Frequent flyer programma's, bonussen, gratis cadeautjes en gepersonaliseerde kortingsbonnen zijn voorbeelden van zulke beloningsinspanningen (Peterson 1995). Punten of zegels sparen bij activiteiten als hotelovernachtingen, bioscoop- en car wash bezoeken zouden consumenten motiveren om trouw te blijven ongeacht de inspanningen of prijspromoties van concurrenten (Sharp en Sharp 1997). Derhalve formuleren we de volgende hypothese:

H₉: Een hoger gepercipieerd gebruik van tastbare beloningen leidt tot een hogere mate van gepercipieerde relatie-oriëntatie van de aanbieder.

2.3. Relatiemarketing succes

Vaak genoemde indicatoren van relatiemarketing succes zijn tevredenheid met de relatie, vertrouwen en commitment ten opzichte van de relatie (Baker, Simpson en Siguaw 1999; Crosby, Evans en Cowles 1990; Doney en Cannon 1997). Sharp en Sharp (1997) stelden expliciet voor om deze attitudinale uitkomsten van relatiemarketing aan te vullen met de gedragsmatige veranderingen die worden bewerkstelligd, wat ten grondslag ligt aan onze keuze om koopgedrag als een additionele parameter op te nemen.

Tevredenheid met de relatie. Tevredenheid met de relatie wordt beschouwd als een belangrijke uitkomst van relaties tussen aanbieders en consumenten (Smith en Barclay 1997). We definiëren tevredenheid met de relatie als de affectieve status van een consument die voortkomt uit de globale waardering van zijn relatie met de aanbieder (Anderson en Narus 1984). Zowel in industriële (bijv., Ganesan 1994) als in consumentenmarkten (bijv., Baker, Simpson en Siguaw 1999) neigen consumenten meer tevreden te zijn met aanbieders die bewust inspanningen doen voor consumenten. Als een gevolg hiervan formuleren we de volgende hypothese:

H₁₀: Een hogere mate van gepercipieerde relatie-oriëntatie van de aanbieder door de consument leidt tot een hogere mate van tevredenheid met de relatie.

Bovendien zijn er redenen om te veronderstellen dat tevredenheid niet uitsluitend afhankelijk is van de perceptie van inspanningen door de aanbieder. Storbacka, Strandvik en Grönroos (1994) stelden dat consumenten die geïnteresseerd zijn in relaties ook de tevredenheid met een relatie eerder als belangrijk ervaren. In onze visie kan deze stelling op twee manieren geïnterpreteerd worden. Ten eerste zouden relatiegeneigde consumenten moeilijker tevreden te stellen kunnen zijn, als een gevolg van hun meer kritische houding ten opzichte van de relatie met de aanbieder. Dit komt overeen met de visie van Kalwani en Narayandas (1995) die stelden dat consumenten die bereid zijn om langdurige relaties aan te gaan het moeilijkst tevreden te stellen zijn. Een tweede mogelijkheid is dat relatiegeneigde consumenten eenvoudiger tevreden te stellen zijn, als een gevolg van hun grote mate van ontvankelijkheid voor inspanningen vanuit de aanbieder. In overeenstemming met deze tweede mogelijkheid, veronderstellen wij dat

consumenten die tevredenheid belangrijk vinden, ook gemakkelijker tevreden te stellen zijn. Als een gevolg hiervan stellen we dat:

H₁₁: Een hogere mate van relatiegeneigdheid leidt tot een hogere mate van tevredenheid met de relatie.

Vertrouwen. Het ontstaan van vertrouwen wordt gezien als een belangrijk resultaat van tweezijdige relaties tussen consument en aanbieder (bijv., Gundlach, Achrol en Mentzer 1995). Evenals Morgan en Hunt (1994) definiëren we vertrouwen als een overtuigd geloof van een consument in de eerlijkheid van de aanbieder ten opzichte van deze consument. Een recente meta-analyse in de context van distributiekkanalen (Geyskens, Steenkamp en Kumar 1999) suggereert dat tevredenheid met de relatie voorafgaat aan vertrouwen, dus we stellen:

H₁₂: Een hogere mate van tevredenheid met de relatie leidt tot een hogere mate van vertrouwen.

Commitment ten opzichte van de relatie. Commitment wordt in het algemeen gezien als een belangrijk resultaat van goede relationele interacties (Dwyer, Schurr en Oh 1987). Wij definiëren commitment ten opzichte van de relatie als de voortdurende wens van een consument om zijn relatie met een aanbieder voort te zetten, vergezeld van een bereidheid om moeite te doen om deze relatie daadwerkelijk te onderhouden (Morgan en Hunt 1994). Relaties gekenmerkt door vertrouwen worden zo sterk gewaardeerd dat de betrokken partijen bereid zijn zich te binden aan zulke relaties. Derhalve geven diverse marketers aan dat vertrouwen commitment positief zal beïnvloeden (bijv., Doney en Cannon 1997), waarvoor reeds sterk empirisch bewijsmateriaal beschikbaar is (bijv., Morgan en Hunt 1994). Daarom formuleren we de volgende hypothese:

H₁₃: Een hogere mate van vertrouwen leidt tot een hogere mate van commitment ten opzichte van de relatie

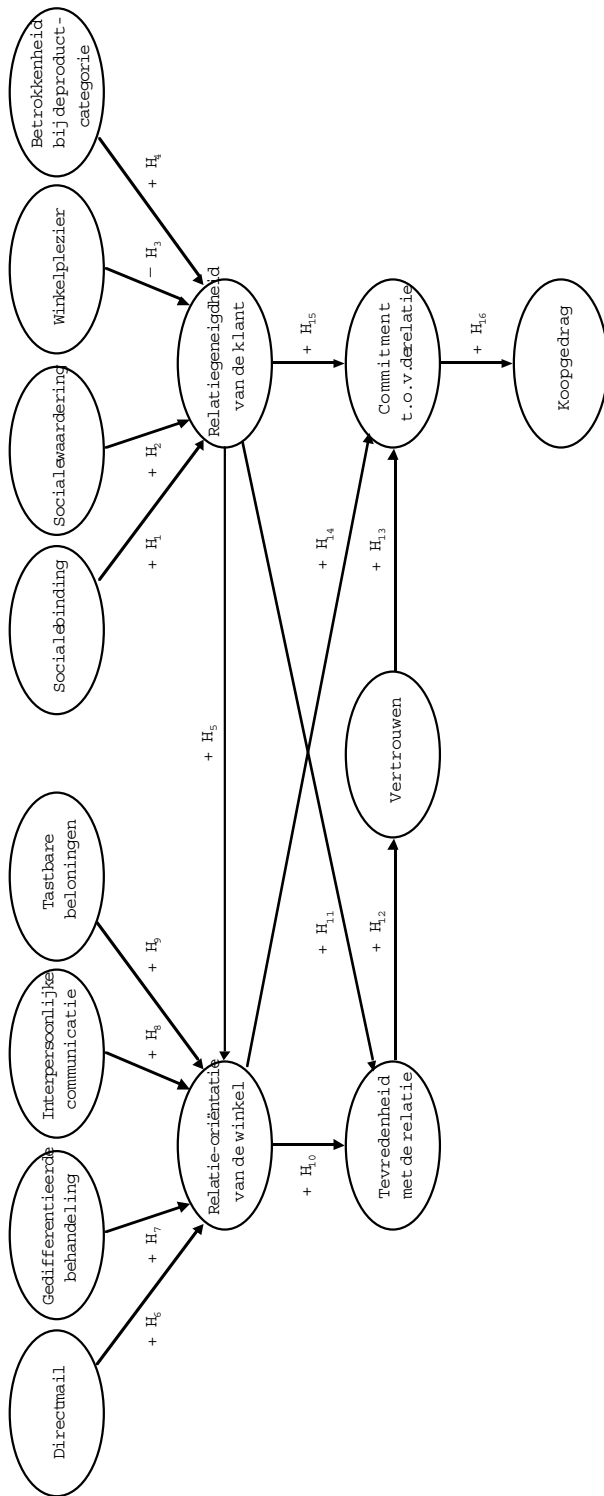
Verder veronderstelden Dwyer, Schurr en Oh (1987) dat er een behoorlijke relationele prestatie noodzakelijk is om commitment te creëren. Zij stelden dat commitment gevoed wordt door het creëren van voordelen voor alle betrokken partijen. In overeenstemming hiermee, ging Bennett (1996) ervan uit dat de sterkte van commitment van een consument bepaald wordt door zijn percepties van de inspanningen die de aanbieder doet. Hieruit vloeit de volgende hypothese voort:

H₁₄: Een hogere mate van gepercipieerde relatie-oriëntatie van de aanbieder leidt tot een hogere mate van commitment ten opzichte van de relatie

Er bestaan echter redenen om te veronderstellen dat het commitment ten opzichte van een relatie niet uitsluitend afhankelijk is van gepercipieerde inspanningen door de aanbieder. Er kan zelfs enige ondersteuning gevonden worden voor de veronderstelling dat relatiegeneigdheid van de consument ook commitment beïnvloedt. Individuele kenmerken zijn reeds herhaaldelijk beschouwd als antecedenten van commitment (Rylander, Strutton en Pelton 1997). Storbacka, Strandvik en Grönroos (1994) gaven bovendien aan dat de belangstelling die een consument in relaties heeft de mate van commitment ten opzichte van een bepaalde relatie kan beïnvloeden. Als een gevolg hiervan stellen we:

H₁₅: Een hogere mate van relatiegeneigdheid leidt tot een hogere mate van commitment ten opzichte van de relatie

Koopgedrag. Moorman, Deshpandé en Zaltman (1993) suggereerden dat consumenten die een hoge mate van commitment ten opzichte van een relatie vertonen een grotere neiging hebben om zich ook zo te gedragen, omdat ze consistent moeten zijn met hun commitment. Ondanks dit gaven Pritchard, Havitz en Howard (1999) in een recente



Figuur 1. Conceptueel model.

studie aan dat het verband tussen commitment en loyaliteit nog steeds weinig empirische aandacht kent. Geïnspireerd door deze ideeën en bevindingen onderzoeken wij de volgende hypothese:

H₁₆: Een hogere mate van commitment ten opzichte van de relatie leidt tot een toename in koopgedrag.

Samenvattend visualiseert figuur 1 ons conceptueel model waarin tevens de veronderstelde effecten zijn opgenomen.

3. METHODE

3.1. Dataverzameling

Twee steekproeven van respectievelijk 338 en 302 respondenten werden samengesteld met het oog op het testen van de hypothesen. Terwijl de eerste steekproef betrekking had op een Nederlands winkelcentrum, bevatte de tweede steekproef bezoekers van een groot Belgisch winkelcentrum. Beide steekproeven werden evenredig verdeeld over de volgende quota criteria: geslacht, leeftijd (jonger dan 25 jaar, 25 tot 40 jaar, 41 tot 55 jaar en boven 55 jaar) en klantaandeel voor de kledingwinkel waarover de vragenlijst werd ingevuld (0-20%, 21-40%, 41-60%, 61-80% en 81-100%). Deze criteria leken ons relevant in het kader van onze studie aangezien veel studies veronderstellen dat ze van invloed zijn op attitudes ten aanzien van winkelen evenals op koopgedrag (bijv., Carman 1970). Tevens hebben we gestreefd naar een gelijke verdeling over tijdstip van interview op de dag (vroeg middag, namiddag, vroeg avond) en dag van interview in de week (woensdag, vrijdag en zaterdag) om op die manier mogelijke vertekeningen te kunnen reduceren. De gegevensverzameling leidde tot inzichten gebaseerd op een breed scala aan soorten kledingwinkels zoals 'discount' winkels, warenhuizen, boetieks en kledingketens. Gemiddeld 37% van de benaderde bezoekers heeft de vragenlijst ingevuld.

3.2. Procedure

We vroegen respondenten eerst of ze ooit zelf een kleding aankoop gedaan hadden. Respondenten die deze vraag met 'ja' beantwoordden, werden verzocht om de namen van vijf winkels te noteren waar ze geregeld hun kleding kopen. Vervolgens gaven respondenten aan welk deel van hun totale kledingbudget ze ongeveer in elk van deze vijf winkels uitgaven (klantaandeel, gemeten op een schaal van 0% tot 100%) en de mate waarin ze zich zelf als vaste consument van elk van die vijf winkels beschouwden (gemeten op een schaal van 1 tot 7). Tot slot selecteerden de interviewers één specifieke winkel uit de opgegeven lijst waarover de resterende vragen beantwoord dienden te worden. Deze keuze was enerzijds gebaseerd op het gewenste percentage van klantaandeel, gebaseerd op de gestelde quotaverdeling en anderzijds werden alleen die winkels opgenomen waarvoor respondenten minimaal een 4 aangaven op de 7 puntsschaal 'in welke mate voelt u zich vaste klant van elk van deze kledingwinkels'. Dit laatste criterium werd gebruikt om van de interne validiteit van de antwoorden te verhogen. Elk construct werd gemeten op basis van 2 of 3 items.

4. RESULTATEN

Om het structurele model te schatten met behulp van LISREL werd de covariantiematrix gebruikt als input voor een maximum likelihood schattingsprocedure. We rapporteren hieronder de resultaten voor het globale model, het meetmodel en het structurele model.

4.1. Evaluatie van het globale model

De chi-kwadraat waarden voor beide steekproeven zijn significant (1,249 met 696 vrijheidsgraden voor de Nederlandse steekproef en 1,200 met 696 vrijheidsgraden voor de Belgische steekproef), iets wat niet ongewoon is bij grote steekproeven (Doney en Cannon 1997). De ratio's van chi-kwadraat ten opzichte van het aantal vrijheidsgraden is 1.79 en 1.72 wat als voldoende beschouwd wordt. CFI, NNFI en GFI zijn 0.93, 0.93 en 0.84 voor de Nederlandse steekproef en 0.92, 0.91 en 0.83 voor de Belgische steekproef. In het algemeen geldt dat de fit indicatoren goed zijn en dat geldt ook voor RMSEA (0.049 respectievelijk 0.049) en SRMR (0.084 respectievelijk 0.070). Gezien de gunstige waarden van deze indicatoren en gezien de solide conceptuele basis van het relatief complexe model, zijn er geen herspecificaties doorgevoerd.

4.2. Evaluatie van het meetmodel

We hebben de kwaliteit van het meetmodel beoordeeld op unidimensionaliteit, convergent validiteit, betrouwbaarheid en discriminant validiteit. Bewijs voor de unidimensionaliteit van elk construct is gebaseerd op een principale componenten analyse die aantoonde dat de items minimaal 0.65 laadden op de veronderstelde component, met een maximale lading van 0.30 op de andere componenten. Convergent validiteit werd ondersteund door een goede fit van het globale model, waarin alle ladingen significant waren ($p < 0.01$) (Hildebrandt 1987). Betrouwbaarheid werd gemeten aan de hand van de samengestelde betrouwbaarheid die voor de meeste constructen groter was dan 0.80. De discriminant validiteit werd getest d.m.v. een serie van geneste confirmatorische factor model vergelijkingen waarin de correlaties tussen de latente constructen gefixeerd werden op 1 (elk van de 84 elementen buiten de diagonaal werden afzonderlijk gefixeerd en het model werd opnieuw geschat). Hieruit bleek dat de chi-kwadraat verschillen inderdaad significant waren voor alle modelvergelijkingen ($p < 0.05$). Verder was het gemiddelde percentage van de verklaarde variantie voor elk construct groter dan 0.50 (met uitzondering van commitment ten opzichte van de relatie in de Nederlandse steekproef, 0.49). Samenvattend kan dus gesteld worden dat beide meetmodellen zuiver zijn, met bewijs voor unidimensionaliteit, convergent validiteit, betrouwbaarheid en discriminant validiteit.

4.3. Evaluatie van het structureel model

Tabel 1 bevat de gedetailleerde resultaten voor beide structurele modellen. Alle significante relaties tussen de latente constructen zijn in de verwachte richting (met uitzondering van het pad van sociale waardering naar relatiegeneigdheid van de consument in de Nederlandse steekproef), wat een sterke ondersteuning biedt voor ons conceptuele model en voor de nomologische validiteit van de constructen.

Tabel 1. Structureel model.

Parameter	Symbool	Nederland (N=338)	België (N=302)
		Schatting	Schatting
Direct mail – relatie-oriëntatie van de winkel	γ_{11}	0.23**	0.27**
Gedifferentieerde behandeling – relatie-oriëntatie van de winkel	γ_{12}	0.10	0.13
Interpersoonlijke communicatie – relatie-oriëntatie van de winkel	γ_{13}	0.35**	0.15
Tastbare beloningen – relatie-oriëntatie van de winkel	γ_{14}	0.18*	0.14
Sociale waardering – relatiegeneigdheid van de klant	γ_{25}	-0.19**	0.07
Sociale binding – relatiegeneigdheid van de klant	γ_{26}	0.12	-0.05
Winkelplezier – relatiegeneigdheid van de klant	γ_{27}	-0.07	0.01
Betrokkenheid bij de productcategorie – relatiegeneigdheid van de klant	γ_{28}	0.54**	0.39**
Relatiegeneigdheid van de klant – relatie-oriëntatie van de winkel	β_{12}	0.18**	0.04
Relatie-oriëntatie van de winkel – tevredenheid met de relatie	β_{31}	0.73**	0.39**
Relatiegeneigdheid van de klant – tevredenheid met de relatie	β_{32}	0.26**	0.43**
Tevredenheid met de relatie – vertrouwen	β_{43}	0.71**	0.72**
Relatie-oriëntatie van de winkel – commitment t.o.v. de relatie	β_{51}	0.21**	0.17**
Relatiegeneigdheid van de klant – commitment t.o.v. de relatie	β_{52}	0.62**	0.35**
Vertrouwen – commitment t.o.v. de relatie	β_{54}	0.30**	0.56**
Commitment t.o.v. de relatie – koopgedrag	β_{65}	0.42**	0.42**
Verklaarde variantie van de endogene constructen	Symbool	Schatting	Schatting
Relatie-oriëntatie van de winkel	ψ_{11}	0.52	0.30
Relatiegeneigdheid van de klant	ψ_{22}	0.23	0.17
Tevredenheid met de relatie	ψ_{33}	0.70	0.36
Vertrouwen	ψ_{44}	0.50	0.52
Commitment t.o.v. de relatie	ψ_{55}	0.79	0.65
Koopgedrag	ψ_{66}	0.18	0.17

(*) $p < 0.05$, (**) $p < 0.01$

Met betrekking tot de antecedenten van relatiegeneigdheid van de consument is alleen betrokkenheid bij de productcategorie in beide steekproeven significant (H_4 geaccepteerd). Voor sociale waardering werd, zoals reeds vermeld, een negatief verband gevonden in de Nederlandse steekproef. De andere antecedenten vertoonden geen verband met relatiegeneigdheid van de consument.

Als we hypothesen H_5 tot en met H_9 bekijken, de antecedenten van relatie-oriëntatie van de winkel, dan blijkt dat de relatiegeneigdheid van de consument een sterke positieve invloed heeft in de Nederlandse steekproef (H_5 gedeeltelijk geaccepteerd). Tussen direct mail en relatie-oriëntatie van de winkel bestond in beide steekproeven een significant, positief verband (H_6 geaccepteerd). Interpersoonlijke communicatie en tastbare beloningen vertoonden alleen in Nederland een sterk positief verband met relatie-oriën-

tatie van de winkel (H_8 en H_9 gedeeltelijk geaccepteerd). Het effect van gedifferentieerde behandeling op relatie-oriëntatie van de winkel was in beide steekproeven niet significant (H_7 verworpen).

Tot slot was er sterke en uniforme ondersteuning voor de hypothesen H_{10} tot en met H_{16} , de structurele verbanden van relatie-oriëntatie van de winkel en relatiegeneigdheid van de consumenten naar de uitkomsten van relatiemarketing en tussen de uitkomsten van relatiemarketing onderling. Al deze paden waren significant en in de verwachte richting. Dus onze conceptuele voorspellingen omtrent de interrelaties tussen relatie-oriëntatie van de winkel, relatiegeneigdheid van de consument, tevredenheid, vertrouwen, commitment en koopgedrag worden empirisch ondersteund.

5. DISCUSSIE

Het belangrijkste doel van dit marktonderzoek was om de specifieke rol van de consument te bepalen in het opbouwen van langdurige relaties in de kledingbranche. In antwoord op recente verzoeken om meer aandacht te schenken aan het klantperspectief (Barnes 1997; Bendapudi en Berry 1997) tonen onze empirische resultaten de cruciale rol van de consument aan in het bewerkstelligen van relatiemarketing succes. Dit blijkt niet alleen uit de sterke invloed van relatiegeneigdheid van de consument, maar ook de percepties van de consument m.b.t. de relatie-oriëntatie van de aanbieder vertonen een significante invloed op tevredenheid en commitment. Een interessante vaststelling is dat tevredenheid relatief sterker wordt beïnvloed door relatie-oriëntatie van de winkel (in de Nederlandse steekproef, in de Belgische steekproef is de invloed ongeveer gelijk). Commitment daarentegen wordt sterker beïnvloed door relatiegeneigdheid van de consument. Op basis hiervan kunnen we veronderstellen dat het moeilijk kan zijn om tevredenheid over de relatie te bewerkstelligen zonder dat klanten relatie-inspanningen van de winkel ervaren. Dit lijkt plausibel aangezien tevredenheid verwijst naar de globale waardering van een relatie met een aanbieder en commitment verwijst naar een voortdurende wens van de consument om de relatie voort te zetten. Vanuit een nomologisch standpunt, zou men inderdaad verwachten dat variabelen gerelateerd aan de aanbieder sterker gerelateerd zijn aan tevredenheid met de relatie, terwijl relatiegeneigdheid van de consument verwacht wordt sterkere verbanden te vertonen met commitment.

Wat betreft de invloed van relatiegeneigdheid van de consument op relatiemarketing succes, zijn onze resultaten tegengesteld aan die van Kalwani en Narayandas (1995), die stellen dat consumenten die relatiegeneigd zijn relatief moeilijker tevreden te stellen zijn. Wel ondersteunen onze data de veronderstelling dat een individueel kenmerk zoals relatiegeneigdheid van de consument van invloed is op commitment (Storbacka, Strandvik en Grönroos 1994).

Mogelijkerwijs kampen bestaande studies op het gebied van relatiemarketing met het ontbreken van het construct relatiegeneigdheid. Onze resultaten tonen immers aan dat de effectiviteit van relatiemarketing strategieën voor een groot deel bepaald wordt door de geneigdheid van consumenten om langdurige relaties aan te gaan. Indien ook in toekomstige studies dit construct niet opgenomen wordt, zou dit kunnen leiden tot vertekende conclusies met betrekking tot de antecedenten en gevolgen van tevredenheid met en commitment ten opzichte van de relatie. Hoewel relatiegeneigdheid vanuit een management oogpunt niet beheerst kan worden door de aanbieder zou segmentatie op basis van relatiegeneigdheid het verwachte markt- en klantaandeel wel kunnen beïnvloeden.

vloeden: relatiegeneigde consumenten vertonen immers een grotere neiging om trouw te blijven aan een bepaalde winkel. Als we de vraag bekijken waarom sommige consumenten geneigd zijn om een relatie aan te gaan, terwijl anderen dit niet zijn, dan zien we dat de betrokkenheid bij de productcategorie een belangrijk antecedent is van relatiegeneigdheid. Dit bevestigt de veronderstelling van King en Ring (1980) dat consumenten bereid zijn om relaties met aanbieders aan te gaan indien hun betrokkenheid voor een bepaalde productcategorie hoog is. Het ondersteunt tevens het idee dat betrokkenheid bij een productcategorie ten grondslag ligt aan individuele kenmerken van consumenten (Beatty, Homer en Kahle, 1988), zoals relatiegeneigdheid van de consument. In tegenstelling tot auteurs die benadrukken dat bevindingen van onderzoek op het vlak van interpersoonlijke relaties vertaald kunnen worden naar commerciële relaties (Shim en Eastlick, 1998) blijken sociale binding en sociale waardering niet ten grondslag te liggen aan de intrinsieke neiging van consumenten om lange-termijn relaties met aanbieders aan te gaan. Mensen die op zoek zijn naar sociale binding en sociale waardering zijn blijkbaar niet noodzakelijkerwijs op zoek naar relaties met winkels in een bepaalde productcategorie. Dit is in tegenstelling tot Ellis (1995) die aantoonde dat mensen die op zoek zijn naar sociale binding ook relaties zoeken met verkooppersoneel. Bovendien werd geen significant verband gevonden tussen mensen die van winkels houden en mensen die relatiegeneigd zijn. Dit komt overeen met Beatty et al. (1996) die stelden dat winkelmotieven verschillen van relatiemotieven. Dus de motieven die consumenten hebben om te winkelen zijn niet noodzakelijkerwijs gerelateerd aan de motieven die mensen hebben om langdurige relaties met winkels aan te gaan.

Een andere factor die van invloed is op het succes van relatiemarketing is de relatie-oriëntatie van de aanbieder. Dit komt overeen met bestaand onderzoek dat aantoonde dat oordelen over prestaties zoals relatie-oriëntatie een belangrijke rol spelen in het beïnvloeden van tevredenheid en commitment (Smith en Barclay 1997). Consumenten die inspanningen van de aanbieder percipiëren lijken hier iets voor terug te willen geven in de vorm van een verhoogde tevredenheid en commitment. Maar op welke wijze kunnen aanbieders deze percepties verbeteren? In overeenstemming met bijv. Howard, Gengler en Jain (1995) en Peterson (1995) toont ons onderzoek aan dat aanbieders die hun Nederlandse consumenten op een persoonlijke manier benaderen en belonen voor hun trouw hier ook de vruchten van zullen plukken in de vorm van verbeterde percepties van de inspanningen op het gebied van klantentrouw, dit geldt niet voor de Belgische consumenten. Dit toont tevens de cruciale rol aan die het winkelpersoneel vervult in haar direct contact met consumenten. Retailers die in staat zijn om hun medewerkers te trainen en motiveren om op een warme en persoonlijke manier om te gaan met consumenten zullen hier voordelen behalen. Als winkelpersoneel aangenaam wordt dient er dus op gelet te worden dat de kandidaat voldoende sociale vaardigheden bezit om de sociale interacties met consumenten te bevorderen (Weitz en Bradford 1999). Dit is vooral van belang gezien de opkomst van geautomatiseerde detailhandel waardoor de mogelijkheden voor sociaal contact in de winkel beperkt worden. In tegenstelling tot onze verwachtingen, hebben we geen verband gevonden tussen gedifferentieerde behandeling en relatie-oriëntatie van de winkel. Dit druist in tegen de algemene opinie dat trouwe klanten op een andere manier behandeld moeten worden dan niet trouwe klanten (Barlow 1992). Tot slot lijken consumenten direct mail inspanningen nog steeds te plaatsen in het kader van relatie-oriëntatie van de winkel. Dus een winkel die haar klanten regelmatig via de post op de hoogte houdt, wordt door de consument gezien als een winkel die belang hecht aan klantentrouw.

Tenslotte, is het niet verbazingwekkend dat het koopgedrag van consumenten ook bepaald wordt door andere factoren (bijv., afstand tot de winkel, assortiment en andere elementen van de retailmix), desondanks wordt respectievelijk 18 en 17 % van de variantie in koopgedrag bepaald door commitment.

6. BEPERKINGEN EN SUGGESTIES VOOR VERDER ONDERZOEK

Enkele beperkingen hebben betrekking op onze dataverzameling en de interpretatie van resultaten. Een eerste beperking heeft betrekking op het negeren van mogelijk belangrijke constructen. Bijvoorbeeld tastbare elementen in de retailmix zoals prijs, promotie, kwaliteit en assortiment zouden toegevoegd kunnen worden als additionele antecedenten van relatie-oriëntatie door de winkel.

Een andere mogelijke tekortkoming van het onderzoek is de zogenoemde ‘common method bias’. Aangezien we slechts één vragenlijst gebruikt hebben om alle constructen te meten, kan de sterkte van de verbanden tussen deze constructen iets vertekend zijn.

Een derde mogelijke tekortkoming heeft betrekking op de wijze waarop koopgedrag is gemeten. Het daadwerkelijk koopgedrag is mogelijk anders dan het door respondenten weergegeven koopgedrag. Helaas was het niet mogelijk om gegevens uit een databestand te gebruiken als input voor het koopgedrag. Het onderzoek zou verbeterd kunnen worden indien er gebruik gemaakt kan worden van feitelijke aankoopgegevens. Tot slot onderkennen we dat onze steekproef van Belgische en Nederlandse consumenten, niet zondermeer vertaald kan worden naar andere contexten. Deze beperkingen kunnen andere onderzoekers inspireren bij het vaststellen van hun toekomstige onderzoeksagenda’s.

LITERATUUR

- Anderson, James C. and James A. Narus, 1984. A Model of the Distributor’s Perspective of Distributor-Manufacturer Working Relationships. *Journal of Marketing* 48, 62-74.
- Baker, Thomas L., Penny M. Simpson, and Judy A. Siguaw, 1999 The Impact of Suppliers’ Perceptions of Reseller Market Orientation on Key Relationship Constructs. *Journal of the Academy of Marketing Science* 27, 50-57.
- Barlow, Richard G., 1992. Relationship Marketing – The Ultimate in Customer Services. *Retail Control* 60, 29-37.
- Barnes, James G., 1997. Closeness, Strength, and Satisfaction: Examining the Nature of Relationships between Providers and Financial Services and Their Retail Customers. *Psychology and Marketing* 14, 765-790.
- Beatty, Sharon E., James E. Coleman, Kristy Ellis Reynolds, and Jungki Lee, 1996. Customer-Sales Associate Retail Relationships. *Journal of Retailing* 72, 223-247.
- Beatty; Sharon E., Pamela Homer, and Lynn R. Kahle, 1998. The Involvement-Commitment Model: Theory and Implications. *Journal of Business Research* 16, 149-167.
- Bellenger, Danny N. and Pradeep K. Korgaonkar, 1980. Profiling the Recreational Shopper. *Journal of Retailing* 56, 77-92.
- Bendapudi, Neeli and Leonard L. Berry, 1997. Customers’ Motivations for Maintaining Relationships With Service Providers. *Journal of Retailing* 73, 15-37.
- Bennett, Roger, 1996. Relationship Formation and Governance in Consumer Markets: Transactional Versus the Behaviourist Approach. *Journal of Marketing Management* 12, 417-436.

- Berry, Leonard L., 1995. Relationship Marketing of Services – Growing Interest, Emerging Perspectives. *Journal of the Academy of Marketing Science* 23, 236-245.
- Biong, Harald and Fred Selnes, 1995. Relational Selling Behavior and Skills in Long-term Industrial Buyer-Seller Relationships. *International Business Review* 4, 483-498.
- Bollen, Kenneth A., 1989. *Structural Equations with Latent Variables*, Wiley, New York.
- Brock, Douglas M., Irwin G. Sarason, Hari Sanghvi, and Regan A.R. Gurung, 1998. The Perceived Acceptance Scale: Development and Validation. *Journal of Social and Personal Relationships* 15, 5-21.
- Carman, James M., 1970. Correlates of Brand Loyalty: Some Positive Results. *Journal of Marketing Research* 7, 67-76.
- Cheek, J.M. and A.H. Buss, 1981. Shyness and Sociability. *Journal of Personality and Social Psychology* 41, 330-339.
- Christy, Richard, Gordon Oliver, and Joe Penn, 1996. Relationship Marketing in Consumer Markets. *Journal of Marketing Management* 12, 175-187.
- Crosby, Lawrence A., Kenneth R. Evans, and Deborah Cowles, 1990. Relationship Quality in Services Selling: An Interpersonal Influence Perspective. *Journal of Marketing* 54, 68-81.
- Dick, Alan S. and Kunal Basu, 1994. Customer Loyalty: Toward an Integrated Conceptual Framework. *Journal of the Academy of Marketing Science* 22, 99-113.
- Doney, Patricia M. and Joseph P. Cannon, 1997. An Examination of the Nature of Trust in Buyer-Seller Relationships. *Journal of Marketing* 61, 35-51.
- Duncan, Tom and Sandra E. Moriarty, 1998. A Communication-Based Marketing Model for Managing Relationships. *Journal of Marketing* 62, 1-13.
- Dwyer, F. Robert, Paul H. Schurr, and Sejo Oh, 1987. Developing Buyer-Seller Relationships. *Journal of Marketing* 51, 11-27.
- Ellis, Kristy, 1995. *The Determinants of the Nature and Types of Customer-Salesperson Relationships in a Retail Setting: An Empirical Study*, doctoral dissertation, University of Alabama.
- Evans, Kenneth R., Tim Christiansen, and James D. Gill, 1996. The Impact of Social Influence and Role Expectations on Shopping Center Patronage Intentions. *Journal of the Academy of Marketing Science* 24, 208-218.
- FEDMA (1998), *Direct Marketing Activities*, Federation of European Direct Marketing Associations, Brussels.
- Forman, Andrew M. and Ven Sriram, 1991. The Depersonalization of Retailing: Its Impact on The “Lonely” Consumer. *Journal of Retailing* 67, 226-243.
- Fournier, Susan, Susan Dobscha, and David Glen Mick, 1998. Preventing the Premature Death of Relationship Marketing. *Harvard Business Review*, 42-44.
- Ganesan, Shankar, 1994. Determinants of Long-Term Orientation in Buyer-Seller Relationships. *Journal of Marketing* 58, 1-19.
- Gassenheimer, Jule B., J. Charlene Davis, and Robert Dahlstrom, 1998. Is Dependent What We Want to Be? Effects of Incongruency. *Journal of Retailing* 74, 247-271.
- Geyskens, Inge, Jan-Benedict E.M. Steenkamp, and Nirmalaya Kumar, 1999. A Meta-Analysis of Satisfaction in Marketing Channel Relationships. *Journal of Marketing Research* 36, 223-238.
- Grönroos, Christian, 1990. Relationship Approach to Marketing in Service Contexts: The Marketing and Organizational Behavior Interface. *Journal of Business Research* 20, 3-11.
- Gruen, Thomas, 1995. The Outcome Set of Relationship Marketing in Consumer Markets. *International Business Review* 4, 447-469.
- Gundlach, Gregory T., Ravi S. Achrol, and John T. Mentzer, 1995. The Structure of Commitment in Exchange. *Journal of Marketing* 59, 78-92.
- Gwinner, Kevin P., Dwayne D. Gremler, and Mary Jo Bitner, 1998. Relational Benefits in Services Industries: The Customer’s Perspective. *Journal of the Academy of Marketing Science* 26, 101-114.
- Hildebrandt, Lutz, 1987. Consumer Retail Satisfaction in Rural Areas: A Reanalysis of Survey Data. *Journal of Economic Psychology* 8, 19-42.

- Hirschman, Elizabeth C. and Morris B. Holbrook, 1982. Hedonic Consumption: Emerging Concepts, Methods and Propositions. *Journal of Marketing* 46, 92-101.
- Howard, Daniel J., Charles Gengler, and Ambuj Jain, 1995. What's in a Name? A Complimentary Means of Persuasion. *Journal of Consumer Research* 22, 200-211.
- Kalwani, Manohar U. and Narakesari Narayandas, 1995. Long-Term Manufacturer-Supplier Relationships: Do They Pay Off for Supplier Firms? *Journal of Marketing* 59, 1-16.
- King, Charles W. and Lawrence J. Ring, 1980. Market Positioning Across Retail Fashion Institutions: A Comparative Analysis of Store Types. *Journal of Retailing* 56, 37-55.
- Kirkpatrick, L.A. and K.E. Davis, 1994. Attachment Style, Gender, and relationship Stability: A Longitudinal Analysis. *Journal of Personality and Social Psychology* 66, 502-512.
- Kumar, Nirmalaya, Lisa K. Scheer, and Jan Benedict E.M. Steenkamp, 1995. The Effects of Supplier Fairness on Vulnerable Sellers. *Journal of Marketing Research* 32, 54-65.
- Metcalfe, Lynn E., Carl R. Frear, and R. Krishnan, 1992. Buyer-Seller Relationships: An Application of the IMP Interaction Model. *European Journal of Marketing* 26, 27-46.
- Mittal, Banwari, 1995. A Comparative Analysis of Four Scales of Consumer Involvement. *Psychology and Marketing* 12, 663-682.
- Moorman, Christine, Rohit Deshpandé, and Gerald Zaltman, 1993. Factors Affecting Trust in Market Research Relationships. *Journal of Marketing* 57, 81-101.
- Morgan, Robert M. and Shelby D. Hunt, 1994. The Commitment-Trust Theory of Relationship Marketing. *Journal of Marketing* 58, 20-38.
- Narver, John C. and Stanley F. Slater, 1990. The Effect of a Market Orientation on Business Profitability. *Journal of Marketing* 54, 20-35.
- Peterson, Robert A., 1995. Relationship Marketing and the Consumer. *Journal of the Academy of Marketing Science* 23, 278-281.
- Pritchard, Mark P., Mark E. Havitz, and Dennis R. Howard, 1999. Analyzing the Commitment-Loyalty Link in Service Contexts. *Journal of the Academy of Marketing Science* 27, 333-348.
- Reichheld, Frederick F. and W. Earl Sasser, 1990. Zero Defections: Quality Comes to Services. *Harvard Business Review* 68, 105-111.
- Rylander, David, David Strutton, and Lou E. Pelton, 1997. Toward a Synthesized Framework of Relational Commitment: Implications for Marketing Channel Theory and Practice. *Journal of Marketing Theory and Practice* 5, 58-71.
- Scheer, Lisa K. and Louis W. Stern, 1992. The Effect of Influence Type and Performance Outcomes on Attitude toward the Influencer. *Journal of Marketing Research* 29, 128-142.
- Sharp, Byron and Anne Sharp, 1997. Loyalty Programs and their Impact on Repeat-Purchase Loyalty Patterns. *International Journal of Research in Marketing* 14, 473-486.
- Sheth, Jagdish N. and Atul Parvatiyar, 1995. Relationship Marketing in Consumer Markets: Antecedents and Consequences. *Journal of the Academy of Marketing Science* 23, 255-271.
- Shim, Soyeon and Mary Ann Eastlick, 1998. The Hierarchical Influence of Personal Values on Mall Shopping Attitude and Behavior. *Journal of Retailing* 74, 139-160.
- Simpson, Jeffrey A., Steven W. Gangestad, and Margaret Lerma, 1990. Perception of Physical Attractiveness: Mechanisms Involved in the Maintenance of Romantic Relationships. *Journal of Personality and Social Psychology* 59, 1192-1201.
- Sirohi, Niren, Edward W. McLaughlin, and Dick R. Wittink, 1998. A Model of Consumer Perceptions and Store Loyalty Intentions for a Supermarket Retailer. *Journal of Retailing* 74, 223-245.
- Smith, J. Brock and Donald W. Barclay, 1997. The Effects of Organizational Differences and Trust on the Effectiveness of Selling Partner Relationships. *Journal of Marketing* 61, 3-21.
- Storbacka, Kaj, Tore Strandvik, and Christian Grönroos, 1994. Managing Customer Relationships for Profit: The Dynamics of Relationship Quality. *International Journal of Service Industry Management* 5, 21-38.
- Weitz, Barton A. and Kevin D. Bradford, 1999. Personal Selling and Sales Management: A Relationship Marketing Perspective. *Journal of the Academy of Marketing Science* 27, 241-254.

6. *Van oude en nieuwe dataverzamelingstechnieken; onderzoek naar verplaatsingsgedrag*

N. KALFS

SAMENVATTING

Informatie over het verplaatsingsgedrag van mensen is voor de overheid van groot belang. De afgelopen jaren is zelfs sprake van een toenemende behoefte aan informatie. Niet alleen wil men weten hoe vaak mensen zich verplaatsen, met welk vervoermiddel over welke afstand ten behoeve van welk doel. Ook is gedetailleerde informatie gewenst over specifieke vervoermiddelen, activiteitenpatronen, routekeuze en snelheden. Deze behoefte aan extra informatie betekent over het algemeen dat de taak van de respondent zwaarder wordt en als gevolg daarvan er een druk ontstaat op de respons. Een respons die in de loop der tijd al duidelijk gedaald is. Reden voor het Centraal Bureau voor de Statistiek en het Ministerie van Verkeer en Waterstaat om te onderzoeken of andere dataverzamelingstechnieken, meer specifiek het Neu Kontiv Design en het gebruik van Global Positioning System, voor een betere respons, hogere kwaliteit van de gegevens en een grotere flexibiliteit kunnen zorgen. In dit hoofdstuk worden de opzet en resultaten van beide dataverzamelingstechnieken besproken.

1. INLEIDING

Informatie over het verplaatsingsgedrag van mensen is voor de overheid, met name voor het ministerie van Verkeer en Waterstaat (V&W) van groot belang. De bron die het meest gebruikt wordt, is het Onderzoek Verplaatsingsgedrag (OVG). Dit onderzoek wordt sinds 1978 jaarlijks uitgevoerd door het Centraal Bureau voor de Statistiek (CBS). Vanaf 1985 is de onderzoeksopzet ingrijpend gewijzigd. Het face-to-face interview werd vervangen door een telefonisch startinterview met daaropvolgend een postaal in te vullen dagboekje. Het OVG kampt sinds de invoering van deze benadering met een gestaag teruglopende respons. Naast deze teruglopende respons, neemt de informatiebehoefte van beleidmakers steeds meer toe. Niet alleen wil men weten hoe vaak mensen zich verplaatsen, met welk vervoermiddel over welke afstand ten behoeve van welk doel. Ook is gedetailleerde informatie gewenst over specifieke vervoermiddelen, activiteitenpatronen, routekeuze en snelheden. Informatie die niet zonder verstoring van de trend met het huidige onderzoek kan worden geleverd.

De ontwikkelingen in respons en informatiebehoefte hebben ertoe geleid dat het CBS in september 1997 een pilot heeft gehouden om te onderzoeken of de onderzoeksopzet van het Neu Kontiv Design (NKD), ontwikkeld door Socialdata uit München en met succes toegepast in diverse andere landen, een substantieel hogere respons, een betere

kwaliteit en meer flexibiliteit zou opleveren dan het toenmalige OVG. De resultaten zijn vergeleken met de resultaten van een OVG controle groep. Op basis van de positieve resultaten is besloten een nieuwe opzet van het OVG conform het NKD te ontwikkelen en testen. Dit nieuwe OVG draait nu ruim 2 jaar en behaalt uitstekende responscijfers. In paragraaf 2 en 3 worden de opzet en de resultaten besproken.

In het OVG en NKD wordt gebruik gemaakt van een vorm van ‘zelfrapportage’. Door middel van een dagboekje wordt gedetailleerde informatie verkregen over de verplaatsingen die een persoon verricht gedurende een bepaalde periode, meestal minimaal 24 uur. Boven een (retrospectieve) vragenlijst heeft de dagboekmethode belangrijke voordelen. Herinneringseffecten en sociaal wenselijke antwoorden komen minder voor in vergelijking met een standaard vragenlijst. De methode heeft echter ook belangrijke nadelen, zeker als meer informatie over het verplaatsingsgedrag is gewenst. Zo kost het bijhouden van een dagboekje behoorlijk wat tijd van respondenten, terwijl het verwerken van de gegevens veel tijd kost van de onderzoekers. Daarnaast heeft de dagboekmethode het nadeel dat de betrouwbaarheid van de antwoorden afhankelijk is van de frequentie waarmee verplaatsingen worden gemaakt en de duur van de verplaatsingen zelf. Infrequente, kortdurende verplaatsingen worden relatief vaker vergeten (Kalfs en van der Waard, 1994). Vergeleken met een retrospectieve vragenlijst is de kwaliteit van een dagboek echter beter.

Een nieuwe ontwikkeling die meer controlemogelijkheden biedt en meer informatie op kan leveren is het gecombineerd gebruik van een Global Positioning System (GPS) en (handheld) computer. Dit systeem is te vergelijken met een combinatie van een people meter in het kijkonderzoek met een computergestuurd interview. Van respondenten die met een dergelijke combinatie aan apparatuur worden uitgerust, kunnen in elke gewenste frequentie automatisch positiegegevens worden verkregen en opgeslagen. Als de positiegegevens worden gekoppeld aan een Geografisch Informatie Systeem (GIS) levert dit informatie op over verplaatsingen, routes en snelheden. Om toepassingmogelijkheden te onderzoeken heeft het Ministerie van Verkeer & Waterstaat een exploratief onderzoek opgezet. In paragraaf 4 van dit hoofdstuk wordt de opzet en de resultaten van dit exploratieve onderzoek beschreven. Tot slot worden in paragraaf 5 zowel het Neu Kontiv Design als de Global Positioning System methode geëvalueerd en wordt tevens aangegeven hoe GPS mogelijk in de toekomst gekoppeld zou kunnen worden aan het Neu Kontiv Design.

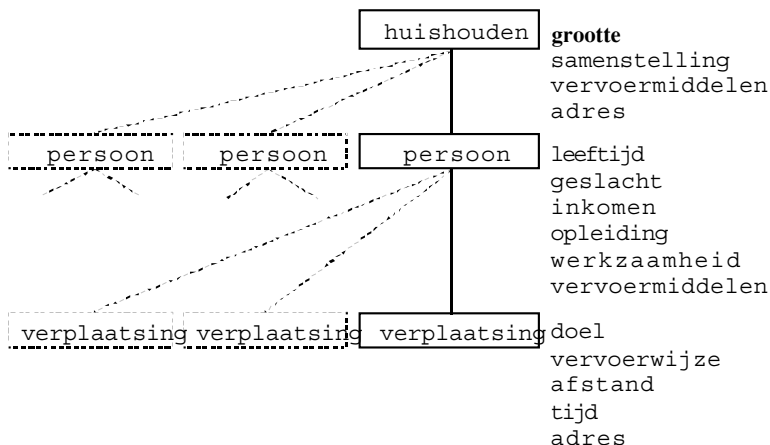
2. ONDERZOEK VERPLAATSINGSGEDRAG (OVG)

2.1. Opzet

Sinds 1978 wordt door het CBS jaarlijks de enquête Onderzoek Verplaatsingsgedrag (OVG) uitgevoerd. Doelstelling van het OVG is het beschrijven van het mobiliteitspatroon van de Nederlandse bevolking. Hiervoor wordt informatie verzameld over verplaatsingen; plaatsen van herkomst en bestemming, gebruikte vervoermiddelen, motieven, reistijden en afstanden (zie figuur 1). Daarnaast wordt aandacht geschonken aan verklarende factoren voor het verplaatsingsgedrag.

OVG

Datastructuur



Figuur 1. Datastructuur OVG.

In 1985 is de onderzoeksmethodiek ingrijpend gewijzigd. Het face-to-face interview werd vervangen door een telefonisch startinterview met daaropvolgend een postaal in te vullen dagboekje. Ook werd nog maar over één dag naar de verplaatsingen gevraagd in plaats van 2 of 3 dagen. Deze wijzigingen leidden tot een trendbreuk in de onderzoeksresultaten. Vergelijking van de gegevens over de periode 1978 tot en met 1984 met de gegevens van 1985 tot op heden zijn daardoor maar zeer beperkt mogelijk. Sinds 1994 zijn nog andere wijzigingen doorgevoerd die niet tot een trendbreuk hebben geleid. Zo is met ingang van 1994 het OVG uitgebreid met de mobiliteit van kinderen, is de steekproef vanaf 1995 verzesvoudigd en is het OVG de afgelopen jaren uitgebreid met vragen over carpoolen en bezit van een OV studenten-kaart. Door deze wijzigingen wordt het microbestand in toenemende mate gebruikt voor het beantwoorden van allerlei beleids- en onderzoeksvragen.

2.2. Respons

Het Onderzoek Verplaatsingsgedrag kampt sinds de invoering van de telefonisch-postale benadering in 1985 met een gestaag teruglopende respons (zie tabel 1). De respons op het CATI-deel bedroeg in de periode 1985-1993 zo'n 76%. In 1994 daalde de respons tot 73% en in 1997 was de respons minder dan 70%. Ook wat betreft het schriftelijke gedeelte (dagboekjes) is sprake van een dalende tendens. Tevens daalt het aantal telefonisch bereikbare huishoudens. Kon in 1985 nog voor 83% van de huishoudens in de steekproef een telefoonnummer worden achterhaald, in 1997 was dat nog maar voor bijna 74% het geval. Op huishoudniveau (complete huishoudens) is de respons inmiddels gedaald tot onder de 32% (85% x 36,8%). In 1998 zet de dalende tendens zich op alle onderdelen verder door.

Tabel 1. Respons OVG 1985-1998.

Jaar	Telefonisch bereikbaar	Respons CATI	Respons Dagboekjes	Respons Totaal*
1985	83,0%	75,0%	81,5%	50,7%
1986	82,3%	76,0%	77,9%	48,7%
1987	82,3%	75,5%	77,6%	48,2%
1988	81,7%	75,0%	72,9%	44,7%
1989	81,4%	77,0%	75,6%	47,4%
1990	81,0%	75,4%	73,2%	44,7%
1991	80,9%	75,6%	75,6%	46,2%
1992	78,1%	76,4%	74,6%	44,5%
1993	77,3%	76,2%	73,1%	43,0%
1994	75,4%	72,6%	75,4%	41,3%
1995	75,1%	72,0%	74,6%	40,3%
1996	73,4%	71,1%	73,8%	38,5%
1997	73,8%	69,9%	71,3%	36,8%
1998	72,7%	68,0%	70,2%	34,7%

* Van het totaal percentage responderende huishoudens is bovendien slechts 85% compleet.

De dalende bereikbaarheid en respons, die mede twijfel oproept ten aanzien van de gewenste representativiteit en vergelijkbaarheid, en de 'veroudering' van de statistische informatie maakten het noodzakelijk mogelijkheden te onderzoeken om de opzet van het OVG te herzien. Per 1 januari 1999 is de onderzoeksmethodiek voor de tweede keer ingrijpend gewijzigd. Op basis van positieve resultaten uit een pilot in september 1997 met het Neu Kontiv Design (NKD), ontwikkeld door Socialdata uit München, is besloten een nieuwe opzet van het OVG conform het NKD te ontwikkelen en te testen (Nieuw OVG). Hiervoor zijn nieuwe vragenlijsten en procedures ontwikkeld gebaseerd op de uitgangspunten van het NKD. De pilot in september had als doel na te gaan of in ons land vergelijkbaar goede resultaten konden worden behaald met het Neu Kontiv Design als in andere landen.

Door een nieuwe uitvoeringsorganisatie is opnieuw proef en parallel gedraaid met het Nieuw OVG. Doel van het proefdraaien was het volledig in de praktijk testen van het nieuwe design en het opzetten van de benodigde projectorganisatie. Het parallel draaien diende om correctiefactoren te berekenen om de gevolgen van de trendbreuk te kwantificeren, zodat de bestaande reeks (vanaf 1985) gecorrigeerd kan worden voor de gevolgen van de invoering van de nieuwe methodiek.

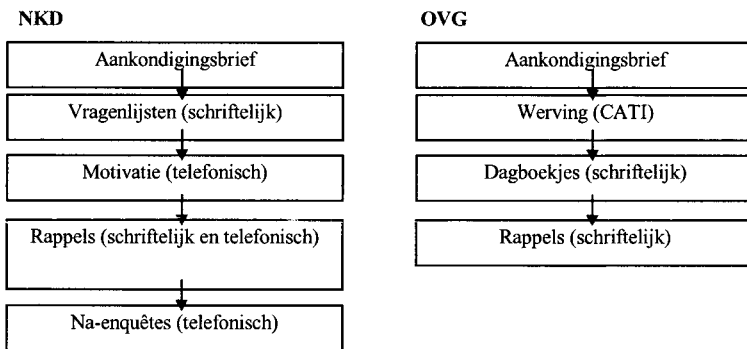
Het Neu Kontiv Design, de opzet en de resultaten van het proef- en parallel draaien met het Nieuw OVG zijn uitvoerig beschreven in de bijdrage 'Nieuw OVG' voor het Colloquium Vervoerplanologisch Speurwerk 1998 (Moritz en van Evert 1998; zie ook De Heer en Moritz 1997, Kadrouch en Moritz 1997). In de volgende paragraaf worden de resultaten samengevat.

3. NIEUW OVG

3.1. Opzet

Het NKD kent als opzet een relatief eenvoudige schriftelijke enquête, met telefonische motivatie van respondenten en eventuele na-enquêtes voor meer gedetailleerde gegevens per subgroep.

Een belangrijk voordeel van een schriftelijke enquête is dat deze wijze van data verzamelen in het algemeen voor respondenten het minst belastend is: men kan de vragenlijst invullen op een tijdstip dat hem of haar het beste uit komt. De basis-enquête bestaat uit een schriftelijke huishoudvragenlijst en een persoonlijk dagboekje waarin over één dag alle verplaatsingen kunnen worden opgegeven. Omdat schriftelijke enquêtes normaliter een slechte respons leveren, wordt de schriftelijke enquête direct gevolgd door een telefonische ronde om mensen te motiveren vragenlijst en dagboekje in te vullen. De telefoon wordt niet ingezet om te enquêteren, maar om te motiveren. Alleen als de gewenste informatie niet schriftelijk is te verkrijgen, wordt telefonisch geënquêteerd. Naast de schriftelijke rappels wordt gelijktijdig telefonisch gerappelleerd (zie figuur 2). Huishoudens zonder telefoonnummer worden alleen postaal benaderd.

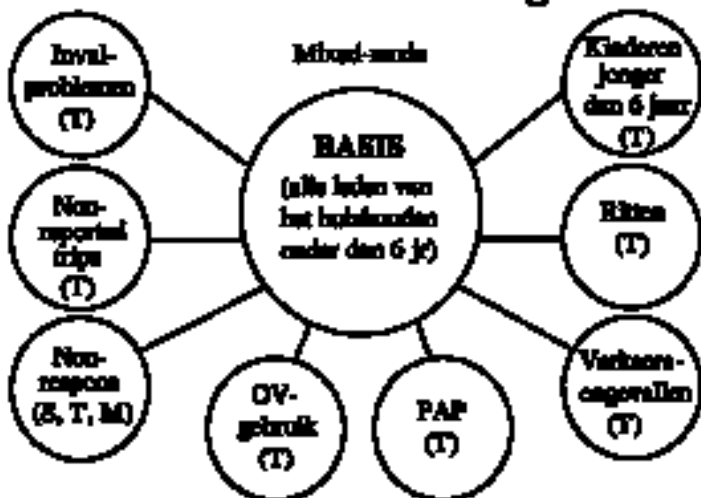


Figuur 2. Onderzoeksopzet NKD en OVG.

De vragenlijsten zijn zo eenvoudig mogelijk gehouden om de respondenten niet lastig te vallen met definities en allerlei vragen die niet voor hen zijn bedoeld. Het uitgangspunt bij het NKD is dat de respondent in eigen terminologie antwoordt. Dit is minder belastend voor de respondent en voorkomt dat voorgecodeerde antwoorden en toelichtingen tot verwarring leiden. In het NKD wordt de belasting van de respondent verschoven naar het bureau. In vergelijking met de OVG-opzet zijn de NKD-vragenlijsten/dagboekjes korter, gemakkelijker leesbaar, en zijn er geen definities, uitgebreide toelichtingen en voorbeelden in opgenomen.

Als vragenlijsten onvolledig, slecht of foutief zijn ingevuld worden telefonisch de missende/correcte gegevens verzameld. Ook worden na-enquêtes gehouden om gegevens van de basis-enquête te completeren (bijv. bij kinderen jonger dan 6 jaar) of bij sub-

Neu KONTIV Design



Figuur 3. Basis- en na-enquêtes (T=telefonisch, S=schriftelijk, M=mondeling).

groepen om extra informatie te verzamelen (bijv. bij Openbaar Vervoer-gebruik). Bij deze zogenaamde satellieten kunnen verschillende dataverzamelingstechnieken worden ingezet (telefonisch, mondeling of schriftelijk). Zie figuur 3.

In tegenstelling tot het OVG kan men stellen dat het NKD het de mensen “makkelijk” maakt om mee te doen en “moeilijk” maakt om af te haken. Dit zou moeten leiden tot een hogere respons.

Het kunnen werken met na-enquêtes biedt nieuwe mogelijkheden en flexibiliteit. Het Nieuw OVG heeft op dit moment vijf standaard na-enquêtes, te weten:

- Invalproblemen; aanvullend telefonisch interview bij respondenten waar missende, slechte of foutieve informatie in de schriftelijke vragenlijst en dagboekje zijn geconstateerd.
- Kinderen jonger dan 6 jaar; aanvullend interview om verplaatsingsgedrag van kinderen jonger dan 6 jaar te achterhalen bij huishoudens die kinderen in die leeftijdsklasse hebben.
- Openbaar Vervoer-gebruik; aanvullend interview bij respondenten waar uit de dagboekjes blijkt dat ze van het Openbaar Vervoer gebruik hebben gemaakt.
- Ritinformatie; aanvullend interview bij respondenten die ketenverplaatsingen hebben gemaakt, d.w.z. een verplaatsing waarbij gebruik is gemaakt van meerdere vervoermiddelen.
- Verkeersongevallen; aanvullend interview bij respondenten die in de vragenlijst hebben geregistreerd betrokken te zijn geweest bij een verkeersongeval.

Deze vijf standaard na-enquêtes worden allemaal telefonisch afgenomen.

De komende jaren zal de flexibiliteit van de nieuwe onderzoeksmethodiek nog meer benut gaan worden. De kwaliteit van de dataverzameling kan bijvoorbeeld verder ver-

beterd worden door standaard na-enquêtes voor vergeten verplaatsingen en non-respons op te nemen. Ook zijn er concrete plannen om het taxigebruik, het sociaal-recreatief verkeer en het internationale personenvervoer te gaan monitoren. Naast deze standaard na-enquêtes (structurele dataverzameling) kunnen ook specifieke na-enquêtes (klantvragen) plaatsvinden om in te spelen op actuele vraagstukken en specifieke informatiebehoefte.

3.2. Respons

In de pilot is de NKD-opzet nauwgezet gekopieerd. De pilot is onder supervisie van Socialdata uitgevoerd. De NKD-steekproef omvatte 1000 adressen in Limburg, Noord Brabant en Gelderland en is aselekt getrokken uit het Geografisch Basis Register. De OVG-controlegroep (zelfde periode en regio) omvatte 1032 adressen. Vervolgens zijn telefoonnummers en naam bijgezocht.

De NKD-pilot leverde in alle opzichten een hogere respons op in vergelijking met de OVG-controlegroep. De totale respons op huishoudniveau bedroeg in het NKD 74% en in het OVG 44% (zie tabel 2). Het aantal bruikbare dagboekjes (responderende personen) is in het NKD 70% hoger dan in de OVG-controlegroep. Ook het aantal huishoudens waarvan complete – dat wil zeggen van alle leden van het huishouden – informatie is ontvangen is in het NKD hoger (90% versus 84% in het OVG).

Tabel 2. Respons NKD pilot 1997.

	OVG controle	NKD pilot
Bruto steekproef	1032	1000
Niet bruikbaar*	18	40
Netto steekproef	1014	960
Met telefoon	811	765
Zonder telefoon	203	195
Respons	446	708
Met telefoon	446	620
Zonder telefoon	–	88
Respons in %	44%	74%
Met telefoon	55%	81%
Zonder telefoon	–	45%

* verhuisd, incorrect adres, leegstaand gebouw.

Het verschil in respons komt enerzijds doordat 45% van de huishoudens waarvan geen telefoonnummer kan worden bijgezocht, respondeert. Anderzijds is de respons in het NKD hoger doordat huishoudens niet alleen schriftelijk kunnen reageren (64%), maar ook door middel van een telefonisch interview (17%) hun verplaatsingsgedrag kunnen opgeven. Vooral ouderen en mensen die weinig verplaatsingen op een dag maken hebben van deze laatste mogelijkheid gebruik gemaakt. De telefonische interviews zorgen ervoor dat deze groepen mensen overtuigd worden mee te doen aan het onderzoek.

Kijken we naar het verplaatsingsgedrag zelf, dan blijkt ook sprake van verschillen. Zo worden in de OVG controle groep per dag gemiddeld meer verplaatsingen gemaakt dan in de NKD pilot groep (zie tabel 3). De verschillen zijn met name te verklaren uit het

feit dat in de pilot NKD meer mensen responderen die zich niet of nauwelijks verplaatsen op de gekozen dagboekdag.

Tabel 3. Verplaatsingskenmerken NKD pilot 1997.

	OVG controle	NKD pilot
Gemiddeld aantal verplaatsingen pppd	3.8	3.2
Afstand pppd	38.6 km	33.3 km
Afstand per verplaatsing	10.2 km	10.4 km
% met 0 verplaatsingen per dag	10.8%	16.6%

Bij de proef (maart en april 1998) en bij het parallel draaien (mei tot en met december 1998) met het Nieuw OVG zijn vergelijkbare goede respons resultaten behaald. Ook de verplaatsingskenmerken zijn vergelijkbaar met de resultaten zoals verkregen uit de pilot. Ten aanzien van de respons is sprake van één verschil met de pilot: de respons van huishoudens waarvan het telefoonnummer niet kon worden bijgezocht is lager (35%). Dit is niet vreemd doordat de proef landelijke dekking heeft en dus ook grote steden omvat. Het is bekend dat in de grote steden de respons lager is dan in andere delen van het land. Desondanks wordt een respons gehaald van ruim 70%.

De hoogte van de respons wordt voor een groot deel bepaald door het aantal reminders. Zowel de huishoudens met telefoon als huishoudens zonder telefoon hebben maximaal 4 schriftelijke reminders gehad. Daarnaast zijn huishoudens waarvan een telefoonnummer bekend is, maximaal 4 keer telefonisch benaderd ten einde de respondenten te motiveren de vragenlijst en dagboekje in te vullen. Kijken we alleen naar de schriftelijke reminders dan zien we dat elke reminder zorgt voor een toename van de respons, zij het dat de toename afneemt met het aantal reminders. Daarbij is sprake van een verschillend effect voor huishoudens die telefonisch kunnen worden benaderd en die waarbij dat niet kan (zie tabel 4). Bovendien blijkt dat langzaam responderende huishoudens een ander verplaatsingspatroon kennen dan snel responderende huishoudens: de reminders halen vooral de minder mobiele mensen over mee te doen aan het onderzoek. Dit is overeenkomstig de literatuur (Brög and Meyburg 1981, Brög and Erl 1999, Brög forthcoming, De Heer and Moritz forthcoming, Richardson et al. 1995, Richardson 2000).

Tabel 4. Respons per mailing Nieuw OVG 1998.

Schriftelijk reminder regiem	Met telefoon	Zonder telefoon
Eerste mailing (incl. vragenlijst)	44%	13%
Eerste reminder	63%	20%
Tweede reminder	74%	25%
Derde reminder (incl. vragenlijst)	78%	33%
Vierde reminder	79%	36%

De responscijfers sinds de integrale invoering van het Nieuw OVG per 1 januari 1999 zijn vergelijkbaar met de resultaten van het NOVG 1998. Overall wordt een respons van 80% gehaald bij huishoudens waarvan het telefoonnummer bekend is en 36% bij

huishoudens waarvan geen telefoonnummer bekend is, hetgeen resulteert in een respons van 71%. Daarmee is de conclusie gerechtvaardigd dat het NKD dus ook met succes toepasbaar blijkt te zijn in Nederland.

4. GLOBAL POSITIONING SYSTEM

4.1. Inleiding

Een nieuwe dataverzamelingstechniek, die meer informatie op kan leveren dan het Onderzoek Verplaatsings Gedrag of Nieuw OVG ooit zal kunnen, is het gecombineerd gebruik van Global Positioning System apparatuur en ('handheld') computers. Hiermee is het mogelijk in elke gewenste frequentie positiegegevens op te slaan. Verplaatsingen presenteren zich zo als een reeks coördinaten. De reeksen coördinaten kunnen vervolgens worden geanalyseerd, waarbij de mogelijkheden van Geografische Informatie Systemen kunnen worden benut. Indien respondenten deze apparatuur bij zich dragen of meenemen opent dit allerlei nieuwe mogelijkheden voor het kennen van het verplaatsingsgedrag. Zo zijn bijvoorbeeld routekeuze en gehanteerde snelheid te bepalen.

De Amerikanen Murakami en Wagner (1997, en artikel in druk) en Guensler en Wolf (1999) deden reeds onderzoek met GPS apparatuur, die was geïnstalleerd in auto's van respondenten. Bij elke rit moesten de respondenten het motief van de verplaatsing in een computer invoeren en gegevens verstrekken over onder andere het aantal medepassagiers. Door het onderzoek op deze wijze in te richten werd verreweg het grootste gedeelte van de verplaatsingen geregistreerd: in Amerika wordt ongeveer 90% van al het vervoer per auto afgelegd (uitgedrukt in kilometers). In Nederland vindt ruim 70% van het vervoer (uitgedrukt in kilometers) plaats per auto. Verplaatsingen met andere vervoermiddelen zijn voor ons dus ook aanzienlijk relevanter. In plaats van de GPS apparatuur in de auto in te bouwen, ligt het meer voor de hand de respondenten de apparatuur bij zich te laten dragen.

In vergelijking met een schriftelijk dagboek biedt het gebruik van GPS potentiële voordelen:

- minder registratiewerkzaamheden voor de respondent
- een hogere kwaliteit van de verplaatsingsgegevens en automatische validatie van de invoer van gegevens
- opslag van de verzamelde gegevens in digitale vorm hetgeen het uitvoeren van kwalitatieve en kwantitatieve analyses vereenvoudigt
- er komt meer informatie beschikbaar; registratie van de afgelegde route vindt automatisch plaats, er is meer bekend over delen van de verplaatsing zoals exacte reistijden, snelheden en richtingen op willekeurige delen van een route of weg.

Nadelen kunnen met name gevonden worden in de bereidheid van respondenten deel te nemen aan een dergelijk onderzoek. De methodiek is privacy gevoelig en de apparatuur mogelijk niet gebruikersvriendelijk.

Momenteel wordt GPS al op vele terreinen toegepast, zowel voor militaire als voor civiele doeleinden. In Nederland zijn op dit moment een groot aantal auto's uitgerust met apparatuur ten einde verkeersinformatie te verschaffen over specifieke delen van

het hoofdwegenet. Onderzoek naar verplaatsingen van de Nederlandse bevolking, waarbij verplaatsingen van alle vervoermiddelen worden geregistreerd, is nieuw. Voor deze toepassingsmogelijkheid is een pilot opgestart. Het doel was op basis van een kleine steekproef inzicht te krijgen in het succes en mogelijkheden die GPS biedt. Een belangrijk onderdeel van het onderzoek was de acceptatie en attitude van respondenten ten opzichte van deze methodiek. Hierbij zijn de volgende vragen aan bod gekomen:

- Hoe is het gesteld met de bereidheid van de respondenten om mee te werken aan het onderzoek? (Dit in verband met privacy en gebruiksgemak.)
- Heeft het dragen van de meetapparatuur invloed op het verplaatsingsgedrag van de respondent?

De opzet en resultaten van de pilot zijn uitvoerig beschreven door Perdok (1997, 2000). In deze paragraaf worden de resultaten samengevat voor zover ze betrekking hebben op de respons en acceptatie en attitude van respondenten.

4.2. Opzet

4.2.1. Ontwikkeling van de apparatuur

Omdat ten aanzien van het gebruik van Global Positioning System apparatuur voor de registratie van het verplaatsingsgedrag een aantal onzekerheden bestond, is vooruitlopend op het project een pre-pilot uitgevoerd (Perdok, 1997). In de pre-pilot bleek het mogelijk om met behulp van de ontwikkelde apparatuur alle typen verplaatsingen te registreren. Alleen ketenverplaatsingen en verplaatsingen met het Openbaar Vervoer leverden enige problemen op. Vooral de apparatuur en software bleken nog voor veel verbetering vatbaar. De belasting voor de respondent was nog aanzienlijk in termen van gewicht, omvang en gebruiksgemak van de apparatuur. Op basis van de pre-pilot is de apparatuur ontwikkeld ten behoeve van het vervolgonderzoek.

Deze apparatuur bestond uit de volgende onderdelen:

- een GPS-ontvanger met antenne
- een ontvanger voor het correctiesignaal (DGPS) met antenne. Door het correctiesignaal in combinatie met GPS te gebruiken kan een hoge nauwkeurigheid in de positiebepaling worden bereikt. De nauwkeurigheid van GPS bedraagt 50 tot 100 meter, van DGPS ongeveer 2 meter.
- een handheld computer (GeoMATE) voor het invoeren van gegevens (vervoermiddel, motief van verplaatsing).

De apparatuur weegt 2 kilo en past in een schouder tas ter grootte van een video camera tas. De antennes voor de ontvangst van GPS en het correctiesignaal waren weggewerkt in de schouderriem van de tas.

De apparatuur is in principe geschikt voor alle verplaatsingen. De kwaliteit van ontvangst is het best wanneer de antennes in de vrije open lucht (uit de buurt van bebouwing) zijn opgesteld. De ontvangstkwaliteit neemt af als de antennes in meerdere of mindere mate worden afgeschermd. Met de antennes in de auto is de ontvangst van het GPS-signaal goed, te meer als de antennes bij de voorruit of achterruit zijn gepositioneerd of indien de schouder tas aan de hoofdsteun van de passagierstoel wordt gehangen. In het openbaar vervoer (bus, trein, tram, metro) is er geen ontvangst van het GPS-signaal of is deze in het algemeen zeer slecht.

De ontvangst van het correctiesignaal is met de gebruikte apparatuur niet optimaal. Om het gebruiksgemak van de apparatuur zo groot mogelijk te maken is gekozen voor een oplossing, waarbij de antenne is weggewerkt in de schouderriem van de draagtas, terwijl in feite een lange sprietantenne is vereist om een optimale ontvangst te krijgen. Bovendien is ervoor gekozen, eveneens in verband met het gebruiksgemak, om de externe antenne bij autoverplaatsingen achterwege te laten. Dit maakt de totale nauwkeurigheid minder groot dan mogelijk is met een voor ontvangst geoptimaliseerde configuratie.

Voor de positiebepaling is een bepaalde opstarttijd nodig. Dit is de tijd tussen het aanzetten van de apparatuur door de respondent en de eerste positiebepaling door het systeem. Deze opstarttijd is niet constant, maar zal variëren als functie van bijv. de dichtheid van de bebouwing. In een omgeving met een hoge dichtheid zal de opstarttijd in het algemeen langer zijn door obstakels, welke in één lijn liggen tussen satelliet en ontvanger.

Naast het inschakelen van de apparatuur, waarna de positiebepaling, datum en tijd automatisch worden opgeslagen, moesten respondenten andere gegevens zelf registreren met behulp van een handheld computer. Hierbij gaat het om motief van de verplaatsing, gebruikt vervoermiddel, aantal passagiers bij autoverplaatsingen, lijnnummer bij OV-gebruik. Deze informatie kon gedurende de verplaatsing worden ingetoetst.

Alle gegevens werden bewaard in de handheld computer. Aan het einde van de onderzoeksperiode zijn de gegevens uit de handheld computer ingelezen in een PC. Vervolgens zijn de verplaatsingsgegevens geanalyseerd met speciaal daartoe ontwikkelde software, waarbij verplaatsingsrecords worden geproduceerd, welke vergelijkbaar zijn met het Onderzoek Verplaatsings Gedrag. Naast gegevens, zoals die ook in het OVG bekend zijn wordt ook informatie over afstand en snelheid afgeleid uit het GPS-signaal.

4.2.2. Werving en instructie van respondenten

Er is een bruto steekproef getrokken van 2000 huishoudens. Deze huishoudens hebben allemaal een introductiebrief ontvangen op het briefpapier van het Ministerie van Verkeer & Waterstaat waarin kort is uitgelegd wat de bedoeling is van het onderzoek. Ook is hierin aangekondigd dat zij telefonisch benaderd zullen worden door het onderzoeksbureau. Na het verzenden van de introductiebrief is er een telefonische screening uitgevoerd met het verzoek om gedurende 4 dagen aan het onderzoek deel te nemen.

In verband met vertragingen in de oplevering van de apparatuur kon het veldwerk pas beginnen in oktober 1998 in plaats van mei 1998. De respondenten zijn op de hoogte gesteld van de vertraging en de reden van de vertraging middels een brief. Nadat de apparatuur beschikbaar was gekomen zijn de respondenten opnieuw benaderd om nadere afspraken te maken over de tijdsperiode, waarin de respondent de apparatuur moest meenemen of meedragen voor alle verplaatsingen. De apparatuur werd bij de respondent gebracht door een medewerker van het onderzoeksbureau. Deze gaf uitleg en een korte demonstratie.

De respondent ontving naast de apparatuur een geschreven instructie, een evaluatiefor-

mulier en een formulier om verplaatsingen te noteren indien de apparatuur weigerde of men vergeten was de apparatuur mee te nemen. Het evaluatieformulier diende om een aantal gebruikskennmerken te evalueren en na te gaan hoe de respondent de deelname aan het onderzoek had ervaren. Naast de genoemde zaken kreeg de respondent een telefoonnummer van de helpdesk. Aan het einde van de eerste registratiedag werd gebeld door een medewerker van het onderzoeksbureau in verband met een controle op de goede werking van de apparatuur. Aan het einde van de registratieperiode werd de apparatuur weer opgehaald door een medewerker. Deze las de gegevens uit, welke waren opgeslagen in de handheld computer.

4.3. Resultaten

4.3.1. De bereidheid om mee te werken

Tijdens het veldwerk is de GPS-apparatuur uitgezet bij circa 150 respondenten. Ca. 100 respondenten zijn geworven via een aselechte procedure. De resterende groep is geworven uit de groep van kennissen en vrienden van personeelsleden van de bij het onderzoek betrokken bedrijven. Dit is gedaan omdat door technische problemen tijdens het veldwerk tijdsproblemen ontstonden. Door de respondenten uit de kring van vrienden en kennissen te betrekken kon met name de tijd tussen de opeenvolgende respondenten worden verkleind. Uit analyses, welke op een aantal respondentkennmerken zijn uitgevoerd, bleken geen fundamentele verschillen tussen de twee groepen. Daarom is besloten om bij de presentatie van de analyseresultaten de twee subgroepen samen te voegen. De respons-analyse is echter gebaseerd op de 100 respondenten, die aselekt zijn geworven. De overige analyses zijn gebaseerd op de gehele groep van 150 respondenten.

In de screening heeft ruim 58% van de benaderde mensen op basis van de bruto steekproef van 2000 toegezegd mee te willen doen aan het onderzoek. Dit is vergelijkbaar met ander screeningsonderzoek waarbij respondenten een computer moeten gebruiken. Als reden om niet mee te doen scoort 'geen zin' hoog, namelijk 28,6%. Ook 'onderzoek vergt teveel tijd' heeft een hoog aandeel: 23,2%. De reden 'vanwege privacy' wordt in 10,3% van de gevallen genoemd. Dit is relatief hoog in vergelijking tot een ander onderzoek naar het verplaatsingsgedrag (MuConsult, 1997), waarin deze reden veel minder vaak werd genoemd (3,2 %). Verder wordt de reden 'te oud' relatief vaak (9,8%) genoemd. Uiteindelijk zijn het vooral jongere, mobiele mannen met een hoog inkomen en een hoge opleiding die bereid zijn deel te nemen aan onderzoek waarin gebruik wordt gemaakt van technologische innovaties. Ook dit stemt overeen met ervaringen bij vergelijkbare onderzoeken (zie bijvoorbeeld MuConsult, 1997).

4.3.2. Evaluatie van het onderzoek door de deelnemers

Over het algemeen is ruim 70% van de deelnemers positief over de deelname; 65% zegt opnieuw bereid te zijn aan een dergelijk onderzoek mee te doen. Een negatief oordeel over de privacy-aspecten van het onderzoek heeft 6% van de respondenten. Dit is niet extreem hoog, maar toch hoger dan in bijvoorbeeld het eerder genoemde vergelijkbare onderzoek (MuConsult, 1997), waar dit 3,8% was. Er is wel een samenhang (hoewel geen perfecte) tussen het oordeel om hernieuwd mee te doen aan het experiment en de oordelen over de tijdsduur, tijdsbeslag en privacy-aspecten. Respondenten die opnieuw willen deelnemen, hebben over deze aspecten vaker een positief oordeel dan de andere respondenten.

Mannen zijn over het algemeen positiever over het experiment dan vrouwen. Vrouwen die meedoen omdat hun man medewerking heeft toegezegd, zijn beduidend minder positief dan vrouwen die hiervoor zelf kiezen. Ook mensen van 50 jaar en ouder zijn over het algemeen minder positief over de verschillende aspecten van het onderzoek. Dit stemt overeen met de bevindingen van de non-responsanalyse, waaruit blijkt dat mannen en jongeren een grotere bereidheid hebben om aan het onderzoek deel te nemen. Naarmate het inkomen hoger is, beoordelen respondenten het onderzoek positiever. Ook stijgt met het inkomen de bereidheid om opnieuw mee te willen doen aan een dergelijk onderzoek.

4.3.3. Gebruiksgemak van de apparatuur

Alhoewel het merendeel van de respondenten positief bleek over de onderzochte gebruikskennmerken, geeft een aanzienlijk percentage respondenten (tussen de 30% en 40%) negatieve oordelen over het gewicht, de grootte, de leesbaarheid van het display en het gebruik van de batterij. Circa 10% van de respondenten vindt de apparatuur te groot bij alle verplaatsingen. Vooral bij verplaatsingen te voet worden gewicht (27%) en omvang (26%) negatief beoordeeld. Dit geldt ook bij verplaatsingen per fiets/brommer/motor en voor winkelverplaatsingen. Grootte en gewicht vormen nauwelijks een probleem bij verplaatsingen per auto (minder dan 1%).

Vrouwen oordelen minder vaak positief (58% t.o.v. 75%) over het gewicht van de apparatuur. Dit geldt ook ten aanzien van de grootte (53% t.o.v. 63%). Respondenten van 65 jaar en ouder zijn relatief minder positief over de verschillende gebruikskennmerken dan jongeren. Het oordeel over de bediening is in zijn algemeenheid (tussen de 85% en 95%) positief, maar het aandeel respondenten met een positief oordeel neemt af met de leeftijd. Het oordeel over de leesbaarheid van het display is vooral bij ouderen minder vaak positief (tussen de 40% en 50%). Ten aanzien van opleiding zijn geen systematische verschillen gevonden bij de beoordeling van de verschillende gebruikskennmerken. De tevredenheid over gewicht, grootte en schouderband neemt toe met het inkomen, terwijl de tevredenheid over de gebruiksduur van de batterij afneemt met het inkomen. Dit zou kunnen samenhangen met een groter aandeel (lange) autoverplaatsingen bij hogere inkomens. Hierbij heeft men minder last van gewicht en omvang van de apparatuur, terwijl de beperkingen van de batterij zich wel doen gelden.

4.3.4. De mogelijke invloed van het onderzoek zelf op het verplaatsingsgedrag

Het verplaatsingsgedrag in een onderzoekssituatie hoeft om verschillende redenen niet overeen te stemmen met het normale verplaatsingsgedrag. In dit onderzoek is het onder andere mogelijk dat:

- respondenten niet willen dat bepaalde verplaatsingen worden geregistreerd
- respondenten het in bepaalde situaties te lastig vinden om de apparatuur mee te nemen
- er problemen ontstaan met de apparatuur.

In tabel 5 is aangegeven in welke mate bovenstaande situaties zich hebben voorgedaan. De tabel laat zien dat er 30 gegevensbestanden verloren zijn gegaan (dit is 20%). Dit is veroorzaakt door mankementen aan de apparatuur tijdens de uitvoering van het veldwerk. Verder blijkt dat de mogelijkheid om een verplaatsing op papier te noteren een belangrijke uitweg heeft geboden in die gevallen waarin men om de een of andere

reden geen gebruik van de apparatuur wilde of kon maken. Als deze mogelijkheid niet geboden zou zijn, zou circa 19% van de verplaatsingen waarschijnlijk niet zijn gerapporteerd.

Tabel 5. Overzicht van de resultaten van het veldwerk.

Grootheid	Aantal
Totaal aantal respondenten	151
Uitval van respondenten (eenheden)	30
Geregistreerde verplaatsingen m.b.v. GPS apparatuur	1456
Geregistreerde ritten m.b.v. GPS apparatuur	1593
Genoteerde verplaatsingen (op papier)	348

Tabel 6 geeft een overzicht van de redenen om verplaatsingen op papier te noteren. Defecte apparatuur is in het merendeel van de gevallen de reden om de verplaatsing op papier te noteren. Echter, ook ‘te belastend om mee te nemen’ heeft een relatief hoog percentage. Deze belasting zou vooral betrekking hebben op fiets-, loop- en Openbaar Vervoer-verplaatsingen en verplaatsingen ten behoeve van het motief winkelen. Privacy speelt een hele kleine rol.

Tabel 6. Overzicht van de redenen om de verplaatsing te noteren en niet te registreren. (% t.o.v. totaal aantal verplaatsingen)

Reden om te noteren	Percentage
Wel aangezet, geen registratie	10
Te belastend om mee te nemen	18
Apparatuur was defect	32
Vergeten te gebruiken	14
Geen zin / geen tijd	10
Batterij niet opgeladen / leeg	4
Foutief vervoermiddel / motief	2
Vanwege privacy	1
Niet ingevuld	9
Totaal	100

Uit het onderzoek komt naar voren dat de bereidheid deel te nemen praktisch vergelijkbaar is met ander onderzoek. Weliswaar zijn er enige verschillen in bereidheid ten gevolge van privacy en leeftijd, maar deze verschillen zijn niet heel erg groot. Verder blijkt dat het belangrijk is nog eens goed naar het gebruiksgemak te kijken. De techniek is nog lang niet uit ontwikkeld, gegeven zowel de hoge uitval van respondenten als de uitval bij verplaatsingen waarbij de apparatuur zijn werk niet deed. De mogelijkheid een verplaatsing te noteren op papier heeft een belangrijke uitweg geboden in die gevallen waarin men om de een of andere reden geen gebruik van de apparatuur kon of wilde maken.

5. SAMENVATTING EN CONCLUSIES

In dit hoofdstuk is onderzocht of andere dataverzamelingstechnieken, meer specifiek het Neu Kontiv Design en het gebruik van Global Positioning System, voor een betere respons, hogere kwaliteit van de gegevens en een grotere flexibiliteit kunnen zorgen in onderzoek. Dit tegen de achtergrond van een toenemende behoefte aan informatie over het verplaatsingsgedrag van mensen en een dalende respons in onderzoek.

Uit de resultaten komt naar voren dat het Neu Kontiv Design (NKD), hetgeen een revitalisering van het reeds lang gehanteerde Onderzoek Verplaatsingsgedrag is, aan alle kenmerken voldoet. Het NKD is met succes toepasbaar gebleken in de Nederlandse situatie. Zo wordt in het Nieuw OVG een respons van 80% gehaald bij huishoudens waarvan het telefoonnummer bekend is en 36% bij huishoudens waarbij geen telefoonnummer bekend is, hetgeen resulteert in een respons van 71%. Ook de kwaliteit van de gegevens is beter. Dit komt met name doordat in het Nieuw OVG meer mensen responderen die zich niet of nauwelijks verplaatsen. Daarnaast is het instrument flexibel omdat het werken met na-enquêtes in de praktijk mogelijk is gebleken.

Uit de Global Position System (GPS) pilot blijkt dat de bereidheid deel te nemen aan een onderzoek waarbij men apparatuur bij zich moet dragen of mee moet nemen die automatisch de verplaatsingen registreert niet echt afwijkt van de bereidheid deel te nemen aan onderzoek (waarin met computers wordt gewerkt). Weliswaar zijn er enige verschillen in de bereidheid ten gevolge van privacy en hoge leeftijd, maar deze verschillen zijn niet heel erg groot. Verder blijkt dat het belangrijk is nog eens goed naar het gebruiksgemak van de apparatuur te kijken. Lang niet iedereen was hierover tevreden. Tot slot is duidelijk geworden dat de techniek op dit moment nog niet is uit ontwikkeld. Dit blijkt uit het feit dat zowel de uitval van respondenten als verplaatsingen waarbij de GPS apparatuur zijn werk niet deed, hoog is. De mogelijkheid om een verplaatsing te noteren op papier is een belangrijke uitweg gebleken in die gevallen waarin de apparatuur het af liet weten. Ondanks alle kinderziektes, zijn wij ervan overtuigd dat in de toekomst gewerkt kan worden met verbeterde apparatuur. Het gebruik van GPS in een na-enquête kan een waardevolle bijdrage leveren aan het Nieuw OVG ten einde aan de behoefte aan extra informatie tegemoet te komen.

LITERATUUR

Brög, W. and A.H. Meyburg (1981) *Consideration of Non-Response Effects in Large Scale Mobility Surveys*. Paper presented at the 60th Annual Meeting of the Transportation Research Board (TRB), Transportation Research Record, No. 807, Washington, DC., 1981.

Brög, W. and E. Erl.(1999) *Systematic Errors in Mobility Surveys*. Paper presented at the ATRF Conference, Perth, 1999.

Brög, W. (forthcoming) *Raising the Standard. Transport Survey Quality and Innovation..* Paper presented at the International Conference "Transport Survey Quality and Innovation", Grainau, 1997.

De Heer, W. en G. Moritz (1997). *Responsproblematiek OVG; een vooronderzoek naar de mogelijkheden voor verbetering*. CBS (interne publicatie).

De Heer, W. and G. Moritz (forthcoming) *Sampling, Response and Weighting. Data-quality problems in Travel Surveys, an international Overview*. Paper presented at the International Confe-

- rence "Transport Survey Quality and Innovation", Grainau, 1997.
- Guensler, R. and J. Wolf (1999). Development of a handheld Electronic Travel Diary for Monitoring Individual Tripmaking Behaviour. *School of Civil and Environmental Engineering, Georgia Institute of Technology, Atlanta.*
- Kadrouch, S. en G. Moritz (1997). *Redesign OVG; verschillen tussen OVG en Neu KONTIV Design.*
- CBS (interne publicatie).
- Kalfs, N. and J. van der Waard (1994) *Kwaliteit van gegevens in tijdsbestedings- en verplaatsingsdagboeken.* CVS. 889-907.
- Moritz, G. en H. van Evert (1998). *Nieuw OVG. Paper t.b.v. het Colloquium Vervoerplanologisch Speurwerk 1998.* Heerlen/Rotterdam, augustus 1998
- MuConsult (1997). *Effecten van dataverzamelingmethoden in onderzoek naar activiteiten en verplaatsingsgedrag.* Eindrapport. Amersfoort.
- Murakami, E. and D. Wagner (1997). *Global Positioning Systems for Personal Travel Surveys.* Final Report for OHIM, OTA and Federal Highway Administration. U.S.A.
- Murakami, E. and D. Wagner (forthcoming). *Using Global Positioning Systems and Personal Digital Assistants for Personal Travel Survey in the U.S.* Paper presented at the International Conference "Transport Survey Quality and Innovation", Grainau, 1997.
- Perdok, J. (1997). *Eindrapport Mogelijkheden Global Positioning System (GPS) voor verkeer en vervoeronderzoek.* Pre-pilot fase. MuConsult, Amersfoort.
- Perdok, J. (1999). *Mogelijkheden Global Positioning System (GPS) voor verkeer en vervoeronderzoek.* Eindrapport. MuConsult, Amersfoort.
- Richardson, A., E. Ampt, A. Meyburg (1995) *Survey Methods for Transport Planning.* Eucalyptus Press, 1995.
- Richardson, A. (2000) *Behavioural Mechanisms of Non-Response in Mailback Travel Surveys.* Paper presented at the 79th Annual Meeting of the Transportation Research Board, Washington, DC.

7. *Total Quality Management bij training en supervisie van interviewers*

A. CARTON

SAMENVATTING

Vanaf het begin van de jaren negentig is de filosofie van integrale kwaliteitszorg niet langer een onbekende binnen het domein van het survey-onderzoek. Het survey-onderzoek kan effectief beschouwd worden als een productieproces waarbij elk onderdeel van het proces – van het conceptualiseren van de vragenlijst tot en met de rapportering van de resultaten – aan kwaliteitsvereisten dient te beantwoorden. Het niet tegemoet komen van een bepaalde fase aan de vooropgestelde kwaliteitsmaatstaven heeft een negatieve impact op de kwaliteit van het “eindproduct” – het formuleren van een antwoord op een centrale probleemstelling – en noopt tot het opzetten van verbeteringsprocessen.

Deze bijdrage beperkt zich tot de kwaliteitsevaluatie van één belangrijk onderdeel van het survey-proces: de dataverzameling door middel van persoonlijke interviews. Meer bepaald worden de principes van integrale kwaliteitszorg toegepast op de training en supervisie van de interviewers. Er wordt gebruik gemaakt van gegevens van het ISPO (Departement Sociologie – K.U.Leuven) om enkele praktijken van een actueel functionerend interviewernetwerk te confronteren met vooropgestelde kwaliteitsnormen. Na deze confrontatie volgen mogelijke voorstellen tot verbetering.

1. INLEIDING

Iets heeft kwaliteit wanneer het voldoet aan eisen. Kwaliteitszorg betreft dan het bewaken en bevorderen dat aan die eisen voldaan wordt (Keursten 1992:1). Sinds enkele jaren is ook binnen het marktonderzoek kwaliteit een belangrijk begrip. Zo bevat het NVMI-Jaarboek van 1992 een artikel met als titel “Kwaliteitscertificatie van marktonderzoek” (Boomsma 1992: 9-19). De auteur besluit zijn uiteenzetting kernachtig als volgt: “Het zou voor ieder marktonderzoeksbureau dat de kwaliteit van producten en diensten wil onderzoeken en daar geld aan wil verdienen, een verplichting moeten zijn om zijn eigen kwaliteit in een erkend ISO 9000 systeem onder te brengen. Voor de grotere bureaus een uitdaging: welk bureau is het eerste gecertificeerde bureau van Nederland?” (1992:19). Inmiddels zijn leden van de Vereniging van Marktonderzoeksbureaus (VMO) ISO gecertificeerd of zijn daarmee bezig.

Binnen survey-onderzoek is de kwaliteit van de interviewer uiterst belangrijk. In 1994 ontwikkelde het Voorlichtingscentrum Uw Mening Telt (UMT) een basispakket voor de opleiding tot face-to-face interviewer. Het doel van de opleiding is het bevorderen

van de kwaliteit van de interviewers en daarmee het bevorderen van de onderzoekskwaliteit. Het examen, dat onder de verantwoordelijkheid van de NVMI-Examencommissie valt, levert het certificaat “Marktonderzoek Interviewer” op. Een andere vorm van opleiding die door UMT wordt verzorgd zijn de “Train de Trainer Dagen”. Deze dagen zijn bestemd voor medewerkers van marktonderzoekbureaus met als doel hen te leren hoe zij interviewers kunnen trainen voor het examen.

Het behalen van ISO-certificatie door marktonderzoekbureaus en certificatie van interviewers spelen uiteraard een belangrijke rol bij de kwaliteitsbeheersing van marktonderzoek. Het tegemoet komen aan Total Quality Management (TQM) gaat echter nog enkele stappen verder. Hoe het concept TQM toepasbaar is binnen het productieproces van het survey-onderzoek in het algemeen en de dataverzameling in het bijzonder, zijn twee aandachtspunten die in de volgende paragrafen centraal staan. Vervolgens illustreren we elementen van TQM bij de kwaliteitsevaluatie bij de training en supervisie van interviewers. Dit mondt uit in een aantal concrete voorstellen ter verbetering van het productieproces van de dataverzameling. We besluiten met een bondig overzicht van enkele “goede” praktijken voor het marktonderzoek en met een terugkoppeling naar het concept TQM.

2. HET CONCEPT TOTAL QUALITY MANAGEMENT BINNEN HET SURVEY-ONDERZOEK

De term “Total Quality Management” is binnen het bedrijfsleven al lang geen onbekende meer. Wil een bedrijf niet van de markt weggeconcentreerd worden, dan lijkt één of andere vorm van kwaliteitsmanagement noodzakelijk. De principes van TQM worden reeds in heel diverse domeinen en sectoren, zowel de profit als de non-profit sector, toegepast. Volgens Biemer en Caspar (1994:309) is deze toepassing echter nog vrij schaars op het vlak van survey-onderzoek. Het lijkt nochtans zinvol om survey-onderzoek in het algemeen en dataverzameling in het bijzonder te beschouwen als een productieproces waarbij TQM kan worden toegepast. Sporen van TQM-toepassingen kunnen reeds teruggevonden worden bij de ontwikkeling van een algemeen kader voor kwaliteitsmanagement binnen de context van nationale instituten voor de statistiek (zie onder meer Colledge en March 1993; Diplo 1997; Morganstein en Marker 1997). Ook binnen de universitaire onderzoekswereld is er de laatste jaren een groeiende belangstelling voor de toepassing van TQM binnen het survey-onderzoek (Lyberg et al. 1997).

In wat volgt vatten we het survey-onderzoek op als een productieproces en geven aan hoe het concept Total Quality Management kan worden toegepast.

De voornaamste fases binnen het productieproces van het survey-onderzoek kunnen samengevat worden als volgt: het conceptualiseren en operationaliseren van de onderzoeksvragen en in het verlengde hiervan de keuze voor een bepaald onderzoeksdesign, het opstellen van de vragenlijst, het samenstellen van de steekproef, de verzameling van de gegevens, de ‘editering’ en codering van de gegevens, het klaarmaken van het definitief databestand en tot slot het analyseren en rapporteren van de gegevens. Een kwalitatief hoogstaand survey veronderstelt dat elk van de onderscheiden fases voldoet aan een aantal kwaliteitsvereisten. Indien de uitvoering van een bepaalde fase niet

beantwoordt aan vooropgestelde kwaliteitsmaatstaven, dan dienen verbeteringsprocessen te worden geïntroduceerd. Het is immers belangrijk de *eindgebruikers* van het product niet uit het oog te verliezen. In het kader van een dataverzameling van een survey-onderzoek zijn de klanten diegenen die de verzamelde gegevens analyseren. Als gebruikers kunnen zij bepaalde criteria hanteren bij de beoordeling van de kwaliteit van de gegevens. Deze kwaliteitsvereisten kunnen betrekking hebben op bijvoorbeeld de leveringstermijn van de gegevens, het algemene percentage van nonrespons, de representativiteit van de gegevens, het voorkomen van item nonrespons op bepaalde vragen, het niveau van sociaal-wenselijke antwoorden op bedreigende vragen, het aantal substantiële antwoorden op open vragen en het eventueel effect van de interviewer. Als de kwaliteit van deze elementen te wensen overlaat, kunnen de eindgebruikers klachten formuleren en/of eventueel geen beroep meer doen op de organisatie die instaat voor de dataverzameling.

Bovenstaande beschrijving bevat reeds enkele principes van TQM. In de loop van dit artikel zullen ook de andere aspecten aan bod komen. Refererend aan enkele belangrijke vertegenwoordigers van kwaliteitsmanagement, onder meer Deming, Juran en Crosby, onderscheiden Colledge en March (1993:157-159) de volgende kernelementen bij TQM:

- (a) De klant staat centraal. De drijvende kracht van kwaliteitsmanagement is immers de tevredenheid van de eindgebruiker.
- (b) Alle processen binnen een organisatie dragen direct of indirect bij aan de kwaliteit van het eindproduct.
- (c) Externe producenten voor de toelevering van bepaalde goederen of diensten worden beschouwd als een uitbreiding van de organisatie.
- (d) De permanente verbetering van alle processen binnen de productiecyclus is belangrijk om tegemoet te komen aan gewijzigde verwachtingen van de gebruiker, om nieuwe technologieën te kunnen integreren en om bepaalde processen efficiënter te laten verlopen. Preventie is immers beter – lees goedkoper – dan het corrigeren achteraf.
- (e) Met het oog op continue verbetering zijn kwaliteitsmaatstaven noodzakelijk. Dit veronderstelt dat de veranderingsprocessen op een of andere wijze kwantificeerbaar zijn.
- (f) Kwaliteitsmanagement is enkel effectief wanneer de top van het bedrijf daadwerkelijk betrokken is en wanneer TQM kan plaatsvinden binnen een geëigende organisatiestructuur.
- (g) Via training en een doorgedreven vorm van communicatie worden alle werknemers van hoog tot laag bewust gemaakt van de noodzaak om kwaliteit na te streven. Werknemers dienen erkend en beloond te worden wanneer bepaalde kwaliteitsmaatstaven worden bereikt.

3. DE DATAVERZAMELING ALS PRODUCTIEPROCES MET TOEPASSING VAN ELEMENTEN UIT TQM

Voor TQM vormt de organisatie die instaat voor de dataverzameling, in concreto een netwerk van interviewers, een cruciale schakel binnen het survey-onderzoek. Onderzoekers oefenen immers een maximale controle uit op de fases voorafgaand aan de

dataverzameling door de “regels van de kunst” te respecteren bij onder meer de opstelling van de vragenlijst en de steekproefsamenstelling. Bijgevolg treedt een interviewer-netwerk dan ook op als een “gebruiker” van de vragenlijst en de steekproef. De eigenlijke verzameling van de gegevens valt echter deels buiten de beïnvloedingssfeer van de onderzoekers daar de kwaliteit van het veldwerk afhankelijk is van de vaardigheid en de bereidheid van interviewers om tegemoet te komen aan hun verwachte taak en rol binnen de fase van de dataverzameling. Uiteraard zal de onderzoeker ook hier een zekere controle proberen te behouden door een aantal kwaliteitsvereisten te stellen aan de organisatie die instaat voor de dataverzameling. Bijgevolg vormt de fase van de dataverzameling op zich een productieproces. De hoofdpdracht van een interviewer-netwerk bestaat uit het verzamelen van geldige en betrouwbare data. Deze hoofdpdracht kan in verschillende deeltaken worden uitgesplitst: de recruitering en selectie van interviewers, de interviewers vertrouwd maken met de principes van de afname van een gestandaardiseerd interview (training), de evaluatie van de interviewers tijdens en na het veldwerk en na verschillende opdrachten (supervisie). Het verzamelen van kwaliteitsdata impliceert dat de verantwoordelijke van een dataverzamelingsorganisatie aan de onderscheiden deeltaken van een interviewernetwerk eveneens kwaliteitsvereisten kan koppelen. Het betreft de volgende basiseisen:

- De selectieprocedure krijgt het label kwaliteit wanneer interviewers zo geselecteerd worden dat zij permanent en vlot inzetbaar blijven.
- Het trainingsprogramma is kwaliteitsvol wanneer het potentiële interviewers de nodige interviewvaardigheden effectief aanleert en wel op die manier dat zij het aangeleerde gedrag adequaat blijven toepassen.
- De evaluatie na elk onderzoeksproject is kwalitatief hoogstaand wanneer de ‘performantie’ van elke interviewer grondig beoordeeld kan worden.
- Regelmatig terugkerende evaluatiemomenten zijn een vereiste om blijvende kwaliteit te garanderen.

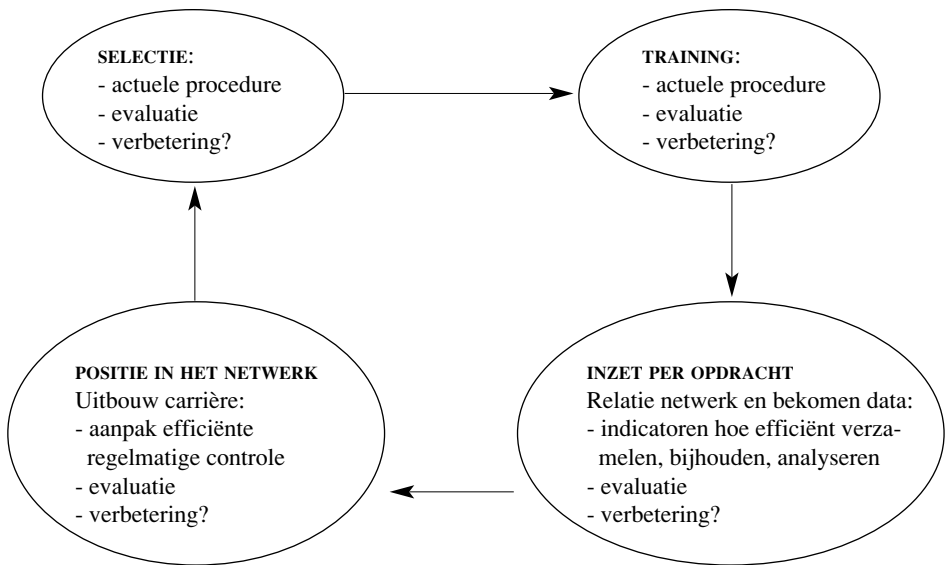
De basiseisen betreffende training, on-the-job training en supervisie van medewerkers blijft uiteraard niet beperkt tot een interviewernetwerk maar zijn uitermate belangrijk geworden in alle bedrijfstakken en dit zowel in Nederland als in België. Door de verbeterde economische situatie en de krapte in de aanbodzijde op de arbeidsmarkt is het momenteel in alle bedrijfstakken moeilijk om nieuwe krachten te werven.

Gegeven dat een dataverzamelingsorganisatie opteert voor een TQM-benadering stelt zich de vraag in welke mate de verschillende deeltaken van de organisatie beantwoorden aan de vooropgestelde kwaliteitsvereisten en of een bijsturing met het oog op een permanente kwaliteitsverbetering al dan niet wenselijk is. Het kunnen beantwoorden van deze vragen veronderstelt dat de verantwoordelijke van een interviewernetwerk over *kwantificeerbare maatstaven* beschikt. De uitwerking van dergelijke kwantificeerbare maatstaven vergt een systematische aanpak (Carton 1999). In navolging van de benadering van Biemer en Caspar (1994) en Demings motto “Plan-Do-Check-Act” lijkt het mij opportuun te opteren voor een stapsgewijze aanpak. Deze aanpak wordt voor elk van de vier deeltaken van een interviewernetwerk in schema 1 voorgesteld:

- In een eerste stap dient de aandacht uit te gaan naar het opstellen van een “status-questionis” van de vier taken (“plan” en “do”). In concreto houdt dit in dat voor wat betreft de selectie- en trainingstaak een beschrijving van de actuele situatie wordt gegeven. Met het oog op de kwaliteitsevaluatie van de interviewers per opdracht die-

nen bepaalde gegevens verzameld en geanalyseerd te worden. De vraag is dan hoe en welke gegevens met betrekking tot de uitvoering van de taken van de interviewers en hun relatie tot het netwerk zo efficiënt en toegankelijk mogelijk verzameld, bijgehouden en geanalyseerd kunnen worden. De vraag hoe de kwaliteitscontrole van het afgeleverde werk van de interviewers op regelmatige tijdstippen zo efficiënt mogelijk georganiseerd kan worden, komt in de vierde taak aan bod.

- Na het opstellen van de actuele situatie met betrekking tot deze vier deeltaken, stelt zich vervolgens in de tweede stap de vraag of de uitvoering van de taken beantwoordt aan de eis van kwaliteit (zie hoger basiseisen). Dit komt neer op een confrontatie tussen de huidige aanpak binnen het interviewernetwerk en de ideale norm van de uitvoering van de deeltaken (“check”).
- Na deze toetsing kunnen de resultaten in een derde stap gebruikt worden bij de formulering van voorstellen ter bevordering van de kwaliteit (“act”).



Schema 1. Uitwerking evaluatieprocedure voor een netwerk van interviewers.

Bij deze driestaps werkwijze mag het *procesmatig* karakter van de werking van een netwerk niet uit het oog verloren worden. De evaluatieprocedure dient rekening te houden met alle aspecten en loopt bijgevolg via selectie, over training naar evaluatie per opdracht met op het einde van de cyclus een beslissing over de positie van de interviewer. Deze verschillende evaluatiemomenten kunnen echter ook telkens teruggekoppeld worden naar het voorgaande niveau. Het procesmatig karakter uit zich nog op een tweede manier. Na het doorvoeren van aanpassingen in de uitvoering van de vier deeltaken met het oog op het bevorderen van de kwaliteit, start de evaluatiecyclus terug van vooraf aan. Dit wijst met andere woorden op het belang van “*continuous quality improvement*”.

Als case study illustreren we de toepassing van principes uit TQM bij twee belangrijke

deeltaken van een interviewernetwerk, de training en supervisie van interviewers. Voor elk van deze twee deeltaken volgt eerst een bondige evaluatie van de concrete praktijk binnen een actueel functionerend interviewernetwerk. Daarvoor maken we gebruik van gegevens van het Interuniversitair Steunpunt Politieke Opinie-onderzoek (ISPO, Departement Sociologie, KULeuven – België) dat over een eigen netwerk van interviewers beschikt.¹ Het ISPO-netwerk wordt ingezet op het verkiezingsonderzoek in Vlaanderen en op verzoek kunnen ook andere wetenschappelijke instellingen gebruik maken van het netwerk voor de afname van face-to-face interviews. Na de confrontatie tussen de actuele praktijk en vooropgestelde kwaliteitsnormen wordt stilgestaan bij mogelijke voorstellen tot verbetering. Deze verbeteringsprocessen vormen op zich een concretisering van “best and good practices” binnen het proces van dataverzameling (AAPOR 1997; Kaase 1999).

4. KWALITEITSEVALUATIE BIJ DE TRAINING VAN INTERVIEWERS

4.1. Inhoudelijke uitbouw van de training

Zoals Fowler en Mangione terecht opmerken: “Virtually no one planning a serious survey would question the value of at least some training for interviewers” (1990:106). Voor hun eerste opdracht volgen de interviewers van het ISPO een algemene interviewtraining van één dag. Deze bijeenkomst valt uiteen in een theoretisch en een meer praktisch gedeelte (Carton 1991). Tijdens het theoretische gedeelte wordt vooreerst het persoonlijke interview gesitueerd. Dit algemeen kader laat de interviewers toe een onderscheid te maken tussen verschillende methodes van onderzoek zodat het eigene van hun opdracht duidelijker wordt. Daarna komen twee belangrijke kenmerken van het persoonlijk interview aan bod: de steekproeftrekking en enkele kenmerken van een voorgestructureerde vragenlijst met onder meer het onderscheid tussen verschillende vraagtypes. Willen onderzoekers komen tot betrouwbare en geldige antwoorden dan dienen zij er zeker van te zijn dat de gestandaardiseerde vragenlijst op een gestandaardiseerde wijze wordt gebruikt door de interviewers. Het vervolg van het theoretisch gedeelte besteedt dan ook uitvoerig aandacht aan de principes die interviewers in acht dienen te nemen bij de afname van een gestandaardiseerd face-to-face interview. Daarbij wordt een onderscheid gemaakt tussen vier onderscheiden fases waarbij telkens de rol en de taak van zowel de interviewer als de respondent worden uitgediept:

- (1) De *voorbereiding op het veldwerk* met onder meer het versturen van introductiebrieven en het verwerven van een grondige kennis van de vragenlijst.
- (2) Het *eerste contact* waarbij het contacteren en het verkrijgen van medewerking een zeer belangrijke opdracht is voor de interviewer. Aandacht gaat naar enkele te volgen richtlijnen bij de eerste contactname en naar adequate interviewerreacties op bepaalde problematische situaties. Bijvoorbeeld hoe dient de interviewer te reage-

¹ Het Interuniversitair Steunpunt Politieke Opinie-onderzoek (ISPO) maakt deel uit van het Programma tot Valoriatie van Federale Sociaal-Economische Databanken, dat onder de bevoegdheid van de Federale Diensten voor Wetenschappelijke, Technische en Culturele Aangelegenheden valt. Prof. dr. J. Billiet is onderzoekspromotor en Prof. dr. M. Swyngedouw onderzoeksdirecteur.

ren wanneer de respondent weigert, niet geïnteresseerd is, geen tijd heeft, bijkomende vragen heeft of opwerpt dat hij/zij zich te oud, te ziek voelt om aan een dergelijk interview mee te werken?

(3) Interviewen omvat meer dan het louter stellen van vragen door de interviewer en het beantwoorden van vragen door de respondent (Loosveldt 1985:97). Tijdens de *afname van het interview* dient de interviewer vooreerst instructies te geven en de vraag te stellen waarbij een aantal richtlijnen in acht worden genomen. Daarna volgt de reactie van de respondent die zowel adequaat (voorbeeld: selectie van één van de voorziene antwoordmogelijkheden) als inadequaat (bijvoorbeeld: weigering om te antwoorden) kan zijn. De interviewer zal dit gedrag beoordelen en reageren in functie van de beoordeling. Deze interviewerreacties zijn bijgevolg complementair aan de instructies. Zij hebben eveneens als doel de taak van de respondent aan te leren en aan te duiden in welke mate hij/zij aan de verwachtingen voldoet. Tot slot gaat de interviewer over tot de registratie van het antwoord.

(4) In de laatste fase, het *afsluiten van het interview*, wordt de respondent bedankt voor zijn/haar medewerking, wordt nog stilgestaan bij eventuele vragen en stellen de interviewers hun zaken terug op orde vooraleer een volgend interview aan te vatten.

Tijdens het tweede dagdeel van de training wordt de theorie geconfronteerd met de praktijk. Met behulp van video-materiaal leren de interviewers via concrete opdrachten verschillende vormen van zowel adequaat als inadequaat interviewer- en respondentengedrag herkennen, benoemen, beoordelen en zo nodig te verbeteren. Gebruikmakend van rollenspelen waarbij de getrainden zowel de rol van interviewer als respondent opnemen, worden tot slot verschillende interviewsituaties ingeoeffend.

Na deze algemene interviewtraining volgt voor de aanvang van elk nieuw onderzoek een meer specifieke training van ongeveer drie uur. Tijdens deze training wordt onder meer achtergrondinformatie gegeven over het onderzoek, wordt de vragenlijst ingevuld door de interviewers en achteraf in detail besproken, worden richtlijnen betreffende de contact- en vervangingsprocedure gegeven en komen praktische afspraken aan bod.

4.2. Evaluatie van de training

In principe voor de eerste interviewopdracht dient elke geselecteerde interviewer, ongeacht eventuele vroegere interviewervaring, de algemene trainingsdag te volgen. We stellen vast dat ongeveer één op vijf van de interviewers deze training niet heeft bijgewoond. Dit kan verklaard worden doordat in bepaalde omstandigheden er minder strikte eisen aan de interviewers gesteld worden met betrekking tot de training. De tendens lijkt te zijn dat naarmate het moeilijker wordt om bepaalde steekproefadressen toe te wijzen en om desalniettemin de dataverzameling binnen de vooropgestelde periode af te ronden, er in het algemeen minder eisen aan de interviewers worden gesteld. Dit wijst alvast op de noodzaak van een voldoende grote reserve van getrainde interviewers verspreid over Vlaanderen. Ook dit is een punt van grote zorg binnen de Nederlandse context.

Daar de voorgestelde algemene interviewtraining in feite neerkomt op een opleiding stelt zich de vraag hoe de waarde, het effect van een opleiding geëvalueerd kan worden. In deze context staat de evaluatie van het product centraal, met name de algemene interviewtraining en het opleidingseffect ervan. Effecten kunnen optreden op het niveau van het leren, op het niveau van het werkgedrag en op het niveau van de organisatie (Spitholt 1989:10-13). Op het leerresultatenniveau wordt nagegaan in hoeverre de

vooraf beoogde leerdoelen bereikt zijn. De onderrichte kennis en vaardigheden worden op een of andere wijze gemeten vooraleer cursisten deze toepassen in het dagdagelijkse werk. Een antwoord op de vraag in hoeverre het gedrag van de ex-cursisten veranderd is in de beoogde richting wordt bekomen via de evaluatie op het werkgedragniveau. Op het organisatieniveau staat de vraag centraal of een door de opleiding veranderd werkplekgedrag daadwerkelijk leidt tot verbetering van de output van een bedrijf.

De vraag in hoeverre de interviewers na het volgen van een training de vooropgestelde leerdoelen bereikt hebben, komt overeen met de evaluatie op het leerresultatenniveau. Binnen de opzet van de training heeft dit type van evaluatie niet plaatsgevonden. Gegeven de vrij omvangrijke groepen van deelnemers (20 tot 30) is het bijvoorbeeld niet mogelijk om elke deelnemer de rol van interviewer te laten opnemen tijdens het rollenspel. Wel is naar de reacties van de deelnemers over de gevolgde training gepeild. In het algemeen wordt de algemene interviewtraining vrij positief beoordeeld. Naargelang het ervaren of onervaren interviewers betreft komt evenwel een genuanceerder beeld tot uiting. Onervaren interviewers evalueren de training positiever dan de ervaren interviewers maar stellen de opportuniteit van het samen trainen van ervaren en onervaren interviewers in vraag.

De conclusie in onze situatie is dat de huidige trainingsopzet dient bijgestuurd te worden. Het is wenselijk dat de training tegemoet komt aan de verwachtingen van een heterogeen publiek en dat er een evaluatie op het leerresultatenniveau kan plaatsvinden.

5. KWALITEITSEVALUATIE BIJ DE SUPERVISIE VAN INTERVIEWERS

Nadat interviewers de algemene interviewtraining gevolgd hebben en aan een concrete opdracht beginnen is het belangrijk een vorm van permanente supervisie te ontwikkelen. Deze supervisie kan plaatsvinden zowel tijdens als na de uitvoering van een opdracht. Bij de uitwerking van een evaluatieprocedure voor wat betreft de taakuitvoering door de interviewers is het belangrijk om rekening te houden met twee verschillende niveaus. Enerzijds is er het taakinhoudelijke niveau en anderzijds het niveau van de manier waarop de interviewer zich verhoudt tot het netwerk. Voor elk van deze evaluaties geven we eerste een korte omschrijving van de aanpak en de resultaten van de concrete praktijk binnen het ISPO-netwerk.

5.1. *Het organisatorische niveau*

Op dit niveau wordt het functioneren van de interviewers in het licht van het al dan niet naleven van bepaalde formele afspraken geëvalueerd. Een netwerk kan de fase van de dataverzameling pas tot een goed einde brengen wanneer het onvoorwaardelijk kan rekenen op de inzet van zijn interviewers. De evaluatie van de relatie van de interviewers tot het netwerk is verbonden met en geoperationaliseerd op basis van gegevens omtrent een aantal verwachtingen ten aanzien van de interviewers tijdens verschillende onderdelen van het organiseren van een dataverzameling. De evaluatie op dit meer *organisatorische niveau* wordt gekenmerkt door bijvoorbeeld aanwezigheid op de specifieke training en het in acht nemen van de veldwerkperiode, het correct bijhouden van eventuele extra formulieren, het afwerken van een vooropgesteld aantal interviews. Interviewers die zich strikt houden aan deze afspraken zijn zeer waardevolle elementen

binnen een netwerk en krijgen dan ook een positieve beoordeling. Enkel een correcte naleving van afspraken kan garant staan voor een efficiënte en effectieve afronding van een onderzoeksproject. Een negatieve beoordeling geldt voor die interviewers die zich niet zo strikt houden aan een aantal evidente, noodzakelijke formele afspraken.

Globaal genomen, over alle 42 projecten heen gedurende de periode 1991-1996, halen 16 % van de interviewers niet het minimum vooropgestelde niveau van kwaliteit. Met andere woorden een groep van meer dan één op zes van de interviewers stelt zich niet correct op tegenover het netwerk. Een vraag die hier dan ook onmiddellijk bij aansluit is die naar de kwaliteit van de verzamelde gegevens door deze en andere meer correcte interviewers. Dit brengt ons bij het tweede luik van de 'interviewerperformantie': het taakinhoudelijke niveau. Dit luik wordt wat meer in detail uitgewerkt.

5.2. Het taakinhoudelijke niveau

Een concretisering van de kwaliteit op het taakinhoudelijke niveau is mogelijk aan de hand van de verschillende deeltaken van de interviewer. Een eerste en zeer belangrijke taak van de interviewer is het contacteren van de respondenten en het verkrijgen van hun medewerking. Een eerste domein van de evaluatie van het feitelijk functioneren van de interviewers vormt bijgevolg het niveau van de behaalde (non)respons. Een ander en eveneens belangrijk aspect van de interviewertaak is het bekomen van betrouwbare en geldige antwoorden. De interviewer bekomt de antwoorden door het stellen van concreet gedrag. De bekomen antwoorden enerzijds en het gedrag anderzijds sluiten aan bij twee klassieke benaderingswijzen van datakwaliteit. Het betreft met name een kwaliteitsbenadering op het niveau van de geregistreerde antwoorden en een kwaliteitsbenadering op het niveau van de gedragscontext waarbinnen die geregistreerde antwoorden ontstaan. Binnen elk van die onderscheiden aspecten van datakwaliteit – (non)respons, de bekomen informatie en het effectieve interviewergedrag – stelt zich de vraag of sommige interviewers beter presteren dan andere of, meer algemeen verwoord, er sprake is van interviewervariantie. Het verklarend onderzoek naar deze drie onderscheiden aspecten van datakwaliteit heeft evenwel uitgewezen dat niet alleen de 'interviewer', bepaald door zijn kenmerken en gedrag, een verklarende factor vormt, maar ook de respondent, het type vragen en zo meer. De analyse van de evaluatie van de ingezette interviewers moet bijgevolg rekening houden met de actor respondent, met kenmerken van het survey design en met kenmerken van het meetinstrument.

5.2.1. Interviewers en behaalde (non)respons

In de overzichtsstudie van Groves en Couper (1998) is er sprake van een tweestapsproces van survey participatie. Interviewers dienen in eerste instantie in contact te treden met de respondenten en eenmaal een persoonlijk contact tot stand is gekomen, dienen de interviewers de medewerking te bekomen. De evaluatie van de interviewers met betrekking tot de behaalde (non)respons kan bijgevolg gebeuren vanuit deze twee verschillende invalshoeken.

Vooreerst stelt zich de vraag in welke mate de interviewers de richtlijnen betreffende het in contact treden met de respondenten naleven. Het respecteren van de ISPO-richtlijnen kan beschouwd worden als een verplichte procedure die door elke interviewer dient gevolgd te worden en dit ongeacht het type respondenten dat de interviewers toegewezen krijgen. Deze evaluatie is mogelijk door informatie verkregen uit contactformulieren die de interviewers voor al de contacten met hun respondenten dienen bij te

houden. Een belangrijke vaststelling bij de vier geselecteerde projecten (Gezondheids-enquête 1994; Verkiezingsonderzoek 1995; Nul- en effectpeiling naar de kennis van de Vlamingen over de federale structuren respectievelijk 1995 en 1996) is alvast dat bepaalde interviewers met een minimum aan inspanningen een maximum aan resultaat boeken. Deze succesvolle interviewers realiseren meestal tijdens het eerste contact een interview of een afspraak die tijdens het tweede contact uitmondt in een interview. Niet alle interviewers respecteren de richtlijn om bij het niet-thuisvinden van de respondent de daarnavolgende contacten te spreiden over verschillende dagen en tijdstippen. En een niet onbelangrijke groep van interviewers neemt bij het eerste contact steeds telefonisch contact op, dit in tegenstelling met de voorkeur voor persoonlijke eerste contacten aan de deur. Deze verschillende evaluaties van het contactproces vertonen onderling weinig systematische samenhang. Verder is er evenmin een systematische samenhang met sociaal-demografische interviewerkenmerken, met het feit of de interviewers al dan niet de algemene interviewtraining hebben gevolgd en met de evaluatie betreffende de relatie tot het netwerk. In die zin laten de achtergrondkenmerken van de interviewers niet toe om goede interviewers te selecteren voor dit aspect. Of het niet strikt naleven van de richtlijnen ook effect uitoefent op het bekomen van survey participatie komt verder aan bod.

In de tweede invalshoek staat het modeleren van de kans op geen contact versus contact en van de kans op weigering versus medewerking centraal. In twee van de vijf geselecteerde projecten (Verkiezingsonderzoek 1991 en 1995) verschillen interviewers significant naargelang het meer of minder bekomen van de reden 'niet te contacteren' en/of de reden 'weigering'. De interviewerkenmerken die voor een reductie zorgen van de vastgestelde interviewervariantie bij de indicator 'weigeringen' liggen gedeeltelijk in de lijn van de onderzoeksliteratuur (Campanelli et al. 1997; Groves en Couper 1998). Geen enkele van de rolafhankelijke interviewerkenmerken zoals leeftijd, geslacht, onderwijsniveau en werksituatie oefenen een effect uit. Het zijn rolafhankelijke kenmerken, met andere woorden, kenmerken die gerelateerd zijn aan de manier waarop de interviewers hun taak uitvoeren, die het meeste effect sorteren. Bij interviewers die overwegend persoonlijk contact opnemen bij het eerste contact stijgt de kans op weigering. Hieruit volgt dat interviewers die overwegend telefonische eerste contacten opnemen niet per definitie minder goede resultaten bekomen. We vinden evenwel geen effect van de vroegere interviewervaring noch van interviewerverwachtingen betreffende de medewerking van de respondenten. Een beperkt aantal interviewers is verantwoordelijk voor de vastgestelde interviewervariantie. Merk tot slot op dat de component 'niet te contacteren' met gemiddeld 5 % over de vijf projecten bijna verwaarloosbaar is in vergelijking met de component 'weigeringen' 22 %. Dit type van nonrespons verdient dan ook nadere aandacht (zie 6.1).

5.2.2. Interviewers en de kwaliteit van de verzamelde informatie

Voor de evaluatie van de kwaliteit van de verzamelde data door de interviewers zijn terug kwantificeerbare maatstaven nodig. Deze maatstaven hebben betrekking op de hoeveelheid en aard van de geregistreerde gegevens.

De kwaliteitsnorm – afwezigheid van interviewereffecten en dit onder controle van relevante respondentkenmerken zoals leeftijd, geslacht, onderwijsniveau, interesse in het onderzoeksonderwerp – wordt in de beide verkiezingsonderzoeken niet gehaald. In het onderzoek naar de kennis van de federale structuren, de nul- en effectpeiling, wordt de norm iets beter gehaald.

In de vijf projecten wordt de interviewervariantie veroorzaakt door een beperkt aantal interviewers.

De uitgevoerde evaluatie levert weinig concrete aanknopingspunten voor het hanteren van andere criteria bij de selectie van interviewers. De interviewerkenmerken die verantwoordelijk werden geacht voor de vastgestelde interviewervariantie blijken immers niet steeds dezelfde te zijn, niet in staat om de effecten weg te verklaren en niet steeds conform de verwachtingen. In tegenstelling met de algemene conclusie uit de onderzoeksliteratuur is het effect van sociaal-demografische kenmerken op de kwaliteit van de gegevens niet altijd volledig afwezig. We vinden effecten van leeftijd, geslacht, opleiding en/of werksituatie in de beide verkiezingsonderzoeken. De resultaten voor wat betreft de rolafhankelijke interviewerkenmerken weerspiegelen beter die van de onderzoeksliteratuur: bepaalde rolafhankelijke kenmerken hebben soms een effect maar het is moeilijk om er een vast patroon in te herkennen. Zo vinden we soms een effect van interviewervaring voor de opname in het netwerk en soms een effect van opgebouwde ervaring tijdens opname in het netwerk. Wat betreft de interviewerwachtingen omtrent de moeilijkheid van afname van bepaalde onderdelen van de vragenlijst vinden we zowel voorbeelden van een “self-fulfilling” als een “self-denying prophecy”. Bij twee attitudeschalen stellen we ook een effect van de eigen opinie van de interviewer vast.

Deze vaststellingen sluiten uiteraard de mogelijkheid niet uit dat andere kenmerken verantwoordelijk kunnen zijn voor de gevonden interviewervariantie, met name persoonlijkheidskenmerken en kenmerken die te maken hebben met het effectieve interviewergedrag. Dit impliceert dat de verzameling van dit type van gegevens expliciet in het onderzoeksdesign dient opgenomen te worden.

Tot slot, de evaluatie op het organisatorische niveau vertoont geen systematische samenhang met de verschillende geconstrueerde kwaliteitsindicatoren. Dit vormt een aanduiding dat het takenpakket van de interviewers heel wat verschillende dimensies omvat die elk op zich dienen geëvalueerd te worden.

5.2.3. Interviewers in het vraag-antwoord proces

De aangeleerde richtlijnen tijdens de algemene interviewertraining vormen het uitgangspunt voor de uitwerking van kwantificeerbare maatstaven met het oog op het evalueren van het effectieve interviewergedrag. Niettegenstaande de eerder beperkte opzet en toepassing van het codeerschema laat de gehanteerde werkwijze toe een antwoord te geven op de vraag of de interviewers de standaard richtlijnen bij de afname van een interview respecteren.

Dit is niet het geval. Specifieke instructies worden niet steeds op een gestandaardiseerde wijze gegeven. De vragen worden niet altijd precies en volledig gelezen zoals ze opgenomen zijn in de vragenlijst. Het is echter belangrijk een differentiëring naargelang vraagtype te maken. Batterijen met uitspraken worden in zowat alle sequenties steeds adequaat voorgelezen. Dit is minder het geval bij de open en gesloten vragen waarbij de antwoordmogelijkheden vervat zijn in de vraag of waarbij verwezen wordt naar het gebruik van een antwoordkaart. Bij het herhalen of accepteren van een antwoord gaan interviewers soms inadequaat te werk en ook het doorvraagdgedrag (probing) is niet steeds neutraal te noemen.

Kortom, de huidige trainingsaanpak standaardiseert en optimaliseert het interviewergedrag niet volledig.

6. VERBETERINGSPROCESSEN

Dit bondig overzicht van de uitgevoerde evaluaties bij twee verschillende deeltaken van een interviewernetwerk, training en supervisie, laat alvast toe te besluiten dat de vooropgestelde kwaliteitsmaatstaven niet steeds gehaald werden. De filosofie van integrale kwaliteitszorg indachtig dienen bijgevolg verbeteringsprocessen te worden geïntroduceerd. Met als basis de uitgewerkte evaluatieprocedure binnen het ISPO-interviewernetwerk, kunnen volgende aanbevelingen met betrekking tot de training en supervisie van interviewers geformuleerd worden.

6.1. Training van interviewers

De centrale vraag is hoe een interviewtraining uitgebouwd dient te worden om tegemoet te komen aan de eis van effectiviteit? Of met andere woorden, hoe kunnen de nodige interviewvaardigheden aangeleerd worden zodanig dat interviewers het aangeleerde gedrag adequaat blijven toepassen?

- Eerst en vooral dient het principe van het verplicht volgen van de algemene interviewtraining vóór een eerste opdracht strikter te worden nageleefd.
- Een gedifferentieerde trainingsaanpak die tegemoet komt aan de verwachtingen en het ervaringsniveau van een heterogeen publiek is een tweede aanbeveling. Bij de ervaren interviewers dient de training te vertrekken vanuit hun eigen interviewpraktijk. Bij de onervaren interviewers wordt bij voorkeur ‘omgekeerd’ te werk gegaan door de theoretische toelichting te laten uitmonden in praktische oefeningen.
- Het is belangrijk om in de beide types van training een evaluatie op het leerresultatenniveau in te passen. Zo is het mogelijk om elke interviewer bij zijn eerste opdracht van nabij op te volgen waarbij de interviewer zijn eerste (twee of drie) interview(s) volledig op band opneemt en hierover uitvoerige feedback krijgt. Conform de Nederlandse situatie kan ook gedacht worden aan de certificatie van interviewers (UMT).
- Tijdens beide trainingstypes dient alvast meer aandacht uit te gaan naar het aspect ‘bekomen van survey participatie’. De focus van onder meer Morton-Williams (1993), Campanelli et al. (1997), Groves en Couper (1998) op wat zich werkelijk afspeelt tussen de interviewer en de respondent tijdens de interactie aan de deur en de belangrijke impact van het interviewgedrag tijdens deze meestal korte introducties, bieden alvast heel wat concrete aanknopingspunten voor een meer adequate voorbereiding van de interviewers in het verkrijgen van medewerking aan een survey.

In verband met deze laatste aanbeveling zijn naar aanleiding van het ISPO-verkiezingsonderzoek 1999 reeds een aantal initiatieven genomen. Zo volgde de helft van de nieuw gerecruteerde interviewers een specifieke trainingsmodule in het bekomen van survey participatie. Deze training is gebaseerd op het werk van Morton-Williams (1993), Groves en McGonagle (1998), National Centre for Social Research (1999). Er wordt onder meer uitvoerig ingegaan op de concepten “maintaining interaction” en “tailoring”. Daarnaast wordt veel aandacht geschonken aan mogelijke overtuigingsstrategieën om weigerachtige respondenten alsnog over te halen tot medewerking. Tijdens de training oefenen de interviewers aan de hand van rollenspelen verschillende contactsituaties in en tot slot gaat heel wat aandacht naar het motiveren van interviewers om een zo groot mogelijke respons te bekomen. Momenteel zijn nog geen gegevens beschikbaar die uitwijzen of de training al dan niet effect heeft geresorteed.

Dat het probleem van nonrespons ook bij Nederlandse marktonderzoekbureaus hoog op de agenda staat, bewijst de bijzondere aandacht van het Voorlichtingscentrum Uw Mening Telt voor het werken “aan de deur en op locatie”. In 1999 heeft UMT een nieuwe module in het opleidingstraject van interviewers voor markt- en opinieonderzoek ontwikkeld. Deze “contactmodule” is bedoeld voor alle nieuwe interviewers en wil de contactuele kwaliteiten van de interviewers verbeteren om daarmee de respons te verbeteren. In de training ligt de nadruk op het eerste contact met de respondent. Hoe spreek je iemand aan, welke rol speelt de interviewer daarin, welke weerstanden kan je verwachten en hoe overwin je die, zijn enkele aspecten die aan bod komen.

6.2. *Supervisie van interviewers*

Hoe de verzamelde informatie in verband met de kwaliteit van de afgewerkte interviews en de relatie tot het netwerk gebruikt kan worden in functie van het opzetten van een kwaliteitsverbetering is de vraag die centraal staat bij de supervisie van de interviewers per project en op achtereenvolgende projecten. Opnieuw enkele aanbevelingen voor de concrete praktijk:

- Vooreerst is het belangrijk om het onderscheid tussen de twee dimensies van de ‘interviewerperformantie’ – evaluatie op het organisatorische en op het taakinhoudelijke niveau – te blijven handhaven.
- Ten aanzien van de interviewers die zich niet correct houden aan de formele afspraken dient strenger te worden opgetreden.
- In tegenstelling met de ‘performantie’ op de organisatorische dimensie kan bij de evaluatie op het inhoudelijk vlak minder snel op de bal worden gespeeld. Dit type van evaluatie kan immers pas plaatsvinden na de levering van het volledige databestand.
- Voor wat betreft de evaluatie van de behaalde (non)respons is het wenselijk dat de huidige contactformulieren zowel naar inhoud als naar toepassing worden uitgebreid (zie Campanelli et al. 1997).
- Na de afronding van het veldwerk moet sneller feedback betreffende de drie onderscheiden dimensies van datakwaliteit, met name (non)respons, kwaliteit van de bekomen data en het effectieve interviewergedrag, aan de interviewers worden bezorgd. Deze aanpak verhoogt meteen de betrokkenheid van de verschillende actoren op het netwerk. Bij onderzoek van het eigen ‘huis’ is dit een haalbare kaart maar de praktijk zal moeten uitwijzen of dit eveneens realiseerbaar is bij opdrachten voor andere wetenschappelijke instellingen.
- Interviewers die goed presteren kunnen extra vergoed worden. Interviewers die het minder goed doen dienen bijgestuurd te worden of desnoods uit het netwerk te worden gezet.
- Met het oog op het opvolgen van de datakwaliteit doorheen de tijd is het verder wenselijk om de verschillende soorten van evaluatiegegevens per project en per interviewer op een gestandaardiseerde en uniforme wijze bij te houden.

7. “BEST AND GOOD PRACTICES”

Naast de eigenlijke data die verzameld worden via het persoonlijke interview met behulp van een gestandaardiseerde vragenlijst, worden dus ook heel wat data over de

data verzameld. Deze metadata laten niet alleen toe om de kwaliteit van de verzamelde data te analyseren maar ook om de kwaliteit van de processen die samengaan met het verzamelen van de data te evalueren (Dippo 1997). Het verzamelen van deze metadata met het oog op de hanteer- en beheersbaarheid vergt dan ook een systematische werkwijze. De procedures die gebruikt worden bij het bekomen van kwaliteitsdata kunnen dan ook verder geconcretiseerd worden in een draaiboek waarin duidelijk wordt wanneer en welke gegevens worden verzameld tijdens de uitvoering van het veldwerk van een bepaald project. Deze procedures komen dan ook neer op “best and good practices” binnen het proces van dataverzameling (AAPOR 1997; Kaase 1999; Morganstein en Marker 1997).

Ter illustratie geven we de “best and good practices” voor markt- en publiek opinieonderzoek van de American Association for Public Opinion Research (AAPOR).

- Stel specifieke doelstellingen voor het onderzoek op.
- Overweeg alternatieve onderzoeksmethodes vooraleer te kiezen voor een survey voor de uitvoering van de dataverzameling.
- Selecteer steekproeven die de onderzoekspopulatie vertegenwoordigen.
- Maak gebruik van onderzoeksdesigns die kosten ten opzichte van baten – geringe foutenmarge – afwegen.
- Zorg ervoor dat de vraagverwoording van de gemeten concepten overeenstemmen met de leefwereld van de onderzoekspopulatie.
- Test vragenlijsten en procedures op voorhand uit om a-priori problemen te detecteren.
- Train de interviewers zorgvuldig in interviewtechnieken en het onderwerp van het onderzoek.
- Voorzie kwaliteitschecks bij elke fase van het onderzoek.
- Streef naar een zo hoog mogelijke medewerking bij de respondenten en dit binnen de grenzen van het ethisch haalbare.
- Maak gebruik van statistische analyse- en rapporteringstechnieken die geschikt zijn voor de verzamelde data.
- Ontwikkel anonimiteitsgaranties voor de respondenten.
- Maak alle gehanteerde methodes van het onderzoek bekend met het oog op evaluatie en replicatie.

Het hanteren van dergelijke beste praktijken voor markt- en sociaal-wetenschappelijk onderzoek vinden we uiteraard niet alleen terug bij de onderzoeksbureaus die deel uitmaken van AAPOR. Zo dienen de leden van de Nederlandse Vereniging van Marktonderzoeksbureaus zich eveneens te houden aan de VMO-kwaliteitseisen voor kwantitatief en kwalitatief marktonderzoek. Deze eisen hebben betrekking op zowel het onderzoeksvoorstel als op de rapportering. Meer bepaald betreft het kwaliteitseisen met betrekking tot de methode (o.a. omschrijving, argumenten voor methodekeuze), de steekproef (o.a. bruto en (beoogde) netto steekproef, beschrijving van de populatie, wijze van steekproeftrekken, responsbevorderende maatregelen), het veldwerk (o.a. lengte vragenlijst, het aantal in te zetten interviewers, wijze van instructie, kwaliteitscontrole, interviewperiode), de verwerking, de rapportage en de condities waaronder het onderzoek wordt uitgevoerd.

Uit dit overzicht komt tot uiting dat de positie van een organisatie die instaat voor de

dataverzameling slechts één, maar wel een belangrijke schakel vormt binnen het survey proces. Op basis van de uitgevoerde evaluatie bij een concreet functionerend netwerk hebben we enkele aanbevelingen voor beste praktijken kunnen formuleren. Deze beste praktijken kunnen slechts in stand worden gehouden en continu verbeterd worden wanneer zij kaderen binnen een TQM-benadering. Vandaar dat we tot slot terug aanknopen bij het concept TQM en aangeven hoe het idee van kwaliteitsbenadering- en bewaking een bruikbare invalshoek vormt voor de concrete praktijk van een interviewernetwerk.

8. TOTAL QUALITY MANAGEMENT: EEN BRUIKBAAR CONCEPT VOOR DE CONCRETE PRAKTIJK VAN EEN INTERVIEWERNETWERK

Uitgaande van de taken van de interviewer bij de afname van een face-to-face interview hebben we verschillende kwaliteitsindicatoren geconstrueerd die toelaten het werk van de interviewers binnen het kader van een netwerk te evalueren. Deze procedure kan uiteraard ook toegepast worden bij andere vormen van dataverzameling, bijvoorbeeld bij CAPI en CATI. Vooraleer een definitief oordeel over de kwaliteit van een bepaalde interviewer te vellen, dient de interviewer in de verschillende fases van het proces van de dataverzameling beoordeeld te worden. Dit duidt op het *procesmatig karakter* van de evaluatie en hoe de kwaliteit op het ene niveau verweven kan zijn met de kwaliteit op een ander niveau. Dit type van evaluatiegegevens betreffende verschillende aspecten van de taak van de interviewer, kan per project worden bijgehouden en geanalyseerd. Op basis hiervan kunnen verbeteringsprocessen worden geïntroduceerd die uiteraard op hun beurt opnieuw zullen dienen te resulteren in een evaluatie. Het aspect van “continuous quality improvement” is hier dan ook zeker op zijn plaats. Een dergelijke analyse is evenwel slechts mogelijk als er bepaalde randvoorwaarden vervuld zijn. Zo dient elke interviewer alle contactpogingen met zijn/haar respondenten bij te houden op een daartoe voorzien contactformulier. Er dienen niet alleen gegevens beschikbaar te zijn van zowel respondenten en van personen die niet willen meewerken, maar ook van de interviewers. Vooraleer de analyse van start kan gaan dient de verzamelde informatie gecodeerd te worden. Kortom, andere elementen waaraan TQM aandacht besteedt, zijn eveneens van toepassing bij de uitwerking van een degelijke evaluatieprocedure:

- Vooreerst het idee van “total employee involvement”. *Werknemers* op alle niveaus dienen in het proces *betrokken* te worden zodanig dat ieder zijn *verantwoordelijkheid* draagt. Inderdaad, de analyse kan pas uitgevoerd worden wanneer elke interviewer zijn taak naar behoren vervult, bijvoorbeeld bij elke contactpoging het contactformulier aanvullen. Contactformulieren met onvolledige informatie worden immers uit de analyse weggelaten waardoor de evaluatie voor een stuk wordt uitgehold. De betrokkenheid van interviewers op het proces van dataverzameling blijkt onder meer uit de training die ze bijwonen vooraleer een interviewopdracht aan te vatten.
- Naast de algemene interviewtraining wordt voor elk nieuw onderzoek een specifieke training opgezet. Tijdens het veldwerk mogen de interviewers niet aan hun lot worden overgelaten, een permanente supervisie is wenselijk. Zo kunnen de eerste afgewerkte vragenlijsten nauwkeurig geëvalueerd worden met schriftelijke feedback voor de interviewers waardoor het mogelijk is om fouten nog tijdens het veldwerk te corrigeren. Tot slot kan ook het meedelen van de behaalde resultaten en de voornaamste onderzoeksconclusies tot een blijvende betrokkenheid van de interviewers bijdragen.

Op deze manier wordt ook tegemoet gekomen aan het aspect *continue opleiding*. Wanneer interviewers op regelmatige tijdstippen geëvalueerd worden, kan dit aanleiding geven tot bijvoorbeeld een extra training rond bepaalde aspecten van het interview. Zo kunnen interviewers die relatief veel weigeringen bekomen in vergelijking met andere interviewers een extra training volgen in verband met het bekomen van survey participatie. Het opleiden van interviewers is dus geen eenmalige gebeurtenis maar vormt een continu proces.

- *Erkenning* en *beloning* volgen wanneer werknemers verbetering van kwaliteit nastreven. Zo zouden interviewers die een zeer goede respons behalen beter vergoed kunnen worden.
- Wil het invoeren van kwaliteitsmanagement een kans op slagen hebben dan is het *engagement* van de *topverantwoordelijken* cruciaal. Alleen wanneer het topmanagement zichtbaar gewonnen is voor het invoeren van kwaliteit kan deze aanpak vruchten opleveren. Dit geldt zeker voor de omgeving waarbinnen het ISPO-interviewer-netwerk opgericht werd. Binnen het Departement Sociologie K.U.Leuven is steeds veel aandacht uitgegaan naar de kwaliteit van survey data (Billiet et al. 1984; 1988; Billiet 1993; Loosveldt 1985).
- De verwachtingen betreffende het verzamelen van kwaliteitsdata dient ook gecommuniceerd te worden naar de interviewers. Dit kan zoals we gezien hebben, via training en supervisie. Dat supervisie en het nauwkeurig ‘monitoren’ van het werk van de interviewers inderdaad heel belangrijk is, blijkt uit de studie van de Leeuw en de Heer (2000). In landen die interviewers strikt ‘monitoren’, is zowel het percentage niet-bereikten als het percentage weigeringen lager in vergelijking met landen die deze veldwerkstrategie niet toepassen. Verder is het ook belangrijk dat tussen de verantwoordelijke van een onderzoeksproject, van een interviewernetwerk en de gebruikers van de verzamelde data een goede uitwisseling bestaat met betrekking tot het kwaliteitsniveau van de data. Pas wanneer deze *communicatie* bestaat kan op een vlotte wijze worden ingespeeld op eventuele tekortkomingen en kunnen verbeteringsprocessen worden ingevoerd.
- Uit dit alles blijkt duidelijk de noodzaak van een eigen *organisatiestructuur* waarbinnen kwaliteitsmanagement werkelijkheid kan worden. Voor het verzamelen en analyseren van gegevens met betrekking tot de kwaliteit van de interviewers en het uitwerken van verbeteringsprocessen dienen nu eenmaal middelen en personeel voorzien te worden.

LITERATUUR

AAPOR, 1997. Best practices for survey and public opinion research.

<http://www.aapor.org/best/index.html>.

Biemer, P. and R. Caspar, 1994. Continuous quality improvement for survey operations: Some general principles and applications. *Journal of Official Statistics* 10, 3, 307-326.

Billiet, J. 1993. Ondanks beperkt zicht: Studies over waarden, ontzuiling en politieke veranderingen in Vlaanderen. Brussel-Leuven: VUBpress-SOI/KULeuven.

Billiet, J., Loosveldt, G. and L. Waterplas, 1984. Het survey-interview onderzocht. Effecten van het ontwerp en gebruik van vragenlijsten op de kwaliteit van de antwoorden. Leuven: SOI/KULeuven.

Billiet, J., Loosveldt, G. and L. Waterplas, 1988. Response-effecten bij survey-research in het Nederlands taalgebied. Leuven: SOI/KULeuven.

- Boomsma, M., 1992. "Kwaliteitscertificatie van marktonderzoek". In: *Jaarboek Nederlandse Vereniging voor Marktonderzoek en Informatiemanagement*, pp.9-19.
- Campanelli, P., Sturgis, P. and S. Purdon, 1997. *Can you hear me knocking: An investigation into the impact of interviewers on survey response rates*. London: Survey Methods Centre at SCPR.
- Carton, A., 1991. *Over interviewen gesproken. Handleiding bij de training van interviewers*. Leuven: ISPO/SOI/K.U.Leuven.
- Carton, A., 1999. *Selectie, training, en evaluatie van interviewers binnen een interviewernetwerk*. Leuven: Garant (proefschrift).
- Colledge, M. and M. March, 1993. *Quality management: Development of a framework for a statistical agency*. *Journal of Business and Economic Statistics* 11, 2, 157-165.
- De Leeuw, E. and W. de Heer, 2000. "Trends in household survey nonresponse: A longitudinal and international comparison". In: R.M. Groves et al (eds.), *Household Survey Nonresponse*. New York: John Wiley (in print).
- Dippo, C., 1997. "Survey measurement and process improvement: Concepts and integration". In: L. Lyberg et al. (eds.), *Survey measurement and process quality*. New York: John Wiley, pp.457-471.
- Fowler, F.J. Jr. and T.W. Mangione, 1990. *Standardized survey interviewing: Minimizing interviewer-related error*. Newbury Park: Sage Publications.
- Groves, R.M. and M.P. Couper, 1998. *Nonresponse in household interview surveys*. New York: John Wiley.
- Groves, R.M. and K. McGonagle, 1998. *A theory-guided training protocol regarding survey participation: Replicated experiments in production settings*. Paper presented at the 9th International Workshop on Household Survey Nonresponse, Bled (Slovenia).
- Kaase, M., 1999. *Deutsche Forschungsgemeinschaft: Qualitätskriterien der Umfrageforschung. Quality Criteria for Survey Research*. Berlin: Akademie Verlag.
- Keursten, P., 1992. "Opleidingsevaluatie in vogelvlucht". In: J.W. Kessels en C.A. Smit (red.), *Opleiders in organisaties. Capita Selecta*, 12. Deventer: Kluwer Bedrijfswetenschappen, pp.1-16.
- Loosveldt, G., 1985. *De effecten van een interviewertraining op de kwaliteit van gegevens bekomen via het survey-interview*. Leuven: Departement Sociologie, K.U.Leuven (proefschrift).
- Lyberg, L., P. Biemer, M. Collins, E. De Leeuw, C. Dippo, N. Schwarz and D. Trawin, 1997. *Survey measurement and process quality*. New York: John Wiley.
- Morganstein, D. and D. Marker, 1997. "Continuous quality improvement in statistical agencies". In: L. Lyberg et al. (eds.), *Survey measurement and process quality*. New York: John Wiley, pp.475-500.
- Morton-Williams, J., 1993. *Interviewer approaches*. Aldershot: Dartmouth Publishing.
- National Centre for Social Research, 1999. *How to improve survey response rates. A guide for interviewers on the doorstep*. London.
- Spitholt, W., 1989. *Effectiviteitsbepaling van bedrijfsopleidingen*. Enschede: Universiteit Twente, Faculteit der Toegepaste Wetenschappen.
- Voorlichtingscentrum Uw Mening Telt (2000). *Handboek: Basisopleiding interviewers en voorbereiding op het examen 'Interviewer Markt- en Opinieonderzoek'* (samenstelling Jenny Josqui, Akwadraat Consultancy).

8. *MOVIEMOD: Een beslissingsondersteunend systeem voor de introductie van nieuwe films**

J. JONKER, B. WIERENGA, J. ELIASHBERG en M.S. SAWHNEY

SAMENVATTING

Het produceren en distribueren van nieuwe films is risicovol. Distributeurs en bioscoopexploitanten hebben daarom baat bij informatie over het succes van een film, voordat de film in de bioscoop draait. In dit artikel wordt een model gepresenteerd, genaamd MOVIEMOD, dat in staat is zulke voorspellingen te leveren door een voorvertoning van de film aan een steekproef van potentiële bezoekers. Daarnaast kan het model ook helpen bij het opstellen van het mediaplan, omdat het in staat is het effect van aanpassingen in het mediaplan te bepalen. Het model is twee keer toegepast in de praktijk: de eerste keer in de Verenigde Staten, en daarna in Nederland. Omdat in Nederland de distributeur en de belangrijkste bioscoopexploitant van de film deelnamen, waren we in staat om het gebruik van MOVIEMOD als decision support instrument te onderzoeken. De voorspelling van MOVIEMOD bij het gebruikte media- en distributieplan was redelijk nauwkeurig (een afwijking van 4% met het werkelijke bezoekersaantal) en deze voorspelling was beter dan die van een aantal benchmark modellen. De resultaten tonen aan dat MOVIEMOD een waardevol marketing instrument kan zijn voor de filmindustrie.

1. INLEIDING

Van oudsher is entertainment een belangrijke industrie in de Verenigde Staten, maar ook in Nederland wordt deze industrie steeds belangrijker. Ondanks het economische belang wordt er pas de laatste jaren wetenschappelijk onderzoek gedaan naar de entertainment industrie. De filmindustrie is een zeer belangrijk onderdeel van de entertainment industrie. Het doel van dit artikel is het ontwikkelen van een model en meetprocedure die gebruikt kunnen worden voor nieuwe films voordat ze in de bioscoop draaien. Het is toegepast bij de introductie van een nieuwe film in de Nederlandse bioscopen. Het voorgestelde model kan een schatting geven van het succes van een film voordat deze in de bioscopen draait, en kan managers helpen bij beslissingen aangaande de marketing en distributie.

De filmindustrie heeft veel baat bij informatie over het succes van een film voordat deze in de bioscoop wordt gebracht aangezien de kosten en de risico's van het maken van een film erg hoog zijn. Voor zestig tot zeventig procent van de films die in de Verenigde Staten wordt gemaakt geldt dat de opbrengsten op de binnenlandse markt lager

* Dit onderzoek is tot stand gekomen dankzij de samenwerking met Pathé Cinemas en RCV Film Distribution.

zijn dan de kosten van het maken van de film (Stevens en Grover 1998). MPAA, de brancheorganisatie in de Verenigde Staten, meldt dat het in 1997 gemiddeld 75 miljoen dollar kostte om een film te maken. Daar komt dan nog gemiddeld 25 miljoen dollar bij voor de marketing en distributie. Dit bewijst dat voorspellingen aangaande het succes van een film niet alleen zinvol zijn voordat er wordt begonnen met de productie, maar ook als de film al gemaakt is en er beslissingen genomen moeten worden aangaande de promotie en distributie.

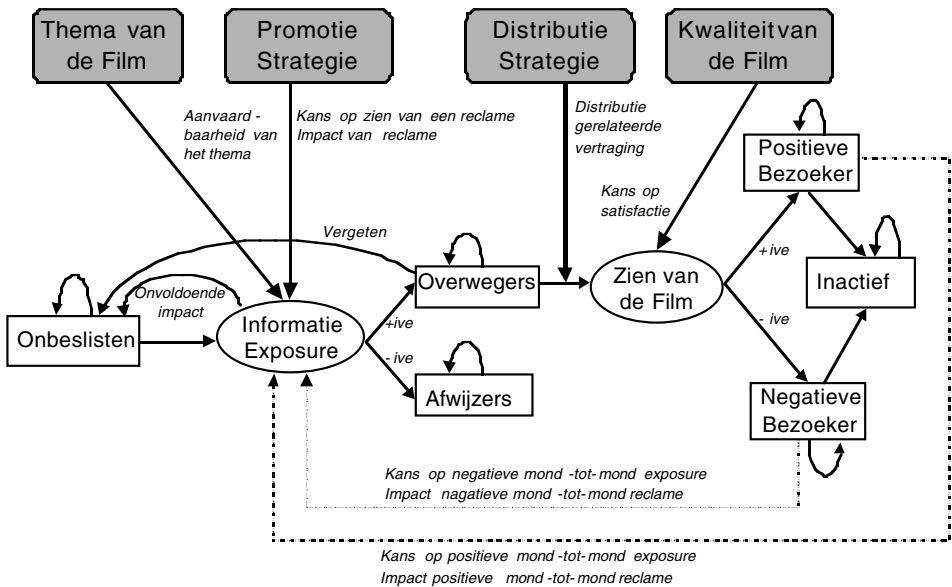
Vooraf in de Verenigde Staten wordt er door de filmindustrie veel (markt)onderzoek gedaan om vast te stellen of een film aan zal slaan bij het publiek. Zo worden verschillende versies van een film aan testgroepen vertoond om vast te stellen welke versie het meest aanspreekt. De definitieve versie wordt voor de release diverse malen vertoond en aan de hand van de uitslagen van de exit polls wordt vastgesteld hoe de marketing strategie eruit moet gaan zien. Voor films die een hoge waardering krijgen kan men rekenen op gunstige mond-tot-mond reclame en dus een wat langere “levensduur” (dit wordt ook wel een platformstrategie genoemd). Maar voor andere films zal men meer moeten vertrouwen op een sterk marketingoffensief waarbij er wordt geprobeerd zoveel mogelijk mensen in de eerste weken naar de bioscoop te krijgen. Omdat men voor deze film een minder sterke mond-tot-mond reclame verwacht, zal het bezoek na de eerste weken sterk afnemen (dit is de “blitz” strategie). Hoewel de filmindustrie dus zelf veel doet om inzicht te krijgen in het potentiële succes van een film, bestaat er grote terughoudendheid ten aanzien van het gebruik van voorspelmodellen en decision support systemen. Dit artikel beschrijft het Moviemod (Movie-Modeling) systeem en is gebaseerd op het paper Eliashberg, Jonker, Sawhney en Wierenga (2000). Voor details wordt verwezen naar dit artikel.

2. MODELBESCHRIJVING

2.1. Gedragsbeschrijving

Moviemod beschrijft in essentie het diffusieproces met betrekking tot informatie van een nieuwe film binnen de relevante consumentenpopulatie. We delen de potentiële populatie van filmbezoekers op in zes gescheiden toestanden (zie Figuur 1). We bekijken in elke periode de proportie van de populatie in elke toestand, beginnend in de eerste periode als de massa-communicatie begint. De toestanden zijn:

1. Onbeslist: Consumenten die nog niets van de film weten of nog geen beslissing hebben genomen.
2. Overwegers: Consumenten die hebben besloten de film te gaan zien nadat ze positieve informatie hebben ontvangen, maar die de film nog niet hebben gezien.
3. Afwijzers: Consumenten die hebben gehoord over de film en hebben besloten om de film *niet* te gaan zien.
4. Positieve Bezoekers: Consumenten die de film hebben gezien, er tevreden over waren en nu positieve mond-tot-mond reclame verspreiden.
5. Negatieve Bezoekers: Consumenten die de film hebben gezien, er niet tevreden over waren en nu negatieve mond-tot-mond reclame verspreiden.
6. Inactieve Bezoekers: Consumenten die de film hebben gezien en er niet meer over praten.



Figuur 1. Gedragsbeschrijving van het consumenten adoptie proces in MOVIEMOD.

2.1.1. Beschrijving van de overgangen

Er wordt verondersteld dat iedereen in het begin bij de Onbeslisten hoort. In elke periode kunnen de Onbeslisten informatie ontvangen via de media of via (positieve of negatieve) mond-tot-mond reclame. De kans dat iemand een media-uiting ziet of hoort hangt af van het marketing (media) plan voor de film. De kans dat iemand positieve of negatieve mond-tot-mond reclame hoort hangt respectievelijk af van het aantal Positieve en Negatieve Bezoekers in de populatie. Daarnaast hangt het ook af van hoe vaak actieve Bezoekers praten met Onbeslisten, en hoe lang een actieve Bezoeker blijft praten over de film.

Iemand die is blootgesteld aan een reclame-uiting kan, afhankelijk van de effectiviteit van de reclame, een overgang maken van de Onbeslisten naar de Overwegers (dit noemen we een positieve exposure) of diegene kan Onbeslist blijven (ineffectieve exposure). Het zien van een reclame kan er ook toe leiden dat iemand van de Onbeslisten een overgang maakt naar de Afwijzers (een negatieve exposure), als er uit de reclame blijkt dat het verhaal of het thema van de film niet aanspreekt. Zo kan iemand die niet van horrorfilms houdt nog zoveel goede informatie horen over een nieuwe horrorfilm, hij of zij zal er nooit heen gaan. Op dezelfde manier kan positieve mond-tot-mond reclame leiden tot een positieve, negatieve of ineffectieve exposure. Voor negatieve mond-tot-mond reclame wordt verondersteld dat deze alleen kan leiden tot een negatieve of een ineffectieve exposure.

Na een positieve exposure gaat men van Onbeslist naar Overweger. Hoewel een Overweger de film kan zien in dezelfde periode als waarin de exposure plaats vond, zal dit in het algemeen niet het geval zijn. Er wordt verondersteld dat de tijd tussen het nemen van de beslissing beïnvloed wordt door persoonlijke factoren (iemand leest op donder-

dag in de krant een artikel over een film en besluit om de film te gaan zien, maar gaat er pas een week later heen) of het kan afhankelijk zijn van de beschikbaarheid van de film. We noemen dit “distributie gerelateerde vertraging” om aan te geven dat de vertraging tussen het besluit en het daadwerkelijke zien van de film in ieder geval gedeeltelijk wordt beïnvloed door de beschikbaarheid van de film. Ten slotte kunnen Overwegers (als de vertraging te lang wordt) vergeten dat ze de film willen gaan zien, en komen dan weer terecht in de groep van Onbeslisten.

Als de Overwegers uiteindelijk de film zien, worden ze Positieve of Negatieve Bezoekers, afhankelijk van hun tevredenheid over de film. Van bezoekers die (on)tevreden zijn wordt verondersteld dat ze een (Negatieve) Positieve Bezoeker worden. Bezoekers stoppen na verloop van tijd met het verspreiden van mond-tot-mond reclame, en worden dan Inactieve Bezoekers.

2.1.2. Impact van de marketing variabelen op de overgangen

De overgangen van de ene naar de andere toestand worden beïnvloed door de volgende marketing variabelen:

- Thema van de film (verhaal, genre, acteurs)
- Promotiestrategie (media budget, marketing plan)
- Distributiestrategie (patroon (blitz of platform), aantal doeken waarop de film gaat draaien, grootte van de zalen)
- Kwaliteit van de film (regie, acteerwerk, cinematografie, etc)

Het thema van de film beïnvloedt de overgang van Onbeslisten naar Afwijzers. De promotiestrategie beïnvloedt de intensiteit waarmee Onbeslisten in aanraking komen met reclame-uitingen voor de film. Distributie heeft invloed op hoe lang het duurt voordat Overwegers Bezoekers worden. Ten slotte bepaalt de kwaliteit van de film met welke kans iemand die de film heeft gezien een Positieve of een Negatieve Bezoeker wordt.

2.1.3. Aannames voor de gedragsbeschrijving

Om het model eenvoudig te houden worden er een aantal aannames gemaakt.

- (1) De verschillende bronnen van informatie worden verondersteld onafhankelijk te zijn.
- (2) Het effect van media en mond-tot-mond reclame is binair (effectief of niet effectief).
- (3) Er wordt geen rekening gehouden met mond-tot-mond reclame voordat de film is uitgebracht.
- (4) Er wordt verondersteld dat Afwijzers niet over de film praten.
- (5) De inhoud van mond-tot-mond reclame is positief of negatief, maar niet neutraal.
- (6) Er wordt geen rekening gehouden met tweede orde effecten zoals gelijktijdige informatie van reclame en mond-tot-mond reclame, of het horen van zowel positieve als negatieve mond-tot-mond reclame in dezelfde periode.

2.2. Model

2.2.1. Interactief Markov model

De overgangen tussen de verschillende toestanden worden gemodelleerd door een Interactief Markov model in discrete tijd (de toestanden worden per week bekeken). De

Markov eigenschap veronderstelt dat de toestand op tijdstip $t+1$ alleen afhangt van de toestand op tijdstip t , en onafhankelijk is van wat er voor tijdstip t is gebeurd.

Een standaard Markov model veronderstelt dat er geen (sociale) interactie is tussen individuen. Deze aanname gaat niet op in ons geval, omdat er mond-tot-mond interactie is tussen Bezoekers en Onbeslisten. Conlinsk (1976, 1978) heeft een model opgesteld dat interactie tussen individuen toestaat, en tegelijkertijd de aangename eigenschappen van een Markov model bezit. Dit wordt een interactief Markov model genoemd.

2.2.2. Notatie

De vector met populatiefrequenties (de fractie van de populatie die in een bepaalde toestand zit) op tijdstip t wordt weergegeven als:

$$Q_t = [O_n_t \ A_t \ O_v_t \ B_t^+ \ B_t^- \ I_t]$$

met

O_n_t = Fractie van Onbeslisten op tijdstip t .

A_t = Fractie van Afwijzers op tijdstip t .

O_v_t = Fractie van Overwegers op tijdstip t .

B_t^+ = Fractie van Positieve Bezoekers op tijdstip t

B_t^- = Fractie van Negatieve Bezoekers op tijdstip t

I_t = Fractie van Inactieve Bezoekers op tijdstip t .

Per definitie is er geen overlap mogelijk tussen de toestanden en is het niet mogelijk dat een persoon niet in één van de toestanden zit. Samen tellen deze populatiefrequenties op tot 1. Ook worden er twee soorten parameters onderscheiden – de filmgerelateerde parameters en de gedragsparameters. De filmgerelateerde parameters verschillen per film. De gedragsparameters geven filmgedrag weer en kunnen per populatie verschillend zijn.

Filmgerelateerde parameters:

- Aanvaardbaarheid van het thema. Kans dat iemand die informatie over de film hoort het thema van de film aansprekend vindt (denk aan horror films).
- Kans op het zien van een reclame. Kans op het zien van tenminste één reclameuiting in de periode $[t, t+1]$. Dit wordt beïnvloed door het mediaplan.
- Kans op satisfactie. Kans dat iemand die de film heeft gezien hem leuk vond en hem aan zal raden aan anderen.
- Impact van positieve mond-tot-mond reclame. Kans dat iemand na het horen van positieve mond-tot-mond reclame besluit om de film te gaan zien.
- Impact van negatieve mond-tot-mond reclame. Kans dat iemand na het horen van negatieve mond-tot-mond reclame besluit om de film *niet* te gaan zien.
- Impact van reclame. Kans dat iemand na het horen van reclame besluit om de film te gaan zien.

Gedraggerelateerde parameters:

- Mond-tot-mond frequentie. Geeft aan hoe vaak mensen per week over films praten.
- Duur van mond-tot-mond. Geeft aan hoeveel weken iemand gemiddeld over films praat.
- Overwegingsduur. Geeft aan hoeveel weken iemand gemiddeld overweegt om een film te gaan zien.
- Distributie gerelateerde vertraging. Geeft aan hoeveel weken het gemiddeld duurt voordat iemand de film gaat zien. Afhankelijk van hoe vaak iemand naar de film gaat en de distributie intensiteit.

De laatste vier parameters beschrijven actief gedrag. We veronderstellen dat de bijbehorende processen (hoe lang praat men over een bepaalde film, hoe vaak praat men over een bepaalde film, etc) exponentieel verdeeld zijn. In de volgende sectie wordt beschreven hoe de parameters geschat worden.

3. OPERATIONALISATIE EN EMPIRISCH TESTEN

MOVIEMOD is tweemaal in de praktijk geïmplementeerd. De eerste toepassing vond plaats in de Verenigde Staten en diende vooral als test. Er werden twee films onderzocht: *Groundhog Day* en *The Cemetery Club*. *Groundhog* is een komedie over een sarcastische weerman die erachter komt dat hij steeds op dezelfde dag wakker wordt, maar dat hij de enige is die zich daarvan bewust is. *The Cemetery Club* is een komedie over drie weduwen die elk op hun eigen manier reageren op de dood van hun partner. De resultaten van de test waren gunstig. MOVIEMOD voorspelde een bruto opbrengst voor *The Cemetery Club* van \$8,3 miljoen, terwijl de werkelijke opbrengst \$ 5,6 miljoen was. Voor *Groundhog Day* werd een opbrengst van \$69,4 miljoen voorspeld en de werkelijke opbrengst was \$ 70,8 miljoen.

Na deze test werd een uitgebreidere implementatie gedaan in Nederland, die in overlegging met de distributeur van de film en de belangrijkste bioscoopexploitant. In deze paragraaf zal kort worden beschreven hoe de parameters zijn gemeten en op welke manier het experiment is uitgevoerd. De details zijn te vinden in Eliashberg *et al.* (2000).

3.1. Schatten van de parameters

3.1.1. Maten voor gedragsparameters

De test begon met het invullen van een enquête over filmbezoekgedrag. De uitkomsten werden gebruikt om de vier gedragsparameters in MOVIEMOD – frequentie van mond-tot-mond reclame, duur van mond-tot-mond reclame, duur van het overwegen, en de distributie gerelateerde vertraging – te bepalen. Eerst werd er aan de respondent gevraagd om zelf een directe schatting te geven, zoals bijvoorbeeld “hoe lang praat u gemiddeld over een film die u hebt gezien?” Er rekening mee houdend dat zelf gerapporteerde schattingen niet altijd even betrouwbaar zijn, werden er meerdere metingen uitgevoerd voor elke parameter en het gemiddelde van deze metingen is gebruikt om tot de uiteindelijke parameterschattingen te komen. Voor elke parameter werd er, naast de directe meting, gebruik gemaakt van een indirecte meting, zoals bijvoorbeeld “wat is de kans dat u twee weken na het zien van de film nog over de film praat?”. Aan de

hand van deze indirecte meting werd er een tweede schatting van de parameter verkregen. De twee schattingen werden gemiddeld om tot de uiteindelijke parameterschatting te komen.

3.1.2. Maten voor filmspecifieke parameters

De parameters die per film verschillend zijn werden geschat door het direct meten van intenties na blootstelling aan filmgerelateerde stimuli. De intenties werden gemeten op een 11-puntsschaal. Zo moesten respondenten op een schaal van 0 (zeker niet geïnteresseerd) tot 10 (zeker wel geïnteresseerd) aangeven hoe groot hun interesse was om de film te gaan zien na de trailer te hebben gezien. De bovenste/onderste vier schalen werden gezien als indicaties voor positieve/negatieve intenties. De aanvaardbaarheid van het thema werd gemeten als de gemiddelde proportie van respondenten die geen negatieve kijkintentie hadden na het zien van een beschrijving van de film. De impactfactoren voor reclame en positieve mond-tot-mond reclame werden gemeten als de proportie van respondenten die zeggen een positieve kijkintentie te hebben na het zien van reclame of het horen van positieve mond-tot-mond reclame, gegeven dat ze het thema aanvaardbaar vonden. De impactfactor van negatieve mond-tot-mond reclame werd geschat door de proportie van de respondenten die zeggen een negatieve kijkintentie te hebben na het horen van negatieve mond-tot-mond reclame. De kans op tevredenheid werd geschat als de proportie van respondenten die na het zien van de film aangaven dat ze de film aan hun vrienden zouden aanraden.

3.1.3. Kans op een media exposure

Om het effect van een media-plan vast te stellen moet de relatie tussen reclame en awareness worden bepaald. Hoeveel mensen zien een bepaalde reclame? Om deze relatie te bepalen is er aan de respondenten gevraagd of ze iets hadden gehoord over zes verschillende films, en zo ja, op welke manier ze deze informatie hadden ontvangen. Voor deze zes films wisten wij, door medewerking van de distributeurs, de mediaplannen. Op die manier waren we in staat de relatie te schatten tussen reclame en awareness. Voor details aangaande deze procedure, zie Eliashberg *et al.* (2000).

3.2. Consumer Measurement Procedure

3.2.1. Opzet

Het experiment werd in Nederland gedaan voor de film *Shadow Conspiracy*, een actie-film/thriller over een complot rond het Witte Huis. Voor het Nederlandse experiment werd er medewerking verleend door de Nederlandse distributeur van *Shadow Conspiracy* en de belangrijkste bioscoop-exploitant in Nederland. Aan de hand van een steekproef uit de Rotterdamse bevolking zou er een voorspelling gemaakt worden van het aantal bezoekers in heel Nederland, en zouden er aanbevelingen worden gedaan ten aanzien van het marketingplan om het resultaat van de film te verbeteren. Een representatieve steekproef van 102 respondenten werd gevraagd deel te nemen aan het experiment dat in Rotterdam werd gehouden op 25 januari 1997. De film werd in Nederland uitgebracht op 8 mei 1997. Omdat het experiment werd gehouden voordat de film in de Verenigde Staten werd uitgebracht (31 januari 1997), hadden de meeste respondenten nog niets over de film gehoord. Dit stelde ons in staat om de meningen van de respondenten te beïnvloeden door een deel positieve en een ander deel van de respondenten negatieve meningen over de film te laten horen.

De procedure bestond uit de volgende stappen:

- 1) Respondenten vulden een enquête in met vragen over filmgedrag. De uitkomsten werden gebruikt om de gedragsparameters te schatten.
- 2) Respondenten kregen reclame-uitingen voor verschillende films te zien, en kregen gesimuleerde mond-tot-mond reclame te horen. De stimuli werden voor de groepen in verschillende volgordes gepresenteerd om rekening te houden met orde-effecten. Na elke stimulus werd de intentie om de film te gaan zien gemeten. Dit werd gebruikt om de impact van de verschillende informatiebronnen te meten en de toegankelijkheid van de film te bepalen.
- 3) Respondenten kregen de film te zien, en vulden daarna een evaluatie-formulier in. Deze werd gebruikt om de kans op satisfactie te bepalen.
- 4) Respondenten gaven ook aan of ze al informatie over verschillende films hadden gehoord, en uit welke bronnen deze informatie was ontvangen. Deze informatie werd gebruikt om de relatie tussen reclame en *awareness* te bepalen.

In totaal kostte het experiment drie uur.

3.2.2. Projectie naar nationale bezoekersaantallen en inkomsten

Het door MOVIEMOD voorspelde bezoekersaantal voor de Rotterdamse populatie werd met behulp van een regressie omgezet in een voorspelling voor heel Nederland. De regressie is gebaseerd op de penetratie van dertig random geselecteerde films in de steekproef van respondenten en het nationaal bezoekersaantal voor deze films. (Kortom, welke relatie is er tussen het bezoekgedrag van de steekproef en het bezoekgedrag van de Nederlandse bevolking.) In de regressie werd er rekening gehouden met de distributie-intensiteit door het aantal schermen waar de films op werden uitgebracht mee te nemen.

3.3. Resultaten en Decision Support Analyse

3.3.1. Parameterschattingen

Het is interessant om de parameterschattingen uit de toepassing in de Verenigde Staten te vergelijken met die verkregen in Nederland. Dit staat weergegeven in Tabel 1 (de definities van de variabelen zijn te vinden in sectie 2.2.2). Groundhog Day was het meest toegankelijke van de drie, met een hoogste kans op een aanvaardbaar thema (0,888). Daarentegen was de impact van reclame minder effectief dan voor Shadow Conspiracy (0,362 vs. 0,490). Dit kan veroorzaakt worden door het feit dat de Nederlandse respondenten een bioscoop-trailer van 90 seconden te zien kregen, terwijl er bij de respondenten in de Verenigde Staten sprake was van een televisie-commercial van 30 seconden. Merk op dat The Cemetery Club niet alleen het minst toegankelijke thema had (0,589), maar ook nog eens een zeer lage kans op satisfactie (0,296). Deze combinatie maakt van deze film een commerciële mislukking. Het is ook interessant om te zien dat negatieve mond-tot-mond reclame meer invloed had in dan positieve mond-tot-mond reclame in de Nederlandse toepassing (0,636 vs. 0,327) vergeleken met de Amerikaanse toepassing waarbij het omgekeerde het geval was (0,451 vs 0,581 voor Groundhog Day en 0,280 vs 0,563 voor The Cemetery Club).

Tabel 1. Parameter schattingen voor de Amerikaanse en Nederlandse implementaties van MOVIEMOD.

Parameter	Nederland	Verenigde Staten	
	<i>Shadow Conspiracy</i>	<i>Groundhog Day</i>	<i>The Cemetery Club</i>
<i>Filmgerelateerde parameters</i>			
Aanvaardbaarheid van het thema	0,708	0,888	0,589
Impact van reclame	0,490	0,362	0,218
Impact van positieve mond-tot-mond reclame	0,327	0,581	0,563
Impact van negatieve mond-tot-mond reclame	0,636	0,451	0,280
Kans op satisfactie	0,550	0,491	0,296
<i>Gedragparameters</i>			
Mond-tot-mond frequentie	1,01 week ⁻¹	2,69 week ⁻¹	
Duur van mond-tot-mond	0,425 week ⁻¹	0,398 week ⁻¹	
Overwegingsduur	0,157 week ⁻¹	0,250 week ⁻¹	
Distributie gerelateerde vertraging (hoge intensiteit)	0,704 week ⁻¹	0,444 week ⁻¹	
Distributie gerelateerde vertraging (lage intensiteit)	0,665 week ⁻¹	0,336 week ⁻¹	

Kijkend naar de gedragsparameters is te zien dat de parameterschattingen redelijk overeen komen. De parameters geven een intensiteit per week weer. Zo is te zien dat de frequentie van mond-tot-mond reclame (aantal keer per week dat men over een film praat) in Nederland lager is dan in de Verenigde Staten (1,01 tegen 2,69). Dit zou verklaard kunnen worden door het feit dat Amerikanen veel meer betrokken zijn bij films (in Amerika gaat men meer dan vier keer zo vaak naar de film dan in Nederland). De inverse van de distributie gerelateerde vertraging geeft aan hoeveel weken het duurt voordat men een film gaat zien, en die is voor Nederland lager dan voor Amerika. Nederlanders gaan dus sneller naar een film dan Amerikanen.

3.3.2. Decision Support Toepassing

Na het schatten van de parameters hebben we MOVIEMOD gebruikt als *decision support tool*. Met behulp van MOVIEMOD hebben we het effect van verschillende media en distributiescenario's bepaald. Vervolgens werd het voorspelde bezoekersaantal berekend voor het toegepast mediaplan. Deze voorspelling kon vergeleken worden met het werkelijke bezoekersaantal, teneinde de voorspelkracht van MOVIEMOD vast te stellen.

De distributeur had een mediaplan voor *Shadow Conspiracy* ontwikkeld, die we twee maanden voordat de film werd uitgebracht ontvingen. Dit was een aantal weken voordat het mediaplan in werking zou treden. Ook ontvingen wij het (verwachte) distributieplan, zoals dat opgesteld was door de belangrijkste bioscoopexploitant. Dit worden de "basis" plannen genoemd. Het mediaplan vermeldt de mate van activiteit per week voor negen verschillende media. We onderscheiden drie niveaus: Laag (L), Middel (M) en Hoog (H). Het mediaplan begint in week 1, en de film wordt vier weken later uitgebracht (in week 5).

Het distributieplan gaf aan dat de film vier weken zou draaien in tien bioscopen: de eerste week in een grote bioscoopzaal, en de volgende drie weken in een kleine. Dit is een gebruikelijk vertoningsplan voor een gemiddelde film. Gebaseerd op deze basisplannen, werd er door MOVIEMOD een landelijk bezoekersaantal van 13.000 voorspeld.

Dit verwachte bezoekersaantal was een teleurstelling voor de distributeur en de exploitant. Zij vroegen ons om MOVIEMOD te gebruiken om te onderzoeken of aanpassingen van het mediaplan zouden kunnen leiden tot een hoger bezoekersaantal. Een aantal uitbreidingen op het basis mediaplan werden doorgerekend. Hoewel er negen media instrumenten gevarieerd konden worden, is er alleen gekeken naar die instrumenten die het meeste effect hadden op het (cumulatieve) bezoekersaantal, en die instrumenten die door de distributeur aangepast konden worden. De volgende vier uitbreidingen werden onderzocht:

- (1) Extra advertenties (met gemiddelde intensiteit) in kranten in week 7.
- (2) Extra artikelen in tijdschriften (met gemiddelde intensiteit) in week 6.
- (3) Extra TV commercials in week 4 en 5.
- (4) Trailers met hoge intensiteit in week 4 en 5.

De voorspelde bezoekersaantallen voor deze aangepaste mediaplannen staan in Tabel 2.

Tabel 2. Voorspellingen voor de vier aangepaste mediaplannen.

	Basis Plan	Aangepaste Plan 1	Aangepaste Plan 2	Aangepaste Plan 3	Aangepaste Plan 4
Voorspelde Cumulatieve Penetratie in Rotterdam (%)	6,17%	6,41%	6,59%	6,56%	7,55%
Voorspelde Bezoekersaantal in Nederland	13.170	13.780	14.240	14.160	16.740

De trailers blijken het meest effectief te zijn (grootste effect op het bezoekersaantal). Hoewel elk van de vier aanpassingen tot meer bezoekers leidt, is het niet noodzakelijk dat elke aanpassing ook leidt tot hogere *winst*. Om dit te berekenen moeten we de kosten van elke uitbreiding weten, en de extra inkomsten die worden gegenereerd door de aanpassing van het mediaplan. Met behulp van informatie over het verdelen van inkomsten tussen distributeur en exploitant werden de extra inkomsten berekend. Dit leverde de volgende resultaten op:

- (1) Extra advertenties in drie nationale kranten kost fl 2.100 en resulteert in een extra opbrengst van fl 3.200 voor de distributeur.
- (2) Extra artikelen in tijdschriften zijn gratis, maar de distributeur moet tijdschriften benaderen en ze overtuigen om de artikelen te plaatsen. Het zou resulteren in een extra opbrengst van fl. 5.700 voor de distributeur.
- (3) Twee keer uitzenden van een TV commercial van 30 seconden in prime time kost

fl 10.000, maar resulteert slechts in een extra opbrengst van fl. 5.200. Deze uitbreiding levert geen extra opbrengst op.

- (4) Het vertonen van extra trailers kost de distributeur geen extra geld. Wel dient de distributeur de exploitant over te halen om meer trailers te vertonen. Deze uitbreiding blijkt het meest winstgevend te zijn, want resulteert in een extra winst van fl. 19.000 voor de distributeur.

Na het presenteren van deze analyse, werd ervoor gekozen om de vierde aanpassing van het mediaplan uit te voeren. Er werden extra trailers vertoond in week 4 en 5, resulterend in een mediaplan waarin de meeste activiteit plaats vindt rond de week van release van de film.

3.3.3. Validatie van de resultaten

Toen *Shadow Conspiracy* werd uitgebracht bleek dat niet alleen het mediaplan afweek van het basis plan, maar ook het distributieplan. Verwacht werd dat de film vier weken zou draaien, maar de film draaide uiteindelijk vijf weken. Ook draaide hij niet één, maar twee weken in een grote zaal. Dit is een indirect gevolg van de aanpassing van het mediaplan: omdat de film het dankzij het aangepaste mediaplan beter deed dan verwacht, werd het distributieplan ook aangepast.

Op basis van deze (echte) plannen, voorspelde MOVIEMOD een bezoekersaantal van ruim 19.352 bezoekers. Dit is bijna 50% hoger dan de uitkomst met de basisplannen. Het werkelijke bezoekersaantal is 18.612, nog geen 5% lager dan het voorspelde aantal.

3.3.4. Vergelijking met benchmark modellen

De voorspelling van MOVIEMOD wordt vergeleken met een aantal benchmark modellen die ook een voorspelling leveren van het bezoekersaantal van een nieuwe film voordat deze in de bioscoop wordt uitgebracht. De eenvoudigste benchmark gebruikt het resultaat van de film in de Verenigde Staten als een voorspeller voor het succes van de film in Nederland. De relatie tussen bezoek in Nederland en bezoek in de Verenigde Staten wordt geschat met behulp van een lineaire regressie (zie ook Smith en Smith (1986)). Omdat films meestal veel eerder in de Verenigde Staten worden uitgebracht dan in Nederland is het mogelijk om op deze manier een schatting te krijgen van het bezoekersaantal voordat de film in Nederland wordt uitgebracht. De relatie werd geschat met behulp van gegevens over dertig films. Het effect van genres wordt expliciet meegenomen. Hiermee wordt er rekening gehouden met het mogelijke verschil in smaak tussen de Verenigde Staten en Nederland. De gevonden relatie is redelijk sterk ($R^2 = 0,84$). De voorspelling op basis van deze relatie was 7.400 bezoekers. Deze voorspelling is 60% lager dan het werkelijke bezoekersaantal van 18.612.

Het tweede benchmark model is gebaseerd op het BOXMOD model dat is ontwikkeld door Sawhney en Eliashberg (1996). Dit model gebruikt de bezoekersaantallen van de eerste twee weken om de totale opbrengsten te voorspellen. Deze voorspelling is gebaseerd op de vorm van de opbrengstcurve die wordt beschreven door drie parameters. De auteurs geven ook een meta-analyse methode die het mogelijk maakt om het (cumulatieve) bezoekersaantal te voorspellen voordat de film wordt uitgebracht. Deze meta-analyse gebruikt historische gegevens voor een aantal films om een relatie te schatten tussen eigenschappen van een film (bijvoorbeeld genre) en BOXMOD parameters. Op

die manier is het dus mogelijk om een schatting te krijgen voor de BOXMOD parameters voor Shadow Conspiracy. Voor de meta-analyse werden vijftien films gebruikt. Dit beperkte aantal wordt veroorzaakt door het feit dat wekelijkse bezoekersaantallen niet publiekelijk beschikbaar zijn in Nederland. De uiteindelijke schatting van BOXMOD is 8,7% lager dan het werkelijke bezoekersaantal.

Een belangrijke eigenschap van MOVIEMOD is de interactiviteit: er wordt rekening gehouden met mond-tot-mond reclame tussen potentiële bezoekers en mensen die de film hebben gezien. We kunnen expliciet onderzoeken of het meenemen van deze interactie leidt tot betere resultaten. Hiertoe stellen we een versie van MOVIEMOD op die geen rekening houdt met mond-tot-mond reclame. Als de bijbehorende voorspelling significant lager is dan het werkelijke bezoekersaantal, wijst dit erop het meenemen van de interactie de voorspellingen verbeterd, en dat de interactieve versie van MOVIEMOD een realistische wiskundige representatie is van het proces om een film te gaan zien. Het voorspelde bezoekersaantal van het model zonder interactie is 17.030, wat 8,5% lager is dan het werkelijke bezoekersaantal. Dit is in overeenstemming met de verwachting dat het weglaten van mond-tot-mond reclame leidt tot een onderschatting van de informatiestromen, en dus ook van de adoptiesnelheid. Verder heeft de interactie versie van MOVIEMOD een lagere voorspelfout (4%) dan de niet-interactie versie.

Tabel 3. Resultaten van MOVIEMOD en de benchmark modellen.

Model	Voorspelling vs. Gerealiseerd
Näïeve model (Resultaat in de VS als voorspeller)	-60%
BOXMOD (Sawhney and Eliashberg 1996)	-8.7%
MOVIEMOD (geen interactie)	-8.5%
MOVIEMOD (met interactie)	+4.0%

De resultaten staan samengevat in Tabel 3. Hieruit blijkt dat MOVIEMOD het beter doet dan de benchmark modellen. Ook is MOVIEMOD in staat om te bepalen *waarom* het bezoekersaantal hoog of laag is, en *wat er gedaan kan worden* om het bezoekersaantal te verhogen. Deze decision support eigenschappen maken van MOVIEMOD een waardevoller voorspel en plannings-instrument dan BOXMOD of de andere benchmarks.

4. SAMENVATTING EN ONDERWERPEN VOOR NADER ONDERZOEK

MOVIEMOD is in staat om een schatting van het succes van een film te bepalen voordat deze in de bioscoop wordt uitgebracht. Daarnaast kan het managers helpen bij het nemen van marketingbeslissingen in de filmindustrie. Hoewel het model nog zeker verbeterd kan worden, laten de resultaten zien dat managers in de entertainment industrie baat kunnen hebben bij kwantitatieve modellen.

MOVIEMOD zou op een aantal punten verbeterd kunnen worden. Zo zou de relatie tussen reclame en awareness verbeterd kunnen worden door bijvoorbeeld gebruik te maken van een panel. Dit zou de relatie betrouwbaarder maken. Ook zou er rekening kunnen worden gehouden met concurrentie. In het huidige model wordt concurrentie meegenomen in de distributie. Een andere manier waarop concurrentie meegenomen zou kunnen worden is door de tijd dat men overweegt een film te gaan zien te laten afhangen van de mate van concurrentie, of door de impact van reclame afhankelijk te maken van concurrentie. Een derde uitbreiding zou een capaciteitsrestrictie kunnen zijn (Jain, Mahajan and Muller 1991). De huidige formulering houdt er geen rekening mee dat de film uitverkocht zou kunnen zijn. Dit kan worden meegenomen in het model door mensen verplicht te stellen in een bepaalde toestand te blijven als de vraag naar de film hoger is dan het aanbod aan stoelen. Ook moet er dan rekening mee worden gehouden dat consumenten die de film niet kunnen zien, minder interesse in de film kunnen krijgen. Tenslotte zouden seizoenseffecten meegenomen kunnen worden. Zo zou de tijd tussen het besluiten om een film te gaan en het feitelijke zien van de film geïndexeerd kunnen worden aan de grootte van de markt per week, gebaseerd op historische data.

De ervaring met Moviemod heeft aangetoond dat ook in een bedrijfstak waar intuïtie en creativiteit dominant zijn, zoals de filmindustrie, zeer nuttige bijdragen kunnen worden geleverd door marktonderzoek gebaseerd op modellen.

LITERATUUR

- Conlinsk, John (1976), "Interactive Markov Chains," *Journal of Mathematical Sociology*, Vol. 4, 157-185.
- Conlinsk, John (1978), "A Stability Theorem for an Interactive Markov Chain," *Journal of Mathematical Sociology*, 6, 163-168.
- Eliashberg, Jehoshua, Jedid-Jah Jonker, Mohanbir S. Sawhney and Berend Wierenga (2000), "MOVIEMOD: An Implementable Decision Support System for Pre-Release Market Evaluation of Motion Pictures", *Marketing Science*, 19, 3, 226-243.
- Jain, Dipak C., Vijay Mahajan and Eitan Muller (1991), 'Innovation diffusion in the presence of supply restrictions', *Marketing Science*, 10, 1, 83-90.
- NFC (2000), Jaarverslag Nederlandse Federatie voor de Cinematografie.
- Sawhney, Mohanbir S. and Jehoshua Eliashberg (1996), "A Parsimonious Model for Forecasting Gross Box-Office Revenues of Motion Pictures", *Marketing Science*, 15, 2, 113-131.
- Smith, S.P. and V.K. Smith (1986), "Successful Movies: A Preliminary Empirical Analysis", *Applied Economics*, 18, 501-507.
- Stevens, Elizabeth L. and Ronald Grover (1998), The Entertainment Glut, *Business Week*, February 16.

9. Modellen voor de effectmeting van sales promotions

H.J. VAN HEERDE

SAMENVATTING

Dit artikel gaat over recent ontwikkelde modellen voor de effectmeting van sales promotions.¹ Met deze modellen kan inzicht worden verkregen in: (1) het effect van de hoogte van een tijdelijke prijskorting op de verkopen van het merk met de korting, (2) het meten van pre- en postpromotionele dips en (3) de opsplitsing van de promotionele verkooppiek in deeleffecten. Het onderzoek maakt gebruik van recente econometrische technieken die worden toegepast op scanninggegevens voor meerdere productcategorieën uit Nederland en de Verenigde Staten. Deze econometrische technieken stellen de gebruiker in staat de netto verkoopeffecten van promoties te bepalen, waarbij de promoties gekarakteriseerd worden door (1) de hoogte van de geboden prijskorting en (2) de vorm van ondersteuning: folder en/of display of geen ondersteuning. Deze netto verkoopeffecten kunnen weer als input dienen voor de ontwikkeling van betere sales promotion strategieën voor fabrikanten en retailers.

1. INLEIDING

Wie tegenwoordig door een supermarkt loopt, struikelt zowat over de sales promotions: “van f 2.19, voor f 1.79”, “drie halen, twee betalen” of “ontvang het aankoopbedrag terug”. Sales promotions worden gedefinieerd als “een marketing-tactiek waarbij men door een tijdelijke verbetering van de prijs/waarde verhouding van een product of dienst tracht tijdelijk extra verkopen te realiseren”. Het gebruik van sales promotions heeft een enorme vlucht genomen in Nederland. Het bedrag dat in Nederland in totaal aan sales promotions wordt uitgegeven is van 1990 tot nu gestegen van 5.1 tot bijna 10 miljard gulden.

Deze toename in het gebruik van sales promotions valt onder andere te verklaren uit het feit dat sales promotions vaak direct een verkoopeffect laten zien, terwijl het effect van bijvoorbeeld thema-reclame veel moeilijker vast te stellen valt. Toch zijn er signalen dat de winstgevendheid van sales promotions niet hoog is en dat veel acties zelfs niet winstgevend zijn. De fluctuaties in de verkopen die promoties met zich meebrengen zorgen voor allerlei extra productie- en logistieke kosten.

¹ Dit artikel geeft een samenvatting van het proefschrift van Harald J. van Heerde, getiteld “Models for Sales Promotion Effects Based on Store-Level Scanner Data”, dat op 22 november 1999 is verdedigd aan de Rijksuniversiteit Groningen.

Het evalueren van sales promotions kan als basis dienen voor een efficiëntere besteding van de promotie-uitgaven. Echter, in de praktijk blijft vaak een grondige effectmeting van sales promotions achterwege. Volgens PriceWaterhouseCoopers wordt in Nederland slechts 1 procent van de sales promotions geëvalueerd. Gezien de hoogte van de bestede bedragen aan sales promotions is het uitermate zinvol de effecten van sales promotions te meten. Er is op dit terrein dan ook al het nodige onderzoek verricht. Blattberg, Briesch en Fox (1995) geven in een overzichtsartikel een samenvatting van empirische generalisaties met betrekking tot de effectmeting van sales promotions. Ze concluderen onder andere dat tijdelijke prijskortingen en ook andere promoties voor aanzienlijke tijdelijke verkoopstijgingen zorgen. Echter, daarmee is niet bewezen dat deze promoties ook succesvol zijn in termen van winstgevendheid. Om dit te bepalen is meer verfijnd inzicht in de werking van sales promotions vereist. Het doel van dit promotie-onderzoek is het ontwikkelen van econometrische modellen om nieuwe inzichten te verkrijgen in de verkoopheffecten van sales promotions. Het onderzoek richt zich op het bepalen van:

- de “deal effect curve”: verband tussen hoogte van de promotionele prijskorting en de verkopen van een merk;
- “pre- en postpromotionele dips”: terugval in de verkopen voor en na promoties;
- de “decompositie” van de promotionele verkooppeik.

Deze onderwerpen zijn door Blattberg, Briesch en Fox (1995) in hetzelfde overzichtsartikel geïdentificeerd als onderwerpen binnen sales promotions onderzoek waarvoor weinig of tegenstrijdige empirische resultaten bestaan. Ook vanuit de praktijk bestaat er veel aandacht voor deze onderwerpen. De opbouw van de rest van dit artikel is als volgt. De gegevens die in de drie onderzoeken worden gebruikt, worden eerst kort beschreven. In de volgende paragrafen wordt achtereenvolgens het onderzoek naar de deal effect curve, de pre- en postpromotionele dips en de decompositie besproken. We besluiten met een conclusie en een korte bespreking van toekomstig onderzoek.

2. GEGEVENS

Voor de drie onderzoeken worden Nederlandse en Amerikaanse scanninggegevens op winkelniveau gebruikt, afkomstig van het marktonderzoeksbureau A.C. Nielsen. Scanninggegevens worden verzameld door het “scannen” van boodschappen bij kassa’s in supermarkten. Ze geven per week informatie over de verkochte hoeveelheden van alle artikelen in een supermarkt en tegen welke prijzen dat is gebeurd. Daarnaast wordt ook gemeten of er sprake was van sales promotions, en zo ja: van welk type. We kiezen dit type data boven huishoudpanel gegevens, die binnen dit onderzoeksgebied veel worden gebruikt. Hierbij worden de aankopen van een panel van huishoudens geregistreerd. De belangrijkste reden voor onze keuze voor scanninggegevens op winkelniveau is dat managers vaak de voorkeur geven aan resultaten van modellen die geschat zijn op basis van gegevens van een hoger aggregatie-niveau dan het huishoudniveau. Daarnaast is de representativiteit van huishoudpanels soms twijfelachtig en is de omvang ervan vaak te klein om modellen te kunnen ontwikkelen voor de meeste productcategoriën in supermarkten. Tot slot zijn scanninggegevens op winkelniveau veel beter verkrijgbaar dan huishoudaankoopgegevens, met name buiten de Verenigde Staten. Scanninggegevens op winkelniveau worden namelijk al automatisch verzameld als bijproduct van het scannen van boodschappen bij kassa’s.

De voor onze onderzoeken gebruikte gegevens hebben betrekking op vier productcategorieën: blikjes tonijn (drie merken), en toiletpapier (zes merken), een drankproduct (zes merken) en een voedselproduct (vier merken). De eerste drie datasets zijn afkomstig uit Amerika, de laatste uit Nederland. De gegevens zijn afkomstig van tussen de 24 en 151 winkels, en beslaan tussen de 52 en 144 weken. Overigens gebruiken we niet alle datasets voor alle onderzoeken.

3. DEAL EFFECT CURVE

3.1. *Achtergrond deal effect curve*

Blattberg, Briesch en Fox (1995) noemen de vorm van de “deal effect curve” een van de belangrijkste onderwerpen binnen sales promotions onderzoek waarover nog nauwelijks informatie bestaat. De deal effect curve geeft het verband weer tussen de hoogte van de geboden tijdelijke prijskorting voor een bepaald merk en de verkopen van dit merk. Kennis van de vorm van deze curve is van belang bij het bepalen van het optimale prijskortingsniveau, optimaal in het termen van winst. Als de curve convex is (toenevende meeropbrengsten) zal de optimale prijskorting wellicht hoger liggen dan bij een concaaf verband (afnemende meeropbrengsten). Meestal nemen onderzoekers een bepaalde functionele vorm aan. Zelden worden alternatieve functionele vormen met elkaar vergeleken. In dit onderzoek passen we een nieuw, flexibel model toe voor het schatten van de deal effect curve.

In de literatuur worden factoren genoemd die de vorm van de deal effect curve bepalen. Zo verwachten Gupta en Cooper (1992) dat er een zekere drempelwaarde zal bestaan. Pas bij prijskortingen hoger dan deze drempelwaarde (bijvoorbeeld: tien procent) gaan de verkopen substantieel omhoog. Verder verwachten zij een zeker verzadigingseffect bij hoge prijskortingsniveaus (bijvoorbeeld: veertig à vijftig procent). Ten eerste vanwege het feit er een maximum bestaat aan hetgeen klanten van een bepaald artikel kunnen en willen opslaan. Ten tweede vanwege het “discounting of discounts effect”: klanten percipiëren te hoge prijskortingen niet meer als geloofwaardig: “de reguliere prijs zal wel omhoog getrokken zijn om de korting hoger te doen lijken”. Bij combinatie van een drempelwaarde en een verzadigingseffect is de deal effect curve S-vormig. Maar juist omdat andere factoren aanleiding kunnen geven tot alternatieve vormen, hebben we een flexibel model nodig.

3.2. *Aanpak deal effect curve*

We gebruiken een semiparametrisch regressiemodel voor het bepalen van de vorm van de deal effect curve. We nemen hiervoor geen bepaalde functionele vorm aan, zoals een lineair verband. We laten de data zelf bepalen hoe de deal effect curve eruitziet. Een semiparametrisch model combineert parametrische en nonparametrische elementen. Het klassieke lineaire regressiemodel is een voorbeeld van een parametrisch model. Hierbij wordt het verband tussen de afhankelijke variabele (y) en de onafhankelijke variabele (x) uitgedrukt met behulp van een regressie-parameter (β) en een storingsterm (u): $y = x\beta + u$. Een nonparametrisch regressie-model daarentegen, laat het verband tussen de afhankelijke en de onafhankelijke variabele (z) volledig vrij: $y =$

$m(z) + u$, waarbij $m(z)$ een onbekende, niet-gespecificeerde functie van z is. Het semiparametrische regressiemodel neemt een nonparametrisch verband aan voor een verzameling van onafhankelijke variabelen (z), maar handhaaft een parametrische structuur voor een andere verzameling (x): $y = m(z) + x\beta + u$. Het parametrische, minder flexibele, equivalent van dit model is: $y = z\gamma + x\beta + u$.

De semiparametrische aanpak vermijdt het probleem dat de praktische toepasbaarheid van volledig nonparametrische modellen belemmert: “the curse of dimensionality”. Dit houdt in dat dergelijke volledig flexibele modellen alleen toepasbaar zijn bij een beperkt aantal onafhankelijke variabelen; bij een toenemend aantal stijgt de benodigde steekproefgrootte vrijwel exponentieel. Parametrische regressiemodellen hebben geen last van dit probleem, maar die zijn weer niet flexibel genoeg. Het semiparametrische model combineert in feite het voordeel van nonparametrische regressie (flexibiliteit) en het voordeel van parametrische regressie (het relatief efficiënt omgaan met waarnemingen).

Voor deze toepassing is de afhankelijke variabele (y) de logaritme van de verkopen van een bepaald merk in een bepaalde week in een bepaalde winkel. De onafhankelijke variabelen waarvoor een nonparametrisch verband is gekozen (de z -variabelen) zijn de “log prijsindices” van het merk zelf en van concurrerende merken. We gebruiken prijsindices (=werkelijke prijs/reguliere prijs) om te corrigeren voor verschillen in reguliere prijzen tussen winkels. Op deze wijze is het verband waarin we met name geïnteresseerd zijn, namelijk, tussen prijs en verkopen (de deal effect curve), volledig vrij gelaten. De overige onafhankelijke variabelen (x) zijn in het parametrische deel ondergebracht. Dit betreffen display- en featurings activiteiten (gemeten als 0/1 variabele; ofwel: dummy-variabele) voor het merk zelf en voor concurrerende merken, alsmede dummy-variabelen voor de specifieke weekeffecten (weekdummies) en winkel (winkel-dummies). Voor een dataset (drankproduct) stoppen we tevens een deel van de prijsvariabelen in het parametrische stuk.

3.3. Resultaten deal effect curve

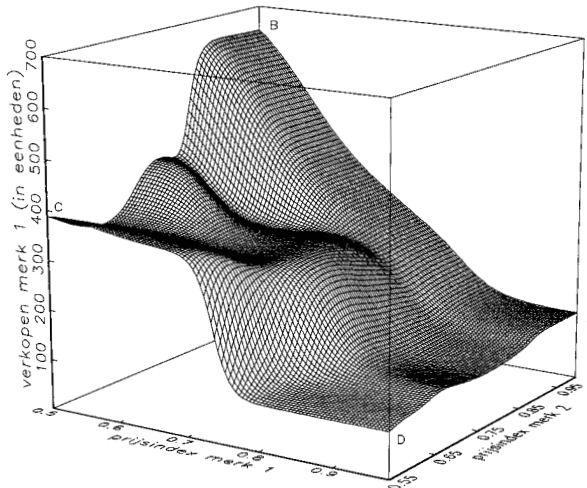
Het model is geschat op basis van wekelijkse waarnemingen van meerdere winkels voor de drie tonijn-merken, de zes drank merken en de vier voedselmerken. De schattingsprocedure is ontleend aan Robinson (1988) en Lee (1996, pp. 203-205). In het bijzonder wordt het nonparametrische gedeelte $m(z)$ geschat met behulp van de Kernel methode (zie ook Härdle 1990). Globaal gesproken komt deze methode op het volgende neer. Voor een bepaalde combinatie van waarden van onafhankelijke variabelen (noem deze combinatie z^*) wordt de waarde van de afhankelijke variabele (y) geschat, net zoals bij lineaire regressie. Deze schatting voor y wordt bij de Kernel methode verkregen door een gewogen gemiddelde te nemen van de waarden die de afhankelijke variabele aanneemt voor waarnemingen van onafhankelijke variabelen (zeg: z^{**}) “in de buurt van” z^* . Het gewicht is omgekeerd evenredig aan de afstand tussen z^{**} en z^* : hoe dichtbij, hoe hoger het gewicht.

We vergelijken in dit onderzoek eerst het semiparametrische model met twee bekende varianten op het standaard parametrische model, zowel op basis van de “fit” als “voorspelkracht”. De “fit” betreft de mate waarin de werkelijke verkopen overeenkomen met de door het model berekende verkopen. De “voorspelkracht” betreft de mate waarin de

werkelijke verkopen voor een andere periode dan die waarvoor het model is geschat, overeenkomen met de door het model voorspelde verkopen voor die periode. Het semi-parametrische model presteert duidelijk beter op fit en voorspelkracht dan de beide parametrische modellen voor alle drie onderzochte productcategorieën. De winst in de voorspelkracht wordt vooral behaald in een reductie van de “bias” (vertekening) en in iets mindere mate in een reductie van de variantie. Dit is in overeenstemming met onze verwachtingen. Verder zien we dat ons semiparametrische model ook veel beter fit en voorspelt in de weken waarin een merk zelf prijspromoties houdt. De door Gupta en Cooper (1992) verwachte drempelwaarden en verzadigingseffecten worden door dit onderzoek empirisch bevestigd.

Ter illustratie van de uitkomsten van het semiparametrische model is figuur 1 opgenomen. Deze figuur toont de niet-lineaire effecten van prijskortingen van twee tonijnmerken (merk 1 en 2) op de verkopen van één merk (merk 1). Op het grondvlak zijn de prijsindices van merk 1 en 2 uitgezet, op de verticale as de verkopen van merk 1. Een prijsindex van 1 wil zeggen: werkelijke prijs is reguliere prijs; geen prijskorting derhalve. Een prijsindex van 0.8 duidt erop dat de werkelijk prijs 8/10 deel van de reguliere prijs is: dit is dus een tijdelijke prijskorting van twintig procent, enzovoort.

Effect van prijskortingen voor merk 1 en 2 op de verkopen van merk 1 (gebaseerd op semiparametrisch model)



Figuur 1. Illustratie niet-lineaire promotie-effecten.

Punt A in figuur 1 toont de situatie dat geen van beide merken een prijskorting heeft: beide prijsindices zijn gelijk aan 1. De lijn A-B in de figuur laat de deal effect curve voor merk 1 zien, waarbij merk 2 op de reguliere prijs zit (prijsindex gelijk aan 1). Duidelijk is dat er een drempelwaarde bestaat: pas bij prijskortingen van meer dan tien procent (prijsindex < 0.9) gaan de verkopen van merk 1 substantieel omhoog. Het verdere verloop van deze deal effect curve is vrij lineair. De lijn A-D in de figuur toont het geringe effect van prijskortingen van merk 2 op de verkopen van merk 1, als dit laatste merk op de reguliere prijs zit. De rest van de figuur toont de flexibele interactie-effecten.

ten tussen gelijktijdige prijskortingen van de beide merken. Zo laat punt C de verkopen van merk 1 zien voor het geval dat merk 1 een 50% prijskorting heeft en merk 2 een 45% prijskorting. De lijn B-C laat zien dat het hoge verkoopniveau dat merk 1 behaalt bij de maximale prijskorting van 50% (punt B), voor een aanzienlijk deel teniet wordt gedaan door een hoge prijskorting voor merk 2 (punt C). Deze bevindingen onderstrepen het belang van een flexibele functionele vorm en flexibele interactie-effecten in het modelleren van prijskortingseffecten. Het is duidelijk dat de gevonden non-lineariteiten zeer moeilijk met een parametrisch model ondervangen zouden kunnen worden.

Tevens vinden we dat de effecten van prijskortingen afhangen van de mate van ondersteuning: met een feature (folder, advertentie), een display (speciaal schap), feature en display of geen van beide. De belangrijkste conclusie is dat we met een flexibel model hebben aangetoond dat de vorm van de deal effect curve duidelijk niet-lineair is.

4. PRE- EN POSTPROMOTIONELE DIPS

4.1. *Achtergrond pre- en postpromotionele dips*

Een van de meest intrigerende paradoxen in sales promotions onderzoek betreft het ogenschijnlijk ontbreken van de zogenaamde postpromotionele dip (Blattberg, Briesch en Fox 1995; Neslin en Schneider Stone 1996). Deze postpromotionele dip is gedefinieerd als de terugval in verkopen na een promotie. Uit analyses van aankoopegegevens van huishoudens is bekend dat consumenten aan voorraadvorming doen ten gevolge van promoties (o.a. Neslin, Henderson en Quelch 1985). Intuïtief zou men daarom kunnen verwachten dat de verkopen van artikelen ondersteund door promotionele acties een terugval laten zien in de week/weken na de actie. Het vreemde is dat deze postpromotionele dip in de praktijk niet of nauwelijks wordt waargenomen in verkoopgrafieken: na de promotie lijken de verkopen weer op het normale niveau te zitten. De paradox luidt dus: hoe kan het dat studies naar effecten van promoties op het niveau van individuele huishoudens overduidelijk wijzen op voorraadvorming, terwijl er op het niveau van winkels geen sprake lijkt te zijn van postpromotionele dips? Het doel van het in deze paragraaf beschreven onderzoek is het ontwikkelen van een econometrisch model dat postpromotionele dips kan schatten op basis van verkoopgegevens op winkelniveau. Hierbij houden we rekening met factoren die ervoor kunnen zorgen dat deze dips moeilijk meetbaar zijn. Neslin en Schneider Stone (1996) geven een overzicht van deze factoren. We hebben deze factoren vertaald in een aantal modelimplicaties. De belangrijkste zijn:

- Het model moet rekening houden met postpromotionele dips die over meerdere weken uitgesmeerd zijn;
- Omdat consumenten kunnen anticiperen op verwachte promoties, kunnen er ook zogenaamde prepromotionele dips bestaan: men wacht met aankopen tot de actie-week. Deze anticipeer-effecten ontstaan doordat klanten, gezien de grote regelmaat waarmee promoties plaatsvinden, een goed inzicht hebben wanneer bepaalde promoties zich voordoen. Het model moet dus ook prepromotionele dips over meerdere weken kunnen schatten;
- De vorm van pre- en postpromotionele dips dient flexibel gemodelleerd te worden. Met andere woorden, de diepte van de dips hoeft niet constant te zijn. Dit fenomeen

wordt mede veroorzaakt door heterogeniteit van huishoudens. Het ene huishouden heeft namelijk een andere aankoopfrequentie en -hoeveelheid dan het andere huishouden. Dientengevolge duurt het voor het ene huishouden langer tot er, na een promotie, weer van het artikel gekocht wordt dan voor het andere huishouden;

- Het model moet concurrentie-effecten opnemen.

4.2. Aanpak pre- en postpromotionele dips

We hebben deze implicaties verwerkt in een “distributed lead and lag regression model”. De afhankelijke variabele is wederom log verkopen. Er zijn vertraagde (“lagged”) log prijsindices als onafhankelijke variabelen opgenomen om postpromotionele dips te modelleren. Met prepromotionele dips wordt rekening gehouden door “toekomstige” (“lead”) prijsindices in het model op te nemen. We nemen vier soorten prijsindices op: prijsindices voor prijspromoties zonder feature of display-ondersteuning, prijsindices voor prijspromoties met alleen feature ondersteuning, prijsindices met alleen display ondersteuning en prijspromoties met ondersteuning in de vorm van feature plus display. Om de vorm van pre- en postpromotionele dips flexibel te modelleren is gekozen voor drie typen modellen:

1. het Almon model. Dit houdt in dat zowel de pre- als postpromotionele effecten worden beschreven door een polynoom als functie van de tijd;
2. het Exponential decay model. Dit houdt in dat het effect in de eerste week vóór of na de promotie het grootst is en vervolgens exponentieel afneemt;
3. het Unrestricted model. Dit houdt in dat de vertraagde en toekomstige effecten van promoties op de huidige verkopen volledig zonder restricties worden gemodelleerd.

Aangezien het niet bekend is hoe lang pre- en postpromotionele effecten duren en hoeveel flexibiliteit noodzakelijk is, hebben we hiervoor een zoekprocedure ontwikkeld. De lengte van de pre- en postpromotionele effecten mag variëren tussen nul (ofwel: geen effect) en zes weken voor alle modellen. De mate van flexibiliteit varieert binnen model (1) en (2); model (3) is al maximaal flexibel. Bij het Almon model (1) wordt de graad van het polynoom gevarieerd. Hiermee wordt de vorm van de dip in de loop van de tijd beschreven. De graad is gelijk aan nul (constante dip), één (lineair stijgende of dalende dip), twee (dip in de vorm van parabool) of drie (dip in de vorm van een derdegraads polynoom). Bij het exponential decay model hebben we de vervalparameter (de mate waarin het effect in de loop van de tijd verdwijnt) gevarieerd tussen 0.1 en 0.9. Een vervalparameter van 0.9 wil zeggen dat na 2 weken 90 procent van het effect na 1 week overblijft; een verval parameter van 0.1 wijst op het overblijven van 10 procent. Zo kwamen we tot een groot aantal modellen per merk. Op basis van het “Akaike Informatie Criterium” is bepaald welk model optimaal is voor een gegeven merk. Dit criterium beloont een goede fit, maar bestraft een groot aantal parameters. Een te groot aantal parameters is niet wenselijk omdat dit leidt tot minder betrouwbare parameterschattingen.

4.3. Resultaten pre- en postpromotionele dips

Het model is toegepast op zowel de tonijn- als de toiletpapier-gegevens. De belangrijkste bevindingen zijn:

- Voor acht van de in totaal negen merken vinden we significante pre- en/of postpromotionele effecten;

Tabel 1. Effect van 20% prijskorting op de verkopen in de promotieweek en op de verkopen in de pre- en postpromotionele perioden.

	Tonijn (1 eenheid = 1 blik)				Toiletpapier (1 eenheid = duizend vellen)			
	Geen ondersteuning	Alleen Folder	Alleen display	Folder en Display	Geen ondersteuning	Alleen Folder	Alleen display	Folder en Display
Huidig verkoop-effect	126	190	222	284	227	542	529	689
Verkoop-effect in pre- en post-promotionele periode	-28	-24	-8	-21	-57	-72	+50	-84
Netto verkoop-effect	98	165	214	263	170	470	579	605
Percentage netto verkooptoename	78%	87%	96%	93%	75%	87%	110%	88%

- Deze dips zijn substantieel ten opzichte van de verkooptoename in de promotionele periode (“de promotionele verkooppiek”). Om tot deze uitspraak te komen hebben we eerst de eenheden die in de promotionele week meer worden verkocht, berekend met het geschatte model. Dit zijn bijvoorbeeld 126 blikjes tonijn voor een prijskorting zonder ondersteuning (zie Tabel 1). Vervolgens hebben we het effect van de actie op de verkopen in de pre- en postpromotionele weken opgeteld (dynamische effecten). Als we hetzelfde voorbeeld aanhouden, komen we op een dynamisch effect van -28 blikjes tonijn. Vervolgens hebben we het netto effect bepaald door de huidige en de dynamische effecten op te tellen. In het voorbeeld levert dit $126 - 28 = 98$ blikjes op. Tenslotte hebben we het netto effect uitgedrukt als percentage van het effect in de promotie-week. Dit komt op $(98 / 126) * 100\% = 78\%$ bij het tonijnvoorbeeld. Over het geheel genomen liggen de netto-effect percentages die lager zijn dan 100%, tussen de 75 en 96%.² Dit is consistent met de resultaten die veelal in huishoudstudies worden gevonden. Deze percentages hangen verder af van de productcategorie en de verschillende vormen van ondersteuning van prijsacties, zoals in Tabel 1 kan worden gezien.
- De postpromotionele dip paradox hield in dat er feitelijk wel postpromotionele dips zijn op huishoudniveau, maar dat deze dips niet terug te vinden zijn in winkelniveau. Aangezien we methoden hebben ontwikkeld die op winkelniveau wél dips kan vinden, die bovendien van een vergelijkbare grootte zijn als die van studies op huishoudniveau, kunnen we stellen dat de postpromotionele dip paradox niet opgaat;
- Omdat de postpromotionele dip van de ene promotie en de prepromotionele dip van de volgende promotie door elkaar heen kunnen lopen, is het niet mogelijk een uitspraak te doen over de relatieve grootte van beide typen dips;

² Er is een netto-effect van 110% bij een prijskorting met alleen display ondersteuning voor toiletpapier. Deze uitschieter kan waarschijnlijk worden verklaard door een na-ijl effect van display: er blijven artikelen zonder korting op de display staan en klanten reageren erop alsof er een aanbieding is.

- Wanneer alleen de effecten van promoties op de verkopen in de promotie-periode beoordeeld worden, krijgt men een vertekend beeld van deze effecten. Er kunnen in de perioden vóór en na de promotie effecten optreden. Neemt men deze effecten niet mee dan worden de effecten van promoties overschat;
- Gegeven de complexiteit van dynamische effecten van promoties lijkt het beter een flexibel model als het Almon model of het Unrestricted model te gebruiken en beter niet een relatief star model als het Exponential decay model. Over het geheel genomen, zit het Unrestricted model het dichtst bij de “beste model” resultaten. Tevens is dit het model dat het eenvoudigst te implementeren is.

5. DECOMPOSITIE

5.1. Achtergrond decompositie

Het decompositie-onderzoek ligt volledig in het verlengde van beide andere onderzoeken van dit proefschrift. We hebben gezien dat een deel van de promotionele verkooppiek bij blikjes tonijn en toiletpapier verklaard kan worden door voorraadvorming- en anticipeer-effecten (post- en prepromotionele dips). De vraag is: waar komt nu eigenlijk de rest van de promotionele verkooppiek vandaan? Hoe kan verklaard worden dat er tijdens een promotie vaak wel vijf maal zo veel wordt verkocht dan normaal? Bij “decompositie” tracht men de promotionele verkooppiek op te splitsen in verschillende bronnen. In de literatuur worden de volgende bronnen onderscheiden:

1. merksubstitutie (“brand switching”);
2. verkorte tussenaankooptijd (“timing acceleration”);
3. een hogere aankoophoeveelheid (“quantity acceleration”). Samen heten 2. en 3. voorraadvorming (“purchase acceleration”);
4. anticipeer-effecten (“anticipatory responses”);
5. winkelsubstitutie (“store switching”);
6. aankopen door het promotie-loyale segment (“deal-to-deal purchasing”);
7. toegenomen consumptie (“increased consumption”);
8. categorie-substitutie (“category switching”).

Merksubstitutie wil zeggen dat men in plaats van het merk dat men van plan was te kopen, een concurrerend merk aanschaft omdat hiervoor een promotionele actie geldt. Voorraadvorming kan op twee manieren plaatsvinden: eerder kopen dan gepland, verkorte tussenaankooptijd, of meer kopen dan gepland, een hogere aankoophoeveelheid (Neslin, Henderson en Quelch 1985). Anticipeer-effecten zijn al in de vorige paragraaf aan de orde gekomen. Winkelsubstitutie wil zeggen dat men het merk in plaats van in de gebruikelijke winkel, in een andere winkel aanschaft omdat het daar in de aanbieding is. Een aantal klanten koopt alleen als er een promotie is: het “promotie-loyale segment”. De verkopen aan deze klanten zijn alleen terug te vinden in de verkooppiek ten gevolge van de promotie en niet in de verkopen in de perioden tussen promoties. Dit segment kan groot zijn: Krishna (1994) vermeldt dat een derde van de huishoudens in Amerika alleen koffie koopt op het moment dat er een koffiemark in de aanbieding is. Sommige klanten hamsteren ten gevolge van promoties en consumeren de extra hoeveelheid vervolgens in een hoger tempo dan normaal: toegenomen consumptie. Hierdoor kopen ze bij het volgende aankoop moment weer net zo veel als normaal. Wansink (1996) stelt dat dit effect zich met name voordoet bij promoties die consumenten

expliciet aanzetten tot een groter aankoopvolume. Bijvoorbeeld: drie voor de prijs van twee, volume-plus acties en dergelijke. Categorie-substitutie, tenslotte, wil zeggen dat men een product uit de ene categorie vervangt door een product uit een andere categorie, dat in de aanbieding is. Naar onze mening zou de decompositie van de extra verkopen als gevolg van een promotie de kern van het sales promotion onderzoek moeten vormen. Zoals al eerder vermeld, zorgen tijdelijke prijskortingen en ook andere promoties voor aanzienlijke tijdelijke verkoopstijgingen. Maar het antwoord op de vraag waar die extra verkopen vandaan komen is cruciaal voor de interpretatie van de promotionele verkooppiek. Immers, wanneer de piek voor 100 procent uit voorraadvorming-effecten zouden bestaan, dan is een herbezinning op het gebruik van sales promotions zeker niet overbodig. Als de verkooppiek voor een aanzienlijk deel verklaard kan worden uit toegenomen consumptie, dan valt continuering van de sales promotions hausse aan te bevelen. Het perspectief van waaruit de decompositie bekeken wordt is ook van groot belang. Zo heeft een fabrikant van een merk onder actie geen problemen met merksubstitutie, terwijl dit voor een winkelier geen netto extra verkochte eenheden oplevert. Anderzijds verkoopt een fabrikant niets extra's op het moment dat winkelsubstitutie plaatsvindt, maar is een winkelier wel weer tevreden met de extra klant; met name als die klant ook de overige boodschappen in die winkel doet. Kortom, decompositie ontleedt in feite het bruto promotionele verkoopeffect en deze ontleding is de basis voor de bepaling van het netto promotionele verkoopeffect voor zowel de fabrikant als de winkelier.

Gupta (1988) heeft een decompositie uitgevoerd waarbij alleen de bronnen merksubstitutie, verkorte tussenaankooptijd en vergrote aankoophoeveelheid worden onderscheiden. Hij heeft op basis van huishoudaankoopgegevens de promotionele verkooppiek voor koffiemerken bestudeerd. Huishoudaankoopgegevens geven de aankopen weer van een vast panel van huishoudens in de loop van de tijd. Gupta kwam uit op 84 procent merksubstitutie, 14 procent verkorte tussenaankooptijd en 2 procent vergrote aankoophoeveelheid. Chiang (1991), Chintagunta (1993), Bucklin, Gupta en Siddarth (1998) en Bell, Chiang en Padmanabhan (1999) hebben een vergelijkbare decompositie uitgevoerd, al besteden de verschillende studies wel aandacht aan verschillende aspecten. Wij doen ook onderzoek naar de decompositie van promotionele verkoopeffecten. Ons onderzoek verschilt met de voorgaande onderzoeken op vijf punten. In de eerste plaats streven wij naar een meer volledige decompositie. De in de vorige alinea beschreven onderzoeken geven naar onze mening geen compleet beeld van de decompositie. Immers, daarin wordt de decompositie beperkt tot slechts drie bronnen en worden impliciet de overige vijf bronnen op nul gesteld. Ons uitgangspunt is anders. Wij zetten geen bronnen a priori op nul, maar stellen de promotionele verkooppiek op 100 procent en proberen de piek zoveel mogelijk te verklaren uit de verschillende bronnen. In de tweede plaats gebruiken wij scanninggegevens op winkelniveau in plaats van huishoudaankoopgegevens, vanwege de redenen die eerder al zijn genoemd. In de derde plaats laten wij de decompositie afhangen van de hoogte van de geboden prijskorting, en laten we dit ook afhangen van de vorm van ondersteuning van de prijskorting (verschil 4). Ten slotte gebruiken we nieuwe nonparametrische methoden (verschil 5).

Vanwege het geaggregeerde karakter van de winkelgegevens waarmee wij de decompositie uitvoeren, is het helaas niet mogelijk alle acht bronnen apart te kwantificeren. Wij kunnen met ons model een decompositie in de volgende drie bronnen verkrijgen:

1. merksubstitutie;
2. dynamische effecten: som van het effect van verkorte tussenaankooptijd, van vergroete aankoophoeveelheid en anticipeer-effecten;
3. categorie-expansie: som van het winkelsubstitutie-effect, de aankopen door het promotie-loyale segment, toegenomen consumptie en categorie-substitutie.

5.2. Aanpak decompositie

Onze aanpak van het decompositie-vraagstuk is in feite een combinatie van het onderzoek naar de deal effect curve en het onderzoek naar de pre- en postpromotionele dips. Het eerste onderzoek heeft ons geleerd dat flexibiliteit in het modelleren van de relatie tussen prijskorting en verkopen geboden is. Het tweede onderzoek bracht aan het licht dat pre- en postpromotionele dips gemodelleerd kunnen worden door vertraagde en toekomstige promotie-variabelen op te nemen. Als we ook de effecten van deze variabelen volledig flexibel zouden willen modelleren met behulp van de traditionele nonparametrische (Kernel) methode, zouden we meer data nodig hebben dan waarover we beschikken: curse of dimensionality. Daarom gebruiken we een nieuwe nonparametrische methode, gebaseerd op het “structured nonparametric regression model” (Linton and Nielsen 1995). Hiermee zijn we toch in staat alle relevante effecten flexibel te schatten.

De kennis uit het onderzoek naar de pre- en postpromotionele dips wordt gebruikt voor de wijze van kwantificeren van dynamische effecten van promoties. Dit doen we met het Unrestricted model. Het decompositie-model bevat uiteindelijk de volgende effecten:

- het effect van een promotie op de eigen verkopen (promotionele verkooppiek);
- het effect van diezelfde promotie op de verkopen van concurrerende merken (merksubstitutie);
- het effect van die promotie op de verkopen van het merk voor en na de promotie (dynamische effecten);

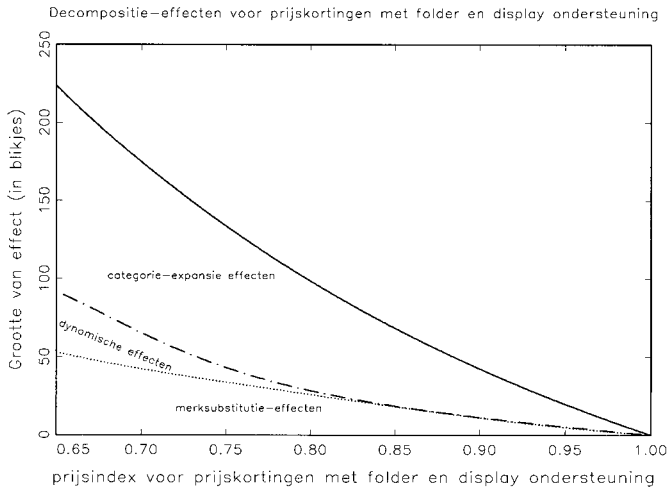
Het categorie-expansie effect wordt vervolgens berekend als het verschil tussen de promotionele verkooppiek en de som van het totale merksubstitutie-effect en de dynamische effecten. Ten slotte worden de drie decompositie-bronnen uitgedrukt als percentage van de promotionele verkooppiek.

5.3. Resultaten decompositie

Dit onderzoek, gebaseerd op de tonijn-dataset, leverde de volgende bevindingen op:

- de grootte van de drie decompositie-effecten (merksubstitutie, dynamische effecten en categorie-expansie) hangt af van (1) de hoogte van de prijskorting en (2) de vorm van ondersteuning die aan de prijskorting wordt gegeven;
- aangezien categorie-expansie effecten vaak groter dan nul zijn, lijkt het onverstandig de decompositie te beperken tot merksubstitutie en voorraadvorming-effecten, zoals in voorgaand decompositie-onderzoek is gedaan;
- het model stelt ons in staat het netto gedeelte van de promotionele verkooppiek te berekenen voor fabrikanten en detaillisten. De conclusies gebaseerd op deze netto effecten kunnen heel anders uitpakken dan wanneer er gekeken wordt naar de bruto verkoopeffecten, die weergegeven worden door de promotionele verkooppieken.

Ter illustratie van de bevindingen wordt in Figuur 2 de decompositie van het effect van variërende prijskortingen met folder en display ondersteuning gegeven. Het grootste deel van de verkooptoeename kan worden verklaard uit categorie-expansie effecten. Daarna komen de merksubstitutie-effecten (cross-brand effects). Pas vanaf een prijskorting van 20 procent (prijsindex kleiner dan 0.80) doen voorraadvorming- (dynamische) effecten zich voor. Het netto effect van deze promotie voor een fabrikant is gelijk aan de som van de categorie-expansie effecten en de merksubstitutie-effecten; het netto effect voor de detaillist wordt gevormd door alleen de categorie-expansie effecten.



Figuur 2. Decompositie van het promotionele verkoopeffect bij variërende niveaus van prijskorting met folder en display ondersteuning.

6. CONCLUSIE EN TOEKOMSTIG ONDERZOEK

In dit artikel worden econometrische modellen besproken die de gebruiker in staat stellen de netto verkoopeffecten van promoties te bepalen, waarbij promoties gekarakteriseerd worden door (1) de hoogte van de geboden tijdelijke prijskorting en (2) de vorm van ondersteuning: folder en/of display, of geen van beide. De netto effecten stellen fabrikanten en detaillisten in staat om betere sales promotion strategieën te ontwikkelen. Met behulp van gegevens over de relevante marges en kosten, kunnen de netto verkoopeffecten worden vertaald in winstgevendheidscijfers.

In toekomstig onderzoek gaan we modellen ontwikkelen voor de effecten van sales promotions, die behulpzaam kunnen zijn bij het implementeren van winkelspecifieke promoties. Deze vorm van micro-marketing wordt mogelijk door het beschikbaar komen van uitgebreide gegevens over de klanten en concurrenten van afzonderlijke winkels.

LITERATUUR

- Bell, D.R., J. Chiang en V. Padmanabhan, 1999. The Decomposition of Promotional Response: An Empirical Generalization. *Marketing Science* 18, nr. 4, 504-526.
- Blattberg, R.C., R. Briesch en E.J. Fox, 1995. How Promotions Work. *Marketing Science* 14, nr. 3, Deel 2 van 2, g122-g132.
- Bucklin, R.E., S. Gupta en S. Siddarth, 1998. Determining Segmentation in Sales Response Across Consumer Purchase Behaviors. *Journal of Marketing Research* 35, 189-197.
- Chiang, J., 1991. A Simultaneous Approach to the Whether, What and How Much to Buy Questions. *Marketing Science* 10, nr. 4, 297-315.
- Chintagunta, P.K., 1993. Investigating Purchase Incidence, Brand Choice, and Purchase Quantity Decisions of Households. *Marketing Science* 12 nr. 2, 184-208.
- Gupta, S. and L.G. Cooper, 1992. The Discounting of Discounts and Promotion Thresholds. *Journal of Consumer Research* 19: 401-411.
- Gupta, S., 1988. Impact of Sales Promotion on When, What and How Much to Buy. *Journal of Marketing Research* 25, 342-355.
- Härdle, W., 1990. *Applied Nonparametric Regression - Econometric Society Monographs*, No. 19. Cambridge: Cambridge University Press.
- Heerde, H.J. van, 1999. *Models for Sales Promotion Effects Based on Store-Level Scanner Data*. Capelle a/d IJssel: Labyrinth Publication.
- Krishna, A., 1994. The Effect of Deal Knowledge on Consumer Behavior. *Journal of Marketing Research* 31, 76-91.
- Lee, M.j., 1996. *Methods of Moments and Semiparametric Econometrics for Limited Dependent Variable Models*. New York: Springer-Verlag.
- Linton, O. en J.P. Nielsen, 1995. A Kernel Method of Estimating Structured Nonparametric Regression Based on Marginal Integration. *Biometrika* 82, nr. 1, 93-100.
- Neslin, S.A., C. Henderson en J. Quelch, 1985. Consumer Promotions and the Acceleration of Product Purchases. *Marketing Science* 4, nr. 2, 147-165.
- Neslin, S.A. en L.G. Schneider Stone, 1996. Consumer Inventory Sensitivity and the Postpromotion Dip. *Marketing Letters* 7, nr. 1, 77-94.
- Robinson, P.M., 1988. Root-N-Consistent Semiparametric Regression. *Econometrica* 56, nr. 4, 931-954.
- Wansink, B., 1996. Can Package Size Accelerate Usage Volume?. *Journal of Marketing* 60, 1-14.

10. Meten wat we allemaal al weten

Non-respons- en intervieweffecten bij de meting van opkomst tijdens de Tweede Kamerverkiezingen van 1998 in het Nationaal Kiezersonderzoek¹

H. VAN DER KOLK²

SAMENVATTING

In het kader van het NKO 1998 is aandacht geschonken aan zowel het non-respons- als aan het intervieweffect. In deze bijdrage beschrijven we de manier waarop dat is gebeurd en gaan we na welk gevolg beide factoren hebben voor de overrapportage van opkomst bij de nationale verkiezingen van 1998. De vraag die in deze bijdrage centraal staat is in hoeverre non-respons- en intervieweffecten een bijdrage leveren aan de verklaring van het (toegenomen) verschil tussen gemeten opkomst en feitelijke opkomst. Naar aanleiding van die vraag worden drie verwachtingen geformuleerd: opkomst hangt samen met het wel of niet meedoen aan enquêtes, voor de verkiezingen benaderd worden voor een interview over politiek leidt tot een hogere stemkans en een intensieve benadering voor de verkiezingen leidt ook tot een hogere stemkans dan een korte telefonische benadering. Hoewel enkele problemen in het design ferme uitspraken naar aanleiding van deze verwachtingen verhinderen, wordt aangegeven dat het verschil in gemeten en feitelijke opkomst niet alleen te wijten is aan non-respons en misrapportage; ook intervieweffecten spelen een rol.

1. PROBLEMEN ROND HET NATIONAAL KIEZERSONDERZOEK

Het Nationaal Kiezersonderzoek (NKO) wordt door sommigen beschouwd als het paradigma van de Nederlandse politicologie. In een groeiend aantal publicaties over veranderingen in de achtergronden van kiesgedrag en opkomst wordt gebruik gemaakt van de gegevens die in het kader van de opeenvolgende NKO's zijn verzameld. Toch is er recentelijk ook veel kritiek geweest op deze onderzoeken (Visscher 1995; Van Goor

1. De gegevens die in deze publicatie zijn gebruikt, zijn verzameld in het kader van het Nationaal Kiezersonderzoek 1998 door Kees Aarts en Henk van der Kolk namens de Stichting Kiezersonderzoek Nederland (SKON). De studie werd mogelijk gemaakt door bijdragen van NWO, de ministeries van Binnenlandse Zaken en Koninkrijksrelaties en Volksgezondheid, Welzijn en Sport, het Sociaal en Cultureel Planbureau, en de universiteiten van Amsterdam en Twente. De data worden in Nederland verspreid door het Steinmetz-archief.

2. Ik wil Jacques Thomassen bedanken voor zijn commentaar bij een eerdere versie van dit artikel en Kees Aarts voor zijn suggesties bij het gebruik van het logit-model in paragraaf 5.

1996). De verschillende kiezersonderzoeken zouden in toenemende mate een slecht beeld geven van de politieke opvattingen en gedragingen van Nederlandse burgers. Een veel besproken aanwijzing voor dit vermoeden is de door het NKO gemeten toegenomen politieke interesse. Ongewogen NKO cijfers geven aan dat de politieke interesse in Nederland, gemeten aan de hand van het aantal mensen dat een positief antwoord geeft op de vraag 'bent u (zeer) in politiek geïnteresseerd?', tussen 1971 en 1994 is gestegen van 53 naar 78 procent. Andere interpretaties van dezelfde cijfers nuanceren dit beeld, maar geven nog steeds een, zij het lichte, toename te zien. Volgens sommige critici kan het echter niet zo zijn dat sinds de jaren zeventig de politieke interesse is toegenomen. Het ligt veel eerder voor de hand dat de politieke interesse, vooral na de eerste helft van de jaren tachtig, is gedaald. Dat zou bijvoorbeeld verklaren waarom de discussie opblaaide over de 'kloof tussen burger en politiek'.

Nu is het natuurlijk mogelijk dat de discrepantie tussen wat we allemaal lijken te weten (afgenomen interesse) en wat we waarnemen met behulp van het NKO (toegenomen interesse) veroorzaakt wordt door de onjuistheid van hetgeen we allemaal lijken te weten. Vaker wordt echter gesteld dat er iets mis is met het NKO. De oorzaak van de gemeten toename van de politieke interesse moet volgens degenen die deze opvatting zijn toegedaan niet worden gezocht in een feitelijke toename van die interesse, maar in de toegenomen non-respons. De non-respons in het NKO is in de afgelopen jaren toegenomen (Van der Kolk 2000). Als inderdaad steeds minder politiek ongeïnteresseerde burgers meedoen aan enquêteonderzoek lijkt de interesse toe te nemen, terwijl dat feitelijk niet het geval is (Visscher 1995).

Voor de veronderstelling dat de waargenomen toename van politieke interesse een gevolg is van non-respons zijn sterke aanwijzingen. Non-respons wordt veroorzaakt door mislukte contactpogingen en door weigeringen. Nu zouden vooral weigeraars sterk afwijken van de rest van de populatie. Weigeringen gebeurt immers voor een belangrijk deel nadat het onderwerp van het interview bekend is gemaakt. In de opeenvolgende NKO's is het aandeel weigeraars onder de non-respondenten sterk toegenomen (Visscher 1995; Smeets 1995).³ Als we veronderstellen dat *alle* weigeraars niet in politiek zijn geïnteresseerd, is de politieke interesse in Nederland in de jaren zeventig eerst toegenomen en daarna in de jaren tachtig gedaald. In de jaren negentig is het niveau stabiel gebleven (Visscher 1995; Van der Kolk 2000: 225). Dit beeld lijkt meer in overeenstemming te zijn met wat we allemaal denken te weten. De gemeten toename van politieke interesse in het NKO zou dan ook niet zozeer wijzen op een werkelijke toename, als wel op een systematisch afgenomen validiteit van de waarneming.

Hoe plausibel de redenering ook is, zij is voornamelijk gebaseerd op niet onderzochte veronderstellingen. We weten immers niet of weigeraars (veel) minder in politiek zijn geïnteresseerd dan degenen die wel hebben meegedaan met het vraagesprek. Vergeten wordt wel eens dat ook onder de respondenten mensen zijn die niet in politiek zijn geïnteresseerd. Zo gaf in 1994, toen het hoogste niveau van politieke interesse werd gemeten, ook ruim 20 procent van de respondenten aan niet in politiek te zijn geïnteresseerd. Uit het hierna beschreven onderzoek rond het NKO 1998 is gebleken dat de ver-

3. Bovendien zeggen weigeraars vaak dat ze niet mee willen doen omdat ze niet geïnteresseerd zijn (Visscher 1995: 30). Uiteraard weten we niet of dat betekent dat ze niet in politiek geïnteresseerd zijn, of dat zij niet geïnteresseerd zijn in een interview van een uur met een onbekende interviewer.

onderstelling dat alle weigeraars niet in politiek zijn geïnteresseerd waarschijnlijk veel te sterk is (Van der Kolk 2000). Correcties van de cijfers uit eerdere NKO's op basis van dat onderzoek wijzen eerder op een stabiel niveau van politieke interesse dan op een daling. Een belangrijk probleem bij onderzoek naar dit onderwerp is echter dat we niet weten wat het werkelijke niveau van politieke interesse is.

Terwijl we van het feitelijke niveau van politieke interesse vrijwel niets weten, is de opkomst en de uitslag van verkiezingen wel bekend. Door te meten wat we al weten (opkomst en stemgedrag), kunnen we het meetinstrument 'ijken'. Op die manier kunnen we proberen iets te zeggen over de respondenten die in geen enkel onderzoek worden benaderd. Komt er uit de vergelijking van voorspelde en gemeten opkomst en uitslag naar voren dat er mogelijk iets schort aan het Nationale Kiezersonderzoek?

Wat de opkomst betreft moeten we deze vraag zeker met 'ja' beantwoorden. Het verschil tussen feitelijke en gerapporteerde opkomst bedroeg tot en met 1986 minimaal 3 en maximaal 8 procent. In 1989 steeg het verschil echter naar 12,3 procent. In 1994 en 1998 bedroeg het verschil 13,5 respectievelijk 17,5 procent. Er lijkt dus een toenemende discrepantie te bestaan tussen de feitelijke opkomst en de opkomst zoals gemeten in het NKO. Ook dit is dus een aanwijzing dat er wat mis is met het NKO. Ook de uitslag van de verkiezingen werd niet in alle jaren precies weerspiegeld in de enquête, hoewel daarbij van een stijgende trend geen sprake is. Deze verschillen wijzen dus niet op toenemende problemen bij het NKO.

Tabel 1. Verschillen tussen feitelijke en gemeten uitslag van Tweede Kamerverkiezingen.

Jaar	Feitelijke opkomst	Gemeten opkomst ¹	Absolute verschil ²	Relatieve verschil ³	Vershil stembusuitslag ⁴
1971	79,1	86,1	7,0	33,5	10,8
1972	83,5	89,4	5,9	35,8	15,5
1977	88,0	91,1	3,1	25,8	12,8
1981	87,0	93,5	6,5	50,0	14,3
1982	81,0	89,0	8,0	42,1	11,7
1986	85,8	93,1	7,3	51,4	8,3
1989	80,3	92,6	12,3	62,4	8,8
1994	78,7	92,2	13,5	63,4	15,1
1998	73,3	90,8	17,5	66,0	11,6

1. Het gaat om het percentage respondenten dat na de verkiezingen opgeeft gestemd te hebben.
2. Dit is het verschil tussen de gemeten opkomst en de feitelijke opkomst.
3. Omdat het absolute verschil wordt beperkt door plafondeffecten, is ook het relatieve verschil weergegeven. Dit is het absolute verschil gedeeld door het percentage mensen dat feitelijk niet heeft gestemd.
4. Dit zijn de gesommeerde absolute verschillen tussen stembusuitslag en enquêtecijfers.

Omdat de discrepantie tussen feitelijke en gemeten uitslag geen duidelijke trend vertoont, is het waarschijnlijk niet zo zinvol naar aanleiding hiervan de mogelijke gevolgen van de toegenomen non-respons te onderzoeken. In deze bijdrage richten we ons daarom verder uitsluitend op het verschil tussen feitelijke en gemeten opkomst. Omdat de feitelijke opkomst bekend is, hopen we op die manier iets meer te weten te komen over het belang van non-respons.

2. VERKLARINGEN VAN HET VERSCHIL TUSSEN FEITELIJKE EN GEMETEN OPKOMST

Non-respons is niet de enige verklaring voor het verschil tussen gemeten en feitelijke opkomst. In de literatuur worden ten minste drie andere oorzaken genoemd van dit verschil: steekproeffluctuaties, misrapportage en intervieweffecten. Steekproeffluctuaties zijn in het kader van het NKO waarschijnlijk het minst belangrijk. De groep geïnterviewde respondenten ligt in alle jaren rond de 2000. Met een N van 2000 en een feitelijke opkomst van 73,3 procent (de opkomst bij de verkiezingen van 1998) is de kans zeer klein dat uitsluitend als gevolg van steekproeffluctuaties het gevonden opkomst percentage meer dan 1 procentpunt afwijkt van 73,3 procent.

Bovendien verklaart de steekproeffluctuatie niet waarom er altijd een *overrapportage* plaatsvindt van opkomst. Dat laatste geldt ook voor de omstandigheid dat in de loop der jaren licht afwijkende steekproefkaders zijn gebruikt voor het NKO (Van der Kolk 2000: 221-222).

Misrapportage, de tweede verklaring van het verschil tussen gemeten en feitelijke opkomst, is in Nederland niet goed te onderzoeken. We kunnen verwachten dat herinneringseffecten in een onderzoek dat vlak na de verkiezingen wordt gehouden geen belangrijke rol spelen. Maar dat sluit niet uit dat kiezers, bijvoorbeeld als gevolg van een sterke sociale norm, een onjuist antwoord geven. In Nederland kunnen we echter niet controleren of respondenten daadwerkelijk zijn gaan stemmen. Uit onderzoek in Zweden komt naar voren dat maar liefst 97 procent van de respondenten het juiste antwoord geeft op de vraag of ze hebben gestemd (in Smeets 1995: 311). Niet geheel onverwacht blijkt uit dat onderzoek ook dat de onjuiste antwoorden vooral, maar niet uitsluitend, voor rekening komen van niet-stemmers die opgeven wel te hebben gestemd. Uit Amerikaans onderzoek komt naar voren dat het verschil tussen gerapporteerde en feitelijke opkomst in dat land groter is (Burden 1999: 13 ev.). In de jaren zeventig en tachtig zou in dat land ongeveer de helft van het verschil tussen feitelijke en gemeten opkomst te wijten zijn aan misrapportage en herinneringseffecten. Misrapportage kan dus waarschijnlijk, in tegenstelling tot steekproeffluctuaties, een deel van het verschil tussen feitelijke en gemeten opkomst verklaren. Omdat de mogelijkheid dit te onderzoeken ontbreekt, concentreren we ons echter verder op de derde en vierde factor: selectieve respons⁴ en intervieweffecten.

In haar studie naar de oorzaken van het verschil tussen de feitelijke en de gemeten opkomst probeert Smeets de vraag of respondenten en non-respondenten verschillen te beantwoorden aan de hand van de moeilijk en makkelijk bereikbare respondenten (Smeets 1995). De respondenten die op het laatste nippertje werden overgehaald worden door haar gebruikt als representanten van de groep die helemaal niet heeft meegedaan aan het onderzoek. Smeets concludeert op basis van haar analyse dat het wel meevalt met de verschillen tussen respondenten en non-respondenten. Daarmee lijkt op het eerste gezicht non-respons het geconstateerde gat tussen gemeten en feitelijke opkomst niet te kunnen vullen. Maar Smeets denkt ook dat alle door haar onderzochte groepen respondenten sterk afwijken van de echte weigeraars. Daarmee wordt non-respons weer geïntroduceerd als belangrijke verklaring voor de geconstateerde discrepantie. Ook

4. Voor een overzicht naar de oorzaken en gevolgen van non-respons wordt verwezen naar Brehm (1993).

wordt daarmee het belang van het aantal contactpogingen als belangrijke (non-)respons variabele gerelativeerd.⁵

Het probleem dat verbonden is met de analyse van het aantal contactpogingen gaat niet op voor de analyse van de verschillen tussen groepen respondenten die niet aan het hele onderzoek hebben deelgenomen. Zo'n analyse is onder andere uitgevoerd door Anker (Anker 1995). Hij richtte zich op de respondenten die uitvielen tussen de eerste en de tweede golf van het onderzoek. Uit dit soort onderzoek komen weliswaar enkele significante verschillen naar voren tussen uitvallers en mensen die ook aan de tweede golf meedoen, maar een groot deel van de uitval blijft onverklaard. Ook de kans dat iemand die niet van plan is te gaan stemmen ook niet meer mee meedoet aan de tweede golf is niet significant kleiner of groter dan de kans dat iemand die wel van plan is te gaan stemmen blijft meedoen aan het onderzoek (Anker 1995: 251). Anker merkt dan ook op basis van dit soort analyses op dat non-respons waarschijnlijk geen grote effecten heeft en dat de uitval niet sterk selectief is (Anker 1995: 252). Nadeel van deze methode en ook van de methode van Smeets is uiteraard dat de sterke selectie van respondenten mogelijk al eerder, voorafgaande aan het eerste gesprek, heeft plaatsgevonden (Van Goor 1996: 169). In dat geval lijken de uitvallers dus niet op non-respondenten, maar veel meer op mensen die wel gewoon hebben meegedaan aan het onderzoek. Het is dus de vraag of dit de juiste methode is om iets te zeggen over non-respondenten (Van der Kolk 2000: 227-228).

De vierde factor die het verschil tussen feitelijke en gemeten opkomst kan verklaren is het intervieweffect. Het hoofdonderzoek van vrijwel elk NKO bestaat uit twee interviews: één voor en één na de verkiezingen. Het is niet ondenkbaar dat respondenten voornamelijk als gevolg van het interview voorafgaande aan de verkiezingen besluiten te stemmen, terwijl ze dat zonder interview niet gedaan zouden hebben. Ook dit intervieweffect zou het verschil tussen gemeten en feitelijke opkomst goed kunnen verklaren. Dit was in het kader van het NKO echter nog niet onderzocht. Op basis van buitenlandse studies lijkt dit effect echter belangrijk genoeg om aandacht aan te geven (Claussen 1969; Kraut & McConahay 1973; Yalch 1976). Dat neem niet weg dat het intervieweffect waarschijnlijk niet het *toegenomen* verschil tussen gemeten en feitelijke opkomst kan verklaren (Burden 1999: 14).

Om zowel het non-respons- als het intervieweffect te onderzoeken is in het kader van het NKO 1998 opnieuw nagedacht over de onderzoeksopzet en benadering van de respondenten. In deze bijdrage beschrijven we hoe dat is gebeurd en gaan we na welk gevolg beide factoren hebben voor de overrapportage bij opkomst bij de nationale ver-

5. Achteraf ligt het ook wel voor de hand dat het aantal contactpogingen geen goede indicatie is voor non-respons. Het aantal contactpogingen dat nodig is om iemand over te halen heeft vooral iets te maken met de groep waarmee niet makkelijk contact kan worden gelegd, bijvoorbeeld omdat ze veel van huis zijn. Een deel van die groep behoort uiteindelijk tot de non-respondentengroep 'geen contact'. De groep waarvan we echter verwachten dat die verschilt van de respondenten zijn nu juist de weigeraars. Mede om die reden is er in het NKO 1998 niet voor gekozen het aantal contactpogingen te registreren. Omdat de onderzoekers de interviewers alle ruimte wilden geven een hoge respons te realiseren leek het niet zinvol ze elke contactpoging (telefonisch, schriftelijk, aan de deur) te laten registreren. Zeker ook omdat de contactpogingen zeer divers zouden zijn (drie keer in gesprek versus drie keer de verkeerde persoon aan de deur).

kiezingen van 1998. De vraag die in de rest van deze bijdrage centraal staat is dus in hoeverre non-respons- en intervieweffecten een bijdrage leveren aan de verklaring van het gevonden verschil tussen gemeten opkomst en feitelijke opkomst.

3. ONDERZOEKSOPZET VAN HET NKO 1998

Om non-respons- en intervieweffecten te onderzoeken zijn in het kader van het NKO 1998 op dezelfde wijze twee afzonderlijke steekproeven getrokken. De eerste steekproef had een omvang van 4207 eenheden. De tweede steekproef heeft een omvang van 1451 eenheden. Deze twee steekproeven verschillen niet significant naar regio, stedelijkheid, leeftijd en geslacht.

Steekproef I werd al voor de verkiezingen benaderd middels een brief en door afzonderlijke interviewers. Uiteindelijk hebben 2101 respondenten uit steekproef I meegedaan aan een face-to-face interview voorafgaande aan de verkiezingen (de eerste golf). Vervolgens hebben 1814 respondenten ook nog eens meegedaan aan een face-to-face interview na de verkiezingen (de tweede golf). Een schriftelijke vragenlijst, die werd achtergelaten bij de 1814 respondenten die deelnamen aan de tweede golf werd teruggestuurd door 1199 personen. De eerste golf, de tweede golf en de schriftelijke vragenlijst worden samen wel het 'hoofdonderzoek' genoemd.

In het kader van een 'non-respons onderzoek' zijn ook alle 2106 non-respondenten uit steekproef I telefonisch benaderd. Allereerst is deze personen gevraagd of zij alsnog wilden meedoen aan een mondeling interview.⁶ Hierop werd door een zeer klein aantal respondenten bevestigend geantwoord. Helaas was de veldwerkorganisatie niet in staat deze mensen ook daadwerkelijk te interviewen. Wel werd hun bereidheid om te worden geïnterviewd geadmistreerd. Degenen die wederom weigerden werd vervolgens gevraagd of ze wel wilden meedoen aan een zeer kort telefonisch interview. Op die manier konden 926 non-respondenten alsnog kort telefonisch worden geïnterviewd.

Omdat de respondenten uit steekproef I allemaal voorafgaande aan de verkiezingen werden geïnterviewd, verwachtten we dat hun opkomstgedrag positief zou worden beïnvloed. Om te onderzoeken of dat daadwerkelijk het geval is, werd steekproef II (1451 eenheden) pas na de verkiezingen benaderd. Dit gebeurde op vrijwel dezelfde wijze als bij steekproef I. Financieel was het echter niet haalbaar na het toezenden van de brief ook interviewers op pad te sturen. De tweede steekproef werd dus telefonisch benaderd. Omdat een groot aantal mensen telefonisch niet kon worden bereikt, werd alle onbereikbare respondenten een briefkaart toegestuurd met daarop enkele vragen. Van de tweede steekproef werden uiteindelijk 712 personen telefonisch geïnterviewd en stuurden nog eens 85 personen een (ingevulde) briefkaart terug.

Om er voor te zorgen dat beide steekproeven zo goed mogelijk te vergelijken zijn, werd respondenten in de tweede steekproef eerst de vraag voorgelegd of ze eventueel bereid

6. Omdat er bij de opzet van het onderzoek voor is gekozen te proberen telefonisch alle non-respondenten over te halen toch mee te doen aan het onderzoek, is er niet slechts een steekproef uit de non-respons groep benaderd. Indien uitsluitend de studie van non-respondenten onze belangstelling had gehad, zou deze optie wellicht de voorkeur hebben gehad.

waren deel te nemen aan een *face-to-face* interview.⁷ Op die manier kon analoog aan het onderscheid tussen respondenten en non-respondenten in de stimulusgroep (steekproef I) een onderscheid worden gemaakt tussen respondenten en non-respondenten in de controlegroep (steekproef II). Daarna volgde een kort interview met vragen over onder meer politieke interesse en het feitelijke stemgedrag.⁸

Door de gekozen opzet kan een onderscheid worden gemaakt tussen vier groepen respondenten in twee verschillende steekproeven. Deze groepen staan ook in tabel 2. De eerste groep bestaat uit de respondenten die geïnterviewd wilden worden in een *face-to-face* interview. De tweede groep bestaat uit de respondenten die dat niet wilden, maar die wel bereid waren telefonisch enkele vragen te beantwoorden. De derde groep bestaat uit respondenten die weigerden mee te werken. De vierde groep bestaat uit degenen waarmee geen contact kon worden gelegd.

Naast deze algemene classificatie kunnen we binnen groep 1 van steekproef I (degenen die tenminste hebben meegewerkt aan de eerste golf) een onderscheid maken tussen (a) degenen die hebben meegedaan aan beide golven en de schriftelijke vragenlijst, (b) degenen die alleen hebben meegedaan aan beide golven maar de schriftelijke vragenlijst niet hebben teruggestuurd en (c) degenen die alleen aan de eerste golf hebben meegedaan, maar die na de verkiezingen niet nogmaals *face-to-face* zijn geïnterviewd. Wel zijn sommige respondenten uit deze groep 1c na de verkiezingen nog telefonisch geïnterviewd over hun stemgedrag.

Tabel 2. Het steekproefdesign.

Groep	Steekproef I (N = 4207)	Steekproef II (N = 1451)
1	A Meegedaan aan de eerste en tweede golf en aan de schriftelijke vragenlijst B Meegedaan aan de eerste en tweede golf C Meegedaan aan de eerste golf	Bereid mee te werken aan een <i>face-to-face</i> interview
2	Uitsluitend geïnterviewd in de non-responsstudie	Niet bereid mee te werken aan <i>face-to-face</i> interview
3	Weigering	Weigering
4	Geen contact	Geen contact

Op basis van het design kunnen we verschillende zaken onderzoeken. In de eerste plaats kunnen we de (voorgenomen) opkomst in de groepen 1a, 1b, 1c en 2 vergelijken.

7. De instructie aan de enquêteurs luidde als volgt: Goedemiddag\avond u spreekt met ... van ITS Nijmegen. Ik zou graag willen spreken met ... Bent u dat, of zou deze persoon even aan de lijn kunnen komen? Kort geleden hebben we u een brief gestuurd met het verzoek om mee te werken aan het Nationaal Kiesonderzoek. Zou u in principe bereid zijn hieraan mee te willen werken?

8. Als respondenten niet wilden meedoen, werd gevraagd naar enkele kenmerken van de non-respondent: We respecteren uw mening. Toch wil ik u nog wat vragen. U zou ons erg helpen als u nu een paar korte vragen zou willen beantwoorden. Dit duurt hooguit 2 minuten. Kan dat?

Op die manier kunnen we iets meer zeggen over de effecten van non-respons op de gemeten opkomst. We verwachten dat de opkomst lager is in groepen die minder bereid zijn zich te laten ondervragen (verwachting 1). Verder kunnen we op basis van dit design iets zeggen over het intervieweffect. In algemene zin verwachten we dat de gemeten opkomst in de controlegroep (steekproef II) lager is dan in de stimulusgroep (steekproef I) (verwachting 2). Meer specifiek verwachten we dat dit effect het sterkst naar voren komt in een vergelijking van de groepen 1. Immers, de respondenten uit steekproef I (groep 1) zijn uitvoerig geïnterviewd, terwijl degenen uit steekproef II die aangaven wel te willen worden geïnterviewd voor de verkiezingen niet zijn benaderd. De stimulus is voor de kort ondervraagde non-respondenten van steekproef I (groep 2) echter veel kleiner dan voor de respondenten. We verwachten daarom een kleiner intervieweffect bij deze groep dan bij groep 1 (verwachting 3).

4. NON-RESPONSEFFECTEN

Voordat we de resultaten van het gehele design in ogenschouw nemen, richten we ons eerst op de resultaten van het hoofdonderzoek en het non-responsonderzoek. Het gaat dus uitsluitend om de resultaten van steekproef I. Gaat verwachting 1 inderdaad op? Stemmen respondenten inderdaad vaker dan non-respondenten? En is er een verband tussen de mate waarin een respondent bereid is mee te werken en de kans dat hij gaat stemmen?

Tabel 3. Gemeten opkomst in steekproef I.

Steekproef I (N = 4207)						
Groep	Hoofdonderzoek (N = 2101)			Non-responsonderzoek (N = 2106)		
	1	2	3	4		
Sub-groep	A Respondenten in twee golven en de sv	B Respondenten in twee golven	C Respondenten in één golf	Telefonisch ondervraag- de weige- raars	Weige- raars	Geen contact
N	1199	615	287	926	1039	141
Intentie ⁹ (%)	94,7	91,4	85,8	85,8	??	??
N	1132	558	261	802		
Opkomst ¹⁰ (%)	92,7	87,2	81,7	??	??	??
N	1199	615	142			

Toelichting: In deze tabel staan gegevens over de gemeten opkomstintentie en gerapporteerde opkomst voor heel steekproef I (N = 4207). In het hoofdonderzoek zijn maximaal 2101 mensen ondervraagd. Een groot aantal van de overige 2106 respondenten (de 'non-respons') is kort telefonisch geïnterviewd. Enkele respondenten gaven geen (valide) antwoord op de gestelde vraag of werden na de verkiezingen (groep 1c) niet meer bereikt. Daarom staan de aantallen waarop de percentages zijn gebaseerd ook in de tabel.

Er blijkt inderdaad een verschil te bestaan tussen respondenten en (verschillende groepen) non-respondenten. De verschillen tussen de onderscheiden groepen zijn significant, substantieel en in de verwachte richting. De opkomstintentie daalt naarmate een respondent minder bereid is mee te doen met het gehele onderzoek van bijna 95 tot iets minder dan 86 procent. Opvallend daarbij is wel dat de opkomstintentie van de onderzochte non-respondenten niet afwijkt van die van respondenten die maar aan één golf hebben deelgenomen. De telefonisch geïnterviewde non-respondenten wijken met andere woorden op dit punt niet sterk af van de uitvallers. Ook de gerapporteerde opkomst vertoont een duidelijke relatie met de mate waarin iemand bereid was mee te doen met het onderzoek.

Probleem blijft natuurlijk dat we van meer dan de helft van de non-respondenten (een kwart van de steekproef) niet weten of ze van plan waren te stemmen.¹¹ Toch willen we op basis van het non-responsonderzoek iets kunnen zeggen over de gehele groep non-respondenten. Daarom concentreren we ons op de groep waarvan we vrijwel niets weten. Wijkt deze groep nu af van de groep wel geïnterviewde non-respondenten?

Een eerste aspect van dit verschil is de opkomst bij de Tweede Kamerverkiezingen van 1998. Laten we eens aannemen dat (1) de respondenten een juist antwoord hebben gegeven (we negeren dus misrapportage), dat ze (2) niet zijn beïnvloed door het interview voorafgaande aan de verkiezingen (we negeren intervieweffecten) en dat (3) de feitelijke opkomst onder de geïnterviewde non-respondenten, net als bij de twee groepen waarvan we wel zowel de opkomstintentie als de gerapporteerde opkomst weten, vier procentpunten lager ligt dan de voorgenomen opkomst. Op basis van deze aannemen moeten we vaststellen dat de opkomst onder de geïnterviewde non-respondenten beduidend hoger ligt dan onder de overige non-respondenten. De feitelijke opkomst in Nederland bedroeg in 1998 immers 73,3 procent. Als de drie veronderstellingen inderdaad kloppen, betekent dit, dat van degenen die door de interviewers op geen enkele manier geïnterviewd zijn slechts 43 procent is gaan stemmen! Dit zou betekenen dat de opkomst onder de telefonisch geïnterviewde non-responsgroep sterk afwijkt van de rest van de opkomst in de rest van de non-responsgroep.¹²

Voordat we nu concluderen dat er *dus* sprake is van een sterk non-responseffect onderzoeken we eerst of de telefonisch geïnterviewde non-respondenten (groep 2) ook op andere punten sterk afwijken van de overige non-respondenten (de groepen 3 en 4). Als dit het geval is, is dat een aanwijzing dat ons onderzoek onder non-respondenten minder valide is. Dat blijkt gelukkig niet het geval. Omdat voor het NKO 1998 gekozen is voor een personensteekproef kennen we de leeftijd en de sekse van alle potentiële respondenten. Uit de analyse blijkt dat de gemiddelde leeftijd en de verdeling naar sekse van respondenten in groep 2 niet significant verschilt van die uit de gecombineerde groepen 3 en 4. Verder zijn de gegevens gekoppeld aan gegevens van Geo-marktpro-

9. De vraag luidt: 'Zoals u misschien weet zijn er in mei van dit jaar verkiezingen voor de Tweede Kamer. Gaat u dan wel of niet stemmen, of weet u dat nog niet?' Degenen die nog niet weten of ze gaan stemmen zijn verwijderd uit de analyse.

10. De vraag luidt: 'Hebt u gestemd bij de Kamerverkiezingen van 6 mei?'

11. En van de helft van de steekproef niet weten of ze feitelijk hebben gestemd.

12. We moeten niet vergeten dat een deel van de non-respons wordt veroorzaakt door bijvoorbeeld vakantie, of verblijf in het buitenland. Een deel van deze activiteiten veroorzaakt ook de lagere opkomst.

fiel.¹³ Deze gegevens hebben betrekking op de geschatte huizenprijzen, inkomen en opleiding in een postcodegebied. Ook uit deze analyse blijkt dat de onderscheiden non-responsgroepen (2, 3 en 4) niet significant van elkaar verschillen.

Terwijl de berekende opkomst onder non-respondenten waarvan we niets afweten sterk afwijkt van de non-respondenten waarvan we wel iets afweten, kunnen we geen andere verschillen ontdekken. Dat is een extra stimulans om nog eens te kijken naar de veronderstellingen waarop deze berekening is gebaseerd. Zo is verondersteld dat respondenten niet worden beïnvloed door het interview voorafgaande aan de verkiezingen. Het ligt echter voor de hand dat dit wel het geval is.

5. INTERVIEWEFFECTEN

Over het effect van het interview op de opkomst bij verkiezingen zijn eerder twee verwachtingen uitgesproken. Allereerst verwachten we dat als gevolg van het interview respondenten vaker stemmen dan vergelijkbare burgers die niet zijn geïnterviewd (verwachting 2). Het gemeten opkomstpercentage in steekproef I zal dus hoger zijn dan in steekproef II. Verder verwachten meer specifiek ook verschillen tussen de onderscheiden groepen 1 en 2. Zo is de non-responsgroep (groep 2) uit steekproef I nauwelijks geïnterviewd voor de verkiezingen. Deze groep zal dus minder worden gestimuleerd te stemmen dan de respondenten die een interviewer over de vloer krijgen om een uurtje over politiek te spreken (groep 1 uit steekproef I) (verwachting 3). Deze verwachtingen onderzoeken we aan de hand van een vergelijking van de twee eerder besproken steekproeven. Eerst gaan we in op het algemene verschil tussen steekproef I en steekproef II en onderzoeken we de tweede verwachting. Daarna maken we in beide steekproeven een onderscheid tussen respondenten en telefonisch geïnterviewde non-respondenten en onderzoeken we de derde verwachting.

5.1. Het globale effect van het interview voorafgaande aan de verkiezingen

Zoals gezegd, hebben we geprobeerd de twee steekproeven op dezelfde manier te benaderen. In tabel 4 staan de resultaten van de benadering van de twee steekproeven. Ondanks de pogingen de beide steekproeven op dezelfde manier te benaderen, doet een eerste inspectie van deze tabel 4 al vermoeden dat een vergelijking tussen beide steekproeven niet zonder meer mogelijk is. Zo blijkt dat in steekproef I 72 procent van de eenheden korter (telefonisch) of langer (face-to-face) is geïnterviewd. In steekproef II werkte maar 55 procent mee aan een (telefonisch) interview. Dit verschil van 17 procentpunten wordt voornamelijk veroorzaakt door het grotere aantal mensen dat onbereikbaar of ziek was (ruim 15 procentpunten). Een combinatie van een persoonlijke benadering en een telefonische benadering verkleint immers de kans dat er geen contact wordt gelegd.

Ondanks dit grote verschil zijn er twee redenen om te vermoeden dat het gevonden verschil in het kader van ons onderzoek naar de tweede verwachting niet onoverkomelijk hoeft te zijn. Zo kunnen we vermoeden dat de groep waarmee geen contact kan worden

13. Voor een precieze beschrijving van de data wordt verwezen naar het codeboek van het NKO 1998 (Aarts et al 1998).

Tabel 4. Twee steekproeven.

Groep	Steekproef I	%	N	Steekproef II	%	N
1	Meegedaan aan de eerste golf	49,9	2101	Bereid mee te werken aan een face-to-face interview	27,6	401
2	Uitsluitend geïnterviewd in de non-responsstudie	22,0	926	Niet bereid mee te werken aan face-to-face interview	27,3	396
3	Geen contact	3,4	141	Geen contact	18,6	270
4	Weigering	24,7	1039	Weigering	26,5	384
	Totaal	100	4207	Totaal	100	1451

gelegd erg lijkt op de gehele populatie.¹⁴ Sommigen van deze groep zouden hebben geweigerd, terwijl anderen zouden wel hebben meegedaan. Van grote verschillen met de responsgroep hoeft echter geen sprake te zijn. Belangrijker is dat een lagere respons in de controlegroep waarschijnlijk leidt tot een *onderschatting* van het intervieweffect. In paragraaf 4 hebben we immers gezien dat respons samenhangt met opkomst. De opkomst onder bereidwillige respondenten is hoger dan onder weinig bereidwillige respondenten. Een lagere respons betekent dan tevens dat we een hogere opkomst meten. Dat betekent weer dat de opkomst in steekproef II feitelijk waarschijnlijk (nog) lager ligt dan we meten.

De verwachting dat er intervieweffecten optreden (verwachting 2) lijkt op het eerste gezicht inderdaad te worden bevestigd. In steekproef I heeft (zie nogmaals tabel 3) ongeveer 90 procent van de geïnterviewde respondenten gestemd. Als we rekening houden met de verwachte opkomst onder geïnterviewde non-respondenten zal dat percentage iets lager uitkomen. In steekproef II, de steekproef die niet is benaderd voorafgaande aan de verkiezingen, ligt dat percentage beduidend lager, op 78 procent. Het lijkt er dus op dat een belangrijk deel het verschil tussen gemeten en feitelijke opkomst kan worden toegeschreven aan het intervieweffect. Toch is deze conclusie alleen gerechtvaardigd op basis van een globale vergelijking. Het lijkt zinniger ook te kijken naar de verschillen tussen de verschillende groepen respondenten.

5.2. Multivariate analyse van opkomst in de twee steekproeven

Voor een meer valide schatting van het intervieweffect moeten we rekening houden met verschillen die zijn ontstaan als gevolg van de benadering van de respondenten. Daartoe proberen we de gemeten opkomst in de gecombineerde twee (random) steekproeven te voorspellen met behulp van alle beschikbare variabelen. Dit zijn naast de variabelen sekse en leeftijd¹⁵ uit het bevolkingsregister, allereerst twee contextuele

14. Deze veronderstelling wordt regelmatig gebruikt in de discussie over de consequenties van non-respons. Op die manier wordt ook het belang van de toename van het aantal weigeraars verdedigd. Toch wijkt in sommige onderzoeken de groep waarmee geen contact kan worden gelegd af van de rest van de populatie (Van Zessen en Sandfort 1991: 31)

15. Omdat leeftijd mogelijk geen lineair effect heeft op de odds-ratio is leeftijd teruggebracht tot drie categorieën. Verschillende analyses gaven aan dat er van een significant leeftijdseffect in dit verband geen sprake was. We hebben daarom afgezien van de presentatie van een meer genuanceerde analyse.

variabelen: stedelijkheid en welstand van de buurt. Stedelijkheid wordt door het CBS geconstrueerd op basis van onder meer adressendichtheid. De welstand van een buurt is een variabele afkomstig uit de gegevensverzameling van Geo-marktprofiel en geconstrueerd op basis van geschat inkomen en huizenprijzen (Aarts et al 1998: 65). De overige variabelen zijn afkomstig uit de face-to-face dan wel telefonische interviews. Naast een vraag naar het opkomstgedrag bij de gemeenteraadsverkiezingen van 1998, waren dat vragen naar politieke interesse, politiek cynisme en politiek zelfvertrouwen.¹⁶

Door alle variabelen op te nemen bij de voorspelling van opkomst schakelen we de verschillen tussen de twee steekproeven die we hiervoor hebben waargenomen zoveel mogelijk uit.¹⁷ Vervolgens kunnen we kijken in hoeverre toevoeging van het onderscheid tussen de twee steekproeven (het intervieweffect), iets toevoegt aan de verklaarde variantie. Hiervoor hebben we gebruik gemaakt van een logistische regressie. De resultaten van deze analyse staan in tabel 5.

Model 1 omvat alle variabelen waarover op individueel niveau informatie beschikbaar is en twee omgevingskenmerken (urbanisatiegraad van de gemeente en welstand van de buurt). Niet geheel onverwacht hebben welstand van de buurt, opleiding, interesse en opkomst bij de laatste gemeenteraadsverkiezingen een significant effect op de kans dat men gaat stemmen bij de Tweede Kamerverkiezingen. De kwaliteit van het geschatte model is overigens niet bijzonder goed. De opkomst in de gecombineerde steekproeven bedraagt 86 procent (N = 3571). Als gevolg van het uitvallen van respondenten in de logistische regressie is het opkomstpercentage onder respondenten 1 procentpunt hoger (N = 2941). Tegen die achtergrond valt de kwaliteit van het model nogal tegen. In het volledige model wordt immers slechts in 88 procent van de gevallen correct voorspeld of iemand wel of niet heeft gestemd. De verklaringskracht van het model is redelijk; 15 procent.

Om interpretatie van de resultaten te vergemakkelijken is van een aantal variabelen de mediaan gehercodeerd tot 0. Twee variabelen zijn opgevat als dichotome variabelen (geslacht en wel of niet gestemd bij de gemeenteraadsverkiezingen van 1998). De 'mediane respondent' (die dus op de meeste variabelen 0 scoort) is, als we ons beperken tot de variabelen die een significant effect hebben op de opkomst bij de Tweede Kamerverkiezingen, tamelijk geïnteresseerd in politiek, heeft de middelbare school (MAVO, HAVO of VWO) afgerond, maar daarna geen opleiding meer gevolgd. Deze respondent is woonachtig in een buurt met een gemiddelde welstand. Als de deze respondent niet heeft gestemd tijdens de gemeenteraadsverkiezingen van 1998, is zijn stemkans slechts 76 procent. Als hij echter wel heeft gestemd tijdens de laatste

16. De vraag naar politieke interesse luidt: 'Bent u zeer geïnteresseerd in politieke onderwerpen, tamelijk geïnteresseerd of niet geïnteresseerd?'. De vraag naar politiek cynisme luidt: 'Kamerleden bekommeren zich niet om de mening van mensen zoals ik (dat is zo, dat is niet zo)'. De vraag naar politiek zelfvertrouwen luidt: 'Soms lijkt de politiek zo ingewikkeld dat mensen zoals ik niet echt kunnen begrijpen wat er speelt' (helemaal mee eens, mee eens, mee oneens, helemaal mee oneens).

17. Omdat we van de non-respondenten niet weten of ze ook daadwerkelijk hebben gestemd, is voor deze groep gebruik gemaakt van de stemintentie. Bij de andere groepen uit de steekproef waarvan we zowel intentie als feitelijke opkomst kennen, wijkt deze iets, maar niet veel af van het feitelijke opkomstpercentage.

Tabel 5. Een multi-variate voorspelling van de gemeten opkomst I.

	Model 1		Model 2	
	B	Exp(b)	B	Exp(b)
Stedelijkheid	-0,01	0,99	-0,03	0,97
Welstand buurt	0,15*	1,16	0,14*	1,14
Leeftijd	0,08	1,08	0,06	1,06
Man i.p.v. vrouw	-0,15	0,86	-0,16	0,85
Opleiding	0,08**	1,09	0,09**	1,09
Wel i.p.v. niet gestemd bij de GR verkiezingen	2,02**	7,53	2,14**	8,49
Politieke interesse	0,77**	2,17	0,78**	2,18
Politiek cynisme	-0,20	0,82	-0,17	0,85
Politiek zelfvertrouwen	0,09	1,09	0,14*	1,15
Constante	1,18**		0,19**	
Wel i.p.v. niet geïnterviewd voor de verkiezingen			1,21**	3,36
Cox & Snell R ²	0,15		0,17	
Correct voorspeld	88%		88%	
N=	2941		2941	

gemeenteraadsverkiezingen stijgt de stemkans sterk tot maar liefst 96 procent. Ook opleiding en politieke interesse hebben een significant positief effect op de kans dat iemand gaat stemmen.

Model 1 houdt geen rekening met de manier waarop respondenten zijn benaderd. Wat gebeurt er nu als we het verschil tussen de beide steekproeven invoeren? Zoals gezegd, verwachten we dat respondenten in steekproef I, die voor de verkiezingen benaderd zijn, vaker gaan stemmen dan de respondenten in steekproef II. Ook als we zoveel mogelijk rekening houden met de verschillen tussen de responsgroepen van beide steekproeven blijkt dit inderdaad het geval. Dit is weergegeven in model II. De kans dat mediane respondent (zie boven) die niet heeft gestemd bij de gemeenteraadsverkiezingen en ook niet is geïnterviewd voorafgaande aan de Tweede Kamerverkiezingen, stemt tijdens de tweede Kamerverkiezingen is 'slechts' 55 procent. Als deze respondent zou zijn geïnterviewd voorafgaande aan de verkiezingen stijgt de stemkans tot 80 procent. Het effect van het interview op de verder mediane kiezer die tijdens de gemeenteraadsverkiezingen van 1998 ook al had gestemd is ook substantieel. De stemkans zonder benadering voorafgaande aan de verkiezingen is 91 procent. Het interview doet de stemkans stijgen tot 97 procent. We kunnen dus concluderen dat er inderdaad een intervieweffect optreedt bij het meten van opkomst in enquêtes die voor de verkiezingen worden gehouden. Het effect blijft bestaan als we rekening houden met de (meetbare) verschillen tussen de groepen respondenten in beide steekproeven.

5.3. Verschillen in het intervieweffect bij verschillende groepen

Nu we verwachting 2 kunnen bevestigen heeft het zin na te gaan of ook de meer specifieke verwachting 3 opgaat. We kunnen immers binnen de twee steekproeven een

responsgroep en een (telefonisch wel geïnterviewde) non-responsgroep onderscheiden. We verwachten dat het intervieweffect bij de responsgroep groter is dan bij deze non-responsgroep.¹⁸ Is het zo dat het verschil in opkomst tussen de twee telefonisch wel geïnterviewde non-responsgroepen (de groepen 2) kleiner is dan tussen de twee responsgroepen (de groepen 1)?

Tabel 6. Gemeten opkomst in de verschillende delen van de steekproef.

		Steekproef I	Steekproef II	
		Interview voorafgaande aan de verkiezingen	Geen interview voorafgaande aan de verkiezingen	Verskil
Responsgroep	Groep 1	90,1 (N = 1956)	78,8 (N = 400)	11,3
Geïnterviewde non-respondenten	Groep 2	85,2 (N = 822)	74,3 (N = 393)	10,9
Totaal		88,7 (N = 2778)	76,5 (N = 793)	11,2

Op basis van een eerste inspectie van tabel 6 lijkt verwachting 3 *niet* op te gaan. Er is in beide gevallen een verschil (zoals we hiervoor al hadden gezien) in de verwachte richting, maar de twee verschillen doen nauwelijks voor elkaar onder.

Er zijn echter ten minste twee redenen om de cijfers uit tabel 6 te relativiseren en daarmee verwachting 3 nog niet direct te verwerpen. In de eerste plaats kennen we van de telefonisch geïnterviewde non-responsgroep uit steekproef I het feitelijke opkomstcijfer niet. In de overall analyses in paragraaf 5.1 en 5.2. hebben we daarom gebruik gemaakt van opkomstintentie. Deze ligt hoger omdat er tussen droom en daad nu eenmaal praktische bezwaren in de weg staan. Dit heeft er wellicht voor gezorgd dat het overall intervieweffect wat werd overschat, maar die overschatting zal gering zijn. Deze opkomstcijfers liggen naar alle waarschijnlijkheid lager dan de hiervoor gebruikte opkomstintentie. Dat betekent dat het aldus gemeten opkomstverschil tussen de groepen 2 waarschijnlijk kleiner dan we hier vaststellen. Als de feitelijke opkomst van de telefonisch geïnterviewde non-responsgroep, net als bij de andere groepen, 4 procentpunt lager ligt dan de voorgenomen opkomst, wordt het verschil ook 4 procentpunt kleiner.

In de tweede plaats stuiten we bij de toetsing van verwachting 3 we op het volgende probleem. De groep respondenten uit steekproef II, de controlegroep, die instemt met een face-to-face interview omvat slechts 28 procent van de totale steekproef II. In steekproef I werd maar liefst 50 procent face-to-face geïnterviewd. Dit verschil kan niet worden verklaard door de grote groep waarmee geen contact kon worden gelegd. Een mogelijke verklaring voor dit verschil is wellicht wel gelegen in de manier waarop in steekproef II werd gevraagd of mensen bereid waren mee te doen met het hoofdonderzoek.¹⁹ Om te voorkomen dat mensen weigerden mee te doen aan het non-respons gedeelte van het stimulus onderzoek konden de telefonische interviewers niet te vasthoudend zijn. Dat zou immers leiden tot volledige non-respons. Daarom werd wellicht eerder geconcludeerd dat iemand niet bereid was mee te werken aan het hoofdonder-

18. Zie voetnoot 17.

19. Zie voetnoot 7.

zoek dan door interviewers die alleen werden betaald als ze iemand bereid vonden mee te werken aan een face-to-face interview.

5.4. Een multivariate analyse van de verschillen in het intervieweffect

Om zoveel mogelijk te controleren voor de omstandigheid dat de samenstelling van de verschillende groepen – waarschijnlijk als gevolg van de benadering – verschilt, analyseren we de twee steekproeven analoog aan de manier waarop we dat hebben gedaan in het kader van de globale vergelijking tussen de twee steekproeven (paragraaf 5.2.). We verwachten daarbij dat de relatie tussen de stimulus (steekproef I, dan wel steekproef II) en de opkomst wordt beïnvloed door het wel of niet behoren tot groep 1 dan wel groep 2. Het verschil in opkomst tussen de groepen 1 van de twee steekproeven zal naar verwachting groter zijn dan het verschil in opkomst tussen de groepen 2 van de twee steekproeven als we rekening houden met alle andere variabelen.²⁰ Oftewel, het effect van de stimulus op het opkomstgedrag zal in groep 1 (de respondenten) groter zijn dan in groep 2 (de non-respondenten). Verder houden we rekening met de mogelijkheid dat respondenten en non-respondenten uit de tweede steekproef niet in dezelfde mate zullen stemmen. De resultaten van de analyse staan in tabel 7.

Model 3, waarin rekening is gehouden met het verschil tussen respondenten en (telefonisch wel geïnterviewde) non-respondenten, wijkt nauwelijks af van model 2. We verwachtten dat telefonisch geïnterviewde non-respondenten uit steekproef I weliswaar iets vaker zouden stemmen dan vergelijkbaar weigerachtige respondenten uit steekproef II (het stimuluseffect), maar dat dit verschil minder groot zou zijn dan de verschillen in opkomst tussen de respondenten (de groepen 1) uit de steekproeven I en II. Dit laatste is echter niet het geval. Het effect op de relatie en het effect op het intercept zijn niet significant en leiden ook niet tot verbetering van het model. Daarmee is de tweede reden om de cijfers uit tabel 4 te relativiseren niet houdbaar gebleken. De eerste reden (het feit dat we voor telefonisch geïnterviewde non-respondenten uit steekproef I de opkomstintentie hebben gebruikt, in plaats van gerapporteerde opkomst) blijft echter staan, waardoor we, gevoegd bij de resultaten van de globale analyse, kunnen concluderen dat er wel degelijk sprake lijkt te zijn van een intervieweffect.

5.5. CONCLUSIE

Het vinden van een intervieweffect betekent dat de conclusie die werd getrokken aan het einde van paragraaf 4 moet worden genuanceerd. We stelden in die paragraaf vast dat de telefonisch geïnterviewde non-responsgroep van steekproef I (groep 2) vaker stemt dan de niet-onderzochte non-responsgroep (de groepen 3 en 4). Daarbij was echter expliciet geen rekening gehouden met het intervieweffect. In deze paragraaf hebben

20. Om het toegevoegde effect van het verschil in stimulus te onderzoeken is een nieuwe dichotome variabele geconstrueerd die de waarde 1 aanneemt als de respondent is geïnterviewd voorafgaande aan de verkiezingen en die de waarde 0 aanneemt in de drie andere gevallen. We verwachten dat het effect van deze nieuwe variabele op de relatieve kansverhouding positief is. Verder is, omdat het intercept kan verschillen, de dichotome variabele of men al dan niet behoort tot groep 1 ook opgenomen.

Tabel 7. Een multi-variate voorspelling van de gemeten opkomst II.

	Model 2 (herhaald)		Model 3	
	B	Exp(b)	B	Exp(b)
Stedelijkheid	-0,03	0,97	-0,03	0,97
Welstand buurt	0,14*	1,14	0,14*	1,14
Leeftijd	0,06	1,06	0,06	1,06
Man i.p.v. vrouw	-0,16	0,85	-0,16	0,85
Opleiding	0,09**	1,09	0,09**	1,09
Wel i.p.v. niet gestemd bij de GR verkiezingen	2,14**	8,49	2,14**	8,47
Politieke interesse	0,78**	2,18	0,77**	2,16
Politiek cynisme	-0,17	0,85	-0,15	0,86
Politiek zelfvertrouwen	0,14*	1,15	0,15*	1,16
Constante	0,19**		0,35**	
Wel i.p.v. niet geïnterviewd voor de verkiezingen	1,21**	3,36	0,98**	2,66
Wel i.p.v. niet bereid mee te doen aan een mondeling interview			-0,31	0,73
Wel mondeling geïnterviewd voorafgaande aan de verkiezingen i.p.v. niet (interactie) (zie voetnoot 20)			0,41	1,52
Cox & Snell R ²	0,17		0,17	
Correct voorspeld	88%		88%	
N=	2941		2941	

we laten zien dat dit effect wel degelijk bestaat. Het gevolg van het bestaan van dit effect is onder meer dat de feitelijke opkomst onder de ‘echte’ non-respondenten (de groepen 3 en 4) minder afwijkt van het stemgedrag van de wel telefonisch geïnterviewde non-respondenten (groep 2) dan we hiervoor hebben berekend. Dat betekent weer dat we de gegevens die we hebben verzameld over non-respondenten in een telefonisch onderzoek wellicht toch kunnen gebruiken om iets te zeggen over de totale groep non-respondenten.

6. DISCUSSIE

De vraag die in deze bijdrage centraal stond, was in hoeverre non-respons- en intervieweffecten een bijdrage leveren aan de verklaring van het gevonden verschil tussen gemeten opkomst en feitelijke opkomst. Er lijkt op het eerste gezicht inderdaad een significant en substantieel non-responseffect te bestaan: naarmate iemand minder bereid is mee te doen met alle onderdelen van het NKO, daalt de kans dat hij of zij gaat stemmen. Deze relatie lijkt zich door te zetten onder de groep non-respondenten waarover, buiten leeftijd en geslacht, nauwelijks aanvullende informatie beschikbaar is. Deze groep moet, als we veronderstellen dat de gemeten opkomst juist is en niet is beïnvloed door het onderzoek zelf, beduidend minder hebben gestemd dan de wel onderzochte groep.

Toch moet rekening worden gehouden met het intervieweffect: het feit dat mensen worden geïnterviewd beïnvloedt wel degelijk hun gedrag. Hoewel het uitgevoerde experiment om dat uit te zoeken niet in alle opzichten is geslaagd, blijkt uit de vergelijking van de responsgroep en de stimulusgroep dat het interviewen een positief effect heeft op de opkomst. Dit betekent weer dat het verschil in opkomst tussen in het geheel niet geïnterviewde non-respondenten en telefonisch wel geïnterviewde non-respondenten waarschijnlijk minder groot is dan we in eerste instantie hadden berekend.

Hoewel er sprake is van een intervieweffect, kan dit effect maar een deel van de veranderingen in het verschil tussen gemeten en feitelijke opkomstcijfers verklaren. Het intervieweffect haalt mensen die anders niet zouden stemmen over wel te gaan stemmen, of herinnert mensen aan de verkiezingen. Er zijn echter geen redenen aan te nemen dat dit intervieweffect de afgelopen decennia is veranderd. Dat betekent dat we met behulp van het intervieweffect wel de discrepantie kunnen verklaren, maar niet de toename van het relatieve verschil tussen gemeten en feitelijke opkomst. Dat betekent waarschijnlijk dat (toegenomen) misrapportage en (toegenomen) non-respons een rol spelen bij de afname van de validiteit van de meting van opkomst in het NKO.

In deze bijdrage is gebleken hoe moeilijk het is te meten wat we allemaal al weten. Toch geeft dit soort onderzoek inzicht in de effecten van non-respons en van het interview voorafgaande aan de verkiezingen. Dat inzicht is onder meer mogelijk omdat we naar de bekende weg vragen. We weten immers wat de populatie heeft gedaan en kunnen de uitkomsten van ons onderzoek daaraan toetsen. Hoe anders is het als we proberen iets te zeggen over de relatie tussen non-response-effecten, intervieweffecten en houdingen en gedragingen van burgers die we niet kennen uit andere bronnen. Als we met andere woorden proberen te meten wat we nog *niet* weten.

LITERATUUR

- Aarts, C.W.A.M., 1999. *Opkomst bij Verkiezingen*, Opkomstrapportage in opdracht van het Ministerie van Binnenlandse Zaken en Koninkrijksrelaties, Directie Constitutionele Zaken en Wetgeving, Augustus 1999, Enschede.
- Aarts, C.W.A.M., H. van der Kolk, M. Kamp, 1998. *Dutch Parliamentary Election Study 1998* Enschede: SKON-Steinmetz (Codeboek en dataset op CD-ROM).
- Anker, H., 1995. *Enkele technische Achtergronden van het NKO 1994*. In: J.J.M. van Holsteyn en B. Niemöller (red.), *De Nederlandse Kiezer 1994*. Leiden: DSWO, pp. 247-252.
- Brehm, J., 1993. *The Phantom Respondents*; Opinion Surveys and Political Representation. Ann Arbor: The University of Michigan Press.
- Burden, B.C., 1999. *Measuring Voter Turnout in National Election Studies*, paper prepared for presentation at the 1999 meeting of the American Political Science Association, Harvard University, Cambridge.
- Clausen, A.R., 1969. *Response Validity*: Vote Report. *Public Opinion Quarterly* vol 32 (4), 588-606.
- Goor, H. van, 1996. *Marginaal: lijk in de kast, non-response in het nationaal kiezeronderzoek*. *Sociologische Gids* 43(3), 166-170.
- Kolk, H van der, 2000. *Bijlage: Nationaal Kiezeronderzoek 1971-1998*. In: J. Thomassen, K. Aarts en H. van der Kolk (red.) *Politieke veranderingen in Nederland 1971-1998*; Kiezers en de smalle marges van de politiek. Den Haag: SDU uitgevers. Pp. 219-233.
- Kraut, R.E., McConahay, J.B., 1973. *How Being Interviewed Affects Voting*: An Experiment. *Public Opinion Quarterly* 37(3), 398-406.
- Smeets, I. 1995. Facing another gap: An exploration of the discrepancies between voting turnout in Survey Research and Official Statistics. *Acta Politica* 30(3), 307-334.

- Visscher, G., 1995. *Kiezersonderzoek op een dwaalspoor: de in politiek geïnteresseerde burger als self-fulfilling prophecy*. Den Haag: SDU uitgeverij.
- Yalch, R.F., 1976. *Pre-election Interview effects on Voter Turnout*. *Public Opinion Quarterly* 40(3), 331-336.
- Zessen, G. van en T. Sandfort, 1991. *Onderzoek naar 'seksualiteit in Nederland*. In: G.van Zessen en T. Sandfort (red.) *Seksualiteit in Nederland, seksueel gedrag, risico en preventie van Aids*. Amsterdam: Swets & Zeitlinger. Pp. 11-33.

11. De kennismaking met cultuur: het eerste bezoek aan bioscoop, musea, theater en concert

I. NAGEL

SAMENVATTING

In dit artikel staat het eerste bezoek aan vijf cultuurvormen centraal. Retrospectieve data over culturele loopbanen laten zien dat *cultuurvormen verschillen* in het moment waarop ze voor het eerst bezocht worden: de meeste mensen maken het eerst kennis met de bioscoop, en gaan als laatste naar een klassiek concert; het eerste bezoek aan cultuur-historische musea, musea voor beeldende kunst en theater vindt daartussenin in willekeurige volgorde plaats. Sterke *verschillen tussen individuen* bestaan er in de leeftijd waarop kennisgemaakt wordt met cultuur. Een cultureel actief ouderlijk milieu en een hoger niveau van voortgezet onderwijs verhogen de kans om kennis te maken met alle vijf kunstdisciplines. De invloed van de ouders is het grootst in het begin van de levensloop maar strekt zich uit tot en met 18-jarige leeftijd. Het effect van schoolniveau begint op de leeftijd waarop voortgezet onderwijs wordt gevolgd en blijft ook op latere leeftijd bestaan. De effecten van kunsteducatie beperken zich voornamelijk tot de leeftijd waarop het kunstonderwijs genoten werd en tot dezelfde kunstdiscipline.

1. INTRODUCTIE

De meeste mensen bezoeken in hun leven ooit wel eens een museum, theater, bioscoop of een klassiek concert. Er is slechts een kleine groep cultuurliefhebbers die deze activiteiten regelmatig onderneemt. Uit eerder onderzoek is bekend wat deze groep onderscheidt van de niet (zo frequente) deelnemers: cultuurparticipanten hebben meestal een hogere opleiding en komen uit gezinnen waarin de ouders ook al cultureel actief waren (Ganzeboom 1989; Knulst 1995; De Haan 1997; De Haan & Knulst 2000). Dat ouderlijk milieu en opleiding zo bepalend zijn voor de participatie in culturele activiteiten op latere leeftijd, wijst erop dat de ongelijkheden in cultuurdeelname al vroeg in het leven hun oorsprong vinden.

Hoe komt de kennismaking met cultuur nu tot stand? Er is in de onderzoeksliteratuur nog niet veel bekend over de eerste schreden in de wereld van kunst en cultuur. Hoewel we weten dat ouders en school een doorslaggevende rol spelen, zijn er over het begin van de culturele loopbaan nog wel wat vragen te beantwoorden die van belang zijn voor het inzicht in ongelijkheden in cultuurdeelname.

Allereerst valt te denken aan *verschillen tussen cultuurvormen in de leeftijd van het eerste bezoek*. Vindt de kennismaking met sommige cultuurvormen in het algemeen vroeg en met andere laat plaats? Aangezien de beginleeftijd volgens sommige auteurs van cruciaal belang is voor de latere participatie in culturele activiteiten (De Jager 1967; Kamphorst & Spruijt 1983), zijn verschillen in de publieksomvang van cultuurvormen mogelijk toe te schrijven aan de leeftijd waarop de meeste mensen in contact komen met die cultuurvormen.

Verschillen in de leeftijd van kennismaking met cultuur zouden bovendien gevolgen kunnen hebben voor ongelijkheden *tussen individuen*. Lang niet alle ouders en scholen onderschrijven het belang van culturele socialisatie in dezelfde mate of achten zichzelf de aangewezen instantie om deze taak tot uitvoering te brengen. Hierdoor maakt niet iedereen even vroeg kennis met cultuur. De leeftijd van kennismaking is daarom een mogelijke oorzaak van verschillen in cultuurparticipatie op latere leeftijd. Eerder gevonden effecten van ouderlijk milieu en opleiding zijn mogelijk het gevolg van het feit dat degenen met cultureel actieve ouders en hoger opgeleiden op jongere leeftijd met cultuur in aanraking komen (De Jager 1967).

In dit artikel onderzoek ik wanneer de eerste kennismaking plaatsvindt met vijf cultuurvormen – bioscoop, cultuur-historisch museum, museum voor beeldende kunst, theater, klassiek concert – en welke rol ouders en school daarbij spelen. De volgende vragen worden beantwoord:

1. Met welke cultuurvormen vindt de kennismaking vroeg en met welke laat plaats?
2. Wie maken vroeg kennis met de verschillende cultuurvormen?

In het vervolg gaan we eerst wat dieper in op de onderzoeksvragen. Na bespreking van de data beschrijven we de resultaten. Het artikel wordt afgesloten met een samenvatting en discussie.

2. HYPOTHESEN

2.1. *De kennismaking met cultuur: verschillen tussen cultuurvormen*

In de samenleving als geheel geldt dat de deelname aan traditioneel hogere cultuur als waardevol gezien wordt. Ouders en school worden als belangrijkste socialiserende instanties verondersteld de normen ten aanzien van (hogere) cultuur en de benodigde culturele kennis tijdens processen van socialisatie over te dragen op de jongere generatie. Een manier om kinderen de waarde van kunst en cultuur bij te brengen is om ze daadwerkelijk mee te nemen naar museum, bioscoop, theater of concertzaal.

Waar zullen ouders en school kinderen het eerst mee naar toe nemen? Aannemelijk is dat ouders en leraren beginnen met de in hun ogen meest toegankelijke cultuurvormen en gaandeweg de minder toegankelijke cultuurvormen toevoegen. Waarschijnlijk laten ze hierbij allerlei overwegingen meespelen, zoals hun inschatting van de benodigde cognitieve vaardigheden en vereiste voorkennis; de bij de instelling gebruikelijke omgangsvormen; de concentratie die opgebracht moet worden; en het tijdstip waarop het bezoek gebracht moet worden. Welke overwegingen ook doorslaggevend zijn, de

vraag is of de socialiserende instanties het eens zijn over een algemene volgorde van toegankelijkheid van de diverse cultuurvormen. Wanneer dat het geval zou zijn, zou de kennismaking met verschillende kunstvormen plaatsvinden in een bepaalde *volgorde* waarbij kinderen het eerst in contact komen met de meest toegankelijke cultuurvormen en gaandeweg geïntroduceerd worden in de minder toegankelijke cultuurvormen. Het bestaan van zo'n vaste volgorde zou het beschrijven van iemands loopbaan sterk vereenvoudigen. Het zal hier onderzocht worden met het cumulatieve schaalmodel van Mokken (Mokken & Lewis 1982).

2.2. De kennismaking met cultuur: verschillen tussen personen

Sommigen zullen vroeg kennismaken met cultuur en anderen laat. Voor nog anderen zullen sommige cultuurvormen zelfs helemaal buiten beschouwing blijven. De vraag is: wie beginnen er vroeg en wie beginnen er laat?

Naarmate ouders en scholen zelf cultureel actiever zijn, is het te verwachten dat zij des te meer belang hechten aan kunst en cultuuroverdracht, en hun kinderen daadwerkelijk vroeger meenemen naar culturele instellingen. De leeftijd waarop ouders en school hun invloed uitoefenen verschilt echter. De verwachting is dat vanaf het moment dat kinderen in de puberteit komen ze zich gaan losmaken van hun ouders en deze hun zeggenschap over hun kinderen langzamerhand verliezen. Ouders zullen daarom de grootste invloed uitoefenen tot hun kinderen een jaar of 12 zijn, daarna zal deze invloed afnemen.

De zeggenschap van scholen bestrijkt een langere periode. Vanaf de kleuterleeftijd tot in ieder geval de leerplichtige leeftijd kunnen scholen leerlingen cultureel gedrag in zekere mate opleggen, door het bijvoorbeeld te koppelen aan een vak, of het als (onderdeel van) een excursie op te nemen. Hoewel alle scholen het belang van kennismaking met hogere cultuur door hun leerlingen wel zullen onderschrijven, is het niet gezegd dat het daadwerkelijk bezoeken van culturele instellingen overal op het programma staat.

De voornaamste invloed van school komt naar verwachting pas naar voren in het voortgezet onderwijs. In het voortgezet onderwijs bestaan namelijk aanzienlijke verschillen in de mogelijkheden die leerlingen krijgen om in contact te komen met cultuuruitingen. Dit hangt in de eerste plaats samen met het niveau van onderwijs. Op havo- en vwo-niveau worden meer culturele activiteiten georganiseerd dan op lagere vormen van voortgezet onderwijs (Van Hoorn e.a. 1997). Ook hangt het cultuurbezoek samen met de kunstvakken die de school aanbiedt, waarin leerlingen onder begeleiding van docenten culturele instellingen bezoeken. De verwachting is dat school – zoals afgemeten aan het niveau en de specifieke kunstvakken in het voortgezet onderwijs – de grootste invloed op het eerste cultuurbezoek zal uitoefenen tussen 12 en 18-jarige leeftijd.

Vanaf een jaar of twaalf is het ook goed mogelijk dat kinderen zelfstandig een bezoek aan een culturele instelling afleggen. Ouders en school zijn als begeleiders niet meer nodig. Integendeel, jongeren gaan misschien wel liever met leeftijdgenoten. Het is daarmee niet gezegd dat de invloed van ouders en school op de kennismaking met cultuur ook afneemt. Behalve dat school cultuurbezoek verplicht kan stellen, stimuleren

ouders en school de cultuurparticipatie van de jeugd ook indirect, door de interesse van de jongeren te wekken voor kunst en cultuur of door hen aan te moedigen alleen of met vrienden te gaan. De school is in dat opzicht ook een ontmoetingsplaats waar in cultuur geïnteresseerde jongeren elkaar kunnen vinden. Wanneer jongeren inderdaad op eigen initiatief – alleen of met vrienden – cultuur gaan bezoeken, zou gesproken kunnen worden van een al geslaagde socialisatie.

3. DATA EN OPERATIONALISATIE

3.1. Data

De hypothesen worden getoetst door analyse van longitudinale data over in oorsprong 1028 respondenten, in leeftijd variërend van 25 tot 38 jaar, die in 1994 retrospectief ondervraagd zijn over hun cultuurparticipatie in het project 'Effecten van Kunsteducatie in het Voortgezet Onderwijs' (Nagel e.a. 1996a), kortweg KUNST94. Het startpunt voor deze dataverzameling vormden 31 scholen voor voortgezet onderwijs, waar informatie verzameld is over oud-leerlingen die daar in de periode van 1975 tot 1985 eind-examen hadden gedaan. Een steekproef van deze oud-leerlingen is in 1994 opgespoord en telefonisch ondervraagd over hun cultuurparticipatie en over hun opleiding, beroep, woonplaats, partner en kinderen. De informatie over receptieve cultuurparticipatie betreft het bezoek aan vijf cultuurvormen: museum voor beeldende kunst, cultuur-historisch museum, theater, bioscoop en uitvoering van klassieke muziek. Naast de oud-leerling zelf is in ongeveer 80 procent van de gevallen één van de ouders ondervraagd. De gevraagde informatie betrof naast achtergrondkenmerken de cultuurparticipatie van beide ouders in de tijd dat de respondent nog thuis woonde. Na selectie op geldige waarnemingen op een aantal variabelen, resteerden 1005 oud-leerlingen en 824 ouders.

De data zijn aanvankelijk verzameld voor onderzoek naar de effecten op de latere cultuurparticipatie van eindexamens in kunstzinnige vakken in het voortgezet onderwijs. De 31 scholen zijn voor het onderzoek geselecteerd omdat zij in de periode tussen 1975 en 1985 één of meer eindexamens in de kunstzinnige vakken aanboden. Bij het trekken van de steekproef van oud-leerlingen van deze scholen zijn twee groepen onderscheiden: oud-leerlingen die examen hadden gedaan in (ten minste) één van vier kunstzinnige vakken (tekenen, handenarbeid, textiele werkvormen, muziek); en oud-leerlingen die wel in dezelfde periode examen hadden gedaan op dezelfde school, maar niet in één van deze vier vakken. Op iedere school is uit elk van de twee groepen een random steekproef getrokken van 26 oud-leerlingen, resulterend in een steekproef van 52 oud-leerlingen per school (dit voor zover er genoeg oud-leerlingen waren; bovendien is op sommige scholen de steekproef wat opgehoogd). De resultaten van dit onderzoek zijn gepubliceerd in Nagel e.a. (1996a, 1997).

3.2. Operationalisatie

3.2.1. Afhankelijke variabele

De te verklaren variabele is de *leeftijd van kennismaking* met vijf cultuurvormen: museum voor beeldende kunst, cultuurhistorisch museum, theater, bioscoop of filmhuis, uitvoering van klassieke muziek. Wanneer respondenten op een eerdere vraag hadden

Tabel 1. De leeftijd van kennismaking met vijf cultuurvormen: percentages eerste bezoek na hot deck imputatie van missende waarden (N = 1005).

	<i>bioscoop</i>	<i>cultuur- historisch museum</i>	<i>museum voor beeldende kunst</i>	<i>theater</i>	<i>klassiek concert</i>
0 – 5 jaar	9	3	4	2	1
6 – 11 jaar	53	27	22	16	8
12 – 18 jaar	35	33	40	44	21
19 jaar en ouder	3	18	12	24	21
geen kennismaking	1	19	22	15	49

Bron: KUNST94 (Nagel e.a. 1997).

aangegeven dat zij een cultuurvorm ooit hadden bezocht, werd voor elke cultuurvorm de vraag gesteld in welk jaar ze deze voor het eerst bezochten (zie tabel 1).

Van respondenten die – tot op het moment van het interview — een cultuurvorm nooit bezocht hadden, ontbreekt de waarde op deze variabele. Om deze ‘gecensureerde’ informatie toch te gebruiken, zijn de analyses uitgevoerd in een discrete gebeurtenissenanalyse (Allison 1984), via een ‘person-period’ bestand, waarbij we de beginleeftijden transformeren in nul-één variabelen: 0 in elk jaar tot de leeftijd waarop de introductie in de betreffende cultuurvorm heeft plaatsgevonden; één op de leeftijd waarop de respondent een cultuurvorm voor het eerst heeft bezocht. Respondenten die een cultuurvorm nooit bezocht hebben krijgen in elk jaar van hun leven een nul toegekend. Op deze manier wordt van respondenten zonder beginleeftijd alle beschikbare informatie gebruikt, namelijk dat hun eerste bezoek in ieder geval niet heeft plaatsgevonden tot op het moment van het interview.

In paragraaf 4.2 wordt een logit-model geschat:

$\log(P(t)/1 - P(t)) = a + b_1x_1 + b_2x_2 + e$, waarbij $P(t)$ de kans is dat een individu kennismaakt met cultuur op tijdstip t , gegeven dat deze kennismaking nog niet heeft plaatsgevonden. Zodra de kennismaking heeft plaatsgevonden is het individu niet meer ‘at risk’ en wordt hij vanaf dat moment buiten de analyses gelaten. De coëfficiënten b_1 en b_2 geven de verandering aan in de logit (log-odds) voor een eenheid verschil in x_1 en x_2 .

3.2.2. Onafhankelijke variabelen

Informatie over de *socialisatie in het ouderlijk milieu* is afkomstig van de ouders zelf. Eén van beide ouders is ondervraagd over zijn of haar eigen participatie in dezelfde vijf cultuurvormen, en over die van de andere ouder in de periode dat de respondent nog bij zijn ouders woonde. Samen met het luisteren naar klassieke muziek thuis, kunstboeken en platen-/cd-bezit vormen deze variabelen de totale cultuurparticipatie van de ouders (alfa = .84) (zie tabel 2).

Doordat niet alle ouders geïnterviewd zijn, is deze variabele beschikbaar voor 824 respondenten. Analyses waarbij de invloed van het ouderlijk wordt bestudeerd hebben alleen betrekking op de 824 respondenten van wie de ouders geïnterviewd zijn. Alleen

Tabel 2. Achtergrondkenmerken: cultuurparticipatie ouders, niveau voortgezet onderwijs, kunsteducatie voortgezet onderwijs, sekse, geboortejaar (in percentages).

ouders van wie ten minste één van beide participeert (N = 824)	
bioscoop	58
cultuur-historisch museum	62
museum voor beeldende kunst	50
theater	67
klassiek concert	40
luisteren naar klassieke muziek	79
kunstboekenbezit ouders	59
platen/cd-bezit ouders	69
niveau voortgezet onderwijs (N = 1005)	
vbo	6
mavo	26
havo	59
vwo	10
examen in kunstzinnig vak (N = 1005)	
beeldend vak	46
muziek	7
vrouw (N = 1005)	64
geboortejaar (N = 1005)	
1956 – 59	5
1960 – 64	48
1965 – 69	48

Bron: KUNST94 (Nagel e.a. 1997).

wat betreft het eerste concertbezoek zijn deze respondenten enigszins selectief: het percentage dat ooit een klassiek concert had bezocht was hoger onder de groep waarvan de ouders geïnterviewd was, terwijl de leeftijd waarop dit voor het eerst plaatsvond lager uitviel.

Informatie over het *niveau van voortgezet onderwijs en kunsteducatie in het voortgezet onderwijs* (tabel 2) is verzameld in de archieven van de 31 deelnemende scholen voor voortgezet onderwijs. De respondenten deden eindexamen op vier niveaus: vbo, mavo, havo, vwo. De variabelen ‘beeldend vak’ en ‘muziek’ geven aan of en op welk schoolniveau de respondent eindexamen heeft gedaan in tekenen, handenarbeid, of textiele werkvormen, respectievelijk het vak muziek. Door de gestratificeerde steekproeftrekking heeft ongeveer de helft van de respondenten examen gedaan in een kunstvak en in de periode daaraan voorafgaand het bijbehorende kunstonderwijs gevolgd. Het percentage *vrouwen* in de steekproef is 64.

Leeftijd is gemeten in jaren en is de enige variabele die over de tijd varieert: het is de observatieperiode in de person-period file. Aangezien de jongste respondenten 25 jaar oud waren op het moment van interview, neemt het aantal waarnemingen na deze leeftijd af. De oudste respondent was 38 toen het interview afgenomen werd. In tabel 2 is het geboortejaar van de respondenten vermeld.

De verschillende eenheden waarin de onafhankelijke variabelen gemeten zijn maken het moeilijk om hun effecten te vergelijken. Om toch vergelijking mogelijk te maken zijn de onafhankelijke variabelen, voor zover dat niet het geval was, gehercodeerd in een 0..1 bereik. De effecten kunnen geïnterpreteerd worden als een vergelijking tussen respondenten met de laagste en de hoogste score op een variabele.

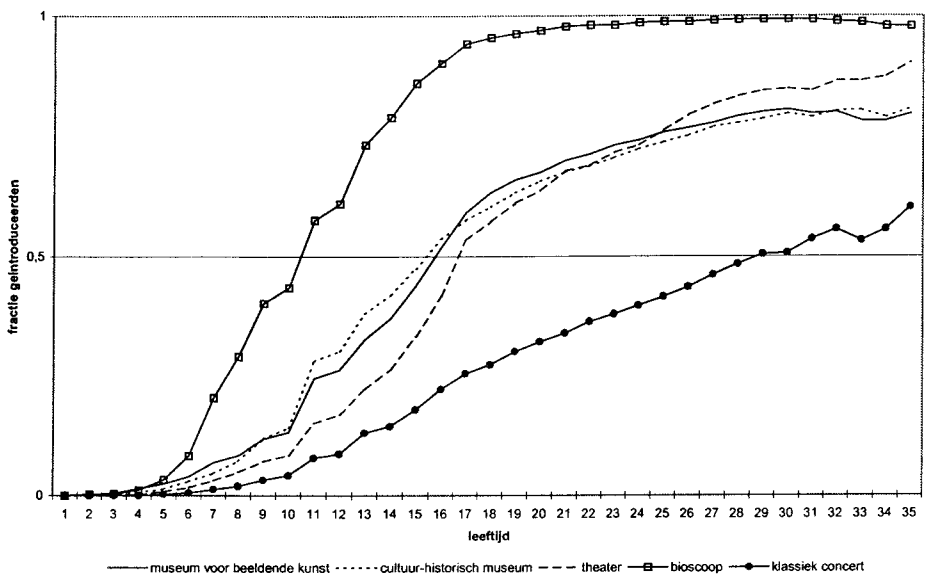
4. RESULTATEN

In deze paragraaf bespreken we de resultaten. De eerste onderzoeksvraag over verschillen tussen cultuurvormen worden in paragraaf 4.1 beantwoord. De tweede onderzoeksvraag gaat over verschillen tussen individuen en is aan de orde in paragraaf 4.2.

4.1. De kennismaking met cultuur: verschillen tussen cultuurvormen

Wanneer en in welke volgorde vindt het eerste bezoek in vijf cultuurvormen plaats? Figuur 1 plot voor elk van de vijf cultuurvormen het deel van de respondenten dat op elke leeftijd ooit heeft deelgenomen. Vanaf twee à drie jaar valt in alle vijf cultuurvormen een lichte groei te bespeuren. Bij oplopende leeftijd zien we de fractie participanten sterker stijgen, en zien we de curven voor de vijf cultuurvormen uit elkaar gaan.

Welke verschillen tussen cultuurvormen laat figuur 1 zien? De links-rechts positie van de curven zegt iets over het aantal respondenten dat de kennismaking met een cultuurvorm op een bepaalde leeftijd achter de rug heeft. Vergelijken we de cultuurvormen in



Bron: KUNST94 (Nagel e.a. 1997)

Figuur 1. Kennismaking met vijf cultuurvormen: fracties respondenten die de cultuurvorm ooit bezocht hebben naar leeftijd (N = 1005).

de leeftijd waarop 50 procent kennis gemaakt heeft, dan zien we dat de curve van bioscoopbezoek de 50-procent lijn het meest links snijdt. Op een leeftijd van 10 jaar is 50 procent van de respondenten wel eens naar de bioscoop geweest. De mediane leeftijden voor het eerste bezoek aan een museum voor beeldende kunst, een cultuur-historisch museum, en theater, verschillen niet veel: 14 en 15 jaar. De leeftijd waarop 50 procent ooit een uitvoering van klassieke muziek heeft bezocht ligt een stuk hoger, op ongeveer 29 jaar. Dit levert een eerste antwoord op de vraag bij welke cultuurvormen de kennis-making relatief vroeg en bij welke relatief laat plaatsvindt. De volgorde van de mediane leeftijden van de vijf cultuurvormen is: bioscoop, cultuur-historisch museum, museum voor beeldende kunst, theater, klassiek concert.

De helling van de curven zegt iets over de groei van het aantal nieuwe participanten op elke leeftijd. Bioscoopbezoek kent tot 15 jaar de sterkste aanwas. Deze begint al op ongeveer 4-jarige leeftijd, en gaat door tot 18 jaar. Op 25-jarige leeftijd is vrijwel iedereen wel eens naar de bioscoop geweest. In de andere cultuurvormen is de groei minder sterk en komt ook wat later op gang. Hierbij lijken museum voor beeldende kunst, cultuur-historisch museum en theater ongeveer dezelfde groei te kennen. De groei in het aantal nieuwe deelnemers bij uitvoeringen van klassieke muziek ligt een stuk lager, maar is na 20-jarige leeftijd wel het hoogst. Bioscoop wordt gevolgd door beide musea en theater, en – op grotere afstand – de uitvoering van klassieke muziek. Op 25-jarige leeftijd heeft nog niet de helft van de respondenten ooit een uitvoering van klassieke muziek bijgewoond. Op grond van het percentage dat op 25-jarige leeftijd ooit een bezoek heeft gebracht, is de volgorde van de cultuurvormen: bioscoop, theater, cultuur-historisch museum, museum voor beeldende kunst, klassiek concert. Dit is een andere volgorde dan die van de mediane leeftijden.

4.1.1. De 'toegankelijkheid' van cultuurvormen op verschillende leeftijden

Om de vraag te beantwoorden of de cultuurvormen geordend kunnen worden op een onderliggend eendimensionaal kenmerk als de 'toegankelijkheid' ervan, gaan we na in hoeverre de vijf items die op elke leeftijd aangeven of een cultuurvorm op dat moment ooit is bezocht een cumulatieve 'Mokken-schaal' vormen (Mokken & Lewis 1982). Het aantal cultuurvormen dat op een bepaalde leeftijd bezocht is zou daarmee een maat zijn voor een onderliggend kenmerk als 'culturele ontwikkeling' van individuen, dat parallel loopt aan de toegankelijkheid van de verschillende cultuurvormen. Daartoe onderzoeken we of deze items (die aangeven of het eerste bezoek heeft plaatsgevonden) de respondenten op dezelfde manier ordenen op grond van hun 'culturele ontwikkeling' (eerste monotonie-eis). Voor elke cultuurvorm moet gelden dat degenen met een grotere 'culturele ontwikkeling' een grotere kans hebben op kennismaking met deze cultuurvorm. De tweede eis, die van 'dubbele monotonie', houdt in dat de volgorde van de items voor iedere respondent hetzelfde is.

Om na te gaan of vijf items op iedere leeftijd een Mokken-schaal vormen, berekenen we op iedere leeftijd Loevinger's coëfficiënt voor homogeniteit H (Mokken & Lewis 1982), die de schaalbaarheid van de items weergeeft. In formule: $H = 1 - E(\text{obs})/E(\text{exp})$, waarin $E(\text{obs})$ en $E(\text{exp})$ het geobserveerde aantal fouten en het verwachte aantal fouten onder onafhankelijkheid weergeven, waarbij een fout een afwijking van de veronderstelde volgorde betekent. De veronderstelde volgorde is in dit geval de volgorde van de mediane leeftijden: bioscoop, cultuur-historisch museum,

museum voor beeldende kunst, theater, klassiek concert. Uit de berekeningen blijkt dat de items van leeftijd 6 tot 22 jaar een redelijke schaal vormen ($H > .40$); op leeftijden hoger dan 22 is de schaal zwak (maar H is wel $> .34$); op leeftijden lager dan 6 jaar is er sprake van een slechte schaal ($H < .30$)(Mokken & Lewis 1982). Het *aantal* bezochte cultuurvormen is dus op elke leeftijd vanaf 6 jaar een maat voor een onderliggend eendimensionaal kenmerk dat wij benoemd hebben als de ‘culturele ontwikkeling’ van een persoon.

4.1.2. Volgorde van introductie over de levensloop

Geldt de volgorde van toegankelijkheid voor alle personen en op alle leeftijden? De eis van dubbele monotonie wordt niet getoetst met Loevingers H . Echter, in figuur 1 zagen we al dat de volgorde van toegankelijkheid van museum voor beeldende kunst, cultuur-historisch museum en theater van leeftijd lijkt af te hangen. De veronderstelde volgorde geldt dus niet voor alle leeftijden.

Dit roept de vraag in welke volgorde de individuele respondenten nu werkelijk hebben kennisgemaakt met de vijf cultuurvormen. Hoe vaak hebben zij het eerste bezoek afgelegd in de volgorde: bioscoop, cultuur-historisch museum, museum voor beeldende kunst, theater, klassiek concert? In totaal kwamen onder de 1005 respondenten 83 verschillende volgordes naar voren. Hierbij zijn overigens nog niet bezochte cultuurvormen gecodeerd alsof ze voldoen aan de veronderstelde volgorde (van de mediane leeftijden). In tabel 3 zijn de meest voorkomende patronen weergegeven. De volgorde van kennismaking verloopt bij 206 respondenten helemaal volgens het verwachte patroon. Dit patroon is ook het meest voorkomende patroon, al neemt het minder dan de helft van de gevallen voor zijn rekening. De zes meest voorkomende volgordes plaatsen wel steeds bioscoop als eerste en concert als laatste. De afwijkingen zitten dus vooral in de volgorde van cultuur-historisch museum, museum voor beeldende kunst en theater. Ook hier stuiten we dus op de slechte schaalbaarheid van de beide soorten musea en theater. Wanneer we deze drie cultuurvormen samen nemen en dus tellen hoe vaak

Tabel 3. Meest voorkomende volgordes van eerste bezoek aan vijf cultuurvormen over de levensloop.

volgorde van bezoek aan vijf cultuurvormen					N	%
bioscoop	c.h. museum	kunstmuseum	theater	concert	206	21
bioscoop	kunstmuseum	theater	c.h. museum	concert	100	10
bioscoop	kunstmuseum	c.h. museum	theater	concert	87	9
bioscoop	c.h.museum	theater	kunstmuseum	concert	63	7
bioscoop	theater	c.h. museum	kunstmuseum	concert	59	6
bioscoop	theater	kunstmuseum	c.h. museum	concert	38	4
bioscoop	concert	c.h. museum	kunstmuseum	theater	38	4
c.h. museum	bioscoop	kunstmuseum	theater	concert	33	3
c.h.museum	kunstmuseum	bioscoop	theater	concert	26	3
bioscoop	c.h. museum	kunstmuseum	concert	theater	21	2
overige patronen					336	33
					1005	100

Bron: KUNST94 (Nagel e.a. 1997)

Tabel 4. Het moment van het eerste bezoek aan vijf cultuurvormen (coëfficiënten logit-analyse).

	<i>bioscoop</i>	<i>cultuur- historisch museum</i>	<i>museum voor beeldende kunst</i>	<i>theater</i>	<i>klassiek concert</i>
leeftijd					
leeftijd	0,92 ***	0,92 ***	0,59 ***	0,80 ***	0,86 ***
leeftijd 2	-0,04 ***	-0,05 ***	-0,02 ***	-0,03 ***	-0,04 ***
leeftijd 3	0,00 ***	0,00 ***		0,00 ***	0,00 ***
seks					
meisje	-0,01	-0,10	-0,05	0,19 *	0,03
cultuurparticipatie ouders	1,14 ***	1,38 ***	1,94 ***	1,86 ***	2,40 ***
0 – 6 jaar	1,37 ***	3,12 ***	3,86 ***	4,60 ***	5,46 ***
6 – 12 jaar	1,41 ***	1,35 ***	2,88 ***	2,77 ***	3,37 ***
12 – 18 jaar	0,82 *	1,34 ***	1,47 ***	1,81 ***	2,71 ***
19 – 25 jaar	0,63	1,18 **	0,41	0,62	1,51 ***
25 jaar en ouder	-2,06	0,68	-0,24	1,13	1,66 **
beeldend vak	-0,09	0,19	1,04 ***	0,15	0,44 **
0 – 6 jaar	-0,23	0,11	0,65	1,05	1,28
6 – 12 jaar	-0,12	0,22	0,53 *	0,38	0,72 *
12 – 18 jaar	0,11	0,22	1,51 ***	0,12	0,37
19 – 25 jaar	0,14	0,05	0,66	0,28	0,38
25 jaar en ouder	0,59	-0,00	0,86	-0,69	0,38
vak muziek	-0,23	-0,09	0,13	0,19	1,08 ***
0 – 6 jaar	-1,00	0,59	0,16	0,74	1,38
6 – 12 jaar	-0,56	-0,50	-0,62	-0,04	0,17
12 – 18 jaar	0,60	0,03	0,38	0,43	1,82 ***
19 – 25 jaar		-0,20	0,65	0,17	0,87
25 jaar en ouder		0,48	0,49	-0,31	-0,22
schoolniveau	0,09	0,50 *	0,55 **	0,55 **	1,15 ***
0 – 6 jaar	-1,24 **	-0,37	-0,09	-3,54 *	-2,93
6 – 12 jaar	-0,22	0,46	0,37	-0,16	0,21
12 – 18 jaar	1,04 ***	0,36	0,64 *	0,82 ***	1,05 **
19 – 25 jaar	1,61 **	0,71	0,72 *	0,49	1,87 ***
25 jaar en ouder	0,74	1,82 **	1,05	1,37 *	1,95 **
constante	-6,03 ***	-9,08 ***	-8,54 ***	-8,28 ***	-10,35 ***
N personen * jaren	9340	14908	14921	15607	19748
N personen	823	824	824	824	824

Bron: KUNST94 (Nagel e.a. 1997)

*** p < ,001; ** p < ,01; * p < ,05

bioscoop als eerste bezocht is en klassiek concert als laatste, dan is het percentage dat aan de verwachte volgorde voldoet een stuk hoger, namelijk 55.

De conclusie luidt dat hoewel de beide soorten musea en theater niet een duidelijke volgorde van toegankelijkheid kennen, bioscoop relatief vroeg in de levensloop voor het eerst bezocht wordt, en klassiek concert relatief laat. Ook kunnen we concluderen dat het eerste bezoek aan musea en theater over het algemeen daartussen in plaatsvindt. Over hun onderlinge volgorde kunnen we echter geen uitspraken doen.

4.2. De introductie in cultuur: verschillen tussen personen

Voor zover de kennismaking met cultuur heeft plaatsgevonden, lijkt het erop dat de meesten het eerst in de bioscoop komen, vervolgens in willekeurige volgorde musea en theater bezoeken en als laatste een klassiek concert bijwonen. Differentiatie treedt naar verwachting wèl op in de *snelheid* waarmee de kennismaking met cultuurvormen zich voltrekt. Wie maken vroeg en wie maken laat kennis met cultuur? De verwachting is dat een cultureel actief ouderlijk milieu, een hoog schoolniveau en kunsteducatie in het voortgezet onderwijs het eerste bezoek aan de vijf cultuurvormen zullen bespoedigen. De ouders zijn naar verwachting vooral in de eerste 12 jaar van invloed, terwijl de school deze taak vanaf 12-jarige leeftijd overneemt en in de periode van 12- tot 18-jarige leeftijd het belangrijkste is.

Tabel 4 laat de resultaten van een logit-analyse zien, gecorrigeerd voor clustering van jaren binnen personen via STATA's optie voor robuuste standaardfouten (Huber 1967). De afhankelijke variabele is de log-odds van de kans op het eerste bezoek in elk jaar, gegeven dat de cultuurvorm nog niet eerder bezocht is. Naast leeftijd en sekse vermeldt de tabel totale effecten van ouders- en schoolkenmerken (vetgedrukt) en daaronder de effecten op verschillende leeftijden. De coëfficiënten geven het verschil in log-odds van de kans op een eerste bezoek tussen degenen met de hoogste en de laagste waarde op een onafhankelijke variabele. De coëfficiënt .19 van de variabele sekse geeft bijvoorbeeld aan dat de log-odds van de kans op een eerste bezoek voor vrouwen in elk jaar .19 groter is dan voor mannen. Dat betekent dat vrouwen op elk moment een $100(\exp(.19)-1) = 21$ procent verhoogde kans hebben om kennis te maken met het theater, gegeven dat ze daar nog nooit geweest zijn. Overigens is er alleen sprake van een sekse-effect bij theater. De kans op kennismaking met de overige cultuurvormen is voor mannen en vrouwen gelijk.

4.2.1. Leeftijd

Leeftijd wordt naast een lineaire term gemodelleerd via een tweede- en een derdegraadspolynoom, behalve bij museum voor beeldende kunst waar alleen de kwadratische functie nog significant van nul verschilde. De leeftijdseffecten geven aan dat de kans om een eerste cultuurbezoek af te leggen stijgt met leeftijd en dat deze het grootst is tussen de 14 en de 18 jaar, afhankelijk van de cultuurvorm. Na deze leeftijd neemt de kans op kennismaking niet meer toe en daalt zelfs licht om later opnieuw te stijgen. Dit geldt niet voor musea voor beeldende kunst, waar de kans op een eerste bezoek niet meer toeneemt.

4.2.2. Ouders en school

Welke kenmerken bepalen wie er vroeg en wie laat kennismaken met cultuur? De positieve effecten van *culturele activiteiten van de ouders* bij alle cultuurvormen duiden erop dat cultureel actieve ouders de kans op kennismaking met alle cultuurvormen vergroten. Naast participatie in dezelfde cultuurvorm verhoogt ook participatie in andere cultuurvormen de kans om een cultuurvorm voor het eerst te bezoeken (niet in tabel). Ouders hebben dus niet alleen invloed doordat ze hun kinderen meenemen naar cultuur. Ook een hoger *niveau van voortgezet onderwijs* vergroot de kans om met cultuur in aanraking te komen, met uitzondering van de bioscoop. De effecten van ouders zijn in alle gevallen groter dan die van het schooltype van het voortgezet onderwijs en het volgen van een kunstzinnig vak. Een cultureel actief ouderlijk milieu vervroegt het eerste bezoek dus sterker dan schoolkenmerken. De effecten van een *beeldend vak* en het *vak muziek* in het eindexamen beperken zich – zoals te verwachten was – tot musea voor beeldende kunst en uitvoeringen van klassieke muziek. Opvallend is wel dat een beeldend vak ook de kans om een klassiek concert te bezoeken vergroot. Dit wijst op eerder gevonden selectie-effecten: leerlingen die een kunstvak kiezen zijn al voor die tijd cultureel actiever (Nagel e.a. 1996a, 1997).

De kans om ooit een klassiek concert te bezoeken staat het sterkst onder invloed van ouders en school, terwijl de minste ongelijkheden optreden bij bioscoopbezoek, waar de effecten van het ouderlijk milieu het kleinst zijn en schoolniveau helemaal geen invloed heeft. In feite lopen de effecten van de ouders en school parallel aan de veronderstelde toegankelijkheid van de cultuurvormen: hoe moeilijker toegankelijk de cultuurvorm, des te sterker hangt de kans op kennismaking af van ouders en school.

4.2.3. Ouders en school op verschillende leeftijden

Cultureel actieve ouders verhogen de kans op een eerste cultuurbezoek het sterkst op jonge leeftijd, maar zijn nog effectief tot 18 jaar; bij cultuur-historisch museum en een klassiek concert oefenen ouders en school ook verder in de levensloop nog invloed uit. Op een leeftijd waarop voortgezet onderwijs gevolgd wordt, van 12 tot 18 jaar, vergroot het *schoolniveau* de kans op een eerste cultuurbezoek (behalve bij cultuur-historische musea). Dit geldt ook voor het eerste bioscoopbezoek (waar het hoofdeffect niet significant was). Leerlingen van hogere schoolniveaus hebben op de leeftijd dat ze voortgezet onderwijs volgen een grotere kans om het eerste bezoek aan de bioscoop te brengen – voor zover ze er nog niet geweest waren – dan leerlingen van lagere schoolniveaus. Opvallend is dat de effecten van schoolniveau, behalve bij de bioscoop, op de leeftijd 12 tot 18 jaar nog kleiner zijn dan de effecten van het ouderlijk milieu. De school heeft de taak om jongeren te laten kennismaken met cultuur op deze leeftijd blijkbaar nog niet volledig overgenomen van de ouders.

De effecten van schooltype strekken zich echter verder uit dan de leeftijd van 12 tot 18. Ook op een leeftijd waarop ze het voortgezet onderwijs al lang verlaten hebben, hebben vwo-leerlingen nog steeds een grotere kans om voor het eerst een bioscoop, museum, theater of klassiek concert te bezoeken. Dit kan wijzen op het feit dat leerlingen van hogere schooltypes ook doorstromen naar de hogere onderwijsniveaus. Een effect van het niveau van voortgezet onderwijs kan daarom een effect van een hoger onderwijstype zijn. Een tweede mogelijkheid is dat het effect van school (en dit geldt ook voor ouders) niet alleen wijst op de culturele uitstapjes die de school organiseert, maar ook

terug te voeren is op socialisatie die blijkbaar op langere termijn zijn vruchten afwerpt. Onder invloed van school (en de ouders) zijn de kinderen cultureel geïnteresseerd geraakt en zijn later op eigen initiatief kunst gaan bezoeken.

De effecten van *kunstzinnige vakken* beperken zich daarentegen tot de leeftijd waarop voortgezet onderwijs is gevolgd. Wel verhoogt het volgen van een beeldend vak de kans op een eerste museum- en concertbezoek op een leeftijd waarop dat vak nog niet aan de orde was. Dit duidt op eerder gevonden selectie-effecten: leerlingen die al cultureel geïnteresseerd waren hebben vaker een kunstzinnig vak gekozen (Nagel e.a. 1996a, 1997).

5. CONCLUSIES

In dit artikel stond het begin van de culturele loopbaan centraal. De eerste onderzoeksvraag betrof de verschillen tussen cultuurvormen. Met welke cultuurvormen wordt relatief vroeg en met welke relatief laat begonnen?

Er blijkt min of meer een vaste volgorde te bestaan waarin kinderen met de verschillende cultuurvormen kennis maken. Kennelijk zijn de socialiserende instanties het eens over de (volgorde van) 'toegankelijkheid' van deze cultuurvormen. Hierbij wordt de bioscoop relatief vroeg, en klassiek concert relatief laat bezocht. Cultuur-historische musea, musea voor beeldende kunst en theater ontlopen elkaar qua toegankelijkheid niet veel. Overigens maakt lang niet iedereen in de eerste 25 jaar van de levensloop kennis met alle cultuurvormen. Op deze leeftijd is vrijwel iedereen wel naar de bioscoop geweest, maar dit geldt niet voor de overige cultuurvormen. Wat betreft het percentage dat er ooit geweest is, wordt de bioscoop gevolgd door museum en theater, terwijl een minderheid een klassiek concert heeft bezocht.

De tweede onderzoeksvraag was wie er vroeg en wie er laat beginnen met cultuur. Een cultureel actief ouderlijk milieu en een hoger niveau van voortgezet onderwijs vergroten de kans om kennis te maken met cultuur. Kunsteducatie in het voortgezet onderwijs verhoogt alleen de kans op een eerste bezoek aan een museum voor beeldende kunst of een uitvoering van klassieke muziek, waarbij de werking tot op grote hoogte discipline-specifiek is. De invloed van de ouders bleek echter wel het grootst te zijn: kinderen van cultureel actieve ouders maken het vroegst kennis met cultuur. Dit betekent dat De Jager (1967) gelijk had met zijn bewering dat cultureel actieve ouders hun kinderen al jong met cultuur in aanraking brengen.

De invloed van de ouders is het grootst in het begin van de levensloop maar beklijft tot en met 18-jarige leeftijd (en bij cultuur-historisch museum en klassiek concert ook nog na die tijd). Het effect van school begint op de leeftijd waarop voortgezet onderwijs wordt gevolgd en blijft aanwezig op latere leeftijd. Al met al blijven ouders en later school dus lang hun invloed houden. Mogelijk wijzen effecten van ouders en school op latere leeftijd op socialisatie-effecten: de ouders en school brengen hun kinderen interesse voor cultuur bij zodat ze later op eigen initiatief naar cultuur gaan, hoewel deze bij school ook kunnen wijzen op een effect van het tertiair onderwijs. De effecten van kunsteducatie beperken zich voornamelijk tot een leeftijd waarop het kunstonderwijs genoten werd.

6. DISCUSSIE

Wat betekenen de gevonden effecten van ouders en school? Over het geheel genomen is kans om met cultuur in aanraking te komen het grootst voor kinderen van cultureel actieve ouders. Dat de effecten van school minder groot zijn, betekent niet dat de school minder effectief zou zijn in het laten kennismaken met cultuur van haar leerlingen. Het effect van school op het eerste bezoek heeft namelijk alleen betrekking op kinderen uit minder cultureel actieve ouderlijke milieus. Alleen leerlingen die nog niet in één of meer cultuurvormen geïntroduceerd zijn, zijn op 12-jarige leeftijd nog 'at risk' om een eerste bezoek af te leggen. Als kinderen met hun ouders al naar de bioscoop, het museum, en theater zijn geweest, kan de school in dat opzicht niets meer toevoegen. Dit is natuurlijk inherent aan wat we onderzoeken: we onderzoeken de eerste keer, en dit kan nu eenmaal maar één keer gebeuren. Om ouders en school te vergelijken in hun effectiviteit om cultuurparticipanten te kweken, zou de lange termijn socialisatie door ouders en school vergeleken moeten worden aan de hand van het niveau van cultuurparticipatie op verschillende leeftijden (Nagel e.a. 1996b).

Hoe beïnvloedt de steekproef de resultaten die we in dit paper gevonden hebben? De gegevens betreffen oud-leerlingen die voortgezet onderwijs hebben gevolgd op scholen waar de mogelijkheid bestond om in een kunstzinnig vak eindexamen te doen. Aangezien lang niet alle scholen voor voortgezet onderwijs in die periode eindexamens in de kunstzinnige vakken aanboden, zijn dit waarschijnlijk scholen waar kunst en cultuur destijds in een hoog vaandel stond. Met de beschrijvende gegevens moeten we daarom voorzichtig zijn. Zo wordt het percentage leerlingen dat via de school kennis heeft gemaakt met cultuur, en op de leeftijd van 12 tot 18 jaar, waarschijnlijk overschat.

In de steekproef heeft de helft van de oud-leerlingen examen gedaan in een kunstzinnig vak, wat groter is dan in een representatieve steekproef van de leerlingen van de betreffende scholen het geval zou zijn geweest. Dit zou consequenties kunnen hebben voor de volgorde waarin de verschillende cultuurvormen de eerste keer bezocht zijn. Musea voor beeldende kunst en uitvoeringen van klassieke muziek zouden daardoor een eerdere plek in de rangorde kunnen hebben gekregen dan in de populatie het geval is. Dit maakt voor klassieke concerten echter niet uit: het eerste klassiek-concertbezoek stond toch al achteraan.

In dit artikel gaat het om het eerste bezoek aan vijf cultuurvormen. Natuurlijk houdt culturele socialisatie meer in dan de eerste kennismaking. Cultuuroverdracht is niet een eenmalige gebeurtenis die ophoudt bij het eerste bezoek. Socialisatie is een continu proces van overdracht van normen en van kennis, waarbij voortgebouwd wordt op de eerder opgedane kennis. Toch is het eerste bezoek belangrijk als element in het socialisatie-proces juist omdat het om het daadwerkelijk bezoek gaat. Dat is tenslotte het beoogde gedrag. Ook is het eerste bezoek belangrijk omdat het een indicator is voor de hoeveelheid en de aard van de socialisatie die men in zijn jeugd ondergaan heeft. De leeftijd van het eerste bezoek staat daarbij voor de hoeveelheid socialisatie: het geeft het aantal (jeugd)jaren aan waarin men nog aan socialisatie-invloeden blootgesteld is. Hoe jonger men het eerste bezoek aflegt des te meer jaren er in de jeugdperiode nog resten om dat nog een keer te doen.

LITERATUUR

- Allison, P.D., 1984. Event history analysis. Regression for longitudinal event data. Sage University Paper series on Quantitative Applications in the Social Sciences, 07-046. Beverly Hills and London: Sage Pubns.
- Ganzeboom, H.B.G., 1989. Cultuurdeelname in Nederland. Assen: Van Gorcum.
- Haan, J. de, 1997. Het gedeelde erfgoed. Een onderzoek naar veranderingen in de cultuurhistorische belangstelling sinds het einde van de jaren zeventig. Het culturele draagvlak 3. Rijswijk: Sociaal en Cultureel Planbureau (cahier 136)
- Haan, J. de; Knulst, W. 2000. Het bereik van de kunsten; een onderzoek naar veranderingen in de belangstelling voor beeldende en podiumkunsten sinds de jaren zeventig. Den Haag: Sociaal en Cultureel Planbureau; Den Haag: Elsevier bedrijfsinformatie (cahier 166)
- Hoorn, M. van; Haanstra, F.; Groot, B. de, 1998. Inventarisatieonderzoek 'Cultuur en school'. Utrecht: LOKV
- Huber, P.J., 1967. The behavior of maximum likelihood estimates under non-standard conditions. Proceedings of the fifth Berkeley Symposium on Mathematical Statistics and Probability 1, pp.221-233
- Jager, H. de, 1967. Cultuuroverdracht en concertbezoek. Leiden: Stenfert Kroese (dissertatie Universiteit Utrecht)
- Kamphorst, T.J.; Spruijt, A.P., 1983. Vrijtijdsgedrag in het perspectief van socialisatie. Een empirisch onderzoek naar samenhang tussen aspecten van socialisatieprocessen op de leeftijd van 5 tot en met 10 jaar vrijetijdsgedrag op de leeftijd van 25 tot en met 30 jaar. Utrecht (dissertatie Universiteit Utrecht)
- Knulst, W.P., 1995. Podiumkunsten in een tijdperk van afstandsbediening: onderzoek naar achtergronden van veranderingen in de omvang en samenstelling van het podiumpubliek sinds de jaren vijftig. Rijswijk: Sociaal en Cultureel Planbureau; Den Haag : VUGA (cahier 17)
- Mokken, R.J.; Lewis, C., 1982. A non-parametric approach to the analysis of dichotomous item responses. Applied Psychological Measurement 6, pp.417-430.
- Nagel, I.; Ganzeboom, H.; Haanstra, F.; Oud, W., 1996. Effecten van kunsteducatie in het voortgezet onderwijs. Amsterdam: SCO-Kohnstamm Instituut
- Nagel, I.; Ganzeboom, H., 1996. De invloed van ouders, school en buitenschoolse kunsteducatie. In: Nagel, I.; Ganzeboom, H.; Haanstra, F., 1996. Cultuurdeelname in de levensloop: de invloed van ouders, school en buitenschoolse kunsteducatie. Utrecht: LOKV (Katernen Kunsteducatie), pp.7-42
- Nagel, I.; Ganzeboom, H.; Haanstra, F.; Oud, W., 1997. Effects of art education in secondary schools on cultural participation in later life. Journal of Art & Design Education 16-3, pp.325-331
- Nagel, I.; Ganzeboom, H.; Haanstra, F.; Oud, W., 1997. Data set: Effecten van kunsteducatie in het voortgezet onderwijs 1994. KUNST94. ICS Code books -31 [KUNST94]. Principal Investigators: Ineke Nagel, Harry Ganzeboom, Folkert Haanstra, Wil Oud. Years of data collection: oktober 1993 – april 1995. Depositor: Department of Sociology, ICS, Utrecht University. Sponsor: Ministerie van OCW, LOKV. Producer: ICS Data Archive. Distributor: NIWI, Steinmetz Archives, postbus 95110, 1090 HC Amsterdam.

12. De opmars van de personal computer in Nederlandse huishoudens

J. DE HAAN, L. VAN DIJK en S. RIJKEN

SAMENVATTING

In dit artikel worden technologische veranderingen in Nederlandse huishoudens (personal computer en internet) besproken. Na een trage start in de eerste helft van de jaren tachtig, is de verspreiding van pc's snel gegaan in de jaren negentig. In 1998 had 58% een pc, tegen 8% in 1985. Sinds 1995 is het aantal internetaansluitingen vervijfvoudigd: van 4% in 1995 naar 21% in 1998. In dat laatste jaar kon in 19% van de huishoudens van e-mail gebruik gemaakt worden. De volgende groepen hebben een achterstand opgelopen in het verspreidingsproces van pc en internet: personen met een laag inkomen (die het meest achter liggen); (alleenstaande) vrouwen; 65-plussers; lager opgeleiden en personen zonder werk.

1. INLEIDING

De digitalisering van de Nederlandse huishoudens gaat snel. Cd-spelers brengen muziek in bits en bytes in steeds meer huizen. De telefoon werd draagbaar, vervolgens mobiel en door deze digitale bemiddeling werden velen op meer plaatsen bereikbaar en nam ook het initiatief tot telefonisch communiceren toe. Huishoudelijke apparatuur is steeds vaker van digitale techniek voorzien zonder dat gebruikers zich daar bewust van zijn. De meest prominente vernieuwing op het terrein van informatie- en communicatietechnologie (ict) is waarschijnlijk de personal computer (pc). De thuis-pc symboliseert als geen ander apparaat dat bezitters een plaats hebben verworven in de informatiesamenleving. Echter lang niet alle Nederlanders bezitten een pc. Hoe groot de veronderstelde impact van de pc ook mag zijn, in veel huishoudens is nog steeds geen pc aanwezig. Voor deze afwezigheid voeren de niet-bezitters uiteenlopende redenen aan. In paragraaf 2 zal eerst ingegaan worden op de verspreiding van de pc en de verschillen tussen bevolkingsgroepen die het gevolg zijn van ongelijke diffusiesnelheid. Daarna komen redenen van niet-bezit aan bod.

Tussen computer bezitters bestaan nog grote verschillen in het gebruik. Deze gebruiksverschillen komen in paragraaf 3 aan bod.

Tot ver in de jaren negentig werd de pc thuis vooral als 'stand alone' gebruikt. Sinds het midden van de jaren negentig groeit het aantal huishoudens met een netwerkaansluiting echter snel. Door de toegang tot internet komt een enorme hoeveelheid informatie beschikbaar. Deze ontwikkeling heeft voldoende gewicht om in paragraaf 4 apart aandacht te schenken aan de aansluiting van Nederlanders op de digitale snelweg. In deze paragraaf wordt het bezit en gebruik van internet en email beschreven.

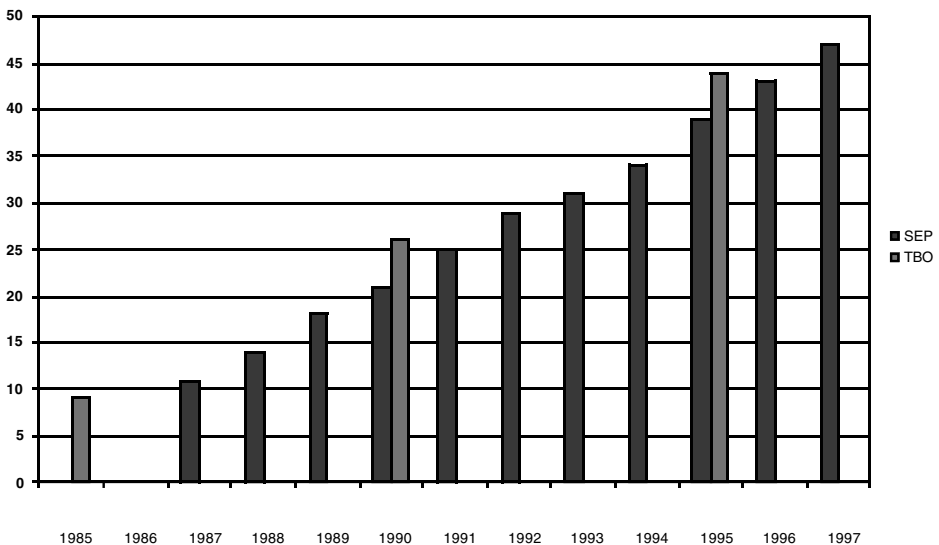
Voor de beschrijving van de opkomst van de pc in Nederlandse huishoudens wordt vooral gebruikt gemaakt van gegevens uit de Tijdsbestedingsonderzoeken (TBO) en van gegevens uit de survey Gebruik Nieuwe Communicatiemiddelen (GNC). De TBO-gegevens verschaffen inzicht in het bezit van de computer en de tijd die mensen thuis achter de computer doorbrengen. Het GNC-onderzoek biedt nadere informatie over de toepassingen waarvoor de pc gebruikt wordt. Kenmerken van beide databestanden worden besproken in het rapport *Digitalisering van de leefwereld* waar dit artikel op gebaseerd is (Van Dijk et al. 2000). Zoals verwacht mocht worden van een rapport over nieuwe media is het ook on line te raadplegen en wel op www.scp.nl.

2. COMPUTERBEZIT IN HUISHOUDENS: VERSPREIDING VAN DE PC TUSSEN 1985 EN 1998

De pc werd begin jaren tachtig in Nederland op de markt gebracht en vond eerst ingang in het bedrijfsleven. Tussen 1985 en 1995 verzesvoudigde het aantal geautomatiseerde werkplekken (WRR 1998:19). De eerste pc's waren duur en gebruikersonvriendelijk. Mede hierdoor kwam de verspreiding van pc's in Nederlandse huishoudens traag op gang. Veel pc's, die vaak door pc-privé projecten in huis verschenen, troffen het droevige lot van verstopping op zolder. Mede door toenemend bedieningsgemak en ondanks gelijkblijvende prijzen schaften steeds meer Nederlanders een personal computer aan voor thuisgebruik. Volgens gegevens van het CBS steeg het computerbezit in Nederlandse huishoudens van 11% in 1987 naar 47% in 1997 (zie figuur 1). In figuur 1 zijn tevens de cijfers van de vijfjaarlijkse tijdsbestedingsonderzoeken (TBO) opgenomen. De TBO-cijfers wijzen op een vergelijkbare groei: in 1985 was 9% van de Nederlandse huishoudens in het bezit van een pc, tegen 44% in 1995 (zie ook: SCP 1996).

Ondanks alle pogingen om gebruikersvriendelijke besturingssystemen te ontwikkelen blijft de pc met een aura van complexiteit omgeven. Door de snelle ontwikkelingen is het bovendien moeilijk om bij te blijven. Waarschijnlijk hierdoor is dit symbool van de informatierevolutie vooral in gebruik genomen door personen die goed met informatie overweg kunnen, de hoger opgeleiden. De gegevens in tabel 1 tonen aan dat al in 1985 de hoger opgeleiden onder de vroege pc-bezitters oververtegenwoordigd waren.¹ Dit verschil tussen hoog en laag opgeleiden is door de jaren heen blijven bestaan. Anno 1998 heeft nog geen kwart van de laagst opgeleiden een pc thuis, terwijl driekwart van de hoogst opgeleiden er een heeft. Ook het pc-bezit van ouderen blijft achter bij dat van jongeren. De complexiteit, prijs en een gebrek aan praktisch nut zullen hier debet aan zijn. Uit tabel 1 blijkt verder dat mannen, personen met kinderen, werkenden en studenten relatief vaak een pc bezitten.

1. De gegevens in tabel 1 zijn uitgedrukt in een percentage van de bevolking van 18 jaar en ouder. Eind 1998 had 58% van de bevolking thuis de beschikking over een pc. Deze presentatie van gegevens wijkt af van die in figuur 1 waar het pc bezit is uitgedrukt in procenten van de huishoudens. In 1997 was in 47% van de huishouden een pc aanwezig. Dat dit laatste cijfer veel lager uitvalt, is toe te schrijven aan het feit dat in huishoudens met veel personen vaker een pc aanwezig is dan in kleine huishoudens.



Bron: SCP (TBO '85-'95); CBS (Sociaal-Economische Panelonderzoek '87-'97).

Figuur 1. Bezit van een home- en/of personal computer in huishoudens, 1985-1997 (in procenten).

Eind 1998 hadden pc-bezitters hun eerste apparaat gemiddeld bijna 7 jaar daarvoor aangeschaft. Al eerder werd aangegeven dat de technologische ontwikkeling doorgaat en dat een computer derhalve snel verouderd. Dit leidt ertoe dat mensen geregeld een nieuwe pc kopen: de pc die eind 1998 in de Nederlandse huiskamers stond was gemiddeld pas 2,5 jaar oud.

De succesvolle verspreiding van een product verloopt volgens Rogers (1962, 1995) via een S-vormige curve. Aanvankelijk is het product bij een select gezelschap aanwezig. Hun voorbeeld wordt door anderen gevolgd tot het moment dat het moeilijk wordt om nog meer mensen voor het product te interesseren. Na verloop van tijd is het nieuwe product namelijk bij zoveel mensen aanwezig dat er van verzadiging gesproken kan worden. De curve zegt niet alleen iets over de verspreiding van een product in de samenleving, maar ook iets over het moment dat een individu het product heeft aangeschaft ten opzichte van het moment waarop anderen dat gedaan hebben. Bij een verdeling van de curve in vijf stadia ontstaat er een typering van de personen die grote bekendheid heeft gekregen. De mensen die als allereersten een product aanschaffen worden door Rogers de 'innovators' genoemd. Deze groep is klein van omvang. Vervolgens volgen de zogenaamde 'early adopters' die het product weliswaar later aanschaffen dan de innovators maar die er eerder bij zijn dan de meesten. Naarmate een product bekender wordt, zijn meer mensen bereid het product aan te schaffen en de derde groep die Rogers onderscheidt is dan de 'early majority': de mensen die even wachten voor ze het product kopen, maar die toch van 'de grote menigte' de eerste zijn die tot koop overgaan. Zij worden gevolgd door de 'late majority' en tot slot de 'laggards': de mensen die als allerlaatsten tot aanschaf van een product overgaan.

Uit de gegevens blijkt dat hoog opgeleiden en mensen met een hoog inkomen op het

Tabel 1. Bezit van een personal computer, naar achtergrondkenmerken van de respondent, personen van 18 jaar en ouder, 1985-1998 (in procenten).

	1985	1990	1995	1998
alle huishoudens	9	28	49	58
man	11	32	55	65
vrouw	7	23	44	52
alleenstaande man	10	23	42	60
alleenstaande vrouw	1	9	19	39
gehuwd/samenw. z. kind	5	19	38	51
gezin j. kind > 14 jaar	13	29	54	74
gezin j. kind \leq 14 jaar	11	41	66	80
20-34 jaar	11	31	55	68
35-49 jaar	13	43	68	76
50-64 jaar	6	16	40	51
65 jaar en ouder	2	4	9	13
lo	4	9	15	21
lbo,ulo,mavo	7	21	36	47
havo,vwo,mbo	12	32	52	68
hbo,univ.	12	44	69	76
inkomen 1e kwartiel	5	13	27	47
2e kwartiel	7	18	32	59
3e kwartiel	10	34	51	71
4e kwartiel	21	48	65	81
werkend	12	36	60	72
AOW gepensioneerd	2	7	18	21
werkloos, arbeidsongeschikt	7	26	37	47
huisvrouw	5	18	37	42
studerend	17	39	76	76

Bron: SCP (TBO 1985, 1990, 1995) , GNC 1998.

gebied van computers oververtegenwoordigd zijn onder de innovators. Dit zijn mensen die voor 1985 een computer aanschafte. Ook onder de early adopters (aanschaf 1985-1989) is een oververtegenwoordiging van mensen met een hogere opleiding of hoger inkomen te zien. Lager opgeleiden en mensen met een laag inkomen zijn oververtegenwoordigd als het gaat om de laggards, de mensen die achterblijven. Ook alleenstaande mannen en huishoudens waarvan de kinderen nu 15 jaar of ouder behoren tot de voorlopers, terwijl samenwonenden zonder kinderen en met name alleenstaande vrouwen zijn oververtegenwoordigd onder de laggards. Veel ouderen geven bovendien aan niet van plan te zijn om een pc aan te schaffen.

Aan mensen die geen computer hebben is gevraagd in hoeverre prijs, vaardigheden, interesse en computergebruik op het werk meespelen in hun beslissing (nog) geen pc aan te schaffen. Voor de helft van de niet-bezitters blijkt de prijs een obstakel te zijn en iets minder dan de helft zegt niet computers geïnteresseerd te zijn. Een derde van de mensen die thuis geen pc hebben, vindt dat zij aan een computer op het werk wel genoeg hebben. Ook zegt een derde niet te weten hoe met een computer omgegaan

moet worden; zij vinden derhalve dat zij over onvoldoende vaardigheden beschikken om een computer te bedienen. Onder de niet-computerbezitters die deze laatste reden opgeven zijn vrouwen, ouderen en lager opgeleiden oververtegenwoordigd.

Een kwart van de mensen die in 1998 geen computer thuis had, heeft dat ooit wel gehad. Aan hen is gevraagd waarom zij de computer weg hebben gedaan. Voor bijna de helft van deze mensen gold dat zij nauwelijks iets met het apparaat deden. Ook werd veroudering van het apparaat vaak genoemd. Tot slot scheidde een redelijk aantal mensen van hun pc vanwege 'familie-omstandigheden' (echtscheiding, vertrek uit ouderlijk huis).

Pc-bezitters verschillen aanzienlijk in het bedrag dat zij aan hun apparaat hebben uitgegeven. Een vijfde van de pc-bezitters betaalde er fl. 3500,— of meer voor. Daar tegenover staat een iets kleinere groep die de pc voor niets kreeg. Dat velen gratis in het bezit van de pc zijn gekomen, heeft verschillende redenen. De meeste kosteloze-pc-bezitters (64%) kregen deze van familie of bekenden, terwijl de rest de pc geheel vergoed kreeg van de werkgever.

Vrouwen, alleenstaanden en mensen met een laag inkomen geven relatief weinig geld uit aan een pc. Hetzelfde geldt voor mensen jonger dan 35 jaar en mensen van 65 jaar en ouder; zij krijgen relatief vaak een computer gratis. Jongeren krijgen vaak een pc van hun werkgever, of in het geval van bijvoorbeeld studenten van hun ouders. Het is aannemelijk dat ouderen vaak een enigszins verouderd toestel van hun kinderen overnemen.

3. GEBRUIK VAN DE PERSONAL COMPUTER: FREQUENTIE EN TOEPASSINGEN

3.1. Gebruiksfrequentie

Het TBO biedt informatie over het gebruik van de computer in de vrije tijd. Van de respondenten van de TBO's die thuis over een pc beschikken, is vanaf 1985 bekend in hoeverre zij in de loop van een week van deze apparatuur gebruik hebben gemaakt. Uit tabel 2 blijkt dat in 1995 23% van de bevolking wekelijks ten minste een kwartier van de vrije tijd achter het computerscherm doorbracht. Tien jaar eerder was dat nog maar 4%. Het wekelijkse gebruik kent daarmee een snellere verspreiding dan de computer zelf. Dit betekent dat computers steeds minder ongebruikt in huis staan. Het percentage pc-bezitters dat wekelijks actief was, steeg van 22% in 1985 naar 46% in 1995. Personen die thuis een internetaansluiting hebben, behoren nog vaker tot de regelmatige computergebruikers. In 1995 maakte 62% van de internetbezitters wekelijks gebruik van zijn of haar computer. De opmars van het computergebruik blijkt eveneens uit het gemiddelde aantal uren dat wekelijkse gebruikers actief zijn. Dit aantal steeg van 3,5 uur in 1985 naar 4,0 uur in 1995.

Voor alle bevolkingsgroepen geldt dat de gebruiksduur per week tussen 1985 en 1995 sterk gestegen is (tabel 2). Dit heeft uiteraard ook te maken met het feit dat per bevolkingsgroep steeds meer mensen een computer gebruiken. Er zijn nog steeds aanzienlijke verschillen tussen bevolkingsgroepen te constateren. Ouderen en vrouwen brengen

Tabel 2. Computergebruik, personen van 12 jaar en ouder, 1985-1995 (in procenten en gem. aantal uren).

	1985	1990	1995
A. Heeft beschikking over computer	19	30	51
B. Gebruikt computer wekelijks in de vrije tijd	4	13	23
C. B in % van A	22	44	46
D. B in % van Internet bezitters	-	-	62
E. Gem. aantal uren computergebruik onder wekelijks actieven	3,5	3,7	4

Bron: SCP (TBO '85-'95).

relatief weinig tijd achter de computer door. Het feit dat zij minder vaak een pc bezitten zal hier debet aan zijn. In gezinnen waar wel een computer aanwezig is, zullen vrouwen deze bovendien niet altijd gebruiken, in ieder geval minder dan hun mannelijke partner. Inwonende kinderen (in de leeftijd tussen 12 tot 19 jaar en schoolgaand of studierend) zijn veruit de grootste gebruikers van de pc. In 1995 zaten zij gemiddeld twee uur per week voor het computerscherm. De enigen die daar enigszins bij in de buurt komen zijn alleenstaande mannen.

Naast de wekelijkse pc-gebruikers zijn er ook mensen die dit minder frequent doen. Om gegevens hierover te krijgen kan beter naar het pc-gebruik voor een ruimere tijdsperiode gevraagd worden. Dit is gedaan in de GNC-enquête. Van alle mannelijke respondenten die een computer thuis hebben, gebruikt 19% de computer één keer per week of minder. Voor vrouwen ligt dit percentage bijna twee keer zo hoog (37%). Dit betekent dat het vragen naar computergebruik per week, zoals in het TBO wordt gedaan, kan leiden tot een onderschatting van het gebruik, met name voor vrouwen.

In de GNC-enquête is ook gevraagd naar het gebruik van de computer door thuiswonende kinderen. Jongens die thuis wonen maken op bijna evenveel dagen gebruik van de computer als volwassen mannen (respectievelijk 16 en 17 dagen per maand) (zie tabel 3). Jongens in de leeftijd van 9 tot 17 jaar zitten vaker achter de computer dan volwassen mannen. Meisjes maken over het algemeen genomen minder gebruik van de computer dan jongens; een uitzondering hierop is de leeftijdsgroep van 6 tot 8 jaar. Het lijkt er echter wel op dat vrouwen hun 'achterstand' zullen inlopen: de verschillen tussen meisjes en jongens zijn over de gehele linie minder groot dan tussen mannen en vrouwen.

3.2. Computertoepassingen

De pc kan voor steeds meer toepassingen gebruikt worden. Tabel 4 geeft voor vijf computertoepassingen weer door hoeveel mensen deze gebruikt worden: tekstverwerking, spreadsheet/databases, spelletjes, reisplanner/CD-foon en tekenprogramma's. Tekstverwerking is de meest gebruikte toepassing, op afstand gevolgd door spelprogramma's. Teknenprogramma's worden het minst gebruikt. Er zijn verschillen tussen bevolkingsgroepen als het gaat om het gebruik van computertoepassingen. Mannen zijn onder de gebruikers van alle computertoepassingen oververtegenwoordigd, met uitzondering van

Tabel 3. Gebruiksfrequentie (in dagen per maand) in huishoudens met een computer voor mannen, vrouwen en inwonende kinderen, naar achtergrondkenmerken, 1998.

	man	vrouw	jongens	meisjes
totaal	17	10	16	13
alleenstaand	18	12		
gehuwd/samenw. z. kind	17	9		
gezin j. kind > 14 jaar	16	8		
gezin j. kind \leq 14 jaar	17	9		
3-5 jaar			7	5
6-8 jaar			12	13
9-11 jaar			20	12
12-17 jaar			20	15
18-27 jaar			16	14
18-34 jaar	18	10		
35-49 jaar	17	11		
50-64 jaar	15	9		
65 jaar en ouder	16	7		
lo	13	10		
lbo,ulo,mavo	14	10		
havo,vwo,mbo	18	8		
hbo,univ.	18	12		

Bron: GNC, 1998.

de spelprogramma's. De verschillen tussen mannen en vrouwen worden voor een deel veroorzaakt door alleenstaande vrouwen. Zij gebruiken minder dan enig ander huishoudtype de spreadsheet, reisplanner/CD-foon en tekenprogramma's. De verschillen in het gebruik van het aantal computertoepassingen tussen leeftijdsgroepen zijn vrij klein. Alleen bij spel en tekenprogramma's zijn ouderen duidelijk ondervertegenwoordigd onder de gebruikers. Hoger opgeleiden vullen hun computergebruik anders in dan lager opgeleiden. Waar hoger opgeleiden vaker te vinden zijn onder de gebruikers van tekstverwerkingsprogramma's en spreadsheets zijn lager opgeleiden oververtegenwoordigd onder de gebruikers van spelletjes. Voor de andere twee toepassingen, reisplanner/cdfoon en tekenprogramma's zijn de verschillen tussen de opleidingsniveaus klein. Voor inkomen zien we een vrijwel identiek patroon als de voor opleiding met uitzondering van de laagste inkomensgroep. De laagste inkomens zijn namelijk oververtegenwoordigd bij gebruik van tekstverwerkingsprogramma's, spelletjes en tekenprogramma's. Deels zal dit komen omdat studenten tot de laagste inkomensgroep behoren; studenten zijn degenen die alle toepassingen – met uitzondering van tekenprogramma's het meest gebruiken.

De hoogst opgeleiden maken het meest gebruik van informatieve CD-roms, terwijl zij duidelijk ondervertegenwoordigd zijn onder de spelletjesspelers. Hierbij moet wel de kanttekening worden gemaakt dat ook van lager opgeleiden die een pc met CD-rom drive hebben iets meer dan de helft weleens informatie van een CD-rom haalt.

Opvallend is het gebruik van CD-roms onder ouderen (50-plus). Zij gebruiken het

Tabel 4. Gebruik van computerprogramma's naar achtergrondkenmerken van de respondent 1998 (in procenten van respondenten met een computer).

	tekst- verwer- king	spread- sheet/ database	spelletjes	cdfoon/ reis- planner	teken- programma's
totaal	86	45	59	46	37
man	89	58	58	55	45
vrouw	81	26	60	32	26
alleenstaande man	89	60	70	55	53
alleenstaande vrouw	94	36	64	30	28
gehuwd/samenw. z. kind	85	45	58	49	36
gezin j. kind > 14 jaar	77	44	39	47	30
gezin j. kind \leq 14 jaar	83	40	58	44	35
18-34 jaar	89	46	74	45	41
35-49 jaar	84	44	54	47	35
50-64 jaar	83	42	36	45	35
65 jaar en ouder	79	46	31	41	16
lo	73	16	78	47	32
lbo,ulo,mavo	72	32	66	42	38
havo,vwo,mbo	88	45	64	43	36
hbo,univ.	95	56	46	50	37
inkomen 1e kwartiel	90	43	74	42	42
2e kwartiel	82	40	62	47	36
3e kwartiel	85	39	57	45	36
4e kwartiel	87	59	46	52	34
werkend	88	47	58	50	37
AOW gepensioneerd	81	46	35	50	22
werkloos, arbeidsongeschikt	78	39	60	32	28
huisvrouw	57	16	58	22	29
studerend	96	53	81	39	51

Bron: GNC 1998.

vaakst educatieve CD-roms en veruit het minst spel-CD-roms. Wat betreft het gebruik van informatieve CD-roms lijken ze veel op de jongere generaties. Hierbij moet wel opgemerkt worden dat het hier om een specifieke groep ouderen gaat: de ouderen die een pc met CD-romdrive hebben. Als ouderen eenmaal een computer hebben aangeschaft dan lijken zij deze voor meer serieuze doeleinden te gebruiken dan jongere computerbezitters.

4. INTERNET EN EMAIL

4.1. Inleiding

Hoewel de basis voor internet al in het einde van de jaren zestig werd gelegd, duurde het tot in de jaren negentig voordat internet een wereldwijd succes werd (Van de Wijn-

gaert 1999).² In mei 1993 werd de eerste Nederlandse internet Service Provider, genaamd XS4ALL, geïntroduceerd. Voor het grote publiek in Nederland kreeg internet bekendheid door het experiment De Digitale Stad dat in januari 1994 van start ging. Toen het World Wide Web (WWW) werd geïntroduceerd, werd gebruik van internet aantrekkelijker, omdat nu ook (bewegende) beelden en geluiden opgevraagd konden worden (Van de Wijngaert 1999).

4.2. Internet en email: aanschaf

De laatste jaren is internet sterk in opkomst. In de periode 1994-1996 verviervoudigde het aantal internetabonnees in Nederland van 150.000 naar 600.000. In 1996 was 60 procent van de abonnementen voor zakelijk gebruik en 40 procent voor particuliere doeleinden (Boumans 1997: bron SUMMO/Multiscope/Trendbox). Uit de vijfde Nationale Internet Monitor van het onderzoeksbureau Pro Active International blijkt dat het aantal internetgebruikers in Nederland is gestegen van 2,7 miljoen in september 1999 naar 4,7 miljoen in maart 2000; de grootste toename in Nederland ooit.³ Deze groei werd in belangrijke mate bepaald door de komst van de gratis internetproviders in 1999.

Om hier een preciezer beeld te geven van de internetgebruikers en het soort gebruik zal hier opnieuw gebruik gemaakt worden van de TBO- en GNC-gegevens. Deze gegevens zijn weliswaar minder actueel, maar bevatten gedetailleerde informatie over de gebruikers en het gebruik. Uit deze gegevens blijkt dat het aantal internetaansluitingen in Nederlandse huishoudens in drie jaar tijd verzevenvoudigd is: van 3 procent in 1995 naar 21 procent in 1998 (tabel 5). In hoeverre de mogelijkheden om van email gebruik te maken zijn toegenomen in die periode kan niet nagegaan worden omdat het TBO van 1995 geen gegevens bevat over emailaansluitingen. In 1998 heeft één op de vijf huishoudens in Nederland email thuis, zo blijkt uit de GNC-gegevens.⁴ Het zijn de inmiddels bekende bevolkingsgroepen die achterblijven als het gaat om het bezit van de internet en email thuis: (alleenstaande) vrouwen, 65-plussers, lager opgeleiden en mensen met een laag inkomen. Door de dalende prijzen voor internet wordt het voor lagere inkomens aantrekkelijker zich de toegang tot de digitale snelweg te verschaffen; wel blijft de prijs van een computer wellicht nog een barrière hiervoor. Alleenstaande

2 De oorsprong ligt bij het Amerikaanse ministerie van Defensie waar een aantal computers aan elkaar gekoppeld werd. Dit netwerk moest gebruikers op verschillende lokaties in staat stellen op computers elders in te loggen. Al snel werd echter ontdekt dat dit netwerk ook vele communicatiemogelijkheden bood zoals het transporteren van bestanden en het versturen van elektronische post (Van den Hooff en Tebbal 1996). Nog steeds zijn genoemde functies belangrijk voor door netwerken gekoppelde computers.

3 Zulke cijfers hebben overigens een brede betrouwbaarheidsmarge. Onderzoeksbureau Trendbox deed eveneens in het voorjaar van 2000 onderzoek naar internet gebruik en kwam tot de conclusie dat al ruim de helft van de Nederlanders online is en 37% wekelijks op internet zit. Dit laatste cijfer (bijna 6 miljoen Nederlanders) ligt nog ruim boven dat van Pro Active International. Ligt de waarheid ergens in het midden?

4 Aangezien meer personen een modem hebben dan toegang tot internet, wordt de netwerktoegang ook nog voor andere doeleinden gebruikt. Gevraagd naar de verschillende gebruiksmogelijkheden van het modem antwoordt twee derde van de modem-bezitters dat zij dit voor internet en email gebruiken. Telebankieren met behulp van een modem wordt door 40 procent gedaan, terwijl één op de vijf het gebruikt voor het inbellen op het computernetwerk van de werkgever.

mannen zijn duidelijk oververtegenwoordigd onder degenen die de digitale snelweg vanaf hun computer thuis kunnen bereiken, evenals mensen met een hbo of universitaire opleiding en mensen met een hoog inkomen.

Internet- en emailbezitters beschikken gemiddeld zo'n twee jaar over deze mogelijkheden. De meeste huishoudens die internet hebben, hadden al wel de beschikking over een computer thuis. Het zijn de early adopters op het gebied van computers die ook als eersten internet aanschafte. Zij deden dit wel geregeld tegelijkertijd met de vervanging van een oude computer. In totaal 46 procent van de huishoudens die een internetaansluiting aanschafte, kochten tegelijkertijd een (andere) computer. In driekwart van deze gevallen ging het om de vervanging van een oude PC terwijl het overige kwart voor het eerst een computer kocht.

Tabel 5. Bezit van emailaansluiting en toegang tot internet, naar achtergrondkenmerken, 1985-1995 (in procenten van Nederlandse huishoudens).

	email		toegang internet		
	1998	1998 alleen PC- bezitters	1995	1998	1998 alleen PC- bezitters
gehele bevolking // computer-bezitters	19	34	4	21	37
mannen	27	42	4	28	44
vrouwen	13	25	3	14	27
alleenstaande man	30	51	5	30	50
alleenstaande vrouw	8	20	1	9	23
gehuwd/samenw. z. kind	18	36	3	20	39
gezin j. kind > 14 jaar	23	31	3	26	35
gezin j. kind \leq 14 jaar	25	32	5	27	35
18-34 jaar	25	36	5	26	38
35-49 jaar	27	35	5	29	39
50-64 jaar	16	31	1	16	32
65 jaar en ouder	3	20	0	2	20
lo	3	12	1	3	16
lbo,ulo,mavo	14	29	2	15	32
havo,vwo,mbo	22	32	3	23	35
hbo,univ.	33	44	7	34	46
1e kwartiel	14	30	2	15	33
2e kwartiel	17	28	2	18	31
3e kwartiel	26	35	2	26	37
4e kwartiel	37	46	6	39	49
werkend	27	37	5	28	40
AOW gepensioneerd	7	29	0	6	28
werkloos, arbeidsongeschikt	15	32	1	15	32
huisvrouw	8	19	2	9	21
studerend	25	35	8	28	39

Bron: SCP (TBO 1995) , GNC (1998).

4.3. Gebruik van internet en email

De digitale snelweg wordt almaar drukker. De dalende kosten dragen eraan bij dat steeds meer mensen steeds langer het internet opgaan. Als men thuis geen internetaansluiting heeft dan kan men ook op andere plaatsen terecht, zoals in openbare bibliotheken of speciale internet café's, op het werk of op school. In de GNC-enquête is aan iedereen gevraagd hoe vaak thuis, op het werk of op school gebruik gemaakt wordt van internet. Nederlanders gaan gemiddeld genomen zes dagen per maand het net op. Email wordt iets vaker gedaan: gemiddeld genomen zo'n 7 tot 8 dagen per maand. Het internet- en emailgebruik is lager onder vrouwen, ouderen, lager opgeleiden en mensen met een laag inkomen (tabel 6). Werkenden en studerende gebruiken internet en email veel vaker dan huisvrouwen, gepensioneerden en werklozen. Bij mannen zijn de verschillen naar leeftijd en opleiding groter dan bij vrouwen.

Tabel 6. Gebruiksfrequentie van internet (in dagen per maand) door mannen, vrouwen en inwonende kinderen naar achtergrondkenmerken 1998 (alle huishoudens).

	man		vrouw	
	internet	email	internet	email
totaal	7	8	3	4
alleenstaand	8	9	4	6
gehuwd/samenw. z. kind	5	6	3	5
gezin j. kind > 14 jaar	5	7	1	4
gezin j. kind ≤ 14 jaar	7	10	2	3
18-34 jaar	10	11	4	5
35-49 jaar	7	9	3	5
50-64 jaar	4	5	2	4
65 jaar en ouder	1	1	0	0
lo	0	2	0	1
lbo,ulo,mavo	4	3	2	3
havo,vwo,mbo	8	9	3	4
hbo, univ.	10	12	5	8
inkomen, 1e kwartiel	6	7	3	3
2e kwartiel	5	6	3	5
3e kwartiel	7	8	3	6
4e kwartiel	10	11	5	7
werkend	8	9	4	7
AOW gepensioneerd	2	1	0	0
werkloos, arbeidsongeschikt	4	4	2	3
huisvrouw	–	–	0	1
studerend	13	12	6	5

Bron: GNC, 1998.

Het internet biedt een grote hoeveelheid informatie over bedrijven, overheidsinstellingen, non-profitorganisaties etc. Ook steeds meer individuen zetten gegevens op internet. Daarnaast kan er via babbelboxen gecommuniceerd worden met mensen over de gehele wereld of kan men bijvoorbeeld recent verschenen Amerikaanse boeken bestel-

len. Maar het meest wordt het internet gebruikt voor spelletjes, het downloaden van muziek en voor sex, althans als wordt afgegaan op de meest gebruikte zoektermen die in de zoekmachines worden ingetikt. De website Searchterms.com houdt deze informatie bij en 'games', 'music' en 'sex' staan bovenaan en dit is al geruime tijd het geval. Wel is 'sex' onlangs van de eerste plaats verdrongen door 'games'. Verder informatie het surfgedrag van internetters kan afgeleid worden uit de bezoektellingen van de websites. Uit onderzoek van Multiscope blijkt dat Nederlandse internetgebruikers het meest op startpagina.nl komen. Ook andere startpagina's, zoals die van Planet Internet, World Online en Microsoft Network zijn populair. Verder worden zoekmachines (Ilse, Altavista) druk bezocht, evenals pagina's met nuttige informatie zoals het spoorboekje en teletekst.

De verschillende gebruiksmogelijkheden betekenen dat gebruikersprofielen aardig uiteen kunnen lopen. Om hier nader inzicht in te krijgen is in de GNC-enquête gevraagd naar zes gebruiksmogelijkheden: zoeken naar informatie, raadplegen van nieuwsgroepen, toegang tot on-line databases, downloaden van software, chatbox, en 'zomaar wat surfen'. Van deze mogelijkheden wordt internet het meest gebruikt voor het zoeken van informatie en zomaar wat surfen. Het raadplegen van nieuwsgroepen, het zoeken van toegang tot on-line databases en het downloaden van software wordt door ruim een derde van de internetgebruikers gedaan. Chatten is het minst populair en wordt slechts door 13 procent van de internetgebruikers gedaan.

Mannen zijn bij alle toepassingen oververtegenwoordigd. Dit geldt bij uitstek voor alleenstaande mannen; zij lijken de meest intensieve gebruikers van de digitale snelweg te zijn en het meest gebruik te maken van alle mogelijkheden die het net te bieden heeft. De verschillen tussen mannen en vrouwen zijn vooral groot als het gaat om het verkrijgen van toegang tot on-line databases en het downloaden van software. Eerder werd duidelijk dat relatief weinig ouderen (vanaf 50 jaar) internet opgaan. Wanneer eenmaal internet gebruikt wordt dan zijn de verschillen tussen leeftijdsgroepen voor de meeste toepassingen echter vrij gering. Uitzonderingen hierop zijn chatten dat beduidend vaker door mensen van 18 tot 35 jaar wordt gedaan. Naarmate mensen ouder zijn, lijken zij bovendien doelbewuster het net op te gaan; zomaar wat surfen wordt minder gedaan door ouderen. Hoger opgeleiden chatten minder en maken minder gebruik van de mogelijkheid om software te downloaden. Lager opgeleiden zoeken juist weer minder toegang tot on-line databases dan hoger opgeleiden.

Voor thuiswonende kinderen is alleen gekeken naar de groep van 12 tot 27 jaar omdat jongere kinderen nauwelijks gebruik maken van internet. Meisjes lijken iets doelgerichter naar informatie te zoeken dan jongens, terwijl jongens iets vaker 'zomaar wat surfen'. Voor zowel jongens als meisjes geldt dat 18-plussers vaker gericht naar informatie zoeken dan jongeren onder de 18 jaar, terwijl ze juist minder zomaar wat surfen. Evenals bij (uitwonende) volwassenen zijn dit de twee meest gebruikte toepassingen. Het grootste verschil met volwassenen is dat zich onder de thuiswonende jongeren en jongvolwassenen een grotere proportie 'chatters' bevindt.

5. SAMENVATTING EN SLOTBESCHOUWING

Begin jaren 80 kwam de personal computer in Nederland op de markt, maar in de privé sfeer duurde het tot in de jaren negentig voordat een aanzienlijk deel van de

bevolking overging tot aanschaf. Aan het einde van de 20e eeuw stond in iets meer dan de helft van de huishoudens een personal computer en het einde aan de groei van het aantal computers is nog niet in zicht.

Nieuwe producten zoals de pc verspreiden zich volgens een trickle-down principe: de hogere statusgroepen – met name de hogere inkomensgroepen – schaffen deze producten eerst aan en de lagere statusgroepen volgen. Personen met weinig financiële draagkracht profiteren relatief laat van nieuwe mogelijkheden. Ook bij de ict-producten duikt een aantal groepen steeds opnieuw op als het niet-bezit ter sprake komt. Tot deze groepen behoren, geordend naar de gemiddelde omvang van hun achterstand: personen in huishoudens met een laag inkomen (die gemiddeld de grootste achterstanden hebben), (alleenstaande) vrouwen, 65-plussers, mensen met een lagere (voortgezette) opleiding en werklozen. Niet expliciet onderzocht, maar wel aannemelijk is dat ook allochtonen hiertoe behoren. Het ligt voor de hand dat bestaande verschillen tijdens de voortgaande verspreiding weer zullen verkleinen. Relevante vragen daarbij zijn hoe lang dit zal duren en, belangrijker nog, hoe erg deze wachttijd is.

In de informatiemaatschappij lijkt het steeds belangrijker om over digitale vaardigheden te beschikken. Maar niet iedereen kan evengoed met computers overweg. Met name ouderen, lager opgeleiden en (alleenstaande) vrouwen hebben een achterstand opgelopen. Zij beschikken minder vaak over een computer, gebruiken deze minder vaak en voor minder (complexe) toepassingen dan respectievelijk jongeren, hoger opgeleiden en mannen. In een maatschappij waarin informatie steeds belangrijker wordt, zou dit tot problemen kunnen leiden. Voor individuele burgers past de vraag in hoeverre mensen door niet-bezit of gebrek aan vaardigheden belemmerd worden in hun sociale en maatschappelijke leven. In hoeverre staat ontoereikend ICT-bezit en -vaardigheden het bereiken van belangrijke maatschappelijke doelen in de weg? Tot deze doelen behoren toegang tot de arbeidsmarkt, aansluiting bij dominante gedrags- en waardepatronen in de samenleving, behoud van sociale contacten en aanraking met de voorzieningen van de verzorgingsstaat.

De arbeidsmarkt vraagt steeds meer digitaalvaardige werknemers. Het zijn de groepen die van ouds een achterstand op de arbeidsmarkt hadden – lager opgeleiden en vrouwen – die minder vaak met computers werken en mede daardoor ook over minder vaardigheden beschikken. Vrouwen, die de laatste jaren hun positie op de arbeidsmarkt hebben versterkt, zouden door deze ontwikkelingen bijvoorbeeld weer – verder – achterop kunnen raken bij mannen. Gezien de doelstellingen van de overheid, waarin onder andere de economische zelfstandigheid van vrouwen centraal staat, lijkt dit een ongewenste ontwikkeling. Daarbij komt dat vrouwen minder dan mannen instromen in de IT-sector, een van de snelst groeiende sectoren van de arbeidsmarkt (Frissen 1999). Een vraag daarbij is ook hoe duurzaam de verschillen tussen mannen en vrouwen zullen zijn. Recente cijfers wijzen al op een opmars van vrouwen op internet. Volgens onderzoeksbureau Trendbox is medio 2000 veertig procent van de frequente gebruikers vrouw. De groei in het eerste half jaar van 2000 van vrouwen lag ruim hoger dan die van mannen. In de Verenigde Staten is in 2000 een onderzoek gehouden (van internet-enquêtebureaus Nielsen/Netratings) waaruit voor het eerst blijkt dat er meer vrouwen op het internet surfen dan mannen. Zulke cijfers wijzen op een inhaalslag van vrouwen. Of de arbeidsmarktparticipatie van vrouwen belemmerd wordt door hun trage start in de digitale wereld is niet bekend, dus kan ook niet gezegd worden of en in hoeverre digitale

vaardigheden samenhangen met toenemende inkomensverschillen en uiteenlopende carrièrekansen tussen mannen en vrouwen.

Een andere groep die achterblijft zijn de lager opgeleiden. Het toenemend belang van digitale vaardigheden betekent dat ook lager opgeleiden voor steeds minder banen in aanmerking kunnen komen. Hierdoor kunnen ongelijkheden op de arbeidsmarkt tussen mensen met een verschillende opleidingsniveau groter worden. Dit zou een ongewenste ontwikkeling zijn.

In het SCP rapport *Digitalisering van de leefwereld* is opgemerkt dat verschillen in ict-bezit en gebruik tussen jongeren en ouderen tot 2002 nog zullen toenemen (Van Dijk et al. 2000). Verwacht werd dat de groei van het pc-bezit en de toegang tot internet het sterkst zou zijn onder jongeren van 18 tot 34 jaar. Het reeds genoemde onderzoek van Trendbox bevestigt dit. Van de jongeren van 16 tot 25 jaar zou zelfs 85 procent al toegang hebben tot internet. De 65-plussers blijven daarentegen steken op zeven procent. De verschillen tussen bevolkingsgroepen zijn nergens zo groot als tussen jongeren en ouderen. Ouderen hebben een extra zetje in de rug nodig om de digitale drempel over te gaan. Het succes van onder meer SeniorWeb bewijst dat als ouderen eenmaal hun schroom overwonnen hebben, zij zeer enthousiaste computer- en internetgebruikers kunnen worden. Vaak zijn deze digitaal actieve ouderen al tijdens hun beroepsloopbaan met computers in aanraking gekomen (Van Kesteren 1999). Maar de meeste ouderen twifelen nog steeds aan het praktische nut van pc en internet. Veel ouderen hebben geen ambities meer op arbeidsmarkt. Voor deze groep is het al dan niet beschikken over digitale vaardigheden dan ook niet instrumenteel voor het verwerven van inkomen. Vanuit dit oogpunt bezien hoeven zij dus niet over de nieuwste technologie te beschikken.

Andere doelen dan arbeidsmarktparticipatie lijken ook en evengoed met andere middelen te realiseren. Voor het contact met anderen en voor de toegang tot overheidsdiensten zijn nog steeds toegankelijke alternatieven aanwezig voor personen die niet over moderne ICT beschikken. Gewenste informatie kan ook via traditionele media verkregen worden en voor de communicatie met anderen zijn alternatieven voorhanden als men geen email of mobiele telefoon heeft. Niet-bezitters zullen dan ook niet snel buiten het maatschappelijke leven komen te staan zolang de traditionele toegang tot informatie en diensten open blijft.

LITERATUUR

Boumans, J., Elektronische informatie diensten en producten. In: J. Hemels, A. Herpers, H. van der Pol en H. Prins (red.) Gids Informatiesector 1997/1998, Den Haag: NBLC / Boekblad, 1997, 191-194.

Dijk, Liset van, Jos de Haan en Susanne Rijken, Digitalisering van de leefwereld; een onderzoek naar informatie- en communicatietechnologie en sociale ongelijkheid; eindrapport, Den Haag: Sociaal en Cultureel Planbureau, 2000.

Frissen, V.A.J., Domeinverkenning 'De informatiesamenleving'. In opdracht van het Ministerie van SZW/DCE. Amsterdam: The Amsterdam School of Communication Research, 1999.

Hooff, Bart van den en Rachid Tebbal, Internet in Nederland. In: Harry Bouwman en Lidwien van de Wijngaert (red.), Multimedia en route; tien notities over multimedia en Internet toepassingen, Amsterdam: Otto Cramwinckel, 1996, 56-76.

Kesteren, Manfred van, Het bezit en gebruik van de personal computer; verslag van een kwalitatief en kwantitatief onderzoek, doctoraalscriptie sociologie Utrecht, 1999.

NIPO. Continue Registratie van Internet-penetratie. Nieuwsbrief, nr 3036, 3 juni 1998. Amsterdam: NIPO, 1998.

Rogers, E.M., Diffusion of Innovations. Fourth Edition. New York: The Free Press, 1995. Eerste editie 1962.

SCP. Sociaal en Cultureel Rapport 1996. Rijswijk / Den Haag: Sociaal en Cultureel Planbureau / VUGA, 1996.

WRR, Staat zonder land. Een verkenning van bestuurlijke gevolgen van informatie- en communicatietechnologie. Den Haag: Wetenschappelijke Raad voor het Regeringsbeleid/SDU Uitgevers, 1998.

Wijngaert, Lidwien van de, Matching media, information need and new media choice, Enschede: Telematica Instituut, 1999.

13. Niet alleen de mening van de klant telt!

*Een onderzoek naar het effect van de mening van gebruikers van marktonderzoekinformatie**

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

SAMENVATTING

Tevreden klanten zijn vandaag de dag van het grootste belang voor organisaties. Dientengevolge worden er door bedrijven vaak grote sommen geld gestoken in klanttevredenheidsonderzoeken. Gezien de hoge kosten die vaak aan dergelijke marktonderzoeken zijn verbonden lijkt een beter begrip van de factoren die maken dat klanttevredenheid informatie door managers gebruikt zal worden van belang. In dit artikel trachten we selectief gebruik van klanttevredenheid gerelateerde marktonderzoekinformatie te verklaren uit de sterkte van de attitude oftewel de mening van een informatie gebruiker ten aanzien van dit onderwerp. De resultaten laten zien dat twee dimensies van attitude sterkte, commitment en embeddedness in het Engels genaamd, zowel direct als indirect tot intenties tot selectief gebruik van tevredenheid gerelateerde informatie zal leiden. Indirecte effecten komen tot stand via cognitieve denkprocessen bij de gebruiker en het selectief beoordelen van de desbetreffende informatie. Een beter begrip van hoe iemands attitude van invloed is op het gebruik van vaak prijzige klanttevredenheid gerelateerde marktonderzoekinformatie zal resulteren in verhoogde acceptatie en effectiever gebruik van dergelijke informatie.

1. INLEIDING

Klanttevredenheid is nog steeds een belangrijke maatstaf in het marketingbeleid van veel ondernemingen. Bij sommige bedrijven staat deze maatstaf wel erg centraal. Zo stelt fabrikant van telecommunicatie apparatuur Motorola in hun missie statement het volgende: 'customer satisfaction is our top priority'. Dit is niet zo vreemd. Steeds weer blijkt namelijk uit onderzoek dat klanttevredenheid in belangrijke mate bijdraagt aan het behoud van klanten en winstgevendheid (Anderson, 1996; Rust et al., 1995; Wetzels et al., 1998). Daarom wordt er door bedrijven veel geld en energie gestoken in het verzamelen van informatie over klanttevredenheid d.m.v. marktonderzoek. Tegelijkertijd vindt er echter ook veel verspilling van geld en energie plaats vanwege het feit dat er niets met de uit onderzoek verkregen informatie gedaan wordt. Tevredenheid blijkt namelijk niet altijd een garantie te zijn voor het blijven van klanten en ontevredenheid leidt niet altijd tot weglopen. Vergelijk het maar met het feit dat in heel wat huwelijken de partners ontevreden zijn, maar toch bij elkaar blijven. Dit leidt vervolgens tot het

* Dit artikel komt voort uit een proefschrift over de effectiviteit van internationaal marktonderzoekinformatie. (Voor meer informatie: M. van Birgelen, 043-3883624).

recentelijk door Mulder (1999) verwoorde cynisme dat stelt dat het allemaal geen moer uit maakt of je klanttevredenheid periodiek meet. En als gevolg hiervan wordt de onderzoekskraan beetje bij beetje dichtgedraaid en wordt klanttevredenheid vaak van het prioriteitenlijstje afgevoerd. Het lijkt daarom belangrijk zicht te krijgen op die factoren die effectief gebruik van informatie over klanttevredenheid bevorderen.

Vaak wordt in het onderzoek naar de effectiviteit van marktonderzoek aandacht besteed aan zaken die het meest voor de hand liggen. Hierbij valt te denken aan de kwaliteit van de inhoud en vorm van de rapportage, de eigenschappen van de onderzoeker, de 'politieke' haalbaarheid van de aanbevelingen, etc. Van dergelijk onderzoek deden we in dit jaarboek recentelijk verslag (van Birgelen et al., 1999). Er is echter weinig bekend over het effect van de mening oftewel attitude van gebruikers van marktonderzoekinformatie. Zo blijkt uit sociaal-psychologisch onderzoek (Jaccard, 1981; Jaccard en Becker, 1985; Pomerantz et al., 1995) dat de sterkte van een attitude ten aanzien van een bepaald onderwerp van doorslaggevend belang is voor de manier waarop mensen zich denken te gaan gedragen. Met betrekking tot het gebruik van marktonderzoekinformatie betekent dit dat als managers in organisaties er heilig van overtuigd zijn dat klanttevredenheid belangrijk is, dan zal dat leiden tot een actiever gebruik van informatie omtrent klanttevredenheid. Dit lijkt een open deur en dat is het waarschijnlijk ook. Echter, achter deze deur zijn er nog een aantal deuren die vooralsnog gesloten bleven. Het gaat hierbij om vragen als "hoe is zo'n sterke overtuiging opgebouwd?", "welke cognitieve processen spelen hierbij een rol?" en "leidt attitude sterkte tot een selectief oordeel over en gebruik van informatie?"

Niet alleen de mening van de klant maar ook die van de informatie gebruiker lijkt daarom van belang te zijn. In dit artikel kijken we derhalve vanuit dit oogpunt naar de effectiviteit van marktonderzoekinformatie en proberen we op eerdergenoemde vragen een antwoord te formuleren. Het artikel is als volgt opgebouwd. Eerst staan we kort stil bij het begrip 'attitude sterkte' om de lezer op de hoogte te kunnen stellen van de resultaten van eerder onderzoek. Vervolgens gaan we na welke directe en indirecte gevolgen een sterke attitude of overtuiging ten aanzien van het belang van klanttevredenheid heeft op intenties om informatie hieromtrent te gebruiken. We bespreken vervolgens de resultaten van een empirisch onderzoek onder gebruikers van marktonderzoekinformatie en gaan na wat de implicaties zijn van de uitkomsten van ons onderzoek.

2. ATTITUDE STERKTE: EEN OVERZICHT

Personen verschillen doorgaans in de manier waarop ze tegen bepaalde zaken aankijken. De sociale en economische psychologie literatuur verwijst naar deze verschillen met de term 'attitude'. Een attitude wordt omschreven als "een psychologische tendens die tot uiting komt bij de evaluatie van een bepaalde entiteit met een zekere mate van voorkeur dan wel afkeur" (vertaald van Eagly en Chaiken, 1993, p. 1). Eigenlijk komt een attitude overeen met een mening over iets of iemand. Attitudes kunnen verschillen in sterkte. Dit fenomeen, ook wel 'attitude strength' in het Engels geheten, is een veel onderzocht concept. Het geven van een eenduidige definitie van attitude sterkte is desondanks niet eenvoudig. In het algemeen heerst er overeenstemming dat sterke attitudes leiden tot selectieve gedachten, niet snel veranderen en voorspelbaar voor gedrag kunnen zijn (Abelson, 1988; Chaiken et al., 1995; Eagly en Chaiken, 1993, 1995; Krosnick et al., 1993; Olson en Zanna, 1993). Krosnick en Petty (1995) geven verder aan

dat attitude sterkte ook van invloed is op de wijze waarop personen informatie over een bepaald onderwerp verwerken.

Het concept attitude sterkte wordt omgeven door verschillende vragen. De belangrijkste vraag luidt: 'Bestaat attitude sterkte uit één enkel construct (ook wel dimensie genoemd) of uit meerdere?' Er bestaan een aantal goede exploratieve studies naar de verschillende constructen achter attitude sterkte (bijv. Abelson, 1988; Krosnick et al., 1993; Lastovika en Gardner, 1979; Raden, 1985). De meeste studies stellen dat attitude sterkte uit meerdere onderliggende constructen bestaat. Pomerantz et al. (1995) vonden bewijs voor twee attitude sterkte dimensies. De eerste dimensie is 'commitment' oftewel iemands betrokkenheid bij een onderwerp. Dit heeft te maken met de mate van overtuiging dat een bepaalde mening over iets of iemand correct is en dat men een extreem standpunt hieromtrent inneemt. De tweede dimensie heet 'embeddedness' oftewel de mate waarin het onderwerp een onderdeel is van iemands diepere persoonlijkheid en heeft te maken met het feit of iemands mening bepaalde innerlijke waarden vertegenwoordigt.

Om commitment concreet te meten gebruikten Pomerantz et al. (1995) voornamelijk items of vragen betrekking hebbende op attitude zekerheid en extremiteit. Zekerheid wordt gedefinieerd als "... de mate waarin een individu zeker is dat zijn of haar mening over iets juist is..." (vertaald van Krosnick et al., 1993, p. 1132). Extremiteit is "... de mate waarin iemands attitude afwijkt van het middelpunt van de positieve versus negatieve dimensie..." (vertaald van Krosnick et al., 1993, p. 1132), in andere woorden de afwijking van een neutrale mening. Embeddedness echter wordt geoperationaliseerd door items over het persoonlijk belang bij, de ego-betrokkenheid bij en het kennisniveau over een bepaald onderwerp (Pomerantz et al., 1995). Persoonlijk belang wordt dan gedefinieerd als "... de mate waarin een individu sterk geeft om en zich persoonlijk bezig houdt met een attitude" (vertaald van Krosnick et al., 1993, p. 1132). Ego-betrokkenheid heeft te maken met hoe sterk iemands attitude gerelateerd is aan de wijze waarop hij zichzelf ziet en het waardesysteem dat hij hanteert (Pomerantz et al., 1995). Tenslotte zien Krosnick et al. (1993, p. 1133) het kennisniveau als "... de hoeveelheid informatie over een bepaald onderwerp die in iemands geheugen verbonden is met een attitude...". Het begrijpen van deze nogal abstracte materie zal waarschijnlijk niet eenvoudig zijn. Daarom zullen we nu verdergaan met een bespreking van enkele gevolgen van attitude sterkte, op basis waarvan we een conceptueel raamwerk van het effect van attitude sterkte op het gebruik van marktonderzoekinformatie in het algemeen en klanttevredenheidinformatie in het bijzonder kunnen ontwikkelen.

3. ATTITUDE STERKTE: GEVOLGEN VOOR DENKEN EN DOEN

Zoals hierboven vermeld heeft onderzoek aangetoond dat sterkere attitudes bijdragen aan het vertonen van gedrag. Fishbein en Ajzen (1975) ontwikkelden de attitude-gedrag relatie verder middels hun befaamde 'theory of reasoned action'. Volgens deze theorie wordt werkelijk gedrag bepaald door een intentie om dat gedrag te vertonen. Deze intentie wordt op haar beurt weer bepaald door iemands attitude ten opzichte van dat gedrag en bepaalde subjectieve normen opgelegd door personen in de omgeving. Ook wij zullen de gedachtegang van de 'theory of reasoned action' volgen door te stellen dat sterkere attitudes ten opzichte van klanttevredenheid en het belang hiervan intenties zullen oproepen om zich conform deze attitude te gedragen oftewel om klant-

tevredenheid gerelateerde marktonderzoekinformatie te gaan gebruiken bij het nemen van beslissingen. Deze intenties zullen dan fungeren als voorspellers van daadwerkelijk gebruik.

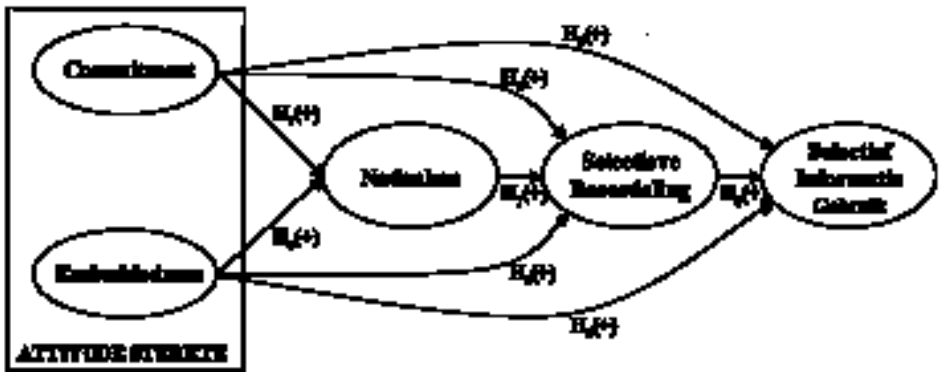
Volgens Pomerantz et al. (1995) moet er voor een beter begrip en de bestudering van attitude sterkte onderscheid worden gemaakt tussen zogenaamde selectieve cognitieve processen en uitkomsten die door weerstand gekenmerkt worden. Krosnick en Petty (1995, p. 3) stellen dat "... attitudes de verwerking en beoordeling van informatie kunnen beïnvloeden". Attitude sterkte kan iemands gedrag dan ook zowel direct als indirect beïnvloeden via selectieve cognitieve denkprocessen en het oproepen van weerstand ten opzichte van niet-attitude conforme standpunten (Eagly and Chaiken, 1995). Voorbeelden van cognitieve processen kunnen zijn selectieve blootstelling aan iets, selectieve aandacht voor iets, selectief nadenken over iets, selectieve perceptie en selectief geheugen (Eagly and Chaiken, 1993). Verder gebruiken Pomerantz et al. (1995) bepaalde maatstaven voor selectieve beoordeling (selective judgment) en attitude polarisatie om iemands weerstand tegen een niet-attitude conform standpunt te meten.

Concreet gezien gebruiken wij bovenstaande ideeën als volgt: het hoe en waarom om een concept als attitude sterkte als voorspeller van gebruik van klanttevredenheidsinformatie te beschouwen komt voort uit dergelijke selectieve cognitieve processen en weerstand ten opzichte van andere standpunten. Sterkere attitudes/meningen ten opzichte van het belang van klanttevredenheid zullen, naast sterkere intenties om klanttevredenheidsinformatie te gebruiken, waarschijnlijk ook resulteren in uitgebreidere denkprocessen over klanttevredenheid als onderwerp. Door hierover na te denken zal er een bepaalde neiging ontstaan om klanttevredenheid gerelateerde informatie selectief te gaan beoordelen of evalueren op basis van een normatief raamwerk dat men heeft ontwikkeld aangaande dit onderwerp. Met andere woorden: er zal een mening ontstaan over waaruit goede klanttevredenheidsinformatie dient te bestaan. Lee et al. (1987) refereren treffend naar dit fenomeen als 'belief perseverance' waarop ze hun conclusie baseren dat onderzoek dat eerdere overtuigingen bevestigt meer gewaardeerd en eerder gebruikt zal worden door besluitvormers. Op basis van het voorgaande zullen we nu overgaan tot de ontwikkeling van een conceptueel raamwerk.

4. EEN CONCEPTUEEL RAAMWERK

In navolging van Pomerantz et al. (1995) stellen wij dat attitude sterkte bestaat uit twee onderliggende dimensies: commitment en embeddedness. We verwachten dat het effect van attitude sterkte op gebruiker intenties om klanttevredenheidsinformatie selectief te gaan gebruiken zowel direct als ook indirect via tussenliggende processen zal optreden. Onze hypothesen kunnen worden samengevat door middel van het conceptuele raamwerk weergegeven in figuur 1.

Commitment ten aanzien van klanttevredenheid zal waarschijnlijk cognitieve denkprocessen over dit thema in gang zetten. Managers zullen bijvoorbeeld eerder gaan nadenken over de positieve gevolgen van het tevredenstellen van klanten voor hun organisatie. Jaccard et al. (1995) stellen ook dat voor het nemen van belangrijke beslissingen de neiging ontstaat om dieper over een onderwerp na te denken. Verder vonden Pomerantz et al. (1995) empirisch bewijs voor een positief effect van attitude commitment op selectieve cognitieve denkprocessen aangaande een onderwerp. Derhalve formuleren we de volgende hypothese:



Figuur 1. Conceptueel raamwerk van gebruiker attitude sterkte.

H_1 : Er bestaat een positieve relatie tussen commitment ten aanzien van klanttevredenheid en het nadenken over klanttevredenheid.

Er is weinig onderzoek bekend waarin expliciet wordt gekeken naar de relatie tussen attitude sterkte en selectieve beoordeling. Pomerantz et al. (1995) echter concludeerden dat de mate van commitment ten aanzien van een onderwerp een positief effect heeft op selectieve beoordeling van onderwerp gerelateerde informatie. We verwachten derhalve dat gebruiker commitment ten aanzien van klanttevredenheid zal bijdragen aan de selectieve beoordeling van klanttevredenheid gerelateerde informatie. Deze beoordeling vindt dan plaats op basis van een normatief raamwerk dat overeenkomt met de attitude en dat stelt waaraan goede klanttevredenheidsinformatie zoal dient te voldoen.

H_2 : Er bestaat een positieve relatie tussen commitment ten aanzien van klanttevredenheid en selectieve beoordeling van klanttevredenheid gerelateerde informatie.

Vervolgens zullen gebruikers die betrokken zijn bij klanttevredenheid waarschijnlijk sterkere intenties hebben om overeenkomstig deze attitude te handelen. Ze zullen informatie die hun mening ondersteunt dat tevreden klanten belangrijk voor de organisatie zijn eerder gaan gebruiken. Deze positieve attitude-gedrag relatie is reeds door verschillende studies bevestigd (bijv. Jaccard, 1981; Jaccard en Becker, 1985; Pomerantz et al., 1995). Deze laatste auteurs vonden ook dat attitude commitment inderdaad intenties om overeenkomstig deze attitude te handelen versterkt.

H_3 : Er bestaat een positieve relatie tussen commitment ten aanzien van klanttevredenheid en intenties om klanttevredenheid gerelateerde informatie selectief te gebruiken.

Over het effect van attitude embeddedness op cognitieve denkprocessen wordt uitgebreid gerapporteerd in zowel de psychologische als de marketing literatuur. Petty en Cacioppo (1979) toonden reeds aan dat persoonlijke betrokkenheid bij en relevantie van een onderwerp (identiek aan wat Pomerantz et al. (1995) omschrijven als ego-betrokkenheid en persoonlijk belang) iemands motivatie om informatie te verwerken verhogen. Sindsdien zijn ook andere studies tot dezelfde conclusie gekomen (bijv.

Greenwald en Leavitt, 1984; Maheswaran en Meyers-Levy, 1990). Tevens blijkt de mate van kennis over iets ook de verwerking van informatie hierover te vergroten (Maheswaran en Sternthal, 1990). Derhalve stellen we dat:

H₄: Er bestaat een positieve relatie tussen embeddedness van klanttevredenheid en het nadenken over klanttevredenheid.

Net als attitude commitment, verwachten we ook dat embeddedness van invloed zal zijn op de selectieve beoordeling van klanttevredenheid gerelateerde informatie door gebruikers. Gebruikers die een diepere innerlijke verbondenheid met het welzijn van hun klanten voelen zullen waarschijnlijk eerder klanttevredenheidinformatie selectief beoordelen tegen de achtergrond van een normatief raamwerk dat ze hebben voor tevredenheid gerelateerde informatie.

H₅: Er bestaat een positieve relatie tussen embeddedness van klanttevredenheid en selectieve beoordeling van klanttevredenheid gerelateerde informatie.

Experimenten hebben aangetoond dat beschikbare kennis, een onderliggend aspect van embeddedness, van invloed is op de vergaring van informatie en het gebruik hiervan (Lee et al., 1999). In navolging van Pomerantz et al. (1995) stellen we dat embeddedness van klanttevredenheid direct zal resulteren in sterkere intenties om tevredenheid gerelateerde informatie selectief te gaan gebruiken: informatie die overeenkomt met de attitude zal eerder gebruikt gaan worden om beslissingen te nemen dan informatie die het belang van tevreden klanten niet ondersteunt.

H₆: Er bestaat een positieve relatie tussen embeddedness van klanttevredenheid en intenties om klanttevredenheid gerelateerde informatie selectief te gebruiken.

Selectieve cognitieve processen over een onderwerp verhogen sterk de mate van selectieve beoordeling van informatie over dit onderwerp (Pomerantz et al., 1995). We verwachten dan ook dat gebruikers die over klanttevredenheid en het belang hiervan voor de organisatie nadenken ('die weten waar ze het over hebben') klanttevredenheid gerelateerde informatie eerder selectief zullen beoordelen op basis van een attitude-overeenkomstig normatief raamwerk wat goede informatie hieromtrent betreft.

H₇: Er bestaat een positieve relatie tussen het nadenken over klanttevredenheid en selectieve beoordeling van klanttevredenheid gerelateerde informatie.

Tenslotte zullen gebruikers die tevredenheid gerelateerde informatie selectief beoordelen op basis van hun persoonlijke attitude geneigd zijn om overeenkomstig deze attitude te handelen. Mede op basis van Hogarth en Makridakis (1988) verwachten we dan ook dat zij sterkere intenties zullen vertonen om informatie die het belang van tevreden klanten onderstreept daadwerkelijk te gaan gebruiken voor het nemen van beslissingen.

H₈: Er bestaat een positieve relatie tussen selectieve beoordeling van klanttevredenheid gerelateerde informatie en intenties om klanttevredenheid gerelateerde informatie selectief te gebruiken.

In de volgende paragraaf zullen we een empirisch onderzoek bespreken dat is uitgevoerd om dit raamwerk in de praktijk te testen.

5. EEN EMPIRISCH ONDERZOEK

5.1. *Onderzoeksontwerp en gegevensverzameling*

Het onderzoek is uitgevoerd onder gebruikers van marktonderzoekinformatie die werkzaam zijn bij een grote Nederlandse fabrikant van kantoormachines. Een Nederlands dienstverleningsonderzoekscentrum heeft een grootschalig internationaal klanttevredenheidonderzoek voor deze fabrikant uitgevoerd, waarbij vooral gekeken is naar after-sales services zoals de aflevering en installatie van producten, klantenondersteuning, service bezoeken en de afhandeling van klachten. Als voorwaarde werd gesteld dat respondenten marketing besluitvormers moesten zijn en op één of andere wijze met de informatie uit het klanttevredenheidonderzoek in aanraking moesten zijn geweest. Respondenten werden geselecteerd op basis van hun functie en taakomschrijving. Functies die gerapporteerd werden varieerden van product managers tot senior marketing managers.

De vragenlijsten werden persoonlijk uitgedeeld aan respondenten door goed voorbereide studenten die als stagiaire in het bedrijf werkzaam waren. Hierdoor kon een zekere mate van controle worden uitgeoefend op de samenstelling van de steekproef wat de validiteit en uiteindelijke respons betreft. In totaal werden er 174 vragenlijsten uitgedeeld. Uiteindelijk resulteerde dit in een respons van 121 respondenten (70%) die aan de criteria voldeden.

De vragenlijst was opgebouwd uit een aantal delen met daarin vragen die de verschillende constructen oftewel variabelen in ons model meten. Er is gebruikt gemaakt van een negen-punts Likert-schaal variërend van 1 = helemaal mee oneens tot 9 = helemaal mee eens. De vragen voor commitment, embeddedness, nadenken en selectieve beoordeling zijn allen gebaseerd op de studie uitgevoerd door Pomerantz et al. (1995). Sommige items moesten echter enigszins aangepast worden voor het specifieke doel van het huidige onderzoek. Intenties van gebruikers om klanttevredenheid gerelateerde informatie selectief te gebruiken werden gemeten door middel van aangepaste items afkomstig van een schaal van Maltz en Kohli (1996). Ter verduidelijking worden in tabel 1 voor de verschillende constructen voorbeelditems gegeven.

Het is belangrijk bij deze voorbeelditems op te merken dat de term 'onderzoek' alleen voorkomt bij het construct 'selectieve beoordeling' aangezien het hier de eisen betreft waaraan klanttevredenheidonderzoek zou moeten voldoen. Voor onderzoekers is dit dus een sleutelconstruct. Verder zijn de onderliggende aspecten van commitment en embeddedness duidelijk te herkennen. Zo wordt voor commitment de attitude zekerheid verwoord door 'Ik ben ervan overtuigd dat ...' en extremiteit door '... moeten ten alle tijden tevreden zijn ...'. Embeddedness daarentegen heeft te maken met het persoonlijk belang ('Persoonlijk vind ik ... zeer belangrijk'), de ego-betrokkenheid ('... is een zakelijke waarde waarmee ik mezelf associeer') en het kennisniveau ('Ik heb een goed idee van ...').

Tabel 1. Betrouwbaarheden en voorbeelditems van de constructen.

Construct	Betrouw- baarheid (α)	Voorbeelditem
Commitment	0.77	Ik ben ervan overtuigd dat ik een goed idee heb van welke after-sales service elementen onze klanten tevreden stellen. Onze klanten moeten ten alle tijden tevreden zijn over onze after-sales service. Het streven naar tevredenheid over onze after-sales service zou één van de topprioriteiten van dit bedrijf moeten zijn.
Embeddedness	0.86	Het tevreden stellen van klanten via het verlenen van goede after-sales service is een zakelijke waarde waarmee ik mezelf associeer. Persoonlijk vind ik tevreden klanten zeer belangrijk. Ik heb een goed idee van wat er zou moeten gebeuren om te voorkomen dat onze klanten ontevreden raken over onze after-sales service.
Nadenken	0.94	Ik denk regelmatig na over hoe dit bedrijf klanten tevreden kan stellen via het verlenen van waardevolle after-sales services. Ik denk vaak na over de voordelen van tevreden klanten d.m.v. het verlenen van excellente service (bijv. verhoogde loyaliteit, positieve mond-tot-mond reclame, enz.). Als er beslissingen over het welzijn van onze klanten genomen moeten worden neem ik zorgvuldig de tijd om hierover na te denken.
Selectieve beoordeling	0.81	Onderzoek dat goed uitgevoerd is dient gedetailleerde informatie te bevatten over kritische after-sales service aspecten die verbeterd kunnen worden. Een overtuigend after-sales service onderzoek zou klachten van klanten moeten aangeven. Een overtuigend after-sales service onderzoek zou alle elementen van het dienstverleningsproces moeten omvatten.
Selectief informatie gebruik	0.89	Informatie die aangeeft wat gedaan kan worden om de tevredenheid van klanten over onze after-sales service te verhogen zal me helpen om een nieuw service beleid te formuleren. Goede klanttevredenheidsinformatie zal mijn begrip verhogen van hoe klanten de kwaliteit van onze after-sales service percipiëren. Ik zal concrete acties ondernemen op basis van informatie over hoe tevreden onze klanten zijn over onze after-sales services.

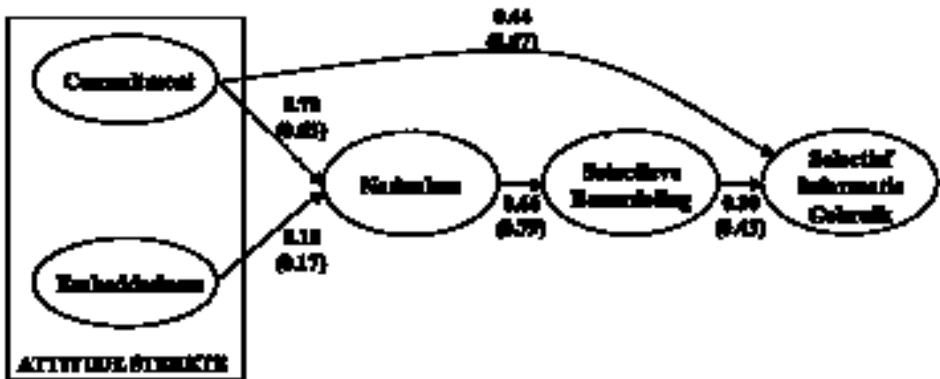
5.2. Validatie van de constructen

Alvorens de hypothesen te toetsen hebben we eerst de validiteit en betrouwbaarheid van onze metingen nader bekeken. Dit is gedaan door middel van een bevestigende factor analyse in LISREL 8. Het voorgestelde meetmodel, bestaande uit vijf constructen en de items die deze constructen meten, bleek in eerste instantie al een goede 'fit' met onze data te hebben. LISREL geeft echter tevens aan of er nog eventuele verbetering in het model mogelijk is. Op basis hiervan werd besloten om één item voor selectief informatie gebruik te verwijderen. De fit van het meetmodel werd hierdoor nog beter, zoals de volgende fit maatstaven aangeven: $\chi^2(109) = 130.31$ ($p = 0.08$), RMSEA = 0.032, GFI = 0.89, AGFI = 0.85, TLI = 0.98, CFI = 0.98.

De convergente validiteit, oftewel de mate waarin items die bij elkaar zouden moeten horen daadwerkelijk overeenstemmen, bleek voldoende hoog te zijn: alle item ladingen waren significant en items laadden op de juiste factor. Ook de discriminante validiteit, oftewel de mate waarin constructen van elkaar verschillen, bleek ook goed te zijn: correlaties tussen de verschillende constructen bleken significant van 1 te verschillen. Verder bleken de meetinstrumenten ook betrouwbaar te zijn: de samengestelde betrouwbaarheidscoëfficiënten varieerden van 0.78 tot 0.92 en de Cronbach's Alpha (zie tabel 1) was voor alle constructen groter dan 0.70 zoals voorgesteld door Nunnally en Bernstein (1994).

5.3. Testen van de hypothesen

Vervolgens hebben we het structurele model getest met behulp van een zogenaamd latente variabele model. Hierbij vormt de variantie-covariantie matrix van de indicatoren/items de input. Het model bleek een goede fit met de data te hebben: $\chi^2(110) = 130.46$ ($p = 0.089$), RMSEA = 0.032, GFI = 0.89, AGFI = 0.85, TLI = 0.98, CFI = 0.98. De bepaling van of een hypothese geaccepteerd of verworpen wordt gebeurt nu aan de hand van de padcoëfficiënten en de bijbehorende t-waarden. De resultaten van deze padanalyse worden weergegeven in figuur 2, waarbij de coëfficiënten zijn gestandaardiseerd en de ongestandaardiseerde coëfficiënten tussen haakjes staan vermeld. Verder zijn alleen de relaties die bij $\alpha = 0.05$ significant zijn opgenomen.



Figuur 2. Eindmodel van gebruiker attitude sterkte.

Zoals uit figuur 2 geconcludeerd kan worden bestaat er een sterk significant positief verband tussen de mate van gebruiker commitment ten aanzien van klanttevredenheid en het nadenken over een onderwerp als klanttevredenheid. Hypothese 1 wordt derhalve empirisch ondersteund door onze studie. Hypothese 2 blijkt echter niet ondersteund te worden: er bestaat geen positieve relatie tussen commitment en de mate van selectieve beoordeling van klanttevredenheid gerelateerde informatie. Wel blijkt er, zoals hypothese 3 stelt, weer een positief effect te bestaan van gebruiker commitment op intenties tot selectief gebruik van informatie die de mening van de gebruiker ten aanzien van klanttevredenheid en het belang hiervan voor de organisatie ondersteunt. Voor de tweede dimensie van attitude sterkte, embeddedness, blijkt slechts één significant verband te bestaan: de mate waarin gebruikers een diepere verbondenheid met het welzijn van klanten voelen is positief van invloed op het nadenken over klanttevredenheid en zaken als de positieve gevolgen hiervan voor de organisatie. Hypothese 4 wordt derhalve bevestigd. Echter, hypothesen 5 en 6 worden niet ondersteund aangezien we geen significant effect van embeddedness op selectieve beoordeling enerzijds en selectief informatie gebruik anderzijds vinden. Ter bevestiging van hypothese 7 vinden we een positief verband tussen het nadenken over klanttevredenheid en de selectieve beoordeling van klanttevredenheid gerelateerde informatie ten opzichte van een normatief raamwerk over waaruit goede informatie dient te bestaan. Tenslotte blijkt selectieve beoordeling van invloed te zijn op intenties om attitude-overeenkomstige klanttevredenheidsinformatie selectief te gaan gebruiken als er beslissingen genomen moeten worden. Hypothese 8 wordt dan ook geaccepteerd.

6. CONCLUSIE

6.1. Interpretatie van de resultaten

De resultaten van deze studie wijzen erop dat de mening van een informatiegebruiker inderdaad mede bepalend is voor het gebruik van marktonderzoekinformatie. Ze verschaffen inzicht in de manier waarop gebruikers beslissingen nemen en hoe informatie hierbij gebruikt wordt. Attitude sterkte en de onderliggende dimensies commitment en embeddedness kunnen resulteren in een subjectieve mening over een onderwerp zoals klanttevredenheid. Via attitude sterkte ten aanzien van klanttevredenheid kan verklaard worden waarom sommige gebruikers klanttevredenheid gerelateerde informatie selectief en eerder zullen gaan gebruiken dan andere. Onze studie geeft ook aan dat zaken als de kwaliteit van de informatie, de wijze waarop deze gepresenteerd wordt, eigenschappen van de onderzoeker, etc. slechts één kant van de medaille betreffen. Een subjectieve mening van een gebruiker kan van grote invloed zijn, ongeacht de kwaliteit van een onderzoek.

Duidelijk is dat de mate van commitment van een gebruiker, iets wat tot uiting komt in de zekerheid dat zijn of haar mening over klanttevredenheid en het belang hiervan juist is en de neiging om dit standpunt in het extreme te verdedigen, een belangrijke rol speelt. Allereerst heeft het een direct effect op gebruiker intenties om tevredenheid gerelateerde informatie selectief te gaan gebruiken. Gebruikers met een sterk standpunt over het belang van tevreden klanten zijn geneigd om dienovereenkomstig te handelen en informatie die dit standpunt ondersteunt daadwerkelijk te gaan gebruiken. Commitment leidt echter ook tot meer cognitieve denkprocessen over het tevredenstellen van

klanten. Betrokken gebruikers zullen uitgebreider nadenken over zaken die belangrijk zijn voor klanten, hoe de organisatie haar klanten tevreden kan stellen en wat voor positieve gevolgen dat zal hebben. Dit kunnen bijvoorbeeld hogere verkopen, trouwere klanten en positieve mond-tot-mond reclame zijn. Het positieve effect van nadenken op selectieve beoordeling van informatie duidt er vervolgens op dat de denkprocessen zich vertalen naar een normatief raamwerk dat gebruikt wordt om klanttevredenheid gerelateerde informatie te beoordelen. Zo'n raamwerk geeft aan wat goede informatie dient te omvatten. Dit zou bijvoorbeeld kunnen zijn dat een onderzoek informatie dient te verschaffen over waar substantiële verbeteringen in de bedrijfsvoering doorgevoerd zouden moeten worden of dat duidelijk naar voren komt waar klanten veelvuldig over klagen. Gebruikers die informatie selectief beoordelen zullen uiteindelijk eerder van plan zijn om informatie dat past in een dergelijk raamwerk, bijvoorbeeld door aan te geven hoe klanten (nog) tevredener gesteld kunnen worden, te gaan gebruiken tijdens het nemen van beslissingen.

Net als gebruiker commitment zal embeddedness, oftewel iemands diepere verbondenheid met klanttevredenheid, ook sterker resulteren in nadenken over dit onderwerp. Het persoonlijk belang dat iemand heeft bij tevreden klanten, de ego-betrokkenheid bij en het kennisniveau over klanttevredenheid zullen de mate waarin hierover nagedacht wordt significant doen toenemen. Opvallend is dat er geen relatie bestaat tussen embeddedness enerzijds en selectieve beoordeling van tevredenheidsinformatie en intenties tot selectief informatie gebruik anderzijds. Een mogelijke verklaring hiervoor kan gezocht worden in de diepere, meer op het innerlijk gerichte, aard van attitude embeddedness. In tegenstelling tot gebruiker commitment, dat meer als een sterk extern gericht standpunt van een gebruiker gezien kan worden, zal de diepere persoonlijke verbondenheid met een onderwerp als klanttevredenheid blijkbaar niet direct intenties tot gedrag teweegbrengen. Het wordt eerst naar een hoger niveau getild door erover na te denken en deze cognities worden dan gebruikt voor de formulering van een soort normatief raamwerk waarmee informatie geëvalueerd wordt.

6.2. Theoretische implicaties

De huidige studie is een eerste poging om attitude sterkte als concept in een marketing besluitvormingssituatie te bestuderen. Er dient derhalve rekening te worden gehouden met enkele theoretische implicaties en beperkingen. Vervolgstudies naar het gebruik van marktonderzoekinformatie, maar ook onderzoek in andere marketing gerelateerde richtingen (denk bijv. aan marketing innovaties waarbij succes vaak sterk afhankelijk lijkt te zijn van de attitudes van de betrokken personen) zouden daarvan kunnen profiteren. Allereerst is het zo dat eerder onderzoek naar de effecten van attitude sterkte zich tot nu toe meestal bezig hielden met zeer controversiële onderwerpen als de doodstraf, abortus en natuurbehoud. Natuurlijk is klanttevredenheid niet zo'n extreem onderwerp en zal het niet zo'n extreme meningen oproepen. Toekomstig onderzoek zou kunnen proberen om extremere attitudes uit te lokken, bijvoorbeeld door zich te richten op een onderwerp als de verdeling van (marketing) budgetten. Ten tweede is klanttevredenheid met after-sales services iets wat verschillend opgevat zou kunnen worden door gebruikers, omdat het een relatief algemeen begrip is. Vervolgonderzoek naar het verband tussen gebruiker attitude sterkte en het gebruik van tevredenheid gerelateerde marktonderzoekinformatie zou zich kunnen richten op eenduidigere zaken zoals specifieke acties die ondernomen kunnen worden om de tevredenheid van klanten

te vergroten, bijv. het investeren in opleiding van service personeel. Verder is de attitude van gebruikers ten opzichte van een bepaald onderwerp natuurlijk slechts één van de factoren die informatie gebruik zullen beïnvloeden. Toekomstig onderzoek zou het concept attitude sterkte moeten combineren met meer instrumentele factoren zoals de inhoudelijke kwaliteit van een onderzoek, de kwaliteit van de presentatie, bepaalde eigenschappen van de marktonderzoeker, bruikbaarheid van de resultaten, etc. Ook zou er gekeken kunnen worden naar factoren waardoor attitude sterkte zelf beïnvloed zou kunnen worden. Hierbij valt bijvoorbeeld te denken aan de mate van ervaring van informatiegebruikers (Perkins en Rao, 1990). Om inzicht in de generaliseerbaarheid van onze resultaten te krijgen zou het onderzoek tevens nog herhaald kunnen worden bij andere organisaties en zou er gekeken kunnen worden of er nog verschillen tussen respondenten bestaan wat de functionele achtergrond betreft. Tenslotte zou vervolgonderzoek wellicht een andere onderzoeksopzet kunnen gebruiken dan de cross-sectionele vragenlijst methode die hier gehanteerd is. Traditioneel onderzoek naar de effecten van attitude sterkte is veelal experimenteel van karakter, waarbij respondenten moeten reageren op materiaal dat wel en dat niet met een attitude overeenkomt. In dit geval zouden gebruikers geconfronteerd kunnen worden met informatie die enerzijds het belang van tevreden klanten onderstreept en anderzijds tegenargumenten hieromtrent geeft.

6.3. Praktische implicaties

Praktisch gezien hebben de resultaten van dit onderzoek enkele implicaties zowel voor marktonderzoekbureaus die regelmatig klanttevredenheid of ander marketing gerelateerd onderzoek voor opdrachtgevers uitvoeren als voor het management van opdrachtgevendende organisaties. Niet alleen inzicht in de mening van de klant maar zeker ook in die van de uiteindelijke gebruikers van marktonderzoekinformatie blijkt namelijk belangrijk te zijn, dit omdat attitudes mede bepalend zijn voor de acceptatie en het gebruik van onderzoekinformatie. Onderzoekers en managers moeten er rekening mee houden dat de sterkte van een mening over een bepaald onderwerp van invloed is op de manier waarop hierover gedacht wordt en vervolgens weer op de evaluatie en het gebruik van informatie hieromtrent. Zowel commitment en embeddedness blijken een significante rol in de acceptatie en het gebruik van klanttevredenheid gerelateerde informatie te spelen. Dit impliceert dat waar mogelijk de mening van informatie gebruikers beïnvloed zou moeten worden door middel van de dimensies van attitude sterkte. Voor zover nodig dienen de personen die daadwerkelijk iets met tevredenheid gerelateerde informatie zouden moeten gaan doen van het (persoonlijk) belang van tevreden klanten bewust te worden gemaakt. Hun betrokkenheid en kennisniveau over de voordelen van tevreden klanten zou verhoogd moeten worden. Marktonderzoekbureaus zouden dit bijvoorbeeld met het management gezamenlijk kunnen realiseren door het organiseren van informatieve sessies over klanttevredenheid en het belang hiervan in termen van verhoogde loyaliteit, winstgevendheid, etc. Ook zouden er in bijvoorbeeld bedrijfsblaadjes gezamenlijke bijdragen gepubliceerd kunnen worden waarin aandacht besteed wordt aan klanttevredenheid, de resultaten van eerder uitgevoerd tevredenheid onderzoek en het effect van tevredenheid op de 'service-profit chain'. Uiteindelijk zal dit leiden tot verhoogde acceptatie en effectiever gebruik van de uit vaak prijzig klanttevredenheidonderzoek verkregen informatie.

Interessant is verder dat de resultaten ook inzicht bieden in de weerstand die vaak

optreedt tegen grotere organisatorische en structurele veranderingen. Op basis van marktonderzoekresultaten worden veelal aanbevelingen gedaan en veranderingen voorgesteld die niet alleen van invloed zijn op het aanbod van producten of diensten maar die soms ook grote neveneffecten hebben. Strategische veranderingen gaan niet zelden gepaard met aanzienlijke organisatorische aanpassingen. Het besteden van aandacht aan de mening en attitudes van de betrokken personen door het onderzoeksbureau en het management tezamen kan hierbij nuttig zijn. Zo is het denkbaar dat de uitkomst van een tevredenheidsonderzoek is dat een bepaalde organisatie te productgericht opereert en dat er te weinig naar de klant wordt gekeken. Het roer zou dus eigenlijk flink omgegooid moeten worden. Wat dan volgens onze resultaten gedaan zou kunnen worden is de algemene mening ten aanzien van tevreden klanten te verbeteren om op die manier de weerstand tegen noodzakelijke maatregelen, zoals bijvoorbeeld onregelmatigere werktijden, te minimaliseren. Gezamenlijke 'campagnes' zouden gevoerd kunnen worden om het belang van klanttevredenheid voor het bedrijfsresultaat en dus voor medewerkers persoonlijk te onderstrepen. De kennis zou verhoogd moeten worden om op die manier een klantgerichtere oriëntatie te creëren die dieper in de organisatie is geworteld. Een hogere mate van commitment en embeddedness van klanttevredenheid zal er tevens voor zorgen dat informatie gebruikers hierover gaan nadenken en een selectief oordeel gaan vormen over wat goede informatie nu eigenlijk zou moeten omvatten. Voor marktonderzoeksbureaus is het derhalve zaak inzicht in deze verwachtingen te krijgen en er voor te zorgen dat de geleverde informatie hier daadwerkelijk aan voldoet.

LITERATUUR

- Abelson, R.P., 1988. Conviction. *American Psychologists* 43, 267-275.
- Anderson, E.W., 1996. Customer satisfaction and price tolerance. *Marketing Letters* 7 (3), 19-30.
- Birgelen, M. van, K. de Ruyter en M. Wetzels, 1999. 'Wat bepaalt de effectiviteit van marktonderzoek informatie? Een gebruikersonderzoek'. In A.E. Bronner, P. Dekker, A.J. Olivier, W.F. van Raaij, M. Wedel en B. Wierenga (eds.), *Recente Ontwikkelingen in het Marktonderzoek*, Jaarboek 1999 van de Nederlandse Vereniging voor Marktonderzoek en Informatiemanagement (pp. 9-23). Haarlem: Uitgeverij de Vrieseborch.
- Chaiken, S., E.M. Pomerantz en R. Giner-Sorolla, 1995. 'Structural consistency and attitude strength'. In: R.E. Petty en J.A. Krosnick (eds.), *Attitude Strength: Antecedents and Consequences* (pp. 387-412). Mahwah, NJ: Lawrence Erlbaum Associates.
- Eagly, A.H. en S. Chaiken, 1993. *The Psychology of Attitudes*. Fort Worth, TX: Harcourt Brace Jovanovich College Publishers.
- Eagly, A.H. en S. Chaiken, 1995. 'Attitude strength, attitude structure, and resistance to change'. In: R.E. Petty en J.A. Krosnick (eds.), *Attitude Strength: Antecedents and Consequences* (pp. 413-432). Mahwah, NJ: Lawrence Erlbaum Associates.
- Fishbein, M. en I. Ajzen, I., 1975. *Belief, Attitudes, Intention, and Behavior: An Introduction to Theory and Research*. MA: Addison-Wesley.
- Greenwald, A.G. en C. Leavitt, 1984. Audience involvement in advertising four levels. *Journal of Consumer Research* 11, 581-592.
- Hogarth, R.M. en S. Makridakis, 1988. 'Forecasting and planning: An evolution'. In: P.E. Earl (ed.), *Behavioural Economics*, Vol. 1, Schools of Thought in Economics Series, No. 6 (pp. 193-216). Aldershot, UK: Elgar/Brookfield, VT: Gower.
- Jaccard, J., 1981. Attitudes and behavior: Implications of attitudes toward behavioral alternatives. *Journal of Experimental Social Psychology* 17, 286-307.

- Jaccard, J. en M.A. Becker, 1985. Attitudes and behavior: An information integration perspective. *Journal of Experimental Social Psychology* 21, 440-465.
- Jaccard, J., C. Radecki, T. Wilson en P. Dittus, 1995. 'Methods for identifying consequential beliefs: Implications for understanding attitude strength'. In: R.E. Petty en J.A. Krosnick (eds.), *Attitude Strength: Antecedents and Consequences* (pp. 337-359). Mahwah, NJ: Lawrence Erlbaum Associates.
- Krosnick, J.A., D.S. Boninger, Y.C. Chuang, M.K. Berent en C.G. Carnot, 1993. Attitude strength: One construct or many related constructs? *Journal of Personality and Social Psychology* 65 (6), 1132-1151.
- Krosnick, J.A. en R.E. Petty, 1995. 'Attitude strength: An overview'. In: R.E. Petty en J.A. Krosnick (eds.), *Attitude Strength: Antecedents and Consequences* (pp. 1-24). Mahwah, NJ: Lawrence Erlbaum Associates.
- Lastovika, J.L. en D.M. Gardner, 1979. 'Components of involvement'. In: J.C. Maloney en B. Silverman (eds.), *Attitude Research Plays for High Stakes* (pp. 53-73). Chicago: American Marketing Association.
- Lee, H., F. Acito en R.L. Day, 1987. Evaluation and use of marketing research by decision makers: A behavioral simulation. *Journal of Marketing Research* 24 (May), 187-196.
- Lee, H., P.M. Herr, F.R. Kardes en C. Kim, 1999. Motivated search: Effects of choice accountability, issue involvement and prior knowledge on information acquisition and use. *Journal of Business Research* 45, 75-88.
- Maheswaran, D. en J. Meyers-Levy, 1990. The influence of message framing and issue involvement. *Journal of Marketing Research* 27, 361-367.
- Maheswaran, D. en B. Sternthal, 1990. The effects of knowledge, motivation, and type of message on ad processing and product judgments. *Journal of Consumer Research* 17, 66-73.
- Maltz, E. en A.K. Kohli, 1996. Market intelligence dissemination across functional boundaries. *Journal of Marketing Research* 33 (February), 47-61.
- Mulder, P., 1999. 'Er verandert geen moer!': Frustraties met continu-klanttevredenheidsonderzoek. *Tijdschrift voor Marketing* (April), 20-24.
- Nunnally, J.C. en I.H. Bernstein, 1994. *Psychometric Theory*. New York: McGraw-Hill, Inc.
- Olson, J.M. en M.P. Zanna, 1993. Attitudes and attitude change. *Annual Review of Psychology* 44, 117-154.
- Perkins, W.S. en R.C. Rao, 1990. The role of experience in information use and decision making by marketing managers. *Journal of Marketing Research* 27 (February), 1-10.
- Petty, R.E. en J.T. Cacioppo, 1979. Issue involvement can increase or decrease persuasion by enhancing message-relevant cognitive response. *Journal of Personality and Social Psychology* 37, 1915-1926.
- Pomerantz, E.M., S. Chaiken en R.S. Tordesillas, 1995. Attitude strength and resistance processes. *Journal of Personality and Social Psychology* 69 (3), 408-419.
- Raden, D., 1985. Strength-related attitude dimensions. *Social Psychology Quarterly* 48 (4), 312-330.
- Rust, R.T., A.J. Zahorik en T.L. Keiningham, 1995. Return on quality (ROQ): Making service quality financially accountable. *Journal of Marketing* 59 (April), 58-70.
- Wetzels, M., K. de Ruyter en M. van Birgelen, 1998. Marketing service relationships: The role of commitment. *Journal of Business and Industrial Marketing* 13 (4/5), 406-423.

14. *On-line kwalitatief marktonderzoek* *'Interviewing the world at a finger tip'*

N. SCHOLL, S. MULDER en R. DRENT

SAMENVATTING

Het Internet heeft ingrijpende gevolgen voor marktonderzoek, voornamelijk met name kwantitatief onderzoek. Voor kwalitatief onderzoek liggen toepassingen minder voor de hand omdat het veel geroemde exploratieve, inzicht biedende karakter juist in een face-to-face situatie tot zijn recht lijkt te komen. In dit artikel wordt een experimenteel onderzoek beschreven naar de validiteit en betrouwbaarheid van internationaal kwalitatief onderzoek via het Internet. Kunnen single interviews en groepsdiscussies via het net 'kwalitatief' genoemd worden? Zo ja, wat zijn dan de verschillen en overeenkomsten in vergelijking met face-to-face onderzoek? De resultaten van face-to-face groepen en singles in Singapore, Zweden en Groot-Brittannië worden vergeleken met de uitkomsten van on-line singles en groepen, uitgevoerd via het Internet, vanuit Nederland. Dit artikel beschrijft de mogelijkheden en onmogelijkheden van on-line kwalitatief onderzoek en gaat in op de toepassingsgebieden waarin dit type onderzoek het meest tot zijn recht komt.

1. ONDERZOEKSDOEL EN METHODE

1.1. Inleiding

Het lijkt erop dat kwantitatief onderzoek via het Internet op dit moment een hogere vlucht heeft genomen dan kwalitatief onderzoek. Dat kan worden verklaard vanuit de onderzoeker en vanuit het instrument. Kwantitatieve onderzoekers zijn door de aard van hun vak geïntrigeerd door instrumenten en methodes en zijn daardoor nieuwsgieriger naar het Internet als onderzoeksinstrument. In het verlengde hiervan ligt de tweede verklaring voor het beperkte gebruik van het Internet door kwalitatieve onderzoekers. Waar de kwantitatieve onderzoeker direct een evident voordeel ziet in de interactiviteit die hij in zijn vragenlijsten kan inbouwen, ervaart de kwalitatieve onderzoeker allereerst een evident nadeel in het verlies aan face-to-face contact met zijn respondent.

Toch is het de moeite waard om ook voor kwalitatief onderzoek vast te stellen of het toegevoegde waarde kan bieden. Het gebruik van het World Wide Web in internationaal onderzoek kan zonder meer snelle resultaten bieden, mogelijk tegen lagere kosten, en daarmee inspelen op de wensen van opdrachtgevers op dit vlak. Voor een internationaal georiënteerd onderzoeksbureau als Research International lag het voor de hand om al in een vroegstadium te zoeken naar de mogelijkheden om het Internet in te zetten bij Internationaal kwalitatief onderzoek.

Levert on-line onderzoek resultaten op die vergelijkbaar zijn met het gangbare face-to-face onderzoek voor wat betreft diepgang, validiteit en betrouwbaarheid? Dit artikel probeert een antwoord te geven op deze vraag, waarbij vooral gekeken wordt naar internationaal onderzoek, omdat hier het voordeel van het Internet, reikwijdte (geen fysieke geografische beperkingen) en snelheid (zowel in dataverzameling als analyse), het grootst zijn.

1.2. Onderzoeksdoelstellingen

De doelstellingen van dit onderzoek zijn om:

- de resultaten van on-line groepsdiscussies en face-to-face groepsdiscussies met elkaar te vergelijken om de sterkten en zwakten van beide methoden vast te kunnen stellen.
- Te bepalen of on-line kwalitatief onderzoek voldoende diepgaand inzicht biedt en in welke specifieke situaties en voor welke onderzoeksdoelstellingen on-line onderzoek het best kan worden ingezet.

Het onderzoek zal ons ook helpen om beter te kunnen begrijpen wat er precies kwalitatief is in het kwalitatieve interviewproces, met name de specifieke rol van visuele en auditieve stimuli in het kwalitatieve interview.

1.3. Methode

Voor dit onderzoek is een kwalitatieve product concept test gedaan. Philips Domestic Appliance Products (DAP) in Groningen heeft ons voor dit onderzoek een concept ter beschikking gesteld: een nieuw type digitale thermometer. Deze infra-rood slaap-thermometer gebruikt dezelfde technologie als de Braun infra-rood oorthermometer maar gebruikt de slaap als locatie van temperatuurmeting in plaats van het oor.

De doelgroep voor dit product bestaat uit ouders van jonge kinderen, waarbij zowel mannen als vrouwen een rol kunnen spelen in het aanschafproces. Als voertaal voor het onderzoek is het Engels gekozen. De keuze voor deze taal, op dit moment de meest wijd verbreide tweede taal, ligt voor de hand, maar in kwalitatief onderzoek is het niet gebruikelijk dat respondenten worden geïnterviewd in een andere taal dan hun moedertaal. De selectie van landen waarin het onderzoek zou moeten plaatsvinden werd hier dan ook sterk door beïnvloed. Daarnaast moest in de te kiezen landen ook het Internet al redelijk zijn geopenetreerd om de selectie voldoende kans van slagen te bieden. De volgende landen werden zorgvuldig geselecteerd om inzicht te bieden in de verschillende aspecten van internationaal on-line marktonderzoek: Groot-Brittannië, Zweden en Singapore:

- Groot-Brittannië werd gekozen als land waarin de respondenten hun moedertaal konden gebruiken
- Zweden als land waarin een groot deel van de bevolking het Engels als tweede taal beheerst, met een cultuur die op veel punten vergelijkbaar is met de Engelse cultuur.
- Singapore, waarin eveneens het Engels goed wordt beheerst, maar waarin de cultuur van oorsprong heel anders is dan die in Groot-Brittannië.

In Figuur 1 staan de verschillende typen onderzoek weergegeven die in elk van de drie landen zijn gehouden.

	<i>Face-to-face</i>	<i>On-line, vanuit Nederland</i>	<i>On-line vanuit locaal land</i>
<i>Groot Britannië</i>	2 groepsdiscussies	1 groep, 8 singles	
<i>Zweden</i>	2 groepsdiscussies	1 groep, 6 singles	
<i>Singapore</i>	2 groepsdiscussies	1 groep	1 groep

Figuur 1. Steekproefindeling.

Op deze manier konden de volgende vergelijkingen worden getrokken:

- on-line* groepsdiscussies vs. *face-to-face* groepsdiscussies;
- on-line groepsdiscussies* vs. *on-line singles*;
- Westerse* vs *Aziatische* cultuur;
- Over de grenzen* (Nederland – Singapore) vs. *binnen hetzelfde land* (Singapore – Singapore);
- on-line moedertaal* (Groot Britannië) vs. *on-line niet moedertaal* (Zweden).

De face-to-face groepsdiscussies en on-line groepsdiscussies duurden twee uur. De on-line groepsdiscussies met de consumenten in elk van de drie landen zijn uitgevoerd vanuit het kantoor van Research International in Rotterdam en werden gehouden door Nederlandse, Engelssprekende interviewers.

Alle gespreksleiders baseerden hun discussie op dezelfde checklist die de volgende onderwerpen bevatte:

- attitudes en gedrag ten aanzien van koorts;
- gebruik en kennis van verschillende typen thermometers;
- sterkten en zwakten van deze thermometers en kenmerken van de ideale thermometer (evaluatie van attributenlijst);
- evaluatie van het thermometerconcept (eerst beschrijving, dan visual);
- image van fabrikanten van huishoudelijke apparaten (gebruik van personaliserings-technieken);
- communicatie, verwacht prijsniveau, koopintenties.

Om een goede vergelijking te kunnen maken tussen de face-to-face en on-line condities, zijn de steekproeven gematched met betrekking tot geslacht, maar ook met kenmerken zoals Engels spreken en het gebruik van internet.

1.4. On-line Interviewen: het gebruik van het chat programma ICQ

Aangezien RI ten tijde van het onderzoek niet beschikte over een eigen chatprogramma en panel, is voor dit onderzoek gebruik gemaakt van het chatprogramma ICQ (voor meer informatie: www.icq.com). Er is voor dit programma gekozen om verschillende redenen:

- Het programma wordt op dit moment door meer dan 30 miljoen mensen gebruikt over de hele wereld, waardoor de *mogelijkheden om te recruter*en groot zijn. Voor dit onderzoek zijn de respondenten on-line geselecteerd, door diegenen die aan het profiel voldeden ('white pages') een ICQ-message te sturen. Dit kostte erg veel tijd en moeite, onder andere vanwege non-respons en het ontbreken van druk om te verschijnen op de afgesproken tijd. Onze verwachting is dat de selectie van respondenten

voor on-line onderzoek in de toekomst meer en meer plaats zal vinden via panels. Naast het feit dat het steeds moeilijker zal worden consumenten te benaderen voor onderzoek biedt een panel een beter zicht op wie er aan het onderzoek deelneemt. Respondenten in dergelijke panels beschikken idealiter over een bepaalde minimale typevaardigheid, zijn getraind in een aantal basale regels met betrekking tot on-line interviews (bijv. gebruik van tekens waarmee emoties kunnen worden aangeduid, de zogenaamde ‘emoticons’) en hebben een korte standaard beschrijving van zichzelf ingevuld die de interviewers voorafgaande aan het interview kunnen doornemen. Op die manier kan ervoor gezorgd worden dat de respondent ‘willing and able’ is om een waardevolle bijdrage te kunnen geven aan het on-line onderzoek, waardoor de kwaliteit van de interviews aanzienlijk beter is en constanter dan in ad-hoc recrutering via het internet.

- de *mogelijkheid tot het sturen van files*, e-mails en site adressen (URLs) maakt het voor onderzoekers mogelijk om visuele stimuli te versturen, een belangrijke voorwaarde voor de meeste concept testen;

Inmiddels zijn de mogelijkheden voor on-line kwalitatief onderzoek verbeterd. Het verdient naar onze mening op dit moment de voorkeur om op de website van het bureau een chatroom aan te bieden met dezelfde functionaliteiten als ICQ, of gebruik te maken van een gespecialiseerde aanbieder op dit vlak. Op die manier hoeft de respondent niet over een chatprogramma te beschikken of een programma te downloaden om toch mee te kunnen doen aan on-line onderzoek.

1.5. Structuur van dit artikel

Het artikel bestaat uit drie delen. Eerst worden onze ervaringen besproken met het doen van on-line kwalitatief onderzoek in de landen vanuit ons Rotterdamse kantoor. Daarna zullen we de verschillen in uitkomsten beschrijven tussen de on-line en face-to-face resultaten. Tenslotte zullen we een framework presenteren van optimale condities voor on-line kwalitatief onderzoek en suggesties doen van onderwerpen die zich goed lenen voor een on-line benadering.

2. DE VOORBEREIDING VAN ON-LINE GROEPSDISCUSSIES EN INTERVIEWS

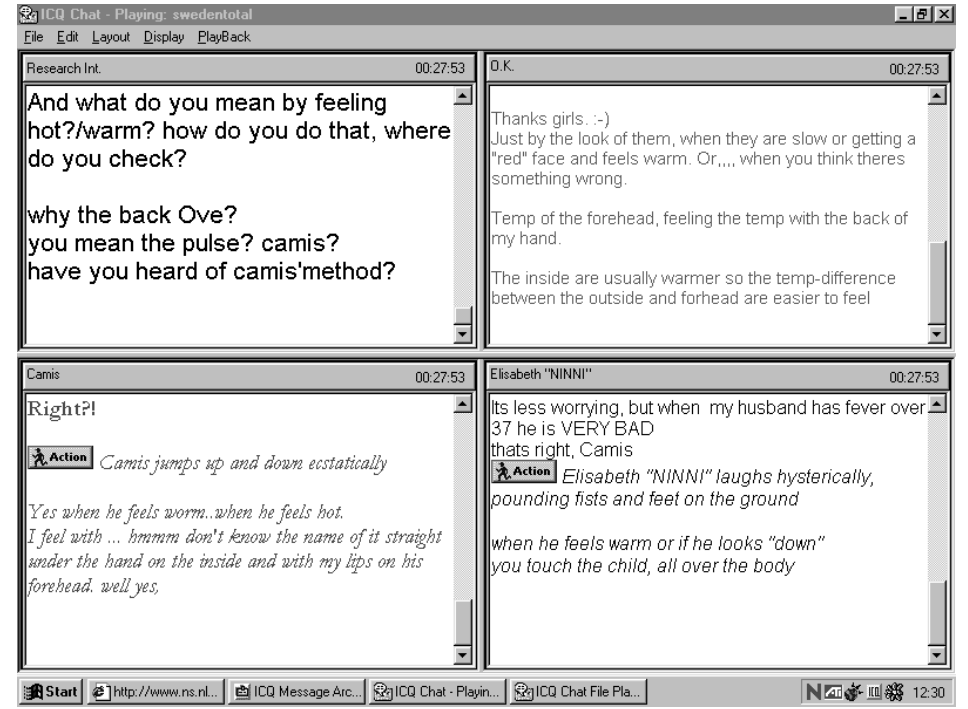
2.1. Het opstarten van de on-line groepsdiscussie

Idealiter worden de respondenten die deelnemen aan een on-line focusgroep ontvangen in een aparte ‘ontvangstruimte’, waarin een systeembeheerder checkt of de juiste respondenten aanwezig zijn en vraagt om een korte persoonsbeschrijving die in het begin van de discussie naar alle respondenten kan worden gestuurd. Als iedereen aanwezig is kan de beheerder de respondenten naar een ander scherm leiden waar de discussieleider hen ‘opvangt’ en de discussie start. Dit voorkomt dat de respondenten één voor één in de chatruimte arriveren en met elkaar beginnen te chatten, waardoor de start van de discussie behoorlijk lang kan duren en rommelig kan verlopen.

2.2. Verschillende opties in de opbouw van het beeldscherm

ICQ biedt verschillende beeldscherminstellingen die het discussieproces direct beïnvloeden. De belangrijkste instellingen zijn:

- één hoofdveld waarin reacties van groepsleden *zinsgewijs* op het beeld verschijnen en zich onder elkaar tot een lijst vormen (zie figuur 2)
- of gesplitste velden waarbij de reacties van respondenten in elk veld real-time *letter voor letter* kunnen worden gevolgd (zie figuur 2).



Figuur 2. Display van een On-line Groep.

Bovenstaande configuratie waarin het beeld in vier tot zes subvelden is gesplitst (afhankelijk van het aantal respondenten) geeft het best het gevoel van een interacterende groep, al kost het enige oefening om alle reacties goed bij te kunnen houden.

3. HET UITVOEREN VAN ON-LINE ONDERZOEK: KENMERKEN VAN HET INTERACTIEPROCES

3.1. Interactieproces tijdens de discussie

Het proces van on-line interviewen verschilt sterk van het face-to-face interviewen. Er zijn verschillen in sfeer, in het sturen van de groep, het ritme van de interactie en het

genereren van dieper inzicht. In dit artikel zullen we ons beperken tot de belangrijkste elementen die de validiteit van de uitkomsten bepalen.

Alle drie de gespreksleiders die de on-line groepsdiscussies leidden klaagden erover dat zij moeite hadden om de ontwikkelingen in de discussie te volgen. Het veld van iedere respondent op het beeldscherm moet snel en continu worden gescand om te zien of er reacties komen en welke reacties er komen, op welke vragen. Om dat te kunnen doen moet de discussieleider gedurende de hele discussie zeer geconcentreerd zijn. Het lezen van reacties kost duidelijk meer moeite dan het horen van reacties, waar het stemgeluid en de richting waaruit het geluid komt duidelijk maakt wie iets zegt. Hoewel de respondenten geneigd zijn om primair op de discussieleider te reageren (er zijn duidelijk minder interacties tussen groepsleden dan in face-to-face groepsdiscussies), heeft de interviewer het gevoel dat de discussie de natuurlijke 'flow' en coherentie mist van de face-to-face groep. Hiervoor zijn verschillende redenen aan te geven die in de volgende paragrafen worden toegelicht.

3.2. Het synchroniseren van vragen en antwoorden

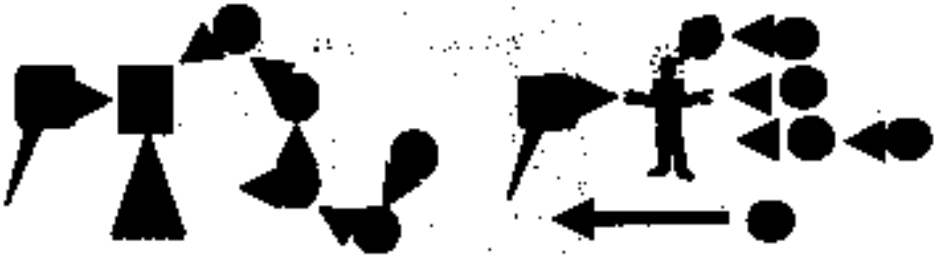
In een face-to-face groep kan de discussieleider goed inschatten wat er gebeurt in de stilte tussen het stellen van een vraag en het eerste antwoord. Zijn er respondenten afgeleid, is de vraag duidelijk over gekomen, komt er wel of geen reactie? Als de discussieleider merkt dat de respondenten geen antwoord weten te geven kan hij zijn vraag verhelderen of nadenken over een andere invalshoek om een onderwerp aan te snijden. Komt er een reactie dan gunt de groep de respondent die begint gewoonlijk voldoende tijd om zijn verhaal duidelijk te maken, alvorens zelf een reactie te geven op de vraag van de interviewer, of de respondent die als eerste heeft gereageerd.

In on-line groepsdiscussies is dit regulerende mechanisme, dat vooral gestuurd wordt door conventies met betrekking tot het zoeken of ontwijken van oogcontact, afwezig. Als een vraag niet aan een specifieke respondent wordt gesteld, maar aan de groep als geheel, dan beginnen doorgaans ook alle respondenten tegelijkertijd te antwoorden. Wanneer een respondent niet snel reageert kan dit twee dingen betekenen: de respondent heeft niets te zeggen of, het antwoord kost meer tijd. Dat laatste heeft belangrijke consequenties, want respondenten willen in het algemeen vermijden dat er een beeldschermstilte ontstaat die het interactieproces in de war gooit en de moderator voor een probleem stelt. Zij zijn dan ook geneigd om snelle en korte antwoorden te geven. Een lang antwoord betekent ook dat andere respondenten moeten wachten tot de respondent die de tijd neemt, klaar is. Ook dit werkt de neiging in de hand om meer antwoorden, maar ook kortere antwoorden te geven dan in een face-to-face situatie. Nadat de discussieleider de antwoorden heeft gelezen en geïnterpreteerd, kan hij een vraag herformuleren of doorvragen, maar dit hele proces duurt aanzienlijk langer dan in een face-to-face situatie, zowel voor singles als voor groepsdiscussies. Ook de interviewer is dus geneigd niet al teveel door te vragen omdat hij daardoor tijd te kort komt om alle onderwerpen aan bod te laten komen.

Een ander probleem in dit verband is dat respondenten die langzaam typen of afgeleid zijn, antwoorden kunnen geven op voorgaande vragen, waar de transparantie van de discussie onder te lijden heeft. De verschillen in interactieprocessen wordt in grafiek 1 nog een keer grafisch weergegeven:

Face-to-face proces

On-line proces



Grafiek 1. Verschillen in synchronisatie tussen on-line en face-to-face groepen.

Het in vier à vijf deelschermen gesplitste beeldscherm geeft de gespreksleider dan ook meer controle dan één groot scherm met alle reacties onder elkaar. Als een respondent zijn antwoord voor de helft getypt heeft kan de gespreksleider de rest van de zin meestal wel gissen en al beginnen zijn eigen reactie daarop te typen. Dit verhoogt de snelheid en geeft meer gelegenheid om door te vragen.

3.3. Grootte van de groep

Respondenten in een kleine groep zijn veel actiever in het geven van antwoorden en hebben bovendien meer aandacht voor elkaars antwoorden, waardoor de interactie tussen respondenten toeneemt in vergelijking met een grote on-line groep. Bovendien is de discussie voor de gespreksleider in een kleine groep gemakkelijker te volgen, wat de mogelijkheid tot doorvragen vergroot. Op grond van bovenstaande overwegingen heeft het onze voorkeur om eerder een groter aantal kleine on-line groepen te laten plaats vinden, dan een klein aantal grote groepen.

3.4. Dynamiek van de discussie

Van de respondenten wordt een redelijke typevaardigheid vereist om een dynamische discussie te kunnen laten ontstaan. Respondenten die niet goed kunnen typen geven kortere antwoorden en waardoor hun aandeel minder waardevol wordt. In een internationale onderzoekscontext, is er naast typevaardigheid ook nog taalbeheersing van belang. De Zweedse respondenten gaven, zoals kon worden verwacht, minder rijke antwoorden dan de Engelse respondenten. De reacties duurden langer en waren korter. Hoewel zowel de gespreksleiders als Zweedse respondenten enthousiast waren over het feit dat zij op een heel open, ontspannen manier communiceerden met vreemden in een ander land, bereikten de gesprekken niet het niveau van diepgang zoals dat gangbaar is in kwalitatief onderzoek. Vooral bij respondenten die het Engels niet goed machtig waren, is er sprake van trage interviews en een relatief oppervlakkige gedachtenuitwisseling. Omdat de respondenten meer tijd nodig hebben om hun antwoorden te formuleren blijft er minder tijd over om door te vragen naar diepere inzichten, waardoor de discussieleider alleen een globaal idee krijg van wat de respondenten willen zeggen. Aan de andere kant, zijn de meeste interviews met Zweedse respondenten van een verras-

send hoog niveau en doen deze zeker niet onder voor de interviews met de Engelse respondenten.

Zowel de gespreksleiders als de respondenten voelen een zekere druk om de discussie gaande te houden en de snelheid van de interactie hoog. De resultaten tonen echter aan dat de discussieleider het tempo juist laag moet houden. In vergelijking met een face-to-face situatie hebben de respondenten immers meer tijd nodig om hun gedachten onder woorden te brengen. Wanneer vragen te snel worden gesteld krijgt de respondent de indruk dat van hem ook snelle en korte antwoorden verwacht worden. Als de respondent meer tijd wordt gegund blijken de antwoorden langer, informatiever en daardoor interessanter te worden.

3.5. Het bereiken van diepere niveaus van inzicht

De nonverbale signalen in face-to-face groepsdiscussies worden niet alleen gebruikt om vragen en antwoorden over de groepsleden te 'distribueren', maar ook om:

- individuele respondenten aan te sporen om meer informatie te geven of toelichtingen te geven op wat ze willen zeggen;
- de spanning in de discussie hoog te houden door dwingende blikken te werpen, of tegenovergesteld: uit te stralen dat de discussie goed verloopt en dat de groep goed presteert.

In de on-line groepsdiscussies blijkt het niet mee te vallen om door te vragen naar diepere betekenissen of motivaties. Sommige groepsleden ervaren het doorvragen als muggenzifterij, een poging van de discussieleider om ingewikkeld te maken wat voor de hand lijkt te liggen. Hier speelt mee dat respondenten niet gewend zijn om introspectie te plegen op hun eigen gevoelsleven ten aanzien van producten en merken. De betrokkenheid bij dit soort zaken is doorgaans beperkt, pas tijdens dit soort onderzoek wordt men zich ervan bewust hoeveel er schuil kan gaan achter een merknaam. In dit onderzoek bleek het nogal wat tijd te kosten om de respondenten in een gemoedstoestand te brengen waarin ze bereid waren om de tijd te nemen om hun eigen innerlijke gevoelens te doorgronden. We zijn ervan overtuigd dat het zoeken naar diepere betekenisniveaus in on-line onderzoek niet onmogelijk is, maar dat het gewoon relatief meer tijd kost. Op dit punt heeft face-to-face onderzoek het voordeel dat de discussieleider beter een zekere spanning kan creëren waarin mensen hun vermogen tot zelfreflectie willen etaleren. Daardoor doen ze beter hun best om diepere betekenissen onder woorden te brengen. On-line onderzoek heeft op dit punt het voordeel dat, als respondenten eenmaal gewend zijn geraakt aan de interviewstijl, de respondent uit vrije wil en nieuwsgierigheid mee doet aan zelfreflectie en associëren, zoals blijkt uit de paragraaf 'het gebruik van associatieve technieken'.

Om te voorkomen dat on-line discussies aan dynamiek en diepgang verliezen bevelen we aan om als maximale discussieduur anderhalf uur aan te houden. Een praktische manier voor de discussieleider om de discussie gefocust te houden is om een goede balans te realiseren in het gebruik van aanmoedigende, en disciplinaire emoticons. Een emoticon is een standaard symbool, soms een woord of zin, waarmee snel en efficiënt een bepaald gevoel kan worden uitgedrukt. Ervaren chatters weten deze emoticons heel soepel te gebruiken en kunnen daarmee een goede, warme sfeer in de groep brengen. Een voorbeeld van een emoticon is de smile: :-) ofwel de glimlach die benadrukt dat je ergens blij mee bent of een rotopmerking goed bedoelt.

3.6. Groepscohesie

Respondenten die ervaring hebben met kwalitatief marktonderzoek zijn van mening dat ze in de on-line groep meer groepscohesie en consistentie ervaren dan bij de gewone chats, maar dat een on-line groep op deze punten toch onderdoet voor de face-to-face groep. Een echte groep geeft toch nog meer een soort plechtig gevoel samen bezig te zijn met het oplossen van een probleem. Hoewel er in de on-line groep minder sprake is van een ‘samen-er-tegen-aan-stemming’, heeft de meer speelse atmosfeer in de on-line groep ook een positieve invloed op de groep. Sociale codes die bepalen wie de informele leider van de groep is (zoals kleding, haardracht, zithouding, gebaren, stemgebruik) ontbreken in een on-line groep. Dit leidt tot een zekere uitbalancerings van de hiërarchie in een groep. De on-line groep is daarom bij uitstek een ‘democratische’ groep waarin verlegen respondenten voldoende ruimte krijgen om deel te nemen op basis van gelijkheid. Er is wel sprake van statusverschil in de groep maar deze zijn uitsluitend gerelateerd aan taal en intellect: woordenschat, gemak waarmee men dingen typt en onder woorden brengt, creativiteit en het hebben van een eigen mening over bepaalde onderwerpen.

Voor de discussieleider betekent dit dat hij minder moeite heeft om de groep uit te balanceren. Het kost meer moeite om de leden betrokken en geconcentreerd te houden op het onderwerp, maar hij hoeft niet stille personen aan te moedigen en dominante personen op hun plaats te zetten. In een face-to-face groep kan het zinvol zijn om verlegen mensen over hun sociaalpsychologische drempels te helpen, in een on-line situatie heeft dat geen zin; de sociaalpsychologische drempel bestaat eenvoudig niet, en respondenten kunnen niet worden aangemoedigd om eloquenter te zijn dan ze zijn. Het enige dat hier helpt is respondenten meer tijd te geven om hun antwoord te formuleren.

3.7. Interactie tussen groepsleden

De on-line groepsdiscussie lijkt minder interactie tussen leden te genereren dan de face-to-face groepsdiscussie. Discussies tussen respondenten komen voor, maar zijn minder duidelijk en coherent. Dit is met name een kwestie van overzicht (het is moeilijk te overzien waar je op moet/kunt reageren) en tijd (de discussie voltrekt zich zo snel dat reageren op een ander niet mogelijk is). Alleen als er interactie wordt gestimuleerd door respondenten persoonlijk aan te spreken met vragen als: “Jim, wat vind jij van het antwoord dat Louise op deze vraag gaf?” begint er iets van interactie te ontstaan. Het is dan ook van belang respondenten uit te nodigen om te reageren op elkaars opmerkingen. Dat kan door een set van ‘probing icons’ te ontwikkelen die de interviewer met een simpele click in op het beeldscherm kan laten vallen:

- “nog andere opmerkingen?”
- “zijn jullie het eens met wat er gezegd is?”
- “kun je verduidelijken wat je bedoelt?”

Tijd is overigens in meerdere opzichten een probleem. In de praktijk bleek dat een face-to-face checklist niet voldoet voor on-line groepen, aangezien er in twee uur niet eens genoeg tijd is om iedere vraag één keer te stellen, laat staan doorvragen daaraan te koppelen. Pas als we ongeveer de helft van de vragen achterwege laten is er genoeg tijd om door te vragen.

3.8. Het gebruik van associatieve technieken

Een belangrijke doelstelling van dit onderzoek is om te inzicht te krijgen in de mogelijkheden om creatieve technieken toe te passen. Deze technieken zijn typerend voor kwalitatief onderzoek omdat er geen directe vragen worden gesteld maar via een omweg. Begrijpen respondenten wat voor soort antwoorden we graag willen horen? Zijn zij in staat en bereid om antwoord te geven op vragen die in eerste instantie op zijn minst vreemd over komen?

De resultaten tonen aan dat respondenten verrassend goed meewerken aan associatieve technieken. We vroegen hen om drie internationaal bekende merken van huishoudapparaten te personaliseren: “Hoe zouden deze merken eruit zien als het mensen waren? Wat voor leeftijd, geslacht, beroep, uiterlijk, hobbies zouden ze hebben? Hoewel deze technieken relatief veel tijd kosten zijn de resultaten interessant en waardevol. De meeste respondenten ervoeren geen probleem bij het beantwoorden van deze vragen en hadden er klaarblijkelijk veel plezier in. Ook bleken deze technieken goed te werken in on-line singles, waar dit bij face-to-face singles moeilijk lukt vanwege de schroom om vrij te fantaseren. Ook hier weer een aanwijzing dat respondenten in een on-line interviewsetting zich veel meer op hun gemak voelen en veel minder drempels ervaren dan in een face-to-face onderzoek. Om een voorbeeld te geven van de reacties op de merk personalisering in een interview met een Engelse vrouw:

Discussieleider: “What image do you have of the Braun woman/man?”

Sue: “Blonde, probably tall, confident, tends to dress in suits, has a German car – a Mercedes or BMW – probably the latter. He will probably have a physical hobby, anything in which he can show off his confidence, so sports.

Discussieleider: “Now about Philips, could you give me a picture of Philips?”

Sue: “Well, this person is more likely to be dark-haired, medium height, again confident, but not overwhelming like his Braun colleague ... umm, think he will have a top of the range car that is not ostentatious en shouts money! I think this person is going to be more a family person.”

De resultaten tonen duidelijk aan dat respondenten niet alleen op productniveau kunnen associëren, maar ook op merk en lifestyle niveau, waardoor dieper inzicht in de merk-persoonlijkheid wordt verkregen.

3.9. Het analyseren van de resultaten

Een unieke ervaring met on-line groepsdiscussies is dat je binnen vijf minuten kunt beschikken over een perfect verslag van wat er in de groep is gezegd. In face-to-face groepsdiscussies zou er twijfel kunnen ontstaan of de aantekeningen tot in detail nauwkeurig zijn, in on-line groepen zijn de aantekeningen foutloos.

Uiteraard ontbreken non-verbale en paralinguïstische cues waardoor de tekst meer gaat leven, maar dit wordt voor een deel gecompenseerd doordat respondenten zelf hun gevoelens proberen in hun teksten te verwerken.

De analyse maakte ons op iets attent waarvan wij ons tot noch toe niet sterk bewust waren. In de analyse van een face-to-face gesprek bouwt de interviewer zich een beeld van de persoon die hij tegenover zich heeft. Kleding, gedrag, houding, taalgebruik, deze geven allemaal een beeld van waar de persoon vandaan komt, het milieu, de regio en wat de ambities zijn van die persoon. Dat beeld blijkt een belangrijke rol te spelen in het interpreteren hoe een respondent iets bedoelt en wat de samenhang is van zijn ant-

woorden. Door een beeld te hebben van de persoon kunnen reacties worden geëxtrapoleerd en ge-intrapoleerd; dingen die niet letterlijk zijn gezegd kunnen gevoeglijk worden aangenomen omdat het past bij de persoon als geheel: “Deze respondent is zo’n type dat geneigd is zich kritischer uit te laten dan dat hij eigenlijk bedoelt, hij zal vast niet bedoeld hebben dat...” Dit is in feite een vorm van min of meer verantwoord ‘hineininterpretieren’ gebaseerd op mensenkennis die van ervaren onderzoekers verwacht mag worden.

In on-line analyses ontbreekt dit holistisch mensbeeld van de respondent dat nodig is om een interviewverslag die consistentie en coherentie te geven. Uiteraard is dit soort basisinformatie essentieel om reacties van respondenten goed te kunnen plaatsen. De onderzoekers waren weliswaar in staat om op basis van de notulen een analyse te maken maar ze voelden zich daarbij minder zeker dan bij face to face onderzoek, waar ze antwoorden achteraf nog kunnen koppelen aan gezichten en stemmen. Wanneer er van iedere respondent in on-line onderzoek een soort ‘videojournaal’ beschikbaar is (electronische foto’s of videofragmenten die vanachter de computer voorafgaande aan het interview kunnen worden geraadpleegd) kan het interview en ook de analyse sterk aan diepgang winnen, doordat de interviewer meer een respondent ‘van vlees en bloed’ voor zich kan zien.

4. VERGELIJKING VAN UITKOMSTEN VAN ON-LINE EN FACE-TO-FACE INTERVIEWS

Kwalitatief on-line onderzoek kan sneller en goedkoper uitgevoerd worden dan face-to-face onderzoek. Weinigen zullen het met deze stelling niet eens zijn. De financiële en praktische voordelen zijn evident, maar zijn de resultaten valide en betrouwbaar genoeg om het gebruik van on-line kwalitatief onderzoek te legitimeren? En biedt on-line onderzoek voldoende diepte, vanuit een kwalitatief oogpunt?

Om die vragen te kunnen beantwoorden moeten we ons eerst afvragen hoe we de validiteit van on-line kwalitatieve onderzoek kunnen bepalen. De validiteit en met name betrouwbaarheid van kwalitatief face-to-face onderzoek is vaak ter discussie gesteld door zowel academische als praktijkonderzoekers. Dit geldt uiteraard ook voor on-line onderzoek. Het feit dat kwalitatief onderzoek in de afgelopen decennia alleen maar aan belang heeft gewonnen wijst er echter op dat de meeste marketeers en marktonderzoekers overtuigd zijn van de relevantie en bruikbaarheid van dit type onderzoek. De methode die is gebruikt in het onderhavige onderzoek is een gestandaardiseerde en geaccepteerde benadering die veel wordt gebruikt in internationale kwalitatieve concept tests. Daarom zullen we een beperkte definitie hanteren van validiteit in die zin dat on-line onderzoek als valide wordt beoordeeld wanneer ze de uitkomsten van de face-to-face resultaten weerspiegelen.

Zoals we in het vorige hoofdstuk konden zien, kleven er aan het on-line interviewen een aantal ernstige bezwaren met betrekking tot belangrijke aspecten, zoals synchronisatie van vragen en antwoorden, snelheid en soepelheid van de conversatie, doorvragen en analyse. Als gevolg van deze problemen en de verschillen in processen lag het voor de hand dat er verschil in uitkomsten zou bestaan tussen beide onderzoekscondities.

4.1. Detailniveau

Wanneer we de rapporten vergelijken, dan geldt dit zeker voor de hoeveelheid informatie die verzameld is, met name het detailniveau van de informatie. In de on-line groepen en interviews moest meer dan een derde van de vragen worden overgeslagen om de gesprekken niet teveel uit te laten lopen. De on-line rapporten bevatten dan ook geen informatie over een aantal interessante en relevante onderwerpen.

Daarnaast boden de on-line interviews minder ruimte om interessante gebieden te exploreren die door de respondenten zelf werden aangereikt, en was er weinig tijd om vraagpunten te herformuleren en in te gaan op details. Het rapport van de face-to-face groepen in Singapore presenteert een lijst van productgerelateerde elementen waarop verbetering mogelijk is die typisch het resultaat is van een diepgaande face-to-face discussie: vergroot het display, maak een back-light in het display, maak een kleurindicatie om de ernst van de koorts weer te geven, zorg voor een high-tec uitstraling, verminder de omvang van het apparaat, zorg voor opslag van data in het apparaat.

4.2. Groot Brittannië, Zweden versus Singapore

De conclusies in de face-to-face rapporten van Groot Brittannië en Zweden lijken veel op elkaar, in tegenstelling tot Singapore. De respondenten in de beide Europese landen tonen eenzelfde soort attitude met betrekking tot het gebruik van thermometers en hun indruk en waardering van het nieuwe concept. De houding van de respondenten in de face-to-face groepen in Singapore is duidelijk anders van aard. Hier komen elementen in naar voren die in de Europese groepen niet genoemd werden. Deze attitudeverschillen zijn ook terug te vinden in de on-line groepen, dezelfde elementen die in de Singaporese face-to-face groep worden genoemd, komen ook in het Singaporese on-line rapport naar voren. De infra-roodtechnologie roept bijvoorbeeld bezorgdheid op over veiligheid, er is behoefte aan geruststelling dat er geen negatieve bijwerkingen zijn. In de Europese rapporten wordt hierover in zowel de on-line groepen als de face-to-face groepen niet gerept. Ook opmerkingen over het design, dat volgens Singaporese respondenten aangepast zou moeten worden om de kindvriendelijkheid te onderstrepen door het uiterlijk minder klinisch te maken zijn in de Europese rapporten niet terug te vinden.

4.3. Bottom line dezelfde uitkomsten

Op het niveau van de overall conclusies en aanbevelingen echter, komen de face-to-face rapporten en on-line rapporten overeen. In alle rapporten wordt de nieuwe thermometer positief geëvalueerd. Het product wordt gezien als een kindvriendelijke thermometer die de meetresultaten snel en accuraat laat zien. In alle rapporten verwachten respondenten dat de prijs van het nieuwe product hoog zal zijn en dat bij respondenten die tevreden zijn over hun huidige thermometer weinig animo is om te switchen naar de nieuwe moderne thermometer. Dit is vooral het geval bij die consumenten die nu al een infra-rood oorthermometer hebben. Sommige respondenten die een klassieke kwikthermometer gebruiken zijn erg gecharmeerd van de gebruiksvriendelijkheid van de thermometer (kan gebruikt worden terwijl het kind slaapt) en denken er serieus over om de nieuwe thermometer te kopen.

Hoewel de belangrijkste resultaten van on-line en face-to-face rapporten overeenko-

men, zeggen de on-line onderzoekers dat zij zich veel minder zeker voelen over hun eigen analyse in vergelijking met face-to-face onderzoek.

Om een metafoor te gebruiken voor de verschillen tussen on-line en face-to-face rapporten; de inhoud van een face-to-face rapport kan vergeleken worden met een foto die gemaakt wordt door een professional die een dure camera gebruikt en uitstekend papier om hoge resolutie foto's te maken van een object in zijn studio. De on-line foto wordt ook gemaakt door een professional, maar hij gebruikt een goedkoop merk camera en heeft geen professionele studio ter beschikking. De professionele foto toont een heel duidelijk beeld met veel details en levendige, scherpe kleuren. De andere foto die met de goedkopere camera wordt gemaakt geeft minder details en de kleuren spreken minder. Hoewel de contouren niet erg scherp zijn kan het object goed worden herkend. De goedkopere camera vervormt niet maar de beelden zijn duidelijk minder krachtig en gedetailleerd. Details moeten worden geraden en ingevuld door de kijker, terwijl de professionele foto niets aan de verbeelding over laat.

5. SUGGESTIES VOOR HET GEBRUIK VAN ON-LINE KWALITATIEVE GROEPEN

Eerder in deze paper is al een aantal suggesties gegeven om de kwaliteit van het on-line interviewen te verhogen. In deze paragraaf willen we de optimale condities bespreken om on-line kwalitatief onderzoek tot zijn recht te laten komen en toepassingsgebieden beschrijven waarin dit type onderzoek een duidelijk toegevoegde waarde kan hebben.

Allereerst kan gesteld worden dat een aantal problemen die we in dit onderzoek tegenkwamen te maken hebben met de vroege fase van ontwikkeling van het Internet en de chatsoftware die is gebruikt. Regelmatige lijnstoringen vormen een groot probleem voor de kwaliteit van het interviewen, maar door de snelle ontwikkelingen zal dit probleem zeer waarschijnlijk al snel tot het verleden behoren.

Het lage tempo in het on-line interviewproces lijkt echter een taai probleem te vormen dat ook in de toekomst nog zal bestaan in on-line onderzoek waarin visuele en auditieve communicatiemogelijkheden ontbreken. De grenzen liggen hier immers niet in de techniek maar in de respondent die grenzen kent in de snelheid waarmee hij zijn antwoorden kan formuleren en typen. Deze inherente eigenschap van on-line interviewen impliceert dat on-line ('chat')onderzoek alleen gebruikt kan worden voor concrete, smalle onderzoeksvragen en niet voor fundamentele inventariserende studies met een brede onderzoeksvraag, gericht op het genereren van een dieper begrip van gedrag, attitudes en imago's.

5.1. On-line kwalitatief onderzoek als snelle 'multi-country check'

Tegenover de sterke beperkingen die on-line kwalitatief onderzoek stelt aan de reikwijdte van het onderzoeksgebied en het detailniveau van de resultaten, staat een indrukwekkende snelheid en groot kostenvoordeel, dat vooral tot uitdrukking komt in internationaal onderzoek. On-line onderzoek is, naar onze mening dan ook vooral interessant voor multi-nationale bedrijven die hun producten wereldwijd op de markt zetten. Zij lopen vaak tegen het probleem aan dat kostbaar kwalitatief onderzoek niet in alle landen kan worden uitgevoerd wat tot strubbelingen kan leiden in de besluitvoering welke landen/locale managers mogen worden bediend en welke niet. On-line onderzoek kan

dan worden ingezet als een relatief goedkope aanvullende check in alle landen waarin het bedrijf opereert, op het duurdere face-to-face onderzoek dat in enkele sleutelmarkten wordt uitgevoerd. Het doel van on-line onderzoek is dan om vast te stellen of er locale culturele of marktverschillen die nopen tot een nuancering van de situatie zoals die in key-markten zijn aangetroffen. Dit onderzoek toont aan dat klanten en onderzoekers zich redelijk veilig mogen voelen dat het on-line kenmerk van deze additionele check niet tot substantieel andere resultaten zal leiden, alleen minder verfijnde resultaten.

5.2. On-line onderzoek als het gaat om gevoelige onderwerpen

Een uniek toepassingsgebied voor kwalitatief on-line onderzoek is onderzoek naar onderwerpen waarover respondenten zich niet gemakkelijk open durven te uiten. Door het grote gevoel van veiligheid en anonimiteit in de on-line setting zal men eerder geneigd zijn om meer het achterste van de tong te laten zien dan in een professionele setting in een onderzoeksbureau. In Engeland heeft on-line onderzoek zijn waarde op dit vlak al bewezen in onderzoek onder criminelen naar hun motivaties voor hun misdaden. Ook onderzoek over onderwerpen zoals seksualiteit, kindermishandeling, belastingontduiking, pesten op school, zinloos geweld en burgerlijke ongehoorzaamheid kunnen met deze techniek wellicht beter worden benaderd dan met face-to-face onderzoek.

6. DISCUSSIE

De uitkomsten van het onderzoek zoals beschreven in dit onderzoek leiden tot een intrigerende conclusie: kwalitatieve onderzoekers die aan dit onderzoek meewerkten zeggen dat het gebrek aan non-verbale communicatie hen hinderde in het begrijpen van de reacties van de respondenten, waardoor ze moeite hadden met de analyse en interpretatie van de uitkomsten. De afwezigheid van non-verbale signalen belemmert niet zo zeer het begrijpen van wat er wordt gezegd, maar breder, wat het gezegde betekent in de context van de lifestyle en persoonlijkheid van de respondent. Het is moeilijk om een holistisch beeld op te bouwen van de respondent waarin uitspraken een logisch en coherent geheel vormen. Dit verklaart waarom de onderzoekers zich onzeker voelden.

Dit fundamentele onderzoek op het terrein van on-line kwalitatief onderzoek lijkt echter te suggereren dat de verschillen in de conclusies tussen on-line en face-to-face onderzoekers klein zijn. Wanneer toekomstig onderzoek zou bevestigen dat de uitkomsten van beiden typen onderzoek onder de zelfde condities gelijk zijn, dan zegt ons dit iets over de rol en het belang van non-verbale stimuli in het proces van analyse in kwalitatief onderzoek. Denkbaar is dat de rol van non-verbale stimuli bij het bepalen van de validiteit en zuiverheid van de kwalitatieve analyse wel eens minder groot zou kunnen zijn dan algemeen wordt aangenomen. Mogelijk wordt kwalitatief inzicht veel eerder verkregen door basisregels die weinig te maken hebben met het kunnen zien van de respondent. Basisregels zoals: respondenten in een onderzoeksetting brengen waarin hij zich prettig voelt en bereid is om zijn meningen vrij te kunnen uiten, open vragen stellen, tijd geven om te antwoorden, doorvragen of een antwoord door de interviewer goed begrepen is, etc. etc.

Deze basale regels, gecombineerd met de analytische vaardigheden die ervaren kwalitatieve onderzoekers in de loop der tijd hebben ontwikkeld, lijken hen in staat te stellen om on-line reacties heel goed te kunnen interpreteren en tot valide conclusies te komen.

Niettemin zouden we onderzoekbureaus willen adviseren om voor breed en exploratief onderzoek niet puur en alleen te vertrouwen op on-line onderzoek en evenmin in lifestyle en communicatie-onderzoek waarin kleine details bepalend kunnen zijn in het ontwikkelen van het juiste marketingbeleid. Dit type onderzoeken vormen nog steeds een belangrijk deel van het domein van kwalitatief onderzoek. Wanneer het gebruikt wordt als een additionele bron van goedkope en snelle marktinformatie in internationaal onderzoek dat veel landen en culturen bestrijkt, dan is het nut en de relevantie van kwalitatief onderzoek evident.

15. Marktonderzoek naar hedonistische consumptie en productervaringen

C. GOOSSENS

'Consumption has begun to be seen as involving a steady flow of fantasies, feelings, and fun encompassed by what we call the experiential view'

(Holbrook en Hirschman 1982:132)

SAMENVATTING

Hoe kan men doelgroepen karakteriseren, keuzegedrag analyseren en productgebruik evalueren als de consument zich laat leiden door genoegens, plezier en een zucht naar sensaties? Veel marketeers kampen met dit probleem. Binnen deze context gaat dit artikel over de drijfveren, het keuzegedrag en de productervaringen van hedonistische consumenten.

1. INLEIDING

Hedonistische consumptie is het gebruik van producten en diensten die genoegens opleveren via de zintuigen, fantasieën creëren en emotioneel opwinden (Sheth, Mittal, en Newman 1999). Enkele voorbeelden zijn:

- Sensorisch genoegens: drank en tabaksgebruik; bubbelbad en sauna; sieraden en parfums; muziekinstallaties, CD's, lichteffecten in disco's, etc.
- Esthetisch genoegens: design producten; museumbezoek; een cursus Griekse kunstgeschiedenis, etc.
- Emotionele beleving: films en soaps op TV; pretpark-, concert- of operabezoek; verjaardagsfeest; cadeaus geven/ontvangen, viering van een jubileum, etc.
- Plezier: sport, spel, recreatie, dansen; vakanties, bars, casino's, zakelijke golfpartij, etc.

Kortom, veel consumentengedrag draait om het ervaren/beleven van genoegens. Pine en Gilmore (1998) heten hun lezers dan ook welkom in 'The Experience Economy' waarbij ze vier ervaringswerelden onderscheiden: amusements-, educatieve-, ontsnapings- en esthetische belevenissen. Bij de eerste twee categorieën is er vooral sprake van het opnemen of absorberen van prikkels, terwijl er bij de laatste twee vormen meer sprake is van een onderdompeling in prikkels en omgevingen. Het traditionele hedonisme zoekt het genoegen vooral in sensaties die gekoppeld zijn aan de zintuigen (smaak-,

geur-, tast-, gehoor- en gezichtsvermogen). Denk bijvoorbeeld aan vorstelijke wijn, verrukkelijke parfums, perfecte audio-visuele apparatuur en andere vormen van luxe en comfort. Tegenwoordig zoekt de hedonist zijn genoegens vooral in nieuwe stimuli, die minder voorspelbaar en comfortabel en meer gevaarlijk zijn. Denk bijvoorbeeld aan 'skydiving', 'bungee jumping', 'hanggliding' en andere gevaarlijke sporten. De moderne hedonist beleeft plezier aan diverse emoties die gepaard gaan met de meest uiteenlopende ervaringen, inclusief angstige en nostalgische gebeurtenissen. Het genoeg wordt hierbij niet primair ontleend aan zintuigelijk genot en comfort, maar veel meer aan avontuur en een speciaal soort van zelfbeheersing. In de traditionele opvatting heeft het 'pleasure principle' betrekking tot het opheffen van onplezierige situaties en het herstellen van een verstoord evenwicht. De moderne hedonist is daarentegen continu op zoek naar stimuli die plezierige ervaringen opwekken (Gabriel en Lang, 1995). Maar wat drijft de almaar zoekende consument?

2. SPANNING EN AFWISSELING

Bij hedonistische consumptie geldt in het bijzonder dat mensen gedragssituaties beoordelen op emotionele voor- en nadelen. Men is in feite opzoek naar een optimaal *spanningsniveau*. Zo zullen consumenten die zich vervelen verlangen naar een zekere sensatie, terwijl ze bij gespannen situaties anticiperen op rust. Beide motivationele krachten worden resp. *arousal-seeking* en *arousal avoiding* genoemd. De term arousal (spanning of opwindning) verwijst naar een verhoogde activatie van het centrale zenuwstelsel, wat gepaard gaat met fysiologische aspecten zoals een verhoogde hartslag of bloeddruk. Volgens Zuckerman zijn er twee soorten arousal, te weten: positieve arousal (i.c. sensation seeking) en negatieve arousal (i.c. anxiety). Voor het meten van het begrip 'spanningsbehoefte' ontwikkelde Zuckerman (1994) de zogenoemde '*Sensation Seeking*' schaal.

Uit onderzoek blijkt bijvoorbeeld dat degenen die hoog scoren op deze schaal het onder andere prefereren om naar *live* muziek te luisteren in een opwindende omgeving, terwijl degenen die laag scoren zo'n optreden liever op de televisie volgen. Verder kijken hoge sensatiezoekers vooral televisie om het nieuws te volgen of om het als achtergrond te gebruiken terwijl ze iets anders doen. Uit een experiment is gebleken dat hoge sensatiezoekers ervan houden om veel te zappen terwijl ze twee of meer programma's simultaan volgen. Lage sensatiezoekers gaan daarentegen helemaal op in één enkel programma. Bij foto's, films, televisie en lectuur tonen de hoog scorende sensatiezoekers meer belangstelling voor morbide en seksuele thema's dan degene die laag scoren. De lage sensatiezoekers vinden deze thema's afschuwelijk en vermijden ze. Dagdromen en fantaseren zijn niet sterk gerelateerd aan sensatiezoeken. Als hoge sensatiezoekers dagdromen hebben ze vaak hedonistische fantasieën over onderwerpen waarbij ze zelf actief zijn, zoals sex, reizen, en hardrijden. Lage sensatiezoekers prefereren een voorspelbare, begrijpelijke, rationele en niet emotionele wereld, met weinig verrassingen en waarin alles duidelijk is. De '*Sensation Seeking*' schaal bestaat uit vier sub-schalen, te weten:

- '*Thrill and Adventure Seeking*', deze schaal heeft betrekking op een voorkeur voor gevaarlijke fysieke activiteiten, zoals bergbeklimmen en diepzeeduiken;
- '*Experience Seeking*', deze schaal heeft betrekking op nieuwe, ongewone ervaringen en een ongebonden leefstijl, zoals verre reizen maken, drugs en mystiek;

- de ‘*Disinhibition*’ schaal, waarin items staan over de behoefte aan ontlasting door bijvoorbeeld wilde feesten, alcohol drinken en gokken;
- en de ‘*Boredom Susceptibility*’ schaal, hierin gaat het om de afkeer van monotonie en routinematigheid.

Het is duidelijk dat deze schalen geschikt zijn voor onderzoek naar consumentenbehoeften. Min of meer analoog aan het ‘*sensation seeking construct*’ zijn er onder andere voor de toeristische sector speciale schalen geconstrueerd. Gebaseerd op de sociologische begrippen ‘romanticism’ en ‘classicism’ is er bijvoorbeeld een index ontwikkeld die aangeeft of mensen emotioneel/rationeel, impulsief/beheerst en risicozoekend/mijdend van karakter zijn. Deze persoonlijkheidsschaal blijkt goed te correleren met het type vakantie dat men boekt (Holbrook en Olney 1995). Verder worden er op Plog’s (1991) continuüm twee extreem verschillende typen vakantiegangers onderscheiden, te weten ‘psychocentrics’ en ‘allocentrics’. Het eerste type wenst een hoge mate van veiligheid en vertrouwde, wat zich o.a. manifesteert in het boeken van volledig verzorgde vakanties, terwijl ‘allocentrics’ de meer avontuurlijke vakantiegangers zijn, die over het algemeen een plezierig en opwindend leven willen leiden. Marktonderzoekers zijn aan de hand van dit soort instrumenten goed in staat om markten te segmenteren, doelgroepen te karakteriseren en consumentengedrag te analyseren. Zo is recentelijk in een Amerikaans onderzoek de markt voor gevaarlijke sporten gesegmenteerd met de begrippen ‘arousal avoidance’ en ‘adventure seeking’ (Shoham, Rose, en Kahle, 1999). Behalve onderzoek met dergelijke schaaltechnieken, wordt het fenomeen dat mensen bewust hoge risico’s nemen om sensaties te ervaren ook via participerende observatie bestudeerd. Een goed voorbeeld is een veldstudie naar ‘skydiving’. In dit onderzoek worden diverse aspecten bestudeerd, zoals de sensatiebeleving, het sportieve element, de rituelen en het prestige van de deelnemers aan deze gevaarlijke sport (Celsi, Rose en Leigh 1993).

Maar waarom zoeken mensen steeds *nieuwe* stimuli? Voor een antwoord op deze vraag wordt vaak verwezen naar de assumpties van Berlyne en Scitovsky: “De behoefte aan stimulerende ervaringen ter verdrijving van verveling vormt een autonome behoefte van het menselijk organisme. Die behoefte wordt optimaal bevredigd door een opvoering van de spanning, als gevolg van toenemende complexiteit in de toestroom van stimuli, afgewisseld door ontspanning, als gevolg van afnemende complexiteit. Bij welke mate van complexiteit dan wel eenvoud een individu een optimum ervaart, is afhankelijk van eerdere ervaringen en intellectuele competentie. Belangrijk in dit verband is de aanname dat bij een geleidelijke gewinning aan een grotere toevoer van stimuli het verassingselement geringer wordt en een grotere behoefte ontstaat aan nieuwe sensaties.” (Knulst 1989:247).

Over het algemeen wordt onderkend dat mensen in bepaalde situaties op zoek zijn naar nieuwe en afwisselende ervaringen, terwijl ze dit in andere situaties bewust vermijden. Enerzijds is er sprake van ‘novelty seeking’, anderzijds van ‘novelty avoiding’. Indien de factoren ‘spanning’ en ‘afwisseling’ (zoeken/vermijden) gecombineerd worden in een 2x2 matrix, dan ontstaat er een kader met vier soorten gedrag waarbij men het volgende beleeft (zie figuur 1): een nieuwe sensatie, bijvoorbeeld jet skiën; een vertrouwde sensatie, zoals mountainbiking; een nieuwe ontspanning, bijvoorbeeld het luxe Thermae 2000; en een vertrouwde ontspanning, zoals de gewone sauna (Goossens, 1996).

	OPSPANNING	
	ZOEKEN	VERMIJDEN
AFWIJSELING	NIJEUWE BEGRIJPE BEGRIJPE	NIJEUWE ONTDEKKINGEN
VERMAAK	VERTROUWDE BEGRIJPE	VERTROUWDE ONTDEKKINGEN

Figuur 1. Motivatiematrix over zoeken naar en vermijden van spanning en afwisseling.

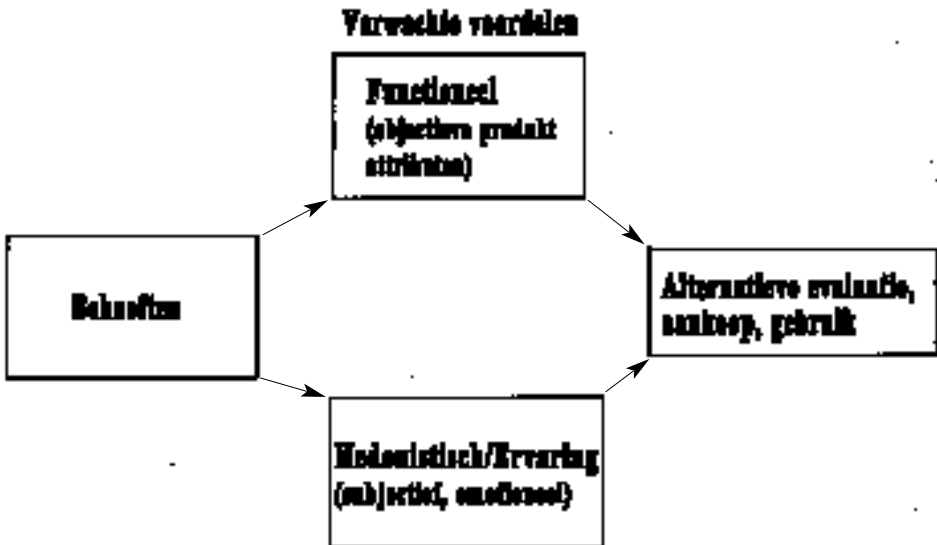
Van Trijp, Hoyer en Inman (1996) stellen dat ‘afwisseling zoeken’ (variety-seeking, exploratory purchase behavior, vicarious exploration, use innovativeness) een vorm is van niet-doelgericht gedrag dat buiten het traditionele ‘informatie-verwerking’ onderzoeksperspectief valt. Dit soort gedragingen worden namelijk eerder door ervarings- en hedonistische motieven verklaard dan door de meer functionele aspecten van consumptie. Baumgartner en Steenkamp (1996) merken dan ook terecht op dat hun conceptualisatie van exploratief consumentenkoopgedrag beperkt is tot de acquisitie van producten en informatie zoeken. Terwijl exploratieve motieven ook betrekking kunnen hebben op *experiential consumption*, zoals bijvoorbeeld esthetische reacties en emotionele responsen tijdens activiteiten met een spelelement en de consumptie van entertainment of kunstobjecten (Holbrook en Hirschman 1982). Een interessant fenomeen bij afwisseling zoeken is dat consumenten, bij het luisteren naar muzieknummers, niet alleen mooie maar ook voor minder geprefereerde nummers kiezen puur omwille van de variatie. Deze keuzestrategie wordt achteraf bekrachtigd door gunstige herinneringen over de afwisselende ervaringen (Ratner, Kahn en Kahneman 1999). Al met al is consumptie in het alledaagse leven geen kwestie meer van overleven, maar veeleer een continue zucht naar genoegens. Hierbij is de consument intrinsiek gemotiveerd in het zoeken van afwisseling en het kiezen van datgene wat de meeste voldoening geeft en het leven plezierig maakt. Bij hedonistisch keuzegedrag streeft men primair naar een aangenaam gevoel: in sommige situaties wordt dit bereikt door spanning en/of afwisseling te zoeken en in andere situaties door dit juist te vermijden. In dit verband kan men kiezen uit vier basale vormen van vermaak (zie figuur 1).

3. KEUZEGEDRAG

De hedonistische consument is op zoek naar amusement, (ont)spanning en plezier, hij wil modieuze producten zo snel mogelijk hebben, koopt ze impulsief en let daarbij niet

in de eerste plaats op de budgettaire mogelijkheden (Ester en Halman 1992). Bij het verzamelen van informatie over hedonistische producten domineren vooral sensorische stimuli. Verder is er vaak sprake van ‘voortdurend informatie zoeken’, waarbij persoonlijke bronnen het meest belangrijk zijn. Bovendien zijn bij het evalueren van hedonistische producten symbolen en verbeelding het meest effectief (Assael 1998). Veel consumenten gebruiken bijvoorbeeld hun visuele verbeelding om goed te anticiperen op de daadwerkelijke consumptie, informatie levendiger te verwerken en betere keuzes te kunnen maken. Denk bijvoorbeeld aan het in gedachten uitbeelden van consumptiepraktijken en het in fantasie opvoeren van een act of rollenspel (*enactive imagery*). Dit ‘doen-alsof’ is een specifieke vorm van mentale verbeelding waarbij de emotionele beleving/ervaring het meest intens is indien de consument de aandacht richt op sensorische informatie (Goossens 2000). Het is dan ook niet verwonderlijk dat behalve objectieve productattributen, zoals de prijs en de functies van een product, ook hedonistische ervaringen evaluatiecriteria zijn die de consument hanteert als een gedragsalternatief gekozen wordt (figuur 2). Omdat de objectieve criteria domineren in het consumentenonderzoek, vormt de subjectieve en hedonistische gedragscomponent een relevante uitbreiding (Engel, Blackwell, en Miniard 1993). Consumenten verlangen naar producten en diensten waar ze zich mee vermaken. De hieraan gekoppelde gevoelens spelen dan ook een essentiële rol bij het keuzegedrag. Veel mensen nemen bijvoorbeeld een beslissing met de vuistregel “Wat-voel-ik-er-voor?”. Dit is vooral het geval bij beslissers met consumptieve motieven in tegenstelling tot instrumentele. Maar ook bij mensen die de neiging hebben om informatie op een visuele en sensorische wijze te verwerken (Pham 1998).

Wat de aard van het keuze-proces betreft wordt er tegenwoordig aandacht besteed aan het ‘affectieve-keuze’ concept. Dit betekent dat men expliciet rekening houdt met het feit dat keuze-processen, bij producten waarvoor de consument een hoge betrokkenheid



Figuur 2. De uiting van motivatie bij aankoop en gebruik van producten: utilitaire en hedonistische voordelen.

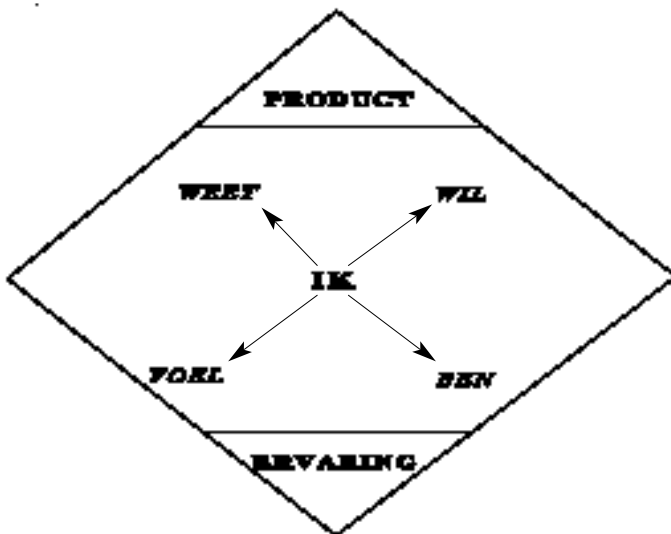
heeft, niet per se verlopen via een uitgebreide (cognitieve) informatieverwerking. Terwijl de traditionele theorieën hier in principe wel van uitgaan. Vooral producten met een hoge expressieve waarde voor persoonlijke- en sociale identiteit worden op een affectieve wijze gekozen (= 'Affective Choice Mode'). Bijvoorbeeld: iemand koopt een sportauto omdat hij/zij hem gewoon mooi vindt en/of iemand kent die hetzelfde merk heeft. Bij dit type productkeuze is er dus sprake van affectieve-betrokkenheid, d.w.z. de consument associeert een marketingstimulus automatisch met een bekend gevoel uit het geheugen dat hoort bij een vergelijkbaar object of een eerdere ervaring. Bij productkeuzes die daarentegen een meer fysiek en objectief nut hebben, is er sprake van een cognitieve-betrokkenheid. De consument verwerkt dan specifieke informatie over een product (zoals de prijs, het benzineverbruik, de ruimte, veiligheid en snelheid van een auto) of denkt uitgebreid na over de inhoud van een boodschap. In dit geval maakt de consument bij zijn keuze gebruik van de 'Information Processing Mode'. Uit een exploratieve studie naar beide typen keuze-processen blijkt dat de zogenaamde 'Affective Choice Mode' met name optreedt als de expressiviteit van het product betrekking heeft op iemands zelf-concept en het innerlijk plezier dat gepaard gaat met de consumptie. Indien de expressiviteit betrekking heeft op het publieke-zelf en het openbaar-vertoon (resp. 'public self' en 'public display') dan is een affectieve keuze minder van toepassing. Verder blijkt dat de 'Affective Choice Mode' en de 'Information Processing Mode' niet tegelijkertijd optreden (Mittal 1994). Maar ze hoeven elkaar nou ook weer niet uit te sluiten. Iemand kan bijvoorbeeld op affectieve wijze voor een sportief automerk kiezen. En vervolgens meer informatie vergaren om het gewenste model te kopen.

4. PRODUCTERVARINGEN

In hun boek "Experience Marketing" stellen O'Sullivan en Sprangler het volgende: *"Most segments of the service industry include an element of an experience within its components. Experiences can be infused into a product, used to enhance a service, or created as an entity into itself"* (1998:2).

Er zijn een paar factoren die een ervaring, of de ervaringscomponent van een product of dienst uniek maken. Een ervaring veronderstelt namelijk: (1) de deelname en betrokkenheid van het individu in de consumptie; (2) het fysiek, mentaal, emotioneel, sociaal, of spiritueel bewogen zijn; (3) een verandering in kennis, vaardigheid, geheugen of emotie teweeggebracht door de deelname; (4) de bewuste perceptie van het samenkomen of doorleven van een activiteit of gebeurtenis; en (5) een inspanning gericht op een psychologische of interne behoefte van de deelnemer. In dit verband kan worden gesteld dat hedonistische consumptie primair gericht is op het ervaren van genoegens ('pleasure'). Nu zijn er twee benaderingen voor het meten van genoegens. Het ene perspectief beschouwt 'pleasure' als een algeheel oordeel. Kenmerkend voor deze benadering is het door Kahneman geïntroduceerde concept '*experience utility*', dat zowel de kwaliteit als de intensiteit van een hedonistische ervaring combineert. Binnen deze onderzoekslijn worden de mechanismen van informatieverwerking en retrospectieve beoordelingen onderzocht. Daarnaast is er een ander perspectief, dat meer de fenomenologie bestudeert van hedonistische ervaringen. Met name in de diensverlenende sector is op deze wijze meer inzicht verkregen over de emoties en het netwerk van sociale relaties die te maken hebben met hedonistische consumptie (LeBel en Dube 1998). Een

dergelijke benadering is interessant omdat het uiteindelijk gaat om de kwaliteit van ervaringen die consumenten met hun producten of merken hebben. Binnen deze context opperde Jerry Olson dat het voor merkentrouw essentieel is om het volgende te weten: “Welke rol speelt je product, je merk in het leven van de consument? Welke emotionele kwaliteiten heeft het, welke ervaringen roept het op?”. Het productervaringsmodel leent zich goed voor dit soort kwesties (Goossens 1999). Een kwalitatieve benadering van dit model bestudeert de fenomenologie van productervaringen. We krijgen zo inzicht in het palet van kennis, behoeften, gevoelens en zelfbeelden die consumenten (zelf) verbinden met productervaringen. Het is bekend dat een product op veel manieren iets voor iemand kan betekenen. Marktonderzoekers gebruiken hiervoor de ‘Betekenis Structuur Analyse’, oftewel de ‘Laddering’-methode. In feite bestudeert men hiermee perceptiestructuren van een consument via instrumentele ‘middel-doelketens’ (zie Asselbergs 1992). Zo kan een automeerk iemands persoonlijkheid weerspiegelen, een kledingstuk kan verwijzen naar goede smaak, een Mars garandeert energie en een glas melk(unie) kan gezondheid betekenen. Maar hoe wordt het een en ander door mensen ervaren? Het productervaringsmodel is expliciet gericht op zaken waarvan iemand zich bewust is. Een ervaring of beleving kan wat dit betreft ruim gedefinieerd worden als de relatie tussen product, lichaam en geest. Via diepte-interviews is het mogelijk om de consument over vier hoofdthema’s te ondervragen (zie figuur 3).



Figuur 3. Het product-ervaringsmodel.

- *Ik Weet vragen:* Allereerst wordt nagegaan wat iemands productkennis is. Dit geeft namelijk een bepaalde mate van betrokkenheid weer. Vervolgens komen de:
- *Ik Wil vragen:* Wat zijn de behoeften en de gewenste ‘benefits’ van het product?. Consumenten verschillen bijvoorbeeld in intrinsieke en extrinsieke motivatie. Bij de eerste vorm gaat het primair om de productervaring. Terwijl extrinsieke motivatie meer gericht is op de consequenties van het productgebruik. Ook ‘benefits’ kunnen

enerzijds betrekking hebben op functionele productattributen en anderzijds op de productervaring of consumptiebeleving.

- *Ik Voel vragen:* Wat betekent de productervaring voor iemand emotioneel gezien?
- *Ik Ben vragen:* Tenslotte is het relevant om te weten of iemand bijvoorbeeld teleurgesteld, tevreden of trots is op zijn product/ervaring? Hierbij is het de bedoeling om na te gaan wat een consument allemaal heeft meegemaakt met een product. Maar ook of dit van betekenis is voor de identiteit, het zelfbeeld en het ego van de persoon.

Antwoorden op deze vier hoofdvragen kunnen de ogen van marketeers openen. Zo kan de ene consument bijvoorbeeld te kennen geven dat hij Mercedes rijdt vanwege economische motieven, zoals de hoge inruilwaarde en de degelijkheid. Bovendien speelt veiligheid een grote rol omdat hij twee jaar geleden ongedeerd een ernstig auto-ongeluk overleefde. Deze man weet dan ook alles van airbags, autogordels, kooiconstructies en dergelijke. Momenteel is hij nog tevreden, maar de kwaliteit van veiligheidsonderdelen zal in de toekomst steeds bepalen welk automerk hij koopt. Een andere consument rijdt bijvoorbeeld ook Mercedes maar dan vanwege de symbolische betekenis en status. Deze persoon voelt zich superieur en trots in z'n auto. De rij-ervaring is voor hem een vorm van vermaak en machtsvertoon. Hij zweert bij z'n merk, koopt stevast een nieuw model en let daarbij vooral op luxe en comfort. Samenvattend kan worden gesteld dat dit soort persoonsgegevens iets zeggen over betrokkenheid, behoeften, productervaringen en merkentrouw. Vandaar dat het nuttig is om er onderzoek naar te verrichten. De consumptiebeleving is in feite een ervaringsbron voor het genereren van bewogen profielschetsen van klanten. Waardoor marketeers hun relatiebeheer en klantenbeleid beter kunnen implementeren.

5. CONCLUSIE

In dit artikel is in vogelvlucht nagegaan hoe men doelgroepen kan karakteriseren, keuzegedrag analyseren en productgebruik evalueren als de consument zich primair laat leiden door genoegens, ervaringen en belevenissen. Binnen dit kader is gesteld dat veel vormen van consumentengedrag een hedonistisch karakter hebben. Gebaseerd op de drijfveren 'spanning' en 'afwisseling' is er een matrix samengesteld die vanuit de producent geredeneerd fundamentele doelgroepen afbakt (figuur 1). Marktonderzoekers kunnen aan de hand van dergelijke schema's de gegevens van verschillende soorten consumenten opslaan en verwerken. Naast de gebruikelijke segmentatievariabelen zoals leeftijd, inkomen en opleiding kan men met schaaltechnieken het *pleasure* motief vaststellen. Bovendien wordt er tegenwoordig aandacht besteed aan hedonistische aspecten van het keuzeproces en productgebruik (figuur 2). Bij het evalueren van hedonistische producten zijn symbolen en verbeelding het meest effectief. Wat dit betreft kiezen consumenten, via sensorische informatie, vaak op hun gevoel. Marketeers moeten hier in hun planning en analyses terdege rekening mee houden. Maar daarmee is de kous nog niet af voor het marktonderzoek. Er is namelijk behoefte aan meer inzicht in productgebruik en de daarbij horende ervaringen (figuur 3). Onderzoek naar deze thematiek verstrekt informatie over het hoe en waarom van consumptie.

LITERATUUR

- Assael, H. (1998). *Consumer Behavior, and marketing action*, 6th Edition, South Western College Publishing, Ohio.
- Asselbergs, W. (1992). Betekenis Structuur Analyse. In: *Recente ontwikkelingen in het Marktonderzoek*. Jaarboek '92-93 van de Nederlandse Vereniging van Marktonderzoekers (99-115), Uitgeverij de Vrieseborch - Haarlem.
- Baumgartner, H. en J-B.E.M. Steenkamp (1996). Exploratory consumer buying behavior: Conceptualization and measurement. *International Journal of Research in Marketing* 13,121-137.
- Celsi, R.L., Rose, R. en T.W. Leigh (1993). An exploration of high-risk leisure consumption through skydiving. *Journal of Consumer Research*, vol. 20 (June), pp.1-23.
- Engel, J.F., Blackwell, R.D. en P.W. Miniard (1993). *Consumer Behavior*, p. 269, Seventh Edition. The Dryden Press, USA.
- Ester, P. en L. Halman (1991). Consumptiestijlen en Waardenorientaties: een typologische reconstructie. In *Recente Ontwikkelingen in het Marktonderzoek*. Jaarboek van de Nederlandse Vereniging van Marktonderzoekers 1990-91, 9 125-147.
- Gabriel, Y. en Lang, T. (1995). *The unmanageable consumer*. Sage publications Ltd, London.
- Holbrook, M.B. en Hirschman, E. C. (1982). The experiential aspects of consumption: Consumer fantasies, feelings, and fun. *Journal of Consumer Research*, vol.9, 132-140.
- Holbrook, M.B. en Olney, T. J. (1995). Romanticism and wanderlust: An effect of personality on consumer preferences. *Psychology & Marketing*, vol. 12, 3, pp. 207-222.
- Goossens, C.F. (1999). De ware reden van trouw. *Tijdschrift voor Marketing*, november, 44-46.
- Goossens, C.F. (1996). Wat beweegt en beleeft de consument?. *Tijdschrift voor Communicatie Wetenschap*, 24, 2, 152-180.
- Goossens, C.F. (2000). Tourism Information and Pleasure Motivation, *Annals of Tourism Research*, Vol. 27 (2) pp.1-21.
- Knulst, W. (1989). Van vaudeville tot video. *Sociale en Culturele Studies*-12, Rijswijk: Sociaal en Cultureel Planbureau, pag. 247.
- Mittal, B. (1994). A study of the concept of 'affective choice mode' for consumer decisions. *Advances in Consumer Research*, vol. 21, pp. 256-263.
- LeBel, J.L. en L. Dube (1998), Understanding Pleasures: Source, Experience, and Remembrances, *Advances in Consumer Research*, vol.25,176-180.
- Mano, H. en Oliver, R.L. (1992). Assessing the Dimensionality and Structure of the Consumption Experience: Evaluation, Feeling, and Satisfaction. *Journal of Consumer Research*, vol.20, 451-466.
- O'Sullivan, E.L. en K.J. Spangler (1998). Experience Marketing. *Strategies for the New Millennium*. Venture Publishing Inc.State College, Pennsylvania.
- Pham, M.T. (1998), Representativeness, Relevance, and the Use of Feelings in Decision Making, *Journal of Consumer Research*, vol. 25, p 144.
- Pine, B.J. en Gilmore, J.H. (1998). Welcome to the Experience Economy, *Harvard Business Review*, July-August, 97-105
- Plog, S.C. (1991). *Leisure travel: making it a growth market... again!* New York: John Wiley.
- Ratner, R.K., B.E. Kahn, en D. Kahneman (1999). Choosing Less-Preferred Experiences for the Sake of Variety. *Journal of Consumer Research*, Vol. 26 June, 1-15.
- Sheth, J., B. Mittal en B.I. Newman. (1999). *Customer behavior*, The Dryden Press, USA.
- Shoham, A., G.M. Rose en L.R. Kahle (1999). Marketing of Risky Sports: From Intention to Action, *Journal of the Academy of Marketing Science*, vol. 26, No.4, 307-321.
- Van Trijp, H.C.M., W.D. Hoyer en J.J. Inman (1996). Why Switch? Product Category-Level Explanations for True Variety-Seeking Behavior, *Journal of Marketing Research*, Vol. 23, 281-292
- Zuckerman, M. (1994). *Behavioral Expressions and Biosocial Bases of Sensation Seeking*. Cambridge University Press, USA.